SET THEORY SET UP

James Holland math.rutgers.edu/~jch258 jch258@math.rutgers.edu

Start Date:2019-07-01End Date:2023-09-30

Contents

	1
1. Logic and Model Theory	4
2. Basic Set Theoretic Concepts	9
3. Well-orders and Ordinals	21
4. Other Well-founded Relations	32
5. Ordinals and Cardinality	40
6. Another Look at Model Theory	59
7. Absoluteness	68
8. The First Inner Models	82
9. Variants of the Axioms	93
10. Exercises for Transitivity	102
Chapter II. Filters, Embeddings, and Extenders*	105
11. Ultrafilters and Logic	105
12. Ultrapowers and Elementary Embeddings	112
13. Introducing Extenders	130
14. External Ultrapowers	152
15. Exercises	165
Chapter III. Relative Definability and Constructibility*	167
16. Relative Constructibility	167
17. Sharp Objects	172
18. Multiple Measures and Comparison	172
19. Extenders and Sequences of Extenders	172
20. Exercises	172
Chapter IV. Descriptive Set Theory*	173
Chapter IV. Descriptive Set Theory* 21. Relevant Topology	173 173
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies	173 173 184
Chapter IV. Descriptive Set Theory* 21. Relevant Topology	173 173 184 198
Chapter IV. Descriptive Set Theory* 21. Relevant Topology . 22. The Boldface Hierarchies . 23. Properties of Sets of Reals . 24. The Lightface Hierarchies .	173 173 184 198 218
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies	173 173 184 198 218 245
Chapter IV.Descriptive Set Theory*21.Relevant Topology22.The Boldface Hierarchies23.Properties of Sets of Reals24.The Lightface Hierarchies25.Properties of the Lightface Hierarchies26.Exercises	173 173 184 198 218 245 269
Chapter IV. Descriptive Set Theory* 21. Relevant Topology . 22. The Boldface Hierarchies . 23. Properties of Sets of Reals . 24. The Lightface Hierarchies . 25. Properties of the Lightface Hierarchies . 26. Exercises . 27. Chapter V.	173 173 184 198 218 245 269 271
Chapter IV. Descriptive Set Theory* 21. Relevant Topology . 22. The Boldface Hierarchies . 23. Properties of Sets of Reals . 24. The Lightface Hierarchies . 25. Properties of the Lightface Hierarchies . 26. Exercises . Chapter V. Games and Determinacy* 27. Fundamentals of Games .	173 173 184 198 218 245 269 271 271
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies 26. Exercises 27. Fundamentals of Games 28. Determinacy and Pointclass Properties	173 173 184 218 245 269 271 271 276
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies 26. Exercises 27. Fundamentals of Games 28. Determinacy and Pointclass Properties 29. Measure, Degrees, and Coding	173 173 184 218 245 269 271 271 276 297
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies 26. Exercises Chapter V. Games and Determinacy* 27. Fundamentals of Games 28. Determinacy and Pointclass Properties 29. Measure, Degrees, and Coding 30. Exercises	173 173 184 198 218 245 269 271 276 297 297
Chapter IV. Descriptive Set Theory* 21. Relevant Topology . 22. The Boldface Hierarchies . 23. Properties of Sets of Reals . 24. The Lightface Hierarchies . 25. Properties of the Lightface Hierarchies . 26. Exercises . Chapter V. Games and Determinacy* 27. Fundamentals of Games . 28. Determinacy and Pointclass Properties . 29. Measure, Degrees, and Coding . 30. Exercises .	173 173 184 218 245 269 271 276 297 297 299
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies 26. Exercises Chapter V. Games and Determinacy* 27. Fundamentals of Games 28. Determinacy and Pointclass Properties 29. Measure, Degrees, and Coding 30. Exercises	173 173 184 218 245 269 271 276 297 297 299 299
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies 26. Exercises Chapter V. Games and Determinacy* 27. Fundamentals of Games 28. Determinacy and Pointclass Properties 29. Measure, Degrees, and Coding 30. Exercises Chapter VI. An Overview of Forcing 31. The Purpose, Motivation, and Terminology of Forcing 32. Examples of Forcing	173 173 184 198 218 245 269 271 276 297 297 297 299 299 310
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies 26. Exercises 27. Fundamentals of Games 28. Determinacy and Pointclass Properties 29. Measure, Degrees, and Coding 30. Exercises Chapter VI. An Overview of Forcing 31. The Purpose, Motivation, and Terminology of Forcing 32. Examples of Forcing 33. More General Topics and Theory	173 173 184 198 218 245 269 271 276 297 297 297 299 310 322
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies 26. Exercises 27. Fundamentals of Games 28. Determinacy and Pointclass Properties 29. Measure, Degrees, and Coding 30. Exercises Chapter VI. An Overview of Forcing 31. The Purpose, Motivation, and Terminology of Forcing 32. Examples of Forcing 33. More General Topics and Theory 34. Iterated Forcing	173 173 184 198 218 245 269 271 276 297 297 297 299 310 322 333
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies 26. Exercises 27. Fundamentals of Games 28. Determinacy and Pointclass Properties 29. Measure, Degrees, and Coding 30. Exercises Chapter VI. An Overview of Forcing 31. The Purpose, Motivation, and Terminology of Forcing 32. Examples of Forcing 33. More General Topics and Theory 34. Iterated Forcing 35. Forcing and Elementary Embeddings	173 173 184 198 218 245 269 271 271 276 297 297 297 299 310 322 333 371
Chapter IV. Descriptive Set Theory* 21. Relevant Topology 22. The Boldface Hierarchies 23. Properties of Sets of Reals 24. The Lightface Hierarchies 25. Properties of the Lightface Hierarchies 26. Exercises Chapter V. Games and Determinacy* 27. Fundamentals of Games 28. Determinacy and Pointclass Properties 29. Measure, Degrees, and Coding 30. Exercises Chapter VI. An Overview of Forcing 31. The Purpose, Motivation, and Terminology of Forcing 32. Examples of Forcing 33. More General Topics and Theory 34. Iterated Forcing 35. Forcing and Elementary Embeddings 36. Exercises	173 173 184 198 218 245 269 271 276 297 297 297 297 299 310 322 337 397

Appen	dix A. 🛛	Computability	Theory			401
1.	Primitive	e Recursion		 	 	401

2. 3. 4.	Computablility The Arithmetical Hierarchy Reducibility and Logic	413 423 432
Appen	dix B. Recursion Theory*	445
1.	The Hyperarithmetical Hierarchy	445
2.	Ordinals and Computability	450
3.	The Length of the Hyperarithmetical Hierarchy	459
4.	Identifying the Hyperarithmetical Hierarchy with Δ_1^1	459
5.	Admissible Sets and Kripke-Platek Set Theory	459
Appen	dix C. Gory Details of the Forcing Relation	461
1.	Gory Detail of the Forcing Relation and Its Definability	461
2.	ZFC in the Generic Extension	466

The Axioms of **ZFC**

1. (Extensionality, Ext) two sets are equal whenever they have the same members:

 $\forall x \; \forall y \; (x = y \leftrightarrow \forall v \; (v \in x \leftrightarrow v \in y)).$

- 2. (Empty set) there is a set \emptyset with no members: $\exists z \forall x (x \notin z)$.
- 3. (Comprehension, Comp) for each x, and for each $FOL(\in)$ -formula $\varphi(v, \vec{w}), \{v \in x : \varphi(v, \vec{w})\}$ exists:

 $\forall w_0 \cdots \forall w_n \ \forall x \ \exists z \ \forall v \ (v \in z \leftrightarrow v \in x \land \varphi(v, \vec{w})).$

- 4. (Pairing, Pair) for any two sets x and y, the pair $\{x, y\}$ exists: $\forall x \forall y \exists z \forall v (v \in z \leftrightarrow (v = x \lor v = y))$.
- 5. (Union, Union) for any family of sets F, there is a set containing the elements of all of those sets:

 $\forall F \exists U \ \forall v \ (v \in U \leftrightarrow \exists x (x \in F \land v \in x)).$

6. (Foundation, Found) for each x, there is a \in -minimal element of x, meaning a member $y \in x$ with no $z \in y$ being in x:

$$\forall x \exists y (y \in x \land \forall z (z \in y \to z \notin x)).$$

- 7. (Infinity, Inf) an infinite set exists: $\exists N(\emptyset \in N \land \forall x (x \in N \to x \cup \{x\} \in N))$.
- 8. (Replacement, Rep) the image of a function over a set is a set: for each $FOL(\in)$ -formula φ ,

 $\forall w_0 \cdots \forall w_n \forall D \left(\forall x (x \in D \to \exists ! y \varphi(x, y, \vec{w})) \to \exists R(y \in R \leftrightarrow \exists x (x \in D \land \varphi(x, y, \vec{w}))) \right).$

- 9. (Powerset, P) for each x, $\mathcal{O}(x)$ exists: $\forall x \exists P \forall v \ (v \in P \leftrightarrow \forall y \ (y \in v \rightarrow y \in x))$.
- 10. (Choice, AC) for any family of non-empty family of non-empty, disjoint sets F, there is a set C which has chosen one element from each $z \in F$:

$$\forall F \ (\emptyset \notin F \land \forall x, y \in F \ (x \cap y = \emptyset) \to \exists C \ \forall x \in F \ \exists ! y \ (y \in x \cap C).$$

Variant Axioms and Axiom Systems –

- i. (Weak pairing, wPair) for any two x, y, there is a z with $x, y \in z$.
- ii. (Weak union, wUnion) for any family F, there is a z with $\forall x \in F \ (x \subseteq z)$.
- iii. (Weak replacement, wRep) the image of a function over a set is contained in a set.
- iv. (Weak powerset, wP) for any x, there is a set containing all subsets of x.
- v. (Collection, Coll) there is a range for a relation with over a given domain: for each $FOL(\in)$ -formula φ ,

 $\forall w_0 \cdots \forall w_n \forall D \left(\forall x \in D \exists y \ \varphi(x, y, \vec{w}) \right) \rightarrow \exists R \ \forall x \in D \ \exists y \in R \ \varphi(x, y, \vec{w}) \right).$

- vi. $(\Sigma_n$ -Comprehension, Σ_n -Comp) for each x, and for each Σ_n -formula $\varphi(v, \vec{w}), \{v \in x : \varphi(v, \vec{w})\}$ exists.
- vii. (Σ_n -Collection, Σ_n -Coll) Coll holds for Σ_n -formulas.
- viii. (Dependent choice, DC) for $R \subseteq X \times X$, if $\forall x \in X \exists y \in X (x R y)$ then there is a sequence $\langle x_n : n \in \omega \rangle$ such that $x_n R x_{n+1}$ for all $n \in \omega$.
- ix. For every $x, y, x \times y$ exists.

- Set Theories

- BST consists of (1)–(6) plus (ix).
- wZF consists of (1), (2), (3), (6), (7), and (i)-(iv). wZFC also adds (10).
- ZF⁻ consists of (1)–(8) plus (v). ZFC⁻ also adds (10).
- $ZF = ZF^- + P$ consists of (1)–(9). ZFC also adds (10).
- $KP = BST Comp + \Sigma_0$ -Coll, i.e. (1), (2), (4), (5), (6), Σ_0 -Comprehension, and Σ_0 -Collection.

Chapter I. Transitivity

Set theory is on one hand the study of collections and their use in mathematics. When mathematicians attempt to make something precise, they do so using collections: functions are viewed as collections of pairs, a line or circle is a collection of points, real numbers are certain collections of rational numbers, and so on. Set theory then serves as a foundational role in mathematics in that questions like "what kinds of things exist?" and "are there any counter-examples to this idea?" become questions about what *sets* exist.

So we use set theory in an attempt to provide a foundation for valid mathematical reasoning, and in doing so, we are unsurprisingly led to ask what is and isn't valid by way of asking what sets do or don't exist. This evolves into the study of statements whose validity is impossible to determine—statements called *independent* of our mathematical principles or axioms. Set theory then becomes the study of *why* these statements are independent.

On the other hand, set theory is the result of a historical process of discovery and definition. The standard axioms we have set as the ultimate foundation of mathematical thought, ZFC, have been formulated and modified by *people* over the last one and half centuries. As such, these principles are not some divine work but instead (supposedly) intuitively clear principles and ideas that hold when thinking about collections. As one studies these principles more and more, the theory of ZFC seems more and more canonical despite its cultured past. So although the precise form the axioms take is not "universal", it's difficult to find *any* intuitively clear principles that aren't already proven by ZFC.

I state these two perspectives because it's important to realize that what principles hold of the real world are not entirely obvious. Indeed, a large chunk of this document is dedicated to why there are lots of principles we simply have no way of knowing the truth about one way or the other if we take ZFC as our only starting point. Nevertheless, we are not doomed to wallow in the weaknesses of ZFC. There are arguments to be made for other, less obvious principles that can have tremendously deep consequences and explanatory power. So we may also act as scientists, having to use our intuitions, imaginations, and available evidence to think about what lies beyond our limited knowledge.

Before getting too deep in the study, I want to give some notation. The only (non-logical) symbol really used in set theory is the membership relation: " $x \in y$ " symbolizes "x is a member of the collection y", or more succinctly, "x is in y". For example, \mathbb{N} is typically used to denote the collection of all natural numbers. So $1 \in \mathbb{N}$ and $4 \in \mathbb{N}$, for example. We can also consider smaller sets. If we can list out all the elements of a set, we may denote the set by enclosing the members in braces: the set of 1, 4, and 8 is $\{1, 4, 8\}$. Note that in general, $x \neq \{x\}$. To see why this is true, consider a more physical analogy: if we take x to be a marble, $\{x\}$ is a bag with one marble in it, whereas x is just the marble itself. The concept is flexible enough to allow us to collect together many things at once, and thinking about statements like the above is where set theory starts.

In practice, one cannot go too deep in set theory without understanding transitive sets. In general, a relation R is called "transitive" iff for all relevant x, y, and z, if x R y and y R z, then x R z. Classic examples of this include equality, and the ordering on the reals <, among many others. In the context of set theory, a collection is called transitive iff the membership relation \in is a transitive relation on it: X is transitive iff $z \in y \in X$ implies $z \in X$. On its own, this property seems unmotivated or perhaps useless, but it plays one of the most fundamental roles in set theory. To hint at an important connection, consider the totality of all sets, denoted here by V. So V is composed just of sets: any member of V is a set and all of their members are sets too, and all of their members, and so on. This collection V is therefore transitive. Transitive collections are then the first candidates for models of set theory: they are an attempt to approximate V.

Transitive sets can also approximate V in truth. In particular, there are lots of statements "absolute" between transitive structures in the sense that they all agree on whether they're true or false. So we will be interested in this kind of absoluteness, as this tells us information about V. In another sense, transitive sets interpret membership correctly. This makes independence results around transitive sets important because the "reason" the statement is independent isn't

the result of misinterpretation. Misinterpretations are relatively easy to come by; formally, ZFC is regarded as a bunch of formulas, and it's impossible for such formulas to uniquely determine what exactly a set is and what properties they have. So with transitive sets, independence results aren't merely due to these weird misinterpretations but instead some deeper facts about sets.

§0A. Philosophy

We begin with the philosophically basic notion of a collection: we take it as immediate that things exist, and that we can consider collections of things as abstract objects. It is in this sense that we mean that these collections "exist", and hence we can take collections of collections, and so forth. We in the real world can then reason about these collections and their properties. The simplest examples of this kind of reasoning comes from Venn diagrams, like the one pictured below.



0A.1. Figure: Example of a Venn-diagram

The first concept we then define is the collection of all sets, the actual set theoretic universe. More precisely, we begin with the sets that are *hereditarily* sets, meaning for each x, every member of x is a set, and all of their members are too, and so on.

— 0A·2. Definition ———

The *universe of sets* is the structure^{*i*} $\mathbf{V} = \langle V, \in \rangle$, where V consists of all (and only) sets, and \in denotes membership.

What exactly should this universe look like? Intuitively, we start with a set with no elements: the empty set, \emptyset . Then, we can take the set of just this object { \emptyset }. Now we have two objects, and we can take collections of these: \emptyset , { \emptyset }, {{ \emptyset }}, and { \emptyset , { \emptyset }}. And we can continue this iterative formation of sets. This iterative conception is at the heart of modern set theory, and I hope to further motivate why it is true through the chapter. But first, we must acknowledge how we will do this: indeed, the question of our base level axioms come into question.

We will go through the chapter introducing principles or *axioms* which are generally seen as statements true of V beyond any doubt. Now we are interested not just in what is true of V, but also what we can *prove* about V from these axioms. In particular, it is not immediately obvious whether certain statements are true or false. If we are to argue that we *cannot* prove nor disprove them, then we need to have agreed upon, intuitively true axioms about V. It is, of course, an open question whether our list of axioms exhausts all intuitively true statements about V. But given the power of the axioms we present, it is difficult to find simple principles that are intuitively obvious but independent of the other axioms.

How we state these axioms is important if we are to have a precise notion of proof or the lack thereof. Usually, mathematical statements are written in a codified version of natural language, where notation replaces the normal words of English, Russian, or whatever other language. This is no different for us, but we rely on notation even more to ensure that we can carry out everything in a formal system using just basic reasoning about finite objects, namely formulas. This then begs the question, what formal system should we use?

ⁱStructures are regarded as ordered lists with the first entry denoting the domain of discourse or "universe", and the other entries denoting the relevant relations and functions we're considering over that universe. Angle brackets are generally used to denote ordered lists in set theory.

Note that there is an important distinction in logic between the reasoning we use in the real world and the reasoning a certain subject allows. For example, we in the real world have the ability to conclude a = c from a = b and b = c. However if we consider only the sentential/propositional connectives there—"and" and "implies"—we cannot make the same conclusions. From the perspective of propositional logic, "a = b and a = c" is no different from "A and B" where A and B are two completely unrelated propositions: the logic no longer considers the meaning of equality, only the meaning of these sentential connectives. To distinguish the two logics, the reasoning we use in the real world is called the *meta-theory* whereas the reasoning a certain subject (like propositional logic) allows is called the *logic system*.ⁱⁱ The reasoning of a logic system is entirely formal, following from strings of symbols, but with the proper setup, it can characterize a portion of the meta-theory, like the simple example of propositional logic.

The more complicated example of first-order logic is where we will state our axioms. This is both because it has the expressive power needed to present the axioms, and because there are a great deal of important results related to it, as we will see in the first section. To give a more cultural reason, first-order logic is not the only logic system one can use to study mathematics, but most other logic systems can be reformulated in terms of set theory with arguments that take place in first-order logic. In fact, second order logic is sometimes called "set theory in sheep's clothing". Generally, first-order logic is the framework in which the results of set theory are given, and results *about* set theory are generally about it in this framework.

Now a priori, there's no guarantee that the world behaves in accordance with the axioms of ZFC (the standard axioms of set theory). The axioms are taken to be intuitively obvious, but in fact, we would need to reject them as part of the meta-theory if it turned out that this system were inconsistent. Furthermore, constructions allowed by ZFC like \mathbb{R} and \mathbb{N} can be called into question if we reject certain axioms like the existence of \mathbb{N} . How then do we regard such statements as " $|\mathbb{R}| > |\mathbb{N}|$ "? Is this a meta-theoretic fact, or is this better regarded as a formula of first-order logic following from certain axioms? There are a few ways to address these concerns. Two major positions are presented here.

One stance is a purely formalist one. This view will neglect to say anything substantial about the meta-theory, taking only the most basic algorithmic reasoning needed for the study of logic for granted. The formalist approach then doesn't connect the reality of the meta-theory with results of axioms like ZFC, and it in some sense ignores whether the theories we study are important at all. No commitments are made for whether the natural numbers \mathbb{N} exist or whether a statement like $|\mathbb{R}| > |\mathbb{N}|$ has any meaning in the meta-theory. But the formalist will deny that the sequence of symbols " $|\mathbb{R}| > |\mathbb{N}|$ " has any actual meaning. Instead, the formalist will view the statements about ' \mathbb{N} ' or ' \mathbb{R} ', for example, as merely symbols algorithmically changed from other symbols collectively called ZFC.ⁱⁱⁱ So the results of theories in the logic system are seen purely as symbolic manipulation with no connection to the meta-theory. At best, a formalist will say the symbols in the logic system can be translated into arguments in the meta-theory where they should have been given in the first place. At worst, a formalist will say the symbols are devoid of content.

Another stance is a platonist one. This view will hold that the results of axioms like ZFC in the logic system do characterize a fragment of the meta-theory—in particular, V. Not only is there a standard meaning of the statement " $|\mathbb{R}| > |\mathbb{N}|$ ", but there is an actual fact of the matter, and we can learn such facts through study of theories in things like first-order logic. By and large, a platonist stance is held by mathematicians that want to claim that their conclusions are actually true and not merely derived from playing with symbols. Indeed most of mathematics is not done through symbolic algorithms like truth tables but instead through intuitions and clever constructions. That said, a platonist stance isn't strictly necessary, since often meta-theoretic arguments can be reformulated as symbolic ones and vice-versa. In this way the two stances are not incompatible.

This work will take more of a platonist stance. More precisely, ZFC is held as a collection of true statements about V, and this is used to reason about ZFC as presented symbolically. Later it will be useful when thinking about independence results to adopt a slightly different outlook where there might be *expansions* of this universe (such expansions would be incompatible with Definition $0 \land 2$). The results we give can be translated into any of these frameworks pretty

ⁱⁱElsewhere in the literature, you might see other words like the *object language, proof system*, or perhaps just *logic* to refer to logic system or how it's written.

ⁱⁱⁱor whatever other foundation they are studied in.

easily, so the view adopted is partially for pragmatic and pedagogical reasons.

Section 1. Logic and Model Theory

We begin with an overview of symbolic logic, because most of the rest of this document will assume some familiarity with the basics of first-order logic, particularly the meaning of \vdash and \vDash , as well as the associated concepts of formulas, sentences, theories, and models or structures. Rather than spend an inordinate amount of time giving the fine details of first-order logic, the reader is referred to any introductory logic text, like [7]. So instead an overview is given with most details omitted. Many students of set theory will already have experience with these concepts. The reader intending to skip this section should just be aware of two things: for a signature or vocabulary σ , the language of first-order logic is written FOL(σ); and if we have a formula with parameters, we say it is a FOLp(σ) or FOLp-formula.

There are two parts to introduce first-order logic as with almost any logic system. Firstly there is a syntactic component ruling what can be said. Secondly there is a semantic component that gives meaning to these formulas. This separation is similar to the separation between the grammar and spelling of English, and the meaning of sentences. There are a number of steps in this introduction. Continuing the natural language analogy, we need to

- 1. determine the alphabet we're using;
- 2. determine how to spell words with this alphabet;
- 3. determine how to "reason" with these words;
- 4. determine the meaning of these words; and
- 5. connect spelling with meaning.

§1 A. The alphabet and its formulas

To start, the alphabet of first-order logic is better regarded as a collection of alphabets that are all variations on a simpler alphabet. In particular, they all share the so called *logical symbols* given below that allow us to make basic formulas that are statements of equality and inequality: " $x \neq y$ ", " $v_3 = v_{10}$ ". From these basic statements—so called *atomic formulas*—we can build up larger formulas using simple rules. For things already determined to be formulas, we can connect them using formula connectives, or quantify them over some variable. So for φ and ψ already formulas, " $(\varphi \land \psi)$ ", " $\neg \varphi$ ", " $\exists x \varphi$ ", and so forth are all formulas too. Although these symbols have no actual meaning introduced, it's useful to have an idea for what they are supposed to represent.

Symbol	'∧'	' ∨'	·_'	·́→'	·↔'	'∃'	'∀'
Usual Meaning	"and"	"or"	"not"	"implies"	"iff" (i.e. "equivalent to")	"there exists"	"for all"
Symbol	'='	<i>'x</i> '	, 'y', 'v ₀	', ' v_1 ', etc.	`(', ')', `,', etc	с.	
Usual Meaning	equality	,	variables		various grammatical marks		

This allows us to build formulas like " $\exists x \exists y (\neg x = y)$ " and " $(x = x \land \neg x = x)$ ". We cannot, however, make ordinary mathematical statements like "x = y + z" or " $\exists z (z \cdot x^{-1} \le z + y)$ " yet. To make such statements we need a bigger alphabet. In particular, we have the concept of a signature or vocabulary to expand the logical symbols with non-logical symbols like '+' or ' \le ' above.

— 1A•1. Definition

A *signature* is a collection of symbols that are divided into constant symbols, relation symbols, and function symbols with the corresponding number of arguments.

The first-order language associated with a signature σ is denoted FOL(σ).

For example, those familiar with some algebra will know that rings and fields generally use a signature of just function symbols: $\{+, \cdot, 0, 1\}$. This expands the signature usually used with groups: $\{\cdot, 1\}$. Partial orders and graphs will use

LOGIC AND MODEL THEORY

only relation symbols for the order and the edges. Most importantly for us, set theory uses the signature with only one element $\{\in\}^{iv}$

The rules for forming formulas change very little from when there were just logical symbols. Essentially, one just needs to respect the number of arguments for the relation and function symbols. So if 'f' is a function symbol with two arguments, you can't write "f(x, y, z)" or "f(t)". The same applies to relation symbols. Building terms $t_1, ..., t_n$ by composing function symbols and variables, we can let relations holding between terms—i.e. strings of the form " $R(t_1, ..., t_n)$ " or " $t_1 = t_2$ "—be the basic building blocks of formulas. Then we can build the rest of the formulas in the same way as before with connectives and quantifiers.

Now we remark that often formulas are written in short-hand, meaning we don't include so many parantheses, and introduce symbols which are defined in the original signature. For example, " $x \subseteq y$ " can be defined by

$$x \subseteq y$$
 iff $\forall z \ (z \in x \to z \in y)$.

Such defined notions affect nothing since they can be replaced by their defining formulas. In general, we're satisfied giving instructions for how to construct a formula as opposed to giving it explicitly. The same principle also holds for proofs. For an explicit example of this, the quantifier ' \exists !' is generally used to mean "there exists a unique". We use " $\exists !x \ \varphi(x)$ " merely as shorthand for " $\exists x \ \forall y \ (\varphi(y) \leftrightarrow x = y)$ ".

§1B. The proofs of formulas

With the notion of formula comes the notion of proof: a means of manipulating formulas. The concept of proof should be fairly familiar at this point. Note that in setting up the proof system, we should be trying to emulate valid reasoning in the meta-theory, though there is no association of meaning with formulas yet. A priori, there's no reason we couldn't allow ourselves to conclude " $\varphi \land \psi$ " from " $\varphi \lor \psi$ "—"both" from "at least one". So there is some careful setup required in what precisely is allowed—so called *logical axioms*. The following is an informal definition, omitting what precisely a logical axiom is.^v

— 1B•1. Definition

Let T be a collection of formulas, and φ a formula. T proves φ , or $T \vdash \varphi$, iff there is a sequence of formulas where every member

1. is a given assumption, i.e. a member of T; or

2. is a logical axiom, e.g. x = x or $(\neg \neg \psi) \leftrightarrow \psi$; or

3. follows from previous ones by given rules of inference, e.g. ψ follows from φ and $\varphi \rightarrow \psi$.

For example, one can prove " $\forall x \forall y (x + y = y + x)$ " from the axioms of peano arithmetic, PA, which are then regarded as given assumptions in the above. A collection of formulas is generally called a *theory*. Note that the statement $T \vdash \varphi$ for "there is a proof of φ from the formulas T" is a meta-theoretic one about the logic system.

And as with formulas, it's rare to give proofs as just a sequence of formulas, because they are hard to read and comprehend. Even when annotated, it's hard to see at a glance that the formulas obey the definition. For example, consider the following tedious proof of the obvious fact that $\varphi \rightarrow \varphi$ for any formula φ .

1. $(\varphi \to (\varphi \to \varphi)) \to$	(from axiom scheme $(\varphi \rightarrow \psi) \rightarrow ((\varphi \rightarrow (\psi \rightarrow \chi)) \rightarrow (\varphi \rightarrow \chi))$
$((\varphi \to ((\varphi \to \varphi) \to \varphi)) \to (\varphi \to \varphi))$	where ψ is $(\varphi \rightarrow \varphi)$ and χ is $\varphi)$
2. $\varphi \to (\varphi \to \varphi)$	(from axiom scheme $\varphi \to (\psi \to \varphi)$ where ψ is φ)
3. $(\varphi \to ((\varphi \to \varphi) \to \varphi)) \to (\varphi \to \varphi)$	(1, 2 and Modus Ponens)
4. $\varphi \to ((\varphi \to \varphi) \to \varphi)$	(from axiom scheme $\varphi \to (\psi \to \varphi)$ where ψ is $\varphi \to \varphi$)
5. $\varphi \rightarrow \varphi$	(3, 4 and Modus Ponens)

So often proofs are given as instructions for creating a proof rather than just a sequence of formulas. This perspective

^{iv}Arguably set theory uses many more symbols, e.g. ' \subseteq ', ' \emptyset ', and so forth. But these can be better regarded as short-hand for statements which use only ' \in ' and '='.

^vMany texts make do with a list of around fifteen axioms, axiom schemes, and rules of inference. So it should be clear why the exact details are omitted here.

is useful when arguing in the meta-theory about proofs in the logic system.

This concludes the syntactic portion of first-order logic, and now we will look towards interpreting these formulas, since thus far formulas are regarded merely as a bunch of markings on paper formed in a certain way.

§1C. The semantics of formulas

Now we will move on to the semantics of first-order logic, looking at how to interpret these formulas and reason from them in the meta-theory. In some sense the goal is to answer "what makes a formula true?". Answering this requires first fixing a context we ask the question in, and then we build up a notion of truth in just the same way we've built up formulas. The explanations given here relate somewhat back to the real world insofar as they assume that structures, relations, functions, and so forth exist. So we might as well assume that we're working in **V**, where we know that these things exist.

Firstly, we have the notion of a structure. This is in some sense where we evaluate truth. For example, when we ask whether the group operation \cdot is commutative, we answer relative to some particular group. The question can be asked of any group, but the answer depends on the group we evaluate in. For a less mathematical example, "how many citizens are there?" can be asked of any particular nation, but the answer depends on the nation. More generally, we can ask questions in a fixed signature, but the answer depends on the structure.

- 1C•1. Definition

Let σ be a signature. A FOL(σ)-structure or model is a pair $\mathbf{M} = \langle M, \varsigma \rangle$ where M is the universe of **M**, and

- 1. For every *n*-place relation symbol R in σ , there is one $R^{\mathsf{M}} \in \varsigma$ with R^{M} a relation on tuples of M; and
- 2. For every *n*-place function symbol f in σ , there is one $f^{\mathsf{M}} \in \varsigma$ with f^{M} a function from tuples of M to M.

Intuitively, ς tells us how the model interprets the symbols of the signature σ , and the members of ς are the *interpretations* of the members of σ . For example, the signature $\sigma = \{\preccurlyeq\}$ has models which are really just any set equipped with a binary relation. For example $\langle \mathbb{N}, < \rangle$ is a $\{\preccurlyeq\}$ -model, and so is any graph $\langle G, E \rangle$ where E is the edge relation of the graph. Under this definition, for any signature σ , any σ -model is also an \emptyset -model where there are no non-logical symbols, and the only statements are about equality.^{vi} In fact for any σ -model is also a δ -model for any $\delta \subseteq \sigma$.

The interpretation of the signature essentially determines truth of the atomic formulas: the structure $(\mathbb{N}, <)$ says that "3 < 2" is false and that "2 < 5" is true. Hence "3 < 2 \lor 2 < 5" is true while "3 < 2 \land 2 < 5" is false for $(\mathbb{N}, <)$.

By following the construction of any given formula, this association of a symbol in σ with the interpretation in ς presents how to tell whether any given formula is true or false in a given structure in the natural way we read formulas. Note that there will always be a fact of the matter in any given structure of whether a formula is true or false in it, even if it isn't possible to determine practically. Explicitly, we have the following definition.

1 C • 2 Definition						
Let σ be a signature with R in σ a relation system.	ymbol. Let φ an	d ψ be FOL(σ)-sentences; and let M a FOL(σ)-model				
with various $m_i \in M$. Write						
$M \vDash "R(m_1, \cdots, m_n)"$	if and only if	$R^{M}(m_1,\cdots,m_n)$ holds,				
$\mathbf{M} \vDash "m_1 = m_2"$	if and only if	$m_1 = m_2,$				
$M\vDash ``\varphi \wedge \psi"$	if and only if	$\mathbf{M} \vDash \varphi \text{ and } \mathbf{M} \vDash \psi,$				
$M\vDash ``\neg \varphi ``$	if and only if	$M\not\models\varphi,$				
$\mathbf{M} \models ``\forall x \ \varphi(x)"$	if and only if	$\mathbf{M} \models ``\varphi(m)$ '' for every $m \in M$,				
$M \vDash ``\exists x \ \varphi(x) "$	if and only if	$\mathbf{M} \models "\varphi(m)$ " for some $m \in M$.				

Implicit in this is the ability to interpret terms in the signature, and this is done exactly as one would expect. For example, the interpretation of " $f(m_1, g(m_2))$ " is just $f^{\mathsf{M}}(m_1, g^{\mathsf{M}}(m_2))$. For a more concrete example, " $3 + (5 \cdot 2)$ " has an interpretation of 13 in the structure of arithmetic $\mathsf{N} = \langle \mathbb{N}, 0, 1, +, \cdot \rangle$.

^{vi}We can still say meaningful things in this language, but mostly this is about the number of things: $\exists x \forall y (x = y)$ will require that there is only one element, for example. Some systems also drop the need for equality, in which case there are no formulas without relation symbols.

LOGIC AND MODEL THEORY

Free variables are left uninterpreted, so this is why we only deal with sentences. Also note that we are mixing formal symbols and non-formal ones, leaving the parameters implicit when needed. It's important to realize that parameters can only be used when we've fixed a particular model. Parameters—like $m_1, m_2 \in M$ in the above—are not symbols in the language, and so cannot be referenced in general. In some sense, parameters are used here merely to build up a notion of truth.

With this concept (or at least the use of (\models)) firmly in place, notation will be slightly abused in the following ways.

```
Let \sigma be a signature, and let \varphi and \psi be formulas, and T a theory all in FOL(\sigma). Let M be a FOL(\sigma)-structure. Write

\mathbf{M} \models T if and only if \mathbf{M} \models \theta for every sentence \theta \in T.

\varphi \models \psi if and only if every \sigma-model M with \mathbf{M} \models \varphi also has \mathbf{M} \models \psi.

T \models \psi if and only if every \sigma-model M with \mathbf{M} \models T also has \mathbf{M} \models \psi.
```

For example, " $(\varphi \land \psi)$ " $\vDash \varphi$, since any model $\mathbf{M} \vDash$ " $(\varphi \land \psi)$ " has $\mathbf{M} \vDash \varphi$ by Definition 1 C • 2.

These definitions comprise all the semantics of first-order logic, and they all take place all in the meta-theory, meaning that $\varphi \models \psi$ if there is a meta-theoretic argument about models of φ . Alternatively, it might be the case that all models of φ also model ψ merely by chance with no intelligible reason behind it. So far this situation hasn't been ruled out. It is up to the next subsection to dispel this possibility.

§1 D. Connecting syntax and semantics

We now have the basic setup for working in mathematics. On the one hand, we can symbolically manipulate our way to various formulas, and on the other, we can argue in the meta-theory about whether certain structures satisfy a given formula. The central question, however, is whether there is any connection between the two, that is, whether " $T \vdash \varphi$ " and " $T \models \varphi$ " have any relationship.

Clearly, we should have set up our proof system to be sound, that is to say that if $T \vdash \varphi$ then $T \models \varphi$. This way we aren't making any "mistakes" in our symbolic manipulations. Proving that any given proof system is in fact sound can be done fairly easily through meta-theoretic arguments about structures. Mostly this amounts to checking that each logical axiom and rule of inference holds in every model.

Quite a striking result in the study of first-order logic is the completeness theorem which says that the converse also holds with our notion of proof.

– 1 D•1. Theorem (Completeness)

Let σ be a signature, and let T be a theory, and φ a formula in FOL(σ). Therefore $T \vDash \varphi$ implies $T \vdash \varphi$.

Proof .:.

Suppose $T \vDash \varphi$, but $T \nvDash \varphi$. This means $T \cup \{``\neg \varphi''\}$ is consistent (assuming the proof system is good), meaning that it doesn't prove a contradiction " $\varphi \land \neg \varphi''$. Note that $T \cup \{``\neg \varphi''\}$ cannot have a model, however, as this model would satisfy T and " $\neg \varphi$ ", contradicting that $T \vDash \varphi$. To get our contradiction, we will construct a model of $T \cup \{``\neg \varphi''\}$ out of syntax.

Call a FOL(σ)-theory *T* complete iff for every FOL(σ)-sentence φ , either φ is in *T*, or " $\neg \varphi$ " is in *T*. By wellordering the FOL(σ)-sentences, we can successively decide whether to put a given sentence in an expansion T_0 or not according to whether the resulting expansion of *T* would be consistent (i.e. put it in if it is, if it's not, then leave it out). Hence we can expand $T \cup \{ \neg \varphi^n \}$ to a theory T_0 which is consistent and complete: just the result of this process.

Now by ordering T_0 and proceeding through each formula one-by-one (i.e. well-ordering T_0), for each existental statement φ , being " $\exists x \ \psi(x)$ " in T_0 , associate a unique constant c_{φ} , and add in the statement " $\psi(c_{\varphi})$ " to the new theory T_1 in the expanded signature σ_1 . Also expand to make sure T_1 is still consistent and complete now

in FOL(σ_1). Repeating this process infinitely many times to take the closure under this propety, we end up with a complete, consistent (assuming the proof system is good) theory T_{ω} in an expanded signature σ_{ω} such that if " $\exists x \ \psi(x)$ " is in T_{ω} , then " $\psi(c)$ " is in T_{ω} for some constant symbol c of σ_{ω} .

Now we construct a model of T_{ω} , which is then still a model of T (by forgetting about the constants of σ_{ω} , we end up with a FOL(σ)-model rather than a FOL(σ_{ω})-model). Firstly, for c a constant symbol of σ_{ω} , consider the equivalence class [c] consisting of all the other constants d such that $T_{\omega} \vdash "d = c"$. This is an equivalence class as T_{ω} is complete (assuming we've set up the proof system correctly). Now consider the structure **M** with universe M being the set of these equivalence classes, and with function interpretations given by

$$f^{\mathsf{M}}([d_1], \cdots, [d_n]) = [d_0]$$
 iff $T_{\omega} \vdash "f(d_1, \cdots, d_n) = d_0"$,

and similarly for relations (again, assuming the proof system is good, this is well-defined). The resulting structure then satisfies $\mathbf{M} \models T_{\omega}$, and so we have a model of $T \cup \{``\neg \varphi''\}$, and so $T \not\models \varphi$.

This identifies the "accidental truth" of being true by chance in all models with the "justified truth" of proof. This also allows us to make conclusions from valid arguments in the meta-theory about models, and conclude that there are syntactic proofs of these results. Most important for our purposes is the fact that if $T \not\vdash \varphi$, then $T \not\models \varphi$. In particular, if T is *consistent*—meaning $T \not\vdash "(\varphi \land \neg \varphi)$ "—then there is a model of T. This connection between finite sequences of formulas and the existence of structures is somewhat surprising considering that structures can be very large. Furthering this relation between the finite and the infinite is the compactness theorem.

Given that proofs are finite, the compactness theorem for proofs can yield important results when paired with Completeness $(1 D \cdot 1)$.

```
— 1 D • 2. Theorem (Compactness) ———
```

```
(ZFC) Let T be a theory. Therefore T has a model if and only if each finite \Delta \subseteq T has a model.
```

Proof .:.

If T has a model, then clearly every finite subset does too. But if T doesn't, then for any formula φ , $T \models$ " $(\varphi \land \neg \varphi)$ ", because no model $\mathbf{M} \models T$. By Completeness (1 D • 1), $T \vdash$ " $(\varphi \land \neg \varphi)$ ". Since proofs are finite, there is some finite subset $\Delta \subseteq T$ which contains all the formulas of T used in proving " $(\varphi \land \neg \varphi)$ ". This finite subset then also has $\Delta \vdash$ " $(\varphi \land \neg \varphi)$ ", and so by soundness, $\Delta \models$ " $(\varphi \land \neg \varphi)$ ". Hence this finite subset of T can't have a model.

These two theorems are very useful for their ability to generate models. As noted above, consistent theories have models which say that they're true. This is the kind of black magic that allows us to form models that satisfy all of the axioms of arithmetic, but aren't just \mathbb{N} . Adding to this black magic is the Löwenheim-Skolem theorem, which is the final theorem we need in the background of first-order logic, and it again allows us to conclude the existence of models with extremely nice properties. The proof of this is basically a more careful version of Completeness (1 D • 1), but we are not yet equipped to prove it without knowing some more set theory. In particular, we require knowledge of infinite cardinals.

We end this section with a bit of notation that will prove useful. In particular, "F0Lp" or "F0Lp(σ)" is used to denote "first-order logic with parameters". Really this is only used in the context of formulas: a formula is F0Lp iff it is of the form $\varphi(\vec{v}, \vec{p})$ for some variables \vec{v} , and some parameters \vec{p} . So this is always made in the context of some (arbitrary) model. For example, the identity element in a group **G** is F0L-definable, meaning definable without parameters. Given an arbitrary element g of the group **G**, g^{-1} is F0Lp-definable: it is the y such that $\mathbf{G} \models "g \cdot y = y \cdot g = 1$ ", i.e. $\mathbf{G} \models "\forall z((g \cdot y) \cdot z = z \cdot (g \cdot y) = z)$ ". The fact that g is used as a parameter here is what makes $g^{-1} \in G$ F0Lpdefinable.

Section 2. Basic Set Theoretic Concepts

Recall the following definition from Subsection 0 A.

- 2.1. Definition The *universe of sets* is the structure $V = \langle V, \in \rangle$, where V consists of all sets, and \in denotes membership.

Recall the notation introduced earlier: we may denote a set by enclosing its members in braces. For example, the set of 0, x, and Abraham Lincoln is $\{0, x, Abraham Lincoln\}$. The empty set is given some special notation: $\{\} = \emptyset$. It's also important to note that repetition in writing is unimportant: $\{x, x, x, x\} = \{x\}$. This is just because all of those xs are the same object and so the only thing in the collection $\{x, x, x, x\} = \{x\}$. And recall that generally $x \neq \{x\}$. Again, a physical analogy with sets is that the braces represent a box and the things in between them represent the contents: $\{x\}$ is a box with x in it while x is just the thing in that box.

The notation of listing out the elements is good enough for sets with only a few members, but things quickly get unwieldy if we want to consider larger collections. If we don't wish to list out explicitly all the elements of a set, we may instead write a description in the following form.

- 2·2. Definition

Let φ be some property, or predicate, or description, etc. Write $\{x : \varphi(x)\}$ for the collection of all x such that φ holds of x, i.e. $\varphi(x)$. For A another collection, we also consider $\{x \in A : \varphi(x)\}$ for $\{x : x \in A \land \varphi(x)\}$.

For example, $\{x : x \text{ is a person}\}$ is the set of all people. Similarly, we can restrict ourselves to a certain domain. For example, $\{x \in \mathbb{N} : x^2 = 1\}$ is the set of all natural numbers that square to 1. As sets are determined by their members (i.e. two sets are the same iff they have the same elements) this set is just $\{1\}$, because the only natural number whose square is 1 is 1 itself (the only other "number" that has a square of 1 is -1 which is not an element of \mathbb{N}). So we have defined a *subset* of \mathbb{N} in that all of $\{1\}$'s elements are in \mathbb{N} : it contains fewer members. We write $x \subseteq y$ to denote that x is a a subset of y, translated as $\forall z (z \in x \rightarrow z \in y)$ in first-order logic.

There are other ways of forming sets. For example, if x is a set, we can consider the *powerset*, the set of all collections formed from elements of x. Formally, $\mathcal{P}(x) = \{t : t \subseteq x\}$. Additionally, we have operations on sets, like union and intersection. These will be formally defined later, but to give a simple example, regarding lines as sets of points, the intersection of two (non-parallel) lines is always the set containing exactly one point. In particular, $L_1 = \{\langle x, y \rangle \in \mathbb{R}^2 : y = 2x + 3\}$ is a line, as is $L_2 = \{\langle x, y \rangle \in \mathbb{R}^2 : y = -x\}$, and their intersection is where the two lines meet, denoted $L_1 \cap L_2$:

$$L_1 \cap L_2 = \{ \langle x, y \rangle \in \mathbb{R}^2 : y = 2x + 3 \land y = -x \} = \{ \langle -1, 1 \rangle \}.$$

Now that we have some basic intuition set up, let's consider the following true statements about V, which are axioms of ZFC.

2.3. Definition (Axioms) (Extensionality) two sets are equal whenever they have the same members: $\forall x \ \forall y \ (x = y \leftrightarrow \forall v \ (v \in x \leftrightarrow v \in y)).$ (Empty set) there is a set \emptyset with no members: $\exists z \ \forall x \ (x \notin z).$ (Comprehension) for each A, and for each FOLp(\in)-formula $\varphi(v), \{v \in A : \varphi(v)\}$ exists: for φ a FOL(\in)-formula, $\forall w_0 \cdots \forall w_n \ \forall A \ \exists z \ \forall v \ (v \in z \leftrightarrow v \in A \land \varphi(v, \vec{w})).$

Extensionality is perhaps the most definition-like axiom, contained in the idea of a set.

- 2.4. Corollary Suppose $\{x\} = \{y\}$. Therefore x = y.

The empty set will provide the base for our universe in the following sense.

— 2•5. Result -

For every set $A, \emptyset \subseteq A$. Moreover, $A \subseteq \emptyset$ implies $A = \emptyset$.

Proof .:.

 $\emptyset \subseteq A$ since every element of \emptyset (of which there are none) is an element of A. Now suppose $A \subseteq \emptyset$. Thus $\forall x \ (x \in A \to x \in \emptyset)$. For each $x, x \in \emptyset$ is false, and thus $x \notin A$, Hence $\forall x \ (x \notin A)$, and therefore A and \emptyset have the same elements: no elements. By extensionality, $A = \emptyset$.

Comprehension^{vii} really is a *scheme*, meaning that for each FOL-formula, we get a different axiom. It is an attempt to formalize the idea of $\{x : \varphi(x)\}$. It's important to realize, however, that the full generality is inconsistent we can only consider the subset $\{x \in A : \varphi(x)\}$ for some set A. The idea is that we can't take arbitrary collections and call them sets, as seen in the following theorem.

– 2•6. Theorem (Russell's Paradox) —

There is no set $\{x : x = x\}$. Equivalently, $\neg \exists s \forall x (x \in s)$.

Proof .:.

If there were such a set, call it V. Now consider by comprehension the subset $a = \{x \in V : x \notin x\}$. By hypothesis, $a \in V$. Now we can question whether $a \in a$ or not. If $a \in a$, then a meets the definition: $a \notin a$, a contradiction. Hence $a \notin a$. But this means that a doesn't meet the definition of a, meaning $a \in a$, again a contradiction. So either way we have a contradiction, and so the hypothesis that V exists is false. \dashv

So comprehension at least says that we can consider (definable) subsets. In some sense, the issue is that the collection of all x is too big to be a set: V is not a set. So comprehension says that if we have a set, then all the subsets are small enough to be sets too.

§2A. A word on classes versus sets

We often want to talk about collections that aren't sets. Russell's Paradox $(2 \cdot 6)$ gives one such example: the collection of all sets, V. There are other, less *ad hoc* collections we will want to consider later, but this raises the question of how do we talk about these things? What is the distinction between "collection" and "set"? The basic idea is that collections *inside* a model are sets. So V is not a set by Russell's Paradox $(2 \cdot 6)$. We can still consider V a collection, though, and in particular, a *definable* collection in that the property of being in V is definable over V (trivially by $x \in V$ iff x = x).

The fact that a collection C is definable allows us to use the axioms of set theory with it like a parameter: we can't necessarily write for example " $\forall x \ (x \in C \to \varphi(x))$ " as a FOLp-formula, but we can write " $\forall x \ (\psi(x) \to \varphi(x))$ " where $C = \{x : \psi(x)\}$. Similarly, the fact that C is definable tells us through Axioms (2 • 3) that $\{y \in A : y \in C\}$ is a set for every set A: $\exists z \ \forall y \ (y \in z \leftrightarrow y \in A \land \psi(y))$.

```
− 2A • 1. Definition 
Let A be a model of set theory. A class of A is a collection C ⊆ A which is F0Lp(∈)-definable, i.e. x \in C iff A \models "φ(x)" for some F0Lp(∈)-formula φ. A class is a proper class iff it is not a set, meaning not in A.<sup>viii</sup>
```

So with V, sets are just things in V and classes are more like concepts that we can define. As a bit of notation, classes will generally be written upright: like 'V', 'L', 'Ord', 'HOD', instead of 'V', 'L', 'Ord', 'HOD'. But this is just a convention for this text, and there isn't a general standard in the field. Often upright boldface is used, and frequently there is no distinction in writing except by the use of majuscule letters.

It's hard to over emphasize that these collections are not necessarily a part of the set theoretic universe V: every set is a class, but not vice-versa. To see this, any set X is $FOLp(\in)$ -definable (by the formula " $x \in X$ "), so all sets are classes,

viialso called separation

^{viii}Technically, A's interpretation of ' \in ' isn't necessarily membership, and so it's better to say that C is a proper class iff it's a class and $A \models$ " $\neg \exists X \forall x \ (x \in X \leftrightarrow \varphi(x))$ " where φ defines C. Basically, for $X \in A$, we might not have that $X = \{x \in A : A \models "x \in X"\}$ because A is misinterpreting membership.

but not vice versa as Russell's Paradox $(2 \cdot 6)$ shows. The point of classes is just to say that while the entire collection isn't in our domain of discourse, the fact that it's definable shows that it still plays nicely with our axioms and we can easily reason about it. As noted before, comprehension tells us that the intersection of a set with a class is a set. So in some sense, every part of a class is a set, although the totality might not be.

There are more complicated understandings of classes that allow more collections than just definable ones. But at that point, we get into the realm of *class theory* rather than set theory. And before learning class theory, one needs to start with a good understanding of set theory. In essence, the typical model of class theory will be one that satisfies a variant of ZFC in an expanded language that has constant symbols for the relevant classes (so at least all definable collections). In this setting, a set is no longer defined to just be a member of the universe but instead something that can be collected together: x is a set iff $\exists y \ (x \in y)$ iff $\{x\}$ exists in the universe.

But under our definition, classes really are just short-hand for formulas. So often results for sets generalize to results for classes just by virtue of classes being definable. That said, it's still important to remember that classes are not always sets, and certain theorems do not always generalize to classes. The basic problem is that we can't quantify over classes in the sense of saying "for all classes such-and-such happens". So often our results about classes are metatheoretic.^{ix}

§2B. Ordered pairs

So far, the set theory presented is relatively uninteresting, because the axioms do not allow us to form sets with new elements: we may only take subsets. Moreover, even if we have these sets, it's not completely clear what the benefit of them is. To motivate things a little more, sets are seen as a foundation of mathematics, both practically, and philosophically. Often, when one needs to make things mathematically precise, it is done using sets.^x So to begin with, we will first show that we can formalize an ordered pair $\langle x, y \rangle$, in that we have a construction where $\langle a, b \rangle = \langle a', b' \rangle$ if and only if a = a' and b = b'. This will allow us to talk about sequences, functions, relations, and so forth. To do this, we need some additional axioms that reflect what's true of **V**.

— 2B•1. Definition (Axiom)

(*Pairing*) for any two sets x and y, the pair $\{x, y\}$ exists: $\forall x \forall y \exists z \forall v (v \in z \leftrightarrow (v = x \lor v = y))$.

- **2B**•**2**. **Definition** For x, y sets, the *ordered pair* of x and y, $\langle x, y \rangle$ is the set {{x}, {x, y}}.

As a side note, if x = y, then $\langle x, y \rangle$ collapses down to $\{\{x\}\}$, since $\{x, y\} = \{x, x\} = \{x\}$ because the two have the same members. Now let's prove the single point of having an ordered pair: that the entries are uniquely determined by the ordered pair.

— 2B•3. Result

Let x, x', y, y' be sets. Therefore $\langle x, y \rangle = \langle x', y' \rangle$ iff x = x' and y = y'.

Proof .:.

Clearly if x = x' and y = y', then $\langle x, y \rangle = \langle x', y' \rangle$. So suppose $\langle x, y \rangle = \langle x', y' \rangle$, meaning that these sets have the same members. The members of these sets are $\{x\}$ and $\{x, y\}$, and $\{x'\}$ and $\{x', y'\}$.

If $x \neq y$ and $x' \neq y'$, then the two-element sets must be equal, and the one-element sets must be equal, implying

^{ix} To give a more precise example of where the distinction is important, Gödel's theorems tell us that ZFC cannot prove the consistency of ZFC. But if we assume we have a model $M \models$ ZFC, then in the metatheory it would seem like M should know that ZFC is consistent because M contains a model of it—after all, the class M is a model. But this isn't true: $M \subseteq M$, but $M \notin M$. In other words, M is unable to talk about M or what infinite collection of axioms it satisfies directly because it is a *class* and so it does not exist in the domain of discourse that M considers. Although M can see that φ holds in M for each formula φ of ZFC, that doesn't mean M understands that M exists as a model.

⁽Nor does it mean that **M** thinks all of those formulas make up ZFC. In particular, there are non-standard models that misinterpret what "finite" means, and thus [as formulas, proofs, and so on are finite strings of symbols] misinterpret what exactly is in our description of "ZFC". In such a model **M**, it's possible for there to be models **W** of the actual ZFC, but **M** doesn't recognize **W** as satisfying *all* of the formulas of ZFC, precisely because it has misinterpreted what exactly ZFC is.)

^xthere are other theories some people put forth as a foundation of mathematics, but their proponents often either defer the serious paradoxical issues for set theory to deal with, or fail to start from philosophically basic notions.

x = x' and $\{x, y\} = \{x', y'\}$. Since we already know x = x', we must have y = y'. If x = y, then $\langle x, y \rangle = \{\{x\}\}$. Hence both elements of $\langle x', y' \rangle$ are equal to this: $\{x\} = \{x, y\} = \{x'\} = \{x', y'\}$, implying that x' = y' = x = y. The same idea holds if x' = y'.

We can also refer to the left and right coordinate of an ordered pair in this way: given an ordered pair z, the leftcoordinate is just the x satisfying $\exists y(\langle x, y \rangle = z)$. In fact, using another axiom, we can restrict the search for such a y to an element of the union of z.

- **2B**•4. Definition (Axiom) (Union) for any family of sets F, there is a set containing the elements of all of those sets: $\forall F \exists U \ \forall v \ (v \in U \leftrightarrow \exists x (x \in F \land v \in x)).$

We denote the union by $\bigcup F$, in this case. For just two sets, write $x \cup y = \{a : a \in x \lor a \in y\}$ rather than the more clumsy $\bigcup \{x, y\}$, which exists by union and pairing. For a concrete example of a union, consider $x = \{1, 2\}$, and $y = \{2, 4, 10\}$. Therefore $x \cup y = \{1, 2, 4, 10\}$. A related concept, which we could already form through comprehension, is the intersection of two sets: $x \cap y = \{a : a \in x \land a \in y\}$. More generally, for a non-empty family, F, the intersection $\bigcap F = \{a : \forall x \in F(a \in x)\}$, which can be written as a subset of each particular $x \in F$. Similarly, we can take complements: $x \setminus y = \{a \in x : a \notin y\}$. Using the same x and y example from before, $x \setminus y = \{1\}$ while $x \cap y = \{2\}$. Note that we have the following trivial facts about intersection, union, and so forth, mostly which just follow from properties of sentential connectives:

- $x \cap x = x, x \cup x = x, x \cup \emptyset = x;$
- $x \cap \emptyset = \emptyset, x \setminus x = \emptyset;$
- $x \setminus (x \cap y) = x \setminus y;$
- $x \cap y \subseteq x$, and if $a \subseteq x$ and $a \subseteq y$, then $a \subseteq x \cap y$;
- $x \cap (y \cap z) = (x \cap y) \cap z$, and similarly for union;
- $(x \cap y) \cup z = (x \cup z) \cap (y \cup z)$, and $(x \cup y) \cap z = (x \cap z) \cup (y \cap z)$;
- if $x \subseteq a$ and $y \subseteq a$, then $x \cup y \subseteq a$;
- if $x \subseteq y$ and $y \subseteq a$, then $x \subseteq a$; and
- $x \subseteq y$ iff $y \cup x = y$ iff $x \cap y = x$ iff $x \setminus y = \emptyset$.
- $x \subseteq y$ implies $\bigcup x \subseteq \bigcup y$.

These also have a related definition, since sets having completely different elements is very useful.

— 2B•5. Definition —

Two sets x and y are *disjoint* iff $x \cap y = \emptyset$. A family of sets F consists of *disjoint sets* or *pairwise disjoint sets* iff $x \cap y = \emptyset$ for all $x, y \in F$.

Now ordered pairs on their own are fine, but we still need to be able to do more with them to do any basic set theory. Obviously using pairing, we can form $\{\langle x, y \rangle, \langle x', y' \rangle\}$. We can also form $\{\langle x, y \rangle, \langle x', y' \rangle, \langle x'', y'' \rangle\}$ using another application of pairing and union:

 $\{\langle x, y \rangle, \langle x', y' \rangle, \langle x'', y'' \rangle\} = \{\langle x, y \rangle, \langle x', y' \rangle\} \cup \{\langle x'', y'' \rangle, \langle x'', y'' \rangle\}.$

We have two potential routes to form arbitrary sets of pairs—excluding finite applications of pairing and union—powerset (with comprehension), and replacement. First we introduce replacement.

− 2B•6. Definition A F0Lp(∈)-formula $\varphi(x, y)$ defines a function over D iff for every $x \in D$ there is a unique y with $\varphi(x, y)$. Symbolically, $\forall x \ (x \in D \rightarrow \exists ! y \ \varphi(x, y))$.

Replacement then says in effect that if we can definably transform elements of a set, then the set of the transformations exist.

— 2B•7. Definition (Axiom) –

(*Replacement*) the image of a function over a set is a set: for each FOL(\in)-formula φ ,

 $\forall w_0 \cdots \forall w_n \forall D \left(\varphi(x, y, \vec{w}) \text{ defines a function over } D \to \exists R(y \in R \leftrightarrow \exists x (x \in D \land \varphi(x, y, \vec{w}))) \right).$

It should be clear that D is the intended domain of the function defined by φ , and R is the range of φ restricted to D. So replacement is saying that R exists: if I can define a function from a set, then the range is a set. So if we consider the function mapping xs and ys to $\langle x, y \rangle$, we get a cartesian product as the range.

```
- 2B•8. Definition
```

The *cartesian product* $A \times B$ of A and B is the set of all pairs from A and B: { $(a, b) : a \in A \land b \in B$ }.

- 2B•9. Result

Let A and B be arbitrary. Therefore $A \times B$ exists^{xi}.

Proof .:.

For each $a \in A$ consider the formula $\varphi(b, p, a)$ which is just that $p = \langle a, b \rangle$. This is of course shortened, but the defining notions can be replaced here. Regardless, it's clear that this defines a function over B, where b maps to $\langle a, b \rangle$ for our fixed $a \in A$. So replacement says that there is some

 $R_a = \{p : \exists b \in B \ \varphi(b, p, a)\} = \{\langle a, b \rangle : b \in B\}.$

This is an individual slice of the cartesian product. So consider the function $\psi(r, a)$ which states $r = R_a$, i.e. $r = \{\langle a, b \rangle : b \in B\}$. (We can do this by taking the even longer formula $\psi(r, a)$ to be $\forall x \ (x \in r \leftrightarrow \exists b \in B \ \varphi(b, x, a))$.) This defines a function over A, and so another application of replacement yields the set

 $P = \{\{\langle a, b \rangle : b \in B\} : a \in A\}.$

Hence $\bigcup \{ \{ \langle a, b \rangle : b \in B \} a \in A \} = \{ \langle a, b \rangle : a \in A \land b \in B \} = A \times B.$

§2C. Relations

The cartesian product is the basis for most of basic set theory, since it allows us to consider relations and fuctions, and thus define sequences, and notions of size. Really, if sets are supposed to be devoid of all structure beyond membership, this idea allows us to put structure back into play, and thus work with more complicated ideas all within set theory.

- 2C·1. Definition A relation is a subset $R \subseteq A \times B$ for some A, B. For any relation R, dom $(R) = \{x : \exists y \ (\langle x, y \rangle \in R)\}$, and similarly, ran $(R) = \{y : \exists x \ (\langle x, y \rangle \in R)\}$.

The existence of the domain and range of *R* can be shown by the union axiom: $x, y \in \bigcup \langle x, y \rangle = \{x, y\} \cup \{x\} = \{x, y\}$ so that $\langle x, y \rangle \in R$ implies $x, y \in \bigcup \bigcup R$. Hence we can take the appropriate subset to define the domain and range. Alternatively, we can use replacement. But resorting to the more basic axioms can be insightful.

Note that then if R is a relation, every subset of R is a relation too. Moreover, the union of relations are relations. Really a relation is just a set R where $z \in R$ implies $z = \langle x, y \rangle$ for some x and y. So the relation doesn't need to be over the same set or have some intuitive reason behind relating elements. Note that for R a relation, we will often write x R y for $\langle x, y \rangle \in R$. Note that we can have the relation defined on three sets (or more) just by having $\langle x, y \rangle \in R$ always having y an ordered pair of some form. We will make this more formal or official later, so for now we focus on binary relations. Again, we get some immediate facts: for R and S relations,

- $\operatorname{dom}(R \cup S) = \operatorname{dom}(R) \cup \operatorname{dom}(S);$
- $\operatorname{ran}(R \cup S) = \operatorname{ran}(R) \cup \operatorname{ran}(S);$
- dom $(R \cap S) \subseteq$ dom $(R) \cap$ dom(S); and
- $\operatorname{ran}(R \cap S) \subseteq \operatorname{ran}(R) \cap \operatorname{ran}(S)$.

Given any relation, we can form the inverse, where we swap all the entries of the ordered pairs:

— 2C•2. Definition

For R a relation, define $R^{-1} = \{(y, x) : (x, y) \in R\}$ to be the *inverse* or *converse* of R.

 \dashv

^{xi}Note that we've also shown that the cartesian product of classes exists as well. In particular, for A and B classes, we have the FOLp-formula defining $A \times B$ by $x \in A \times B$ iff $\exists y \exists z \ (x = \langle y, z \rangle \land y \in A \land z \in B)$.

The existence of R^{-1} can be shown through a variety of methods, notably replacement. Note that this behaves exactly as one would expect:

- 2C·3. Result Let R be a relation. Therefore R^{-1} is a relation, and $(R^{-1})^{-1} = R$. Moreover, for S a relation, $(R \cap S)^{-1} = R^{-1} \cap S^{-1}$.

Proof .:.

Clearly R^{-1} , as a set of ordered pairs, is a relation. Moreover, $(R^{-1})^{-1} = \{\langle x, y \rangle : \langle y, x \rangle \in R^{-1}\} = \{\langle x, y \rangle : \langle x, y \rangle \in R\} = R$. To see that the inverse of an intersection is the intersection of the inverses, let $\langle y, x \rangle \in (R \cap S)^{-1}$. Therefore $\langle x, y \rangle \in R \cap S$ and so $\langle y, x \rangle \in R^{-1}$ and $\langle y, x \rangle \in S^{-1}$. Similarly, if $\langle y, x \rangle \in R^{-1} \cap S^{-1}$, then $\langle x, y \rangle$ must be in both R and in S, so that $\langle y, x \rangle \in (R \cap S)^{-1}$. So the two sets have the same elements, and so must be equal.

One of the most important kinds of relations is a partial order, notable mostly for the notion of transitivity.

- 2C·4. Definition

Let R be a relation. Write x R y for $\langle x, y \rangle \in R$. We say R is a realtion over X iff dom $(R) \cup \operatorname{ran}(R) = X$.

- *R* is *transitive* iff $\forall x \forall y \forall z (x R y \land y R z \rightarrow x R z)$.
- *R* is symmetric iff $\forall x \forall y (x R y \leftrightarrow y R x)$.
- *R* is antisymmetric iff $\forall x \ \forall y \ (x \ R \ y \land y \ R \ x \rightarrow x = y)$.
- *R* is *total* iff $\forall x \forall y (x, y \in \text{dom}(R) \cup \text{ran}(R) \rightarrow (x R y \lor x = y \lor y R x))$.
- *R* is *reflexive* iff $\forall x \ (x \in \text{dom}(R) \cup \text{ran}(R) \rightarrow \langle x, x \rangle \in R)$.
- *R* is a *partial order* iff it is transitive, and antisymmetric.
- *R* is *linear* iff it is transitive, antisymmetric, and total.

A relation *R* is called a *strict order* if $\langle x, x \rangle \notin R$ for all *x*.

We now get some very easy results about various relations that the reader should check to confirm their intuitions.

- The identity relation $id_A = \{(x, x) : x \in A\}$ is symmetric and antisymmetric.
- *R* is symmetric iff $R = R^{-1}$.
- If *R* is a linear order then $R \cap (A \times A)$ is a linear order for any set *A*.
- If *R* is antisymmetric, and dom(*R*) \cup ran(*R*) has more than one element, then $R^{-1} \neq R$.
- If *R* and *S* are reflexive, then $R \cup S$ is reflexive.
- If *R* and *S* are antisymmetric, then $R \cup S$ is antisymmetric.
- If R is antisymmetric, and $S \subseteq R$, then S is antisymmetric.
- If *R* is transitive, then R^{-1} is transitive.
- If R is a partial order, then $R \cup id_{dom(R)\cup ran(R)}$ is a reflexive partial order.
- If R is a partial order, then $R \setminus id_{dom(R)\cup ran(R)}$ is a strict partial order.

The relations which are of fundamental importance to set theory are well-founded relations, and equivalence relations.

A relation R is well-founded iff for every subset $X \subseteq \text{dom}(R) \cup \text{ran}(R)$, there is an R-minimal element of X, meaning an $x \in X$ with no $y \in X$ with y R x.

When we investigate well-founded linear orders. It turns out that they are *canonical* in the sense that they are all initial segments of each other (up to isomorphism). We will investigate well-founded relations later on. For now, consider some terminology regarding equivalence relations.

- 2C•6. Definition

A relation R is an *equivalence relation* iff R is reflexive, symmetric, and transitive. An *equivalence class* of R is a $X \subseteq \text{dom}(R) \cup \text{ran}(R)$ such that x R y for every $x, y \in X$. For $x \in \text{dom}(R) = \text{ran}(R)$, write $[x]_R$, the *equivalence class of* x, for $\{y \in \text{dom}(R) : x R y\}$. For X an equivalence class of R, a *representative* of X is an $x \in \text{dom}(R)$ such that $X = [x]_R$. For X a set, a partition is a set P such that $\forall x (x \in X \to \exists ! Y(Y \in P \land x \in Y))$ and $\forall Y \in P(Y \subseteq X)$. For example, id_X is an equivalence relation over X with $[x]_{=} = \{x\}$ for all $x \in X$. But an equivalence relation is more general than equality. But in essence, an equivalence relation still acts like it in the following sense.

- 2C·7. Result — For R an equivalence relation and $x, y \in \text{dom}(R), x R y \text{ iff } [x]_R = [y]_R$.

Proof .:.

If $[x]_R = [y]_R$, then by reflexivity, $y \in [y]_R = [x]_R$ implies x R y. So suppose x R y. If $a \in [x]_R$ then x R a. By symmetry, a R x. Since x R y, symmetry yields that a R y and symmetry again yields y R a, i.e. $a \in [y]_R$. Thus $[x]_R \subseteq [y]_R$. The same argument shows $[y]_R \subseteq [x]_R$. Therefore x R y implies $[x]_R = [y]_R$. \dashv

- **2C**•8. Corollary For *R* an equivalence relation, $[x]_R = [y]_R$ or $[x]_R \cap [y]_R = \emptyset$ for all $x, y \in \text{dom}(R)$.

Proof .:.

Suppose $a \in [x]_R \cap [y]_R$. By transitivity and symmetry, $x R a \wedge a R y$ implies x R y so that $[x]_R = [y]_R$. \dashv

Hence, the set of equivalence classes partitions the domain of R.

- 2C•9. Corollary For R an equivalence relation, $\{[x]_R : x \in \text{dom}(R)\}$ is a partition of dom(R).

Conversely, partitions give rise to equivalence classes, and thus equivalence relations and partitions can be seen as the same thing.

- 2C•10. Result

Let *X* be a set and let *P* be a partition of *X*. Therefore the relation $R = \{ \langle a, b \rangle \in X \times X : \exists Y \in P (a \in Y \land b \in Y) \}$ is an equivalence relation over dom(R) = X.

Proof .:.

Symmetry is immediate by the commutativity of \wedge . Since each $x \in X$ has some $Y \in P$ with $x \in Y$, reflexivity is true of R, and this shows dom(R) = X. So it suffices to show transitivity. Suppose $x, y \in Y \in P$ and $y, z \in Y' \in P$. As a partition, there is only one $Y'' \in P$ with $y \in Y''$ so that Y = Y' and thus $x, y, z \in Y \in P$, which yields x R z.

The main point of equivalence classes is just that they give a new notion of equality by considering the equivalence classes instead of the equivalence relation so directly. This allows us to say things like "x and y are the same modulo R". Similarly, it allows us to define other relations so long as they respect the equivalence relation. In doing this, note that often the equivalence class $[x]_R$ will have multiple elements: $[x]_R = [y]_R$ although $x \neq y$. So if we are to make a definition about $[x]_R$ that makes reference to x, we need to ensure that this gives the same thing if we were to choose y instead as our representative.

– 2C•11. Result

Let \approx be an equivalence relation on X, and let $R \subseteq X \times X$. Suppose x R y iff x' R y' for $x \approx x'$ and $y \approx y'$. Therefore the relation $R_{/\approx}$ over $X_{/\approx} = \{[x]_{\approx} : x \in X\}$ defined by

$$[x]_{\approx} R_{\approx} [y]_{\approx}$$
 iff $x R y$

is well-defined, meaning independent of the choice of representatives.

Proof .:.

Suppose $[x]_{\approx} = [x']_{\approx}$ and $[y]_{\approx} = [y']_{\approx}$. Therefore $x \ R \ y$ iff $x' \ R \ y'$, meaning $[x]_{\approx} \ R_{/\approx} \ [y]_{\approx}$ iff $x' \ R \ y'$. \dashv

This is the idea from algebra that allow us to "mod out" by an equivalence relation, like via the orbits induced by other groups or ideals of a ring, generating a new group or ring. There are many applications, which we will see later.

§2D. Functions

We introduce some simple notions about functions pictured below. In particular, the notion of hitting every element in the range, and the notion of "doubling up": sending two elements to the same place.

- 2D • 1. Definition A *function* is a relation $f \subseteq A \times B$ such that for each $x \in A$ there is exactly one $y \in B$ with $\langle x, y \rangle \in f$. We write $f : A \to B$, and y = f(x) in this case.

• We call $f : A \to B$ injective iff $f(x) \neq f(y)$ for all $x \neq y$ in A.

- We call $f : A \to B$ surjective iff ran(f) = B.
- We call $f : A \rightarrow B$ bijective iff it is injective and surjective.

We also call such functions injections, surjections, or bijections. Note that a function being surjective depends on how we regard it: obviously $f : dom(f) \to ran(f)$ is surjective for *any* function f. Clearly if f is a function $f : A \to B$ and $B \subseteq C$, then we can also regard $f : A \to C$ which may no longer be surjective. So surjectivity is only ever referenced when the *co-domain*—the object to the right of the arrow—is specified. Note also that in this text, instead of ran(f) for the "range of f", we will write im(f) for the "image of f". This is merely a personal preference to distinguish relations from functions.

Occasionally, we might reference a function by the notation $x \mapsto f(x)$. For example, $x \mapsto \{x\}$ is a function and $(x \mapsto \{x\})(\emptyset) = \{\emptyset\}$. This is mostly done to avoid introducing too many letters, especially if the function is only going to be referenced a few times. The domain of this function isn't clear from the notation, and usually is left to context. For example, $x \mapsto x^2 + 2$ is *not* an injective function with the usually assumed domain of real numbers, \mathbb{R} (whatever this might be in our set theoretic framework, since we haven't defined it yet). But if we restrict our domain just to positive real numbers, the resulting function is injective.



2D•2. Figure: An injection f, surjection g, and bijection h

Because the objects we deal with in set theory are sets—in particular, sets that are *hereditarily* sets, meaning all their members are also sets, and the same holds for them too—we need to make the distinction between the "pointwise image" of a function as opposed to the "value" of a function. To motivate the example, consider the set $A = \{a, b, \{a\}\}$ and a function f with domain A. In general, there is a difference between f(a), $f(\{a\})$, and $\{f(a)\}$. But sometimes we do want to consider the set of values of a function, like $\{f(a)\}$. Similarly, sometimes we want to take a function, but restrict our attention to a smaller subset of its domain. To denote the difference, we have the following definition.

2D·3. Definition Let $f : A \to B$ be a function over sets A, B. Let $X \subseteq A$. Write the *pointwise image* of f under X as $f"X = \{f(x) : x \in X\}$. Write the *restriction* of f to X as $f \upharpoonright X = \{\langle a, b \rangle \in f : a \in X\} = f \cap (X \times \operatorname{im} f)$.

So in the example above, $f''\{a\} = \{f(a)\}$ while $f(\{a\}) \neq f''\{a\}$. Note that $\operatorname{im}(f) = f'' \operatorname{dom}(f)$. Since restriction allows us to chain our domain, $\operatorname{dom}(f \upharpoonright X) = X$; we can also write $f''X = \operatorname{im}(f \upharpoonright X)$. We also have the following operations on functions: composition and inverses (which might not be functions).

- 2D•4. Definition Let $f : A \to B$ be a function over sets A, B. Let f^{-1} be the relation $\{\langle b, a \rangle : \langle a, b \rangle \in f\} \subseteq B \times A$. For $g : B \to C$, the composition $g \circ f$ is defined by $\{\langle a, c \rangle : \exists b \in B(f(a) = b \land g(b) = c)\}$.

It should be clear that $g \circ f$ is also a function, now from dom(f) to im(g).



2D.5. Figure: Example of composition

Functions are fundamental to mathematics, as they are a means of transformations. More than functions, really, the importance is placed on the properties of functions. Most graphs of most functions will be set-theoretic haze: just a bunch of points with no discernible relationship between the points beyond satisfying the definition of being a function. So most applications will care about functions that preserve certain relationships. These are typically called homomorphisms, embeddings, and so on. We have already defined one such property: preserving inequality, or injectivity. But the key thing for now is to recognize that functions can be interpreted in purely set theoretic terms.

Let me take a moment to talk further about bijections, injections, and surjections. When letting their sheep out to graze, one technique that shepards used to make sure all sheep were accounted for was to pick up a pebble every time a sheep left. Then a pebble was dropped for every sheep that returned. So if there were any left over pebbles, there were sheep left out. Stated in terms of functions, there was a function f: sheep \rightarrow pebbles which was *injective*—two different sheep get two different pebbles—and *surjective*—every pebble corresponds to a sheep—and hence bijective. Going back to the example, this means we have the same number of pebbles as sheep, and we have confimed this *without counting*. So bijections really form a notion of size between two sets: we merely rename the elements via the bijection. For a very simple example, consider $\{a, b, c\}$ and $\{\alpha, \beta, \gamma\}$. Renaming $a `\alpha', b `\beta'$, and $c `\gamma'$, we get $\{a, b, c\}$ should have the same number of elements as $\{\alpha, \beta, \gamma\}$, which it clearly does, and we did this without directly counting both and then seeing that the two numbers line up.



2D.6. Figure: Counting without and with numbers

In some sense, counting just adds a third set of numbers, and then considers bijections to the numbers as a means of counting each set. So to remove the middle-man of numbers—which we have not yet introduced in set theoretic terms yet—we have the following definition.

- 2D.7. Definition Let A and B be sets. Write $A =_{size} B$ iff there is a bijection $f : A \rightarrow B$.

Ideally, we'd like to say the *cardinality* of A and B are the same. But without further technology in the form of ordinals, we have no means of saying this. Instead, we will say that the cardinality of a set A is the *class* of $\{B : A =_{size} B\}$. We also have a notion of order on these equivalence classes in the following sense.

- **2D**•8. Definition Let A and B be sets. Write $A \leq_{\text{size}} B$ iff there is an injection $f : A \rightarrow B$.

For example, $A \subseteq B$ has $A \leq_{\text{size}} B$. Note that this is in essence the only way to have a size less than or equal to a set

Ch I §2 E

in the following sense.

- 2D•9. Result

$$A <_{\text{size}} B$$
 iff there is some $C =_{\text{size}} B$ with $A \subseteq C$.

Proof .:.

To see this, note that if $A \leq_{\text{size}} B$, then the injection $f : A \to B$ witnessing this has $f''A \subseteq B$. So take $C = (B \setminus f''A) \cup A$, where clearly $A \subseteq C$. Ostensibly, $C =_{\text{size}} B$ since it seems we can consider the function $F : C \to B$ defined by

$$F(b) = \begin{cases} b & \text{if } b \in B \setminus f''A \\ f(a) & \text{if } a \in A. \end{cases}$$

The only issue with this is that $A \cap (B \setminus f^*A)$ might not be empty, which would make the above ill-defined. But assuming $A \cap B = \emptyset$, then F is a bijection. To remove the assumption $A \cap B = \emptyset$, consider instead $C = ((B \setminus f^*A) \times \{\emptyset\}) \cup (A \times \{\{\emptyset\}\})$. with $F(\langle b, \emptyset \rangle) = b$ and $F(\langle a, \{\emptyset\}) = f(a)$. This yields the appropriate bijection.

We will see later that $A \leq_{\text{size}} B$ and $B \leq_{\text{size}} A$ implies $A =_{\text{size}} B$, as suggested by the notation. But the long proof of this isn't instrumental to us for now. What's important is the notion of bijection giving a notion of size.

We have the following easy properties of size and bijections. Note that " $f : A \to B$ " is not just a statement that $f \subseteq A \times B$, but that f is a function with f defined on all of A (so dom(f) = A) and im(f) $\subseteq B$.

- If $f : A \to B$ and $g : B \to C$ are injective, then $g \circ f : A \to C$ is injective.
- If $f: A \to B$ is surjective, and $g: B \to C$ is surjective, then $g \circ f: A \to C$ is surjective.
- If $f : A \to B$ and $g : B \to C$ are bijections, then $g \circ f : A \to C$ is a bijection;
- equivalently, if $A =_{\text{size}} B$ and $B =_{\text{size}} C$ then $A =_{\text{size}} C$.
- If $f : A \to B$ is a bijection, then $f^{-1} : B \to A$ is a bijection;
- equivalently, $A =_{\text{size}} B$ iff $B =_{\text{size}} A$ for all A and B.
- If $f : A \to B$ is injective, then $f : A \to \text{im } f$ is a bijection;
- equivalently, $X =_{\text{size}} f''X$ for $f : A \to B$ injective with $X \subseteq A$.

All of this has been done without the notion of counting, but the benefit of being able to count is that it opens up a new theory of "numbers". So we will return to the notion of size or *cardinality* later, after we have introduced the ordinals. But now we should have a basic intuition for functions and size.

§2E. Transitive sets

Let's take a moment to look at so-called "transitive" sets. In some sense, this is a misnomer, since it is not the set that is transitive, but the membership relation.

```
- 2E•1. Definition
```

A set *x* is *transitive* iff membership into *x*, meaning $\{\langle a, b \rangle : a \in b \land (b \in x \lor b = x)\}$, is transitive.

So x being transitive is the same as saying $a \in b \in x$ implies $a \in x$. Equivalently, $b \in x$ implies $b \subseteq x$.^{xii} In some sense, this means that transitive xs not only contain various a with $a \in b \in x$, but that we go all the way down to the basis of the universe: \emptyset . This is partially shown in Figure 2 E • 2.

But to prove this, we need an additional axiom. In another sense, x being transitive means that the structure (x, \in) is a *submodel* of V: they both interpret \in in the same way. As a result of this, we get some nice model-theoretic results. Below is just one example of this showing that transitive sets have nice *absoluteness* properties that we will consider later.

^{xii}Of course, we cannot have a set where $\forall b \ (b \subseteq x \rightarrow b \in x)$ by the same reasoning as in Russell's Paradox (2 • 6).

BASIC SET THEORETIC CONCEPTS



2E-2. Figure: The membership relation compared to a transitive set

- **2E·3.** Result Let X be transitive. Let $a, b \in X$. Therefore $X = \langle X, \in \rangle \vDash a \subseteq b$ " iff $V \vDash a \subseteq b$ ".

Proof .:.

To say that $a \subseteq b$ is just short-hand for $\forall y \ (y \in a \to y \in b)$. Since **X** and **V** interpret \in the same way, if $y \in X$, **X** \models " $y \in a \to y \in b$ " iff **V** \models " $y \in a \to y \in b$ ". Since $X \subseteq V$, y ranges over more sets in V than in X: if **V** \models " $a \subseteq b$ ", then **X** \models " $a \subseteq b$ ". The other direction, if **V** \models " $a \not\subseteq b$ ", then there must be some element $y \in a$ with $y \notin b$. But since X is transitive with $a, b \in X$, $y \in a \in X$ implies $y \in X$. Hence **V** \models " $y \in a \land y \notin b$ " implies **X** \models " $y \in a \land y \notin b$ ", because they interpret \in in the same way. But then **X** \models " $a \not\subseteq b$ ". \dashv

Finding examples of transitive sets and examples of non-transitive sets is easy. In particular,

- 1. \emptyset is transitive. { \emptyset } is transitive.
- 2. If x is transitive, then $x \cup \{x\}$ is transitive (any element $b \in x \cup \{x\}$ is still a subset since $b \subseteq x \subseteq x \cup \{x\}$).
- 3. Writing 0 = Ø, 1 = {0}, and 2 = {0, 1}, then from the above, 0, 1, 2, and {0, 1, 2} are transitive, but {1}, {0, 2}, and {2} are not.
- 4. If x is transitive and $y \subseteq x$, then $x \cup \{y\}$ is transitive.

Now we introduce the axiom of foundation. To motivate the axiom, it's difficult to think of a set which could be an element of itself. Considering a more physical picture, you can't place a box (completely) inside itself—the concept wouldn't make any sense. Indeed, Russell's Paradox $(2 \cdot 6)$ partly goes through because we consider that the collection of everything that exists is an element of itself. This would suggest we should assume $\forall x \ (x \notin x)$ as an axiom. This would rule out some direct approaches, but we could still code the counter-intuitive situations through other loops: $x \in y$ and $y \in x$, for example.

The axiom of foundation rules out loops of arbitrary length, and has a great number of consequences. Intuitively, the idea can be motivated as above, but it can also be motivated though the iterative conception of what a collection is: namely, collections are built up of smaller things that have come before in a certain sense. This will turn out to be equivalent to the axiom. Explicitly, foundation merely states that membership is well-founded.

- 2E•4. Definition (Axiom)

(Foundation) for each x, there is a \in -minimal element of x: $\forall x \exists y \ (y \in x \land \forall z \ (z \in y \rightarrow z \notin x))$.

– 2 E • 5. Corollary —

Assume the axiom of foundation. Therefore:

1. We never have $x \in x$.

2. In fact, there are no finite loops $x_0 \in x_1 \in \cdots \in x_n \in x_0$.

- 3. If $x \neq \emptyset$ is transitive, $\emptyset \in x$ is the \in -minimal element of x.
- 4. *x* is transitive iff $x \cup \{x\}$ is transitive.

BASIC SET THEORETIC CONCEPTS

Proof .:.

- 1. Suppose $x \in x$. By foundation, there is a \in -minimal element of $\{x\}$, which must be x. So any $y \in x$ has $y \notin \{x\}$ by minimality. But $x \in x$ has $x \in \{x\}$, so we have a contradiction.
- 2. Consider the set $\{x_0, \dots, x_n\}$, which exists by finite applications of union and pairing. This has no \in -minimal element, since any x_i has $x_{i-1} \in x_i$ for i > 0 or else $x_n \in x_i$ for i = 0.
- 3. If x is transitive, then every element $y \in x$ is a subset of x. Hence if $y \neq \emptyset$ is \in -minimal, then there is some $z \in y \in x$, which yields $z \in x$ and $z \in y$, contradicting the minimality of y. Hence any \in -minimal element must be \emptyset .
- 4. We know that x being transitive implies x ∪ {x} is transitive. For the other direction, if x ∪ {x} is transitive, then any a ∈ b ∈ x ∪ {x} must have either a ∈ x or a = x. But a cannot equal x without us having a finite loop: either x ∈ b ∈ x or x ∈ b = x. Hence a ∈ b ∈ x ∪ {x} requires a ∈ x. This clearly implies that x is transitive since a ∈ b ∈ x ∪ {x} implies a ∈ x.

Important for later is the idea that any set is contained in a transitive set, which should seem rather clear: just continually add in the elements missing. To formalize this, however, we need some more ideas in general: the natural numbers. In general, we need ideas which will take the form of ordinals. In particular, we need a better idea of how to talk about *rank*. If \emptyset is the base of the universe, then $\{\emptyset\}$ is just above it, and so has a rank one higher. Similarly, collections built from these like $\{\emptyset, \{\emptyset\}\}$ and $\{\{\emptyset\}\}$ are a rank higher than that. This is the iterative concept we will explore: $\{\emptyset\}$ comes "before" $\{\emptyset, \{\emptyset\}\}$ because it has a lower rank.

§2F. Formula abbreviations

We will often make abbreviations to our formulas to change their domain of discourse. For example, instead of writing " $\forall x \ (x \in A \to \varphi)$ ", we will write " $\forall x \in A \varphi$ ". Similarly, instead of " $\exists x \ (x \in A \land \varphi)$ ", we will write " $\exists x \in A \varphi(x)$ ". These are standard translations of the more natural language ways of phrasing the formulas: "for all x in A, $\varphi(x)$ is true" and "there is an x in A such that $\varphi(x)$ is true". We may also do this with other properties. For example, " $\forall x < a \varphi(x)$ " stands for " $\forall x \ (x < a \to \varphi(x))$ ". Mostly this just serves to simplify formulas and make them easier to read, which we have already done with other abbreviations like ' \subseteq ', ' \cup ', and so forth.

Section 3. Well-orders and Ordinals

We will primarily be working with well-orders. Ordinals themselves are the "canonical" well-orders in that they are well-ordered by membership. They will also be special transitive sets, giving some credence to the axiom of foundation, since these canonical examples of transitive sets are well-ordered.

- $3 \cdot 1$. Definition - A relation *R* is a *well-order* iff *R* is linear, and well-founded.

We will see later that all well-orders on their domain and range are isomorphic to ordinals with the membership relation. First we must figure out what the ordinals are, and what properties they have.

§3A. Introducing ordinals

- **3A**•1. Definition A set α is an *ordinal* iff α is transitive, and \in is a strict well-order of α .

Note by foundation that \in is well-founded on any set α . In the absence of the axiom of foundation, the requirement that \in be well-founded isn't redundant. In the absence of foundation, we could have a set $x = \{x\}$ which is clearly well-ordered by \in , but this isn't a strict order: $x \in x$.^{xiii} For the remainder of this section, we will not assume the axiom of foundation to show that the ordinals behave the same regardless. The well-founded property of membership on ordinals is used extensively in the arguments below. In essence, the results say that collection of ordinals themselves is linearly ordered by \in , rather than just each individual ordinal.

- **3A**•**2**. **Result** - Let α , β be ordinals. Therefore,

1. Any $\gamma \in \alpha$ is an ordinal.

2. $\alpha \in \beta \lor \alpha = \beta$ is equivalent to $\alpha \subseteq \beta$.

3. $\alpha \in \beta, \beta \in \alpha$, or $\alpha = \beta$.

4. $\alpha \cup \beta$ is an ordinal.

Proof .:.

- For $\delta \in \alpha$, suppose $y \in x \in \delta$. We know $y, x \in \alpha$. Since α is linearly ordered by \in , it follows that either $\delta \in y$ or $y \in \delta$. Clearly $\delta \in y$ is impossible by well-foundedness. Hence $y \in \delta$ verifies that δ is transitive. Anti-symmetry follows from antisymmetry on α : $\gamma \subseteq \alpha$. Similarly, totality follows from the totality on α .
- Clearly if α ∈ β or α = β then α ⊆ β by transitivity. So suppose α ⊆ β for α an ordinal, but that the conclusion fails: α ≠ β and α ∉ β. Without loss of generality, take β as the least failure in the sense that for each α' ∈ β, α ⊆ α' implies α ∈ α' or α = α' (to do this, take any ordinal β₀ witnessing the failure, and then consider the subset {β ∈ β₀ : β has it fail} and thus take a minimal element β by well-foundedness of ∈ on ordinals).

Consider $\beta \setminus \alpha$ as a subset of β . Since β is well-ordered by \in , there is a least element $\alpha' \in \beta \setminus \alpha$. Now suppose $\gamma \in \alpha$. Clearly $\gamma \in \alpha'$ by totality of \in on β . Hence $\alpha \subseteq \alpha'$. By minimality of $\beta, \alpha \in \alpha'$ or $\alpha = \alpha'$. Therefore $\alpha \in \beta$, a contradiction.

xiiiEnsuring the well-order is strict gets rid of these degenerative cases in the absence of foundation. But it also allows for the usual arguments to go through. A typical argument will be to consider the \in -least counter example α and conclude that for every $\beta \in \alpha$, β has the property we're after. This doesn't work if $\alpha \in \alpha = \{\alpha\}$ because we're critically assuming $\beta \neq \alpha$, and this is why \in being strict is important (although it's not very important).

Let α be fixed. Let β be an ordinal with α ∉ β, α ≠ β, and β ∉ α. Without loss of generality, take β as the least failure in the sense that for each α' ∈ β, α ∈ α', α = α' or α' ∈ α (to do this, just take any ordinal β₀ witnessing the failure, and then consider the subset {β ∈ β₀ : β has it fail} and thus take a minimal element β by well-foundedness).

Clearly if $\alpha \subseteq \alpha' \in \beta$ for any $\alpha' \in \beta$, then (2) yields that $\alpha \in \beta$. So then $\alpha' \in \alpha$ for every $\alpha' \in \beta$. But then $\beta \subseteq \alpha$ so that $\beta \in \alpha$ or $\beta = \alpha$ by (2), again, a contradiction.

• That $\alpha \cup \beta$ is transitive is immediate: any $y \in \alpha \cup \beta$ has $y \in \alpha$ or $y \in \beta$. So if $x \in y$, then $x \in \alpha$ or $x \in \beta$ and hence $x \in \alpha \cup \beta$. Well-foundedness follows from the property holding on α and on β : for any subset X, $\alpha \cap X$ has a minimal element α_X and $\beta \cap X$ has a minimal element β_X , and one of these must be minimal for $\alpha \cup \beta$. Antisymmetry is trivial. Totality follows from (1) and (3).

Some easy examples of ordinals can be gotten from Subsection 2 E. In particular, \emptyset is an ordinal, and we have the following result.

- **3A**•**3**. **Result** - Let α be an ordinal. Therefore $\alpha \cup \{\alpha\}$ is an ordinal.

Proof .:.

We know by Corollary 2 E • 5 that $\alpha \cup \{\alpha\}$ (or rather the memberhsip relation on it) is transitive. So all that suffices to be shown is antisymmetry, and totallity of \in . Since antisymmetry is vacuously true for well-founded relations, as in Corollary 2 E • 5, we only need to show totality. But this follows from Result 3 A • 2: all elements of $\alpha \cup \{\alpha\}$ are ordinals, and so can be related by \in .

In particular, for \emptyset , we have $\{\emptyset\}$, $\{\emptyset, \{\emptyset\}\}$, and so on as ordinals. To make the notation a bit nicer, we will use the extremely suggestive notation below.

3A•4. **Definition** For α an ordinal, write $\alpha + 1$ for $\alpha \cup \{\alpha\}$. Write $\beta < \alpha$ for $\beta \in \alpha$. Write 0 for \emptyset .

Hence 0, 1 = 0 + 1, 2 = 1 + 1, 3 = 2 + 1 are all ordinals. Note further that then every ordinal $\alpha = \text{pred}_{<}(\alpha)$ so that, for example, $5 = \{0, 1, 2, 3, 4\}$ (which has five elements). Note that the use of "+1" is appropriate here as a kind of successor operation.

– 3 A • 5. Corollary –

Let α be an ordinal. Therefore there is no ordinal β between α and $\alpha + 1$.

Proof .:.

Obviously, $\alpha \in \beta \in \alpha + 1$ requires $\beta = \alpha$ or $\beta \in \alpha$. Since $\alpha \notin \alpha$ by well-foundedness, we must have $\beta \in \alpha$, contradicting antisymmetry and that $\alpha + 1$ is an ordinal.

So far we are able to construct $n = 1 + \cdots + 1$ (*n* additions of 1) for each natural number *n*. But (provably) we can't show that the set of all of these ordinals exists from the axioms thus far. To do this, we must introduce the axiom of infinity: that there exists an infinite set of these.

– 3A•6. Definition (Axiom)

(*Infinity*) The set of natural numbers (or a set containing them) exists: $\exists N \ (\emptyset \in N \land \forall x \in N \ (x \cup \{x\} \in N))$.

The definition isn't able to properly say that the set of natural numbers exists without the notion of an ordinal. So we have to note the following result to then define the set of natural numbers. Clearly the result follows from foundation, but to get better acquainted with ordinals, we don't resort to this fact.

- **3A**•**7**. Theorem - For any non-empty set X of ordinals,

- sup $X := \bigcup X$ is an ordinal, and $\bigcup X \ge \alpha$ for each $\alpha \in X$;
- inf $X := \bigcap X$ is an ordinal, and $\bigcap X \le \alpha$ for each $\alpha \in X$;
- inf $X \in X$ so that inf $X = \min X$ is the minimum element of X.

Proof .:.

- It's clear that $\sup X \ge \alpha$ for each $\alpha \in X$. To see that $\sup X$ is an ordinal, transitivity follows from the transitivity of each ordinal in $X: x \in y \in \sup X$ has $y \in \alpha$ for some α and hence $x \in \alpha \subseteq \sup X$ implies $x \in \sup X$. Antisymmetry is trivial, and totality follows easily from Result 3 A 2.
- It should be clear that then $\inf X \leq \alpha$ for each $\alpha \in X$. To see that $\inf X$ is an ordinal, if $y \in x \in \alpha$ for each $\alpha \in X$, then $y \in \alpha$ for each $\alpha \in X$ so that $\inf X$ is transitive. Antisymmetry is again trivial, and totality is again easy to see as $\inf X$ is still a set of ordinals by Result $3 \land 2$.
- Since every $\alpha \in X$ has $\alpha \leq \sup X$, it's easy to see that $\alpha < \sup(X) + 1$ so that $X \subseteq \sup(X) + 1$. As \in is well-founded on $\sup(X) + 1$, it follows that X has a minimal element min X, which is an ordinal. As $\bigcap X \leq \min X$, by (2), it suffices to show min $X \subseteq \bigcap X$. But this is clear: every element $\alpha \in X$ has min $X \subseteq \alpha$ so that min $X \subseteq \bigcap X$. Hence min $X = \bigcap X \in X$.

Thus far, we've only seen ordinals where $\sup X = \max X \in X$ or else $X = \emptyset$. But this won't always be true in general. In fact, there is a whole class of ordinals where this is false. Such ordinals are called limit ordinals, and in fact all ordinals can be broken down into limits or successors (or 0). As a hint of what to come, the set of natural numbers will be a limit ordinal, and in fact the least such.

- 3A•8. Definition

Let $\alpha \neq 0$ be an ordinal. α is a *successor* ordinal iff $\alpha = \beta + 1$ for some ordinal β . α is a *limit* ordinal iff $\alpha = \sup \alpha$.

This classifies all ordinals.

- **3A**•9. Theorem Let α be an ordinal. Therefore $\alpha = 0$, or $\alpha = \sup(\alpha) + 1$, or $\alpha = \sup \alpha$.

Proof .:.

Let $\alpha \neq 0$. If $\alpha = \sup \alpha$, then for each $\beta < \alpha$, there is an $\gamma < \alpha$ with $\beta < \gamma$. In particular, by Corollary 3 A • 5, $\beta + 1 < \alpha$. So it's easy to see that $\alpha = \sup \alpha$ is equivalent to $\forall \beta < \alpha \ (\beta + 1 < \alpha)$. So if $\alpha \neq \sup \alpha$, there is some $\beta < \alpha$ with $\beta + 1 \neq \alpha$. Thus $\beta < \alpha \leq \beta + 1$ so by Corollary 3 A • 5, $\alpha = \beta + 1$. But then for every $\gamma < \alpha, \gamma \leq \beta$, implying $\beta = \sup \alpha$ and thus $\alpha = \sup(\alpha) + 1$.

Let's now collect the major properties of ordinals that we know so far.

3A • 10. Theorem For all ordinals α, β, 1. α is a set of ordinals; 2. α = 0, α is a successor ordinal, or α is a limit ordinal; 3. 0 is the least ordinal; 4. the ordinals are well-ordered by ∈; 5. α ∪ β = max(α, β); 6. α ∩ β = min(α, β); 7. α ≤ β iff α ⊆ β (although not all sets x ⊆ β are ordinals);

8. $\inf \alpha \leq \sup \alpha \leq \alpha < \alpha + 1$, and for $\alpha > 0$, $\inf \alpha < \sup \alpha$.

Well-orders and Ordinals

Proof .:.

- 1. follows from Result $3 \mathbf{A} \cdot 2(1)$.
- 2. follows from Theorem 3 A 9.
- 3. follows from Definition 3 A 4: $\gamma < 0$ implies $\gamma \in \emptyset$, which is always false.
- 4. has linearity follow from transitivity and Result 3 A 2 (2). To show well-foundedness, let X be a nonempty set (or class) of ordinals. Taking $\alpha \in X$ yields that $X \cap (\alpha + 1) \subseteq \alpha + 1$ which then has a least element $\beta \in X \cap (\alpha + 1)$. Any least element $\gamma \in X$ must have $\gamma \leq \alpha$ and thus $\gamma \in X \cap (\alpha + 1)$ so that β is the least element of X.
- 5. follows from Result $3 \mathbf{A} \cdot 2 (4)$ and (2).
- 6. follows from Theorem 3 A 7.
- 7. follows from Result $3 A \cdot 2 (4)$.
- 8. follows from Theorem 3 A 7.

Now formally, we've defined a well-order to be a certain kind of set, which would make (4) false: the collection of all ordinals doesn't constitute a set. But it's easy to see what is meant by \in well-ordering the ordinals (just the defining conditions without the additional requirement that the relation— \in here—be a set).

- 3A•11. Result (Burali-Forti Paradox)

 $\neg \exists s \ \forall x \ (x \text{ is an ordinal} \rightarrow x \in s)$. Informally, the collection Ord of all ordinals is not a set. In particular, there is no largest ordinal.

Proof .:.

There is no largest ordinal, since the largest ordinal α has $\alpha + 1 > \alpha$ by the reasoning above: $\alpha + 1 = \alpha$ implies $\alpha \in \alpha$, contradicting well-foundedness (even a set { α } has no least element, since a least element β requires $\forall z \in \beta \ (z \notin \alpha)$, which isn't true for $\beta = \alpha$).

To show that Ord can't be a set, by Theorem 3 A • 10, \in well-orders Ord. Since each $\alpha \in$ Ord is transitive, it follows that $\alpha \subseteq$ Ord and hence Ord is transitive. Therefore Ord is an ordinal. But then Ord is the largest ordinal, contradicting the idea above. \dashv

Let's return to the idea of natural numbers. Notice that by our classification, every natural number is a successor ordinal, and in particular is of the form $0 + 1 + \cdots + 1$ for some (natural) number of +s.

- **3A** • **12**. **Definition** Write ω for the least limit ordinal, the set of natural numbers.

To see why ω should be the set of natural numbers, note that the supremum of the natural numbers must be a limit ordinal: *n* is a natural number implies n + 1 is too, so if $n < \sup \mathbb{N}$ then $n + 1 < \sup \mathbb{N}$, meaning $\sup \mathbb{N}$ is a limit ordinal. Moreover, $\sup \mathbb{N}$ must be the least limit ordinal, since every $n < \sup \mathbb{N}$ is a natural number, which means it's either a successor or 0. So this implies $\omega = \mathbb{N}$, but we haven't shown that ω actually exists, yet.

- **3A** \cdot **13**. **Result** - The set ω , the least limit ordinal, exists.

Proof .:.

Let N be as in the axiom of infinity. Take the subset $N' = \sup\{\alpha \in N : \alpha \text{ is an ordinal}\}$ so that N' is an ordinal. We need to show that $\omega \leq N'$. If N' has a limit ordinal below it, then clearly ω is least by definition. So if N' has no limit ordinals below it, we want to show that $N' = \omega$.

Let $\alpha \in N'$ be the least such that $\alpha \in \omega \setminus N'$. As ω is the least limit ordinal, α must be a successor or 0. If $\alpha = 0$, then $0 \in N'$ by hypothesis that $1 \in N$ so $0 < 1 \le N'$. If $\alpha = \beta + 1$, then $\beta \in \omega$. By the minimality of β , $\beta \in N'$ so that by the hypothesis on N, $\beta + 1 = \alpha \in N$ and hence $\alpha + 1 \in N$ so that $\alpha < N'$, a contradiction. Therefore $\omega \subseteq N'$. But then as ordinals, $\omega < N'$ or $\omega = N'$. Since N' has no limit ordinals below it, $N' = \omega$. \dashv

Well-orders and Ordinals

It would seem that the reasoning alone gives the existence of ω , but really the idea only only characterizes ω . We still need the existence of such an N as in the axiom of infinity to ensure the existence of ω .

With the natural numbers at our fingertips, we can show that ω satisfies all the usual properties that we want, namely the axioms of peano arithmetic, PA. To do this, we need a notion of addition and multiplication of ordinals. To do this, we need a better way of defining operations on ω .

As a side note, we have a characterization of ω in meta-theoretic terms (able to be reached from 0 by finite applications of adding 1). What we've done now is show that in **V**, this coincides with the characterization of ω as the least limit ordinal. This formal characterization, however, isn't necessarily the set of natural numbers. Consider the following from model theory: in the language FOL(\in , c) where c is a constant symbol, the theory of set theory adjoined with " $\omega > c > 1 + \cdots + 1$ " (n times) for each real-world natural number $n \in V$ yields a theory T_n that is consistent assuming that set theory is consistent (just interpret c as n + 1 in **V**). Hence every finite subset of the theory $T = \{\varphi : \varphi \text{ in } T_n \text{ for some } n \text{ a natural number}\}$ is consistent so that T itself has a model by Compactness (1 D • 2). But in this model $\mathbf{M} \models T$, we have $\mathbf{M} \models "\omega > c^{\mathbf{M}} > 1 + \cdots + 1$ " (n times) for each real-world natural number $n < \omega$. So $\omega^{\mathbf{M}}$ can't be the same as ω in the real-world **V**. All of this is to say that we must be careful about using our *intuitive*, *meta-theoretic* characterization of ω to *formally* prove things about it from set theory. To ensure that we can prove all of the intuitive properties of ω *formally*, we resort to the principle of induction.

§3B. Finitary recursion and induction

Recall the defining property of ω : if $0 \in \omega$, and $n \in \omega$ then $n + 1 \in \omega$ (and this is all there is in ω). In particular, this yields the following result, called the principle of induction.

- 3B•1. Theorem (Induction on ω) -

Let $\varphi(x)$ be a FOLp(\in)-formula. Suppose $\varphi(0)$ and $\varphi(n) \rightarrow \varphi(n+1)$. Therefore $\forall n \in \omega \varphi(n)$.

Proof .:.

Consider the set $X = \{n \in \omega : \neg \varphi(n)\}$ and suppose $X \neq \emptyset$. This has a least element $x \in X$. Note that $x \neq 0$ by the hypothesis. Since ω is the least limit ordinal, $x = \sup(x) + 1$ is a successor. But by minimality, $\varphi(\sup(x))$ holds and so $\varphi(\sup(x) + 1)$ holds, contradicting that $\sup(x) + 1 = x \in X$.

Really, this is just a consequence of ω being well-ordered. But this reflects the properties of arithmetic that ω should have. The key thing here is that by specifying what happens at 0, and what happens at successor stages, we can define something on all of ω . This idea is referred to as *recursion*.

The formal statement of recursion is long and clunky. So to better understand it, we give some examples. Firstly, we would normally define addition by n by $f_n(x) = x + 1 + \dots + 1$ where we add 1 ns. The issue with this is that this definition is informal and meta-theoretic, in some sense. It's not clear how we would define this function purely in terms of set theory without resorting to "n-times". Surely for each x this makes sense, but the map sending $n \mapsto f_n$ isn't so obviously well defined (consider non-standard models with different ω s). To get around this, for each $x < \omega$ consider the map defined by $f_x(0) = x$ and $f_x(n+1) = f_x(n) + 1$ for all $n < \omega$. Using induction, any functions that satisfy this agree everywhere so this defines f_x on all $n < \omega$. Moreover, intuitively, this f_x satisfies $f_x(n) = x + n$.

Once we have f_x for each $x < \omega$, we can consider the map sending $\langle x, n \rangle$ to $f_x(n)$. This map, call it '+', sends $\langle x, n \rangle$ to x + n in the usual sense.

To define this whole process more formally, what we're doing is specifying what happens at the start, and then what happens at successor stages. So we are given functions f and g, and we define the function h starting with f(0), and finding the next values based on g and the previous value: for $n < \omega$,

$$h(0) = f(0)$$
$$h(n+1) = g(n, h(n)).$$

So to calculate h(2), we start with h(0) = f(0), and then calculate h(1) = g(1, f(0)), and then calculate h(2) = g(2, g(1, f(0))). In principle, we could then keep going to define h(4), h(5), and so on, meaning h(n) will be some particular number for each n. This means the function h is determined by these conditions in the sense that it is the unique function satisfying them. Formally, we have the following theorem. The proof of this theorem is very technical, and long, and not terribly illuminating, mostly just making precise and formal the intuitive idea of "starting and 0 and defining what happens next determines it on all of the natural numbers". It is included for those interested in the precise details, but for those uninterested, it can be skipped.

$-3B\cdot 2$. Theorem (Recursion on ω)

Let f with $0 \in \text{dom}(f)$ be a function. Let g be a function from ordered pairs with the first entries being natural numbers: $\omega = \text{dom}(\text{dom}(g))$. Therefore, there is a unique function h where $\text{dom}(h) = \omega$ and for all $n < \omega$,

$$h(0) = f(0)$$
$$h(n+1) = g(n, h(n)).$$

Proof .:.

To show existence, we proceed by induction to show that for each $n \in \omega$, $\varphi(n, h)$ defines a unique function h_n , which is supposed to represent $h \upharpoonright n$. Once we do this, we pull together all of the h_n s to define h.

Consider the formula $\psi(n, h)$ given formally below:

$$\operatorname{dom}(h) = n < \omega \land \forall k < n \left(\begin{array}{c} (k = 0 \land \langle 0, f(0) \rangle \in h) \\ \lor \exists ! v \exists m \ (k = m + 1 \land \langle m, v \rangle \in h \land \langle k, g(m, v) \rangle \in h) \end{array} \right).$$

Informally, $\psi(n, h)$ says

h is a function with domain *n* and obeys the recusive definition up to *n*.

One may easily check the following facts:

1. if $\psi(n, h)$, then h is a function;

2. if $\psi(n, h)$ and $m \leq n$, then $\psi(m, h \upharpoonright m)$; and

3. if $\psi(n, h)$ for $n = n^* + 1$, then $\psi(n + 1, h \cup \{\langle n, g(n^*, h(n^*)) \rangle\})$.

We want to now show that for each $n < \omega$, there is exactly one h with $\psi(n, h)$. This will allow us to use replacement to collect all of these approximations to the h of the theorem together.

```
- \text{ Claim 1} - \frac{1}{\forall n < \omega} \exists ! h \ \psi(n, h).
```

Proof .:.

There are two parts to this: the existence of h, and the uniqueness of h. Existence holds by induction: since $h_0 = \emptyset$ exists trivially, and h_{n+1} satisfying $\psi(n + 1, h_{n+1})$ exists by (3) above. So induction shows that for each $n < \omega$, there exists such an h where $\psi(n, h)$.

To show there is at most one *h* with $\psi(n, h)$, let $n + 1 < \omega$ is the least where this fails (it vacuously holds for n = 0). Thus we have two functions $h_0 \neq h_1$ where $\psi(n + 1, h_0)$ and $\psi(n + 1, h_1)$. Note by (2) above, $\psi(n, h_0 \upharpoonright n)$ and $\psi(n, h_1 \upharpoonright n)$ hold. So by the minimality of $n + 1, h_0 \upharpoonright n = h_1 \upharpoonright n$. So the only place the two functions can differ is at $n: h_0(n) \neq h_1(n)$. But in satisfying ψ , we must have that for k = n = m + 1, $\langle k, g(m, h_i(m)) \rangle \in h_0, h_1$, i.e. $h_0(n) = g(n, h_0(m)) = g(n, h_1(m)) = h_1(n)$, a contradiction. \dashv

Thus by replacement, we have the set $\{h_n : n \in \omega\}$ where $\psi(n, h_n)$ for each $n < \omega$. Therefore $\bigcup_{n \in \omega} h_n = h$ is a function with domain ω , and for each $n < \omega$, h satisfies $\psi(n, h \upharpoonright n)$. Thus $h(0) = (h \upharpoonright 1)(0) = f(0)$ and $h(n + 1) = (h \upharpoonright n + 2)(n + 1) = g(n, (h \upharpoonright n + 2)(n)) = g(n, h(n))$, showing that h shows the existence of such a function as in the theorem statement.

Now for uniqueness, suppose $h' \neq h$ also satisfied the hypothesis. Therefore for each $n < \omega$, $\psi(n, h' \upharpoonright n)$ holds so that uniqueness of the parts yields $h' \upharpoonright n = h \upharpoonright n$ for each $n < \omega$. Hence $h'(n) = (h' \upharpoonright n + 1)(n) = (h \upharpoonright n + 1)(n) = h(n)$ for each $n < \omega$. Thus h' = h. The above theorem isn't actually given in its fullest generality: we are allowed more variables. As long as the order we proceed through the tuples in is well-founded, we are guaranteed the result by the same idea as above. Another example would be to consider building a *tree* of finite length sequences of 0s and 1s. We can proceed by the above idea to define $h(\tau)$ for any τ in the tree by breaking down into cases: defining $h(\tau^{(1)})$ and $h(\tau^{(1)})$ for arbitrary τ gives a definition to h generally: $h(\langle 0, 1, 0, 0, 1 \rangle)$ is given by looking at $h(\emptyset)$, then looking at $h(\emptyset^{(1)})$, then $h(\langle 0 \rangle^{(1)})$, and so on. The proof of the existence and uniqueness of h is exactly the same.

— 3 B•3. Corollary

Let X be a set, and let T be the tree of finite sequences of elements of X. Let f with $\emptyset \in \text{dom}(f)$ be a function. Let g be a function of the form $g: X \times T \times A \to B$ for some sets $A \subseteq B$. Therefore there is a unique function $h: T \to A$ where for all $\tau \in T$ and $x \in X$,

$$\begin{split} h(\emptyset) &= f(\emptyset) \\ h(\tau^{\frown}\langle x \rangle) &= g(x,\tau,h(\tau)). \end{split}$$

Going beyond this is more difficult, because it's unclear how to deal with limit stages with just the above information. So we must consider the *transfinite* versions of these.

§3C. Transfinite recursion and induction

The existence of limit ordinals is incredibly powerful, as it allows us to form larger and larger ordinals beyond just ω . To go further, we need a better way of defining or constructing these ordinals. To do this, we use the notion of transfinite recursion and induction. Intuitively, $\omega + 1$, $\omega + 2$, $\omega + 3$, and so on have all been defined. If we wish to define $\omega + \omega$, we could do this as the least limit ordinal after ω , but this clumsy characterization isn't sustainable to define $\alpha + \beta$ for general ordinals $\alpha + \beta$. To do this, we use the characterization of ordinals into 0, successors, and limits. If we specify the definition at 0, at successors, and at limits, we will have defined it everywhere. The idea of transfinite recursion makes this explicit.

Again, first we have the fundamental property that allows us to do this: transfinite induction. The idea was already noticed in Theorem $3 \land 10$ (4). But to make it explicit, we have the following theorem.

— 3C•1. Theorem (Transfinite Induction) –

Let $\varphi(x, \vec{w})$ be a FOL(\in)-formula with \vec{v} parameters. Suppose $\varphi(\alpha, \vec{v})$ holds whenever $\forall \beta < \alpha \ \varphi(\beta, \vec{v})$. Therefore for every ordinal $\alpha, \ \varphi(\alpha, \vec{v})$.

Proof .:.

Otherwise, take α the least such that $\neg \varphi(\alpha, \vec{v})$). Thus for every $\beta < \alpha, \varphi(\beta, \vec{v})$. Hence by hypothesis $\varphi(\alpha, \vec{v})$, a contradiction.

This also applies to the natural numbers, but stated this way allows us to incorporate limit ordinals. If we had simply left the same sort of statement as in Induction on ω (3 B • 1), we wouldn't necessarily have the result for ω , much less all ordinals α . In particular, consider the property of being 0 or a successor ordinal. Clearly this holds for 0 and if it holds for α , it holds for $\alpha + 1$. But this never allows one to reason their way to the limit ordinals: only successors of successors and so on.

To make the notion of transfinite recursion formal, we need three functions specifying what happens at stage 0, what happens at successor stages, and what happens at limit stages. This idea of breaking down into cases proceeds in precisely the same way as in Recursion on ω (3 B • 2). But there is a slightly easier way to state it formally. Rather than breaking down into more and more cases with more and more classifications, the main idea of recursion is just that we can calculate the next value from the previous ones. So the value at ω should be determined by the values on all $n < \omega$. Stated formally, this yields the much more compact version below.

— 3C•2. Theorem (Transfinite Recursion) -

Let α be an ordinal. Let f be a function, writing $f(x) = \emptyset$ for $x \notin \text{dom}(f)$. Therefore there is a unique function g with domain α such that for all $\beta < \alpha$, $g(\beta) = f(g \upharpoonright \beta)$.

Proof .:.

Assuming existence, uniqueness follows easily by induction on α . For α the least such where this fails, there are then functions $g \neq g'$ where $g(\beta) = f(g \upharpoonright \beta)$ and $g'(\beta) = f(g' \upharpoonright \beta)$ for all $\beta < \alpha$. But by minimality of α , $g' \upharpoonright \beta = g \upharpoonright \beta$ so that $g(\beta) = f(g' \upharpoonright \beta) = g(\beta)$, meaning $g(\beta) = g'(\beta)$ for all $\beta < \alpha$, and thus g = g'.

To show existence, proceed as in Recursion on ω (3 B • 2). In particular, consider the formula $\psi(\beta, g)$ which says that g is a function with domain β and $\forall \gamma < \beta$ ($g(\gamma) = f(g \upharpoonright \gamma)$). By induction on β , we can show $\exists ! g \ \psi(\beta, g)$. To see this, let β be least where this fails. Hence for each $\gamma < \beta$, $\exists ! g \ \psi(\gamma, g)$. By replacement we get a set { $g_{\gamma} : \gamma < \beta$ }. One can easily see that $\psi(\gamma, g)$ implies $\psi(\delta, g \upharpoonright \delta)$ for any $\delta < \gamma < \beta$. Hence the union $g = \bigcup \{g_{\gamma} : \gamma < \beta\}$ is a function with domain β , and one can easily check that for each $\gamma < \beta$, $g(\gamma) = f(g \upharpoonright \gamma)$. Uniqueness follows from the uniqueness of each $g_{\gamma} = g \upharpoonright \gamma$ as in Recursion on ω (3 B • 2). \dashv

The idea above actually extends to Ord in the sense we can get define an output on every ordinal α . Although we won't get a g such that Ord \subseteq dom(g) (since g needs to be a set), we can still define what the output will be at any given α by considering the resulting function with domain $\alpha + 1$. Uniqueness ensures that this output doesn't vary with the change in domain. So it makes sense to say that this defines a function on all of Ord, even though only the approximations to this function exist. Formally, we might say $\varphi(\alpha, y)$ holds iff $\exists g$ (g is a function with dom(g) = $\alpha + 1 \land \forall \beta < \alpha + 1(g(\beta) = f(g \upharpoonright \beta)) \land y = g(\alpha)$). The reasoning above tells us that $\forall \alpha \in \text{Ord } \exists ! y \varphi(\alpha, y)$. So this is the sense in which we have defined a function on all of Ord.

§3D. A word on sequences and functions

Although much of this section has been stated in terms of functions, it's perhaps most intuitive to think of functions from ordinals as sequences: for each entry in a sequence, there is a subsequent entry, and there should always be a least point in the sequence where something happens. In most other branches of math, the only sequences that appear are those of length ω , or else finite.

- 3D•1. Definition

A sequence is a function f with dom(f) as an ordinal (or dom(f) = Ord, in which case f is a class). The length of a sequence is its domain.

This notion of a sequence is incredibly important if we want to define functions with more than just finitely many inputs. Thus far, if we wanted a function from tuples in A, B, and C to D, we'd need to consider $f : A \times B \times C \to D$. The introduction of sequences allows us to consider tuples instead as sequences: (a, b, c) can be identified with the function $f : 3 \to A \cup B \cup C$ where f(0) = a and f(1) = b and f(2) = c, identifying each entry with *where* it is in the tuple. And we can generalize this, allowing us to talk about infinite products.

- 3D•2. Definition

Let I be a set, and suppose $\{A_i : i \in I\}$ is a family of sets. Therefore the *cartesian product* $\prod_{i \in I} A_i$ is the set of functions $f : I \to \bigcup_{i \in I} A_i$ such that $f(i) \in A_i$ for each $i \in I$.

In particular, for α an ordinal, we write $A^{\alpha} = \prod_{\beta < \alpha} A$, generalizing $A^n = A \times \cdots \times A$ (*n* times) for $n < \omega$. Note that the finite product of non-empty sets is non-empty. That infinite products of non-empty sets are non-empty is equivalent to an axiom yet to be introduced. We will have no need of it for now, but it should be noted.

Really, the inherent notion of a sequence just comes from any well-order. So we should investigate further what wellorders exist. As it turns out, the ordinals will exhaust all the well-orders in V.

§3E. The model theory of well-orders

We have defined what it means for a structure $A = \langle A, R \rangle$ to be a well-order: R well-orders $A = \text{dom}(R) \cup \text{ran}(R)$. This property, however, is not expressible in first-order logic alone. To see this, we use compactness and the existence of ω .

- 3E•1. Result

Let σ be a signature with a binary relation symbol R. Let T be a FOL(σ)-theory such that T contains the axioms of partial orders. Therefore, if T has an infinite, well-ordered model, then T has an ill-founded (i.e. not well-founded) model. Hence being a well-order isn't FOL-expressible.

Proof .:.

Let $A \models T$ be an infinite well-order. Using Recursion on ω (3 B • 2), define an infinite R^A -increasing sequence: let a_0 be the R^A -minimal element of A. By well-foundedness, for n + 1, let a_{n+1} be the R^A -least element $a \in A$ such that

$$\forall i \leq n \ (\mathbf{A} \vDash "a_i \ R^{\mathbf{A}} \ a").$$

Now in the expanded signature $\mathcal{L} = \{R\} \cup \{c_n : n \in \omega\}$ with new constant symbols c_n for each $n < \omega$, consider the theory $T' = T \cup \{c_{n+1} \ R \ c_n : n < \omega\}$. Note that any model of T' is ill-founded. Since any model of T' is also a model of T, it suffices to show that T' is consistent. To do this, we use Compactness (1 D • 2).

For any finite subset $\Delta \subseteq T'$, there is a largest $n < \omega$ where c_n occurs in a formula of Δ . Taking N to be this, we can interpret Δ in the expansion A' of A where c_n is interpreted as a_{N-n} : $A \models ``a_{N-(n+1)} R a_{N-n}$ '' so clearly $A' \models ``c_{n+1} R c_n$ ''. Since $A' \models T$, it follows that $A' \models \Delta$, and thus Δ is consistent. As Δ was arbitrary, T' is consistent. By Completeness (1 D • 1), there is a model $B \models T'$ which then models T, but is ill-founded. \dashv

So the property of being a well-order is a property of the the set theoretic universe. Depending on the (non-V) model of set theory, certain sets may or may not be well-founded, because the models don't have the set witnessing the ill-foundedness. This is a weakness of first-order logic, but it is no challenge to the legitimacy of the concept. Really, this idea just expresses the inadequacy of first-order formulas to properly characterize these notions. This is a common part of logic, as even group theory is subject to the limitation: the property of being a cyclic group isn't first-order expressible, for example. This is merely something we must live with.

Clearly, however, being a well-order is preserved under isomorphisms. In fact, our goal here will be to show that the ordinals are the *canonical* well-orders in the sense that every well-order is isomorphic to a particular ordinal (under membership). To do this, we proceed in a similar way as when we introduced ordinals. Before this, we introduce some definitions that should be familiar from model theory.

- 3E•2. Definition

Let $\mathbf{A} = \langle A, \langle A \rangle$ and $\mathbf{B} = \langle B, \langle B \rangle$ be structures where $\langle A \rangle$ and $\langle B \rangle$ are relations.

• A function $f : A \to B$ is a homomorphism iff $a <_A a' \to f(a) <_B f(a')$ for every $a, a' \in A$.

• A function $f : A \to B$ is an *embedding* iff $a <_A a \leftrightarrow f(a) <_B f(a')$ for every $a, a' \in A$ and f is injective.

• A function $f: A \to B$ is an *isomorphism* iff f is an embedding, and f is surjective.

If A is a linear order, an *initial segment* of A is a substructure with universe $\text{pred}_{\leq_A}(a_0) = \{a \in A : a \leq_A a_0\}$ for some $a_0 \in A$.

So for each ordinal α , pred_{\in}(α) = α . Now we consider the following result about well-orders. Note that for $X \subseteq A$ and $<_A \subseteq A \times A$, we continue to write $\langle X, <_A \rangle$ for the sake of readability when really we mean $\langle X, <_A \cap (X \times X) \rangle$. Note that if **A** is a well-order, then its initial segments are well-orders too.

- 3E•3. Lemma

Let $\mathbf{A} = \langle A, \langle A \rangle$ be a well-order. Let $a \in A$. Write $\operatorname{pred}_{\langle A}(a)$ for $\{x \in A : x \langle A a \}$. Therefore $\langle \operatorname{pred}_{\langle A}(a), \langle A \rangle$ is a well-order.

This can be seen just by noting that all of the properties are inhereted from the well-order on A: transitivity, antisymmetry, and totality all hold since we're taking all variables in $\text{pred}_{<_A}(a)$, and well-foundedness also clearly holds, since

we're taking a subset of $\operatorname{pred}_{<_4}(a)$. In fact, for any subset $X \subseteq A$, $\langle X, <_A \rangle$ is well-founded if $\langle A, <_A \rangle$ is.

- **3E**•4. Lemma Let $\mathbf{A} = \langle A, \langle A \rangle$ be a well-order. Therefore, $\mathbf{A} \ncong \langle \operatorname{pred}_{\langle A}(a), \langle A \rangle$ for any $a \in A$.

Proof .:.

Let $f : A \to A_{<a}$ be an isomorphism where $\mathbf{A}_{<a} = \langle \operatorname{pred}_{<A}(a), <_A \rangle$. Consider the subset $X = \{x \in A : f(x) \neq x\}$. Note that X is non-empty, since $a \in X$, for example: $a \notin A_{<a}$ cannot be in the image of f. Consider the $<_A$ -least element $x \in X$ so that f" $\operatorname{pred}_{<_A}(x) = \operatorname{pred}_{<_A}(x)$, but $f(x) \neq x$. By injectivity, it follows that $f(x) \neq_A x$ and thus $f(x) >_A x$ by totality. Since f is an isomorphism, there must be some $x' \in A$ where x = f(x'). But then $\mathbf{A}_{<a} \models "f(x) >_A f(x')$ " requires $\mathbf{A} \models "x >_A x'$ " as an embedding. But by minimality, this implies $f(x') = x' \neq x$, a contradiction.

Using this, we get the following, which will allow us to show that any two well-orders can be compared in the sense that they are either isomorphic to each other, or to an initial segment. In particular, when we restrict an isomorphism to an initial segment, we get an isomorphism between initial segments.

— 3E•5. Lemma

Let $\mathbf{A} = \langle A, \langle A \rangle$ and $\mathbf{B} = \langle B, \langle B \rangle$ be well-orders. Let $f : A \to B$ be an isomorphism. Therefore, for any $a \in A$, $f \upharpoonright \text{pred}_{\langle A}(a)$ is an isomorphism between $\langle \text{pred}_{\langle A}(a), \langle A \rangle$ and $\langle \text{pred}_{\langle B}(b), \langle B \rangle$ for some $b \in B$.

Proof .:.

Write $\mathbf{A}_{<a}$ for $\langle \operatorname{pred}_{<A}(a), <_A \rangle$ and similarly for $b \in B$. Let $a \in A$ be $<_A$ -least such that the result fails. Let b be the least element of $B \setminus f''A_{<a}$. We will show that $f''A_{<a} = B_{<b}$, and thus that $f \upharpoonright A_{<a}$ is an isomorphism between $\mathbf{A}_{<a}$ and $\mathbf{B}_{<b}$.

By minimality, $B_{<b} \subseteq f^{"}A_{<a}$, so suppose the reverse doesn't happen: there is some $a_0 \in A_{<a}$ with $f(a_0) >_B b$. As an isomorphism, there is some $a' \in A$ with b = f(a') so that $\mathbf{B} \models "f(a_0) >_B f(a')$ ". As an embedding, this means $\mathbf{A} \models "a_0 >_A a'$ " so that $a' \in A_{<a}$, contradicting that $b \notin f^{"}A_{<a}$.

- **3E**•6. Lemma Let $\mathbf{A} = \langle A, \langle A \rangle$ and $\mathbf{B} = \langle B, \langle B \rangle$ be two well-orders. Suppose $\mathbf{A} \ncong \mathbf{B}$, and \mathbf{B} is not isomorphic to an initial segment of \mathbf{A} . Therefore there is a unique $b_0 \in B$ with $\mathbf{A} \cong \langle \text{pred}_{\langle B}(b_0), \langle B \rangle$.

Proof .:.

Write $A_{<a}$ for $\langle \text{pred}_{<A}(a), <_A \rangle$ and similarly for $b \in B$. Uniqueness clearly holds by Lemma 3 E • 4: $A \cong B_{<b_0}$ and $A \cong B_{<b_1}$ implies $B_{<b_1} \cong B_{<b_0}$. So if $b_1 \neq b_0$, then $b_0 <_B b_1$ or $b_1 <_B b_0$, and we contradict Lemma 3 E • 4 in either case.

Now suppose existence fails. Without loss of generality, let **A** be minimal in the following sense: for every $a \in A$, there is a unique $b \in B$ such that $\mathbf{A}_{< a} \cong \mathbf{B}_{< b}$. (Otherwise just choose the least $a \in A$ where this fails, and consider the structure $\mathbf{A}_{< a}$ instead. This new structure still has **B** not isomorphic to an initial segment, nor isomorphic to it as a whole.) So let $f = \{\langle a, b \rangle : \mathbf{A}_{< a} \cong \mathbf{B}_{< b}\}$ be the function such that $\mathbf{A}_{< a} \cong \mathbf{B}_{< f(a)}$. Note that f must be injective since if $x <_A y$, then $\mathbf{A}_{< x}$ is an initial segment of $\mathbf{A}_{< y}$: thus $\mathbf{B}_{< f(x)} \cong \mathbf{A}_{< x} \not\cong \mathbf{A}_{< y} \cong \mathbf{B}_{< f(y)}$ by Lemma 3 E • 4.

— Claim 1

f is an embedding. Given that f is already injective, we mean $x <_A y \to f(x) <_B f(y)$ for all $x, y \in A$.
Proof .:.

Otherwise, $f(x) \ge_B f(y)$ so that $A_{<y} \cong B_{< f(y)}$ is an initial segment of $B_{< f(x)} \cong A_{<x}$. Composing the isomorphisms, we get that $A_{<y}$ is isomorphic to an initial segment of $A_{<x}$, contradicting Lemma 3 E • 4. Explicitly, take $f_y : A_{<y} \to B_{< f(y)}$ and $f_x : B_{< f(x)} \to A_{<x}$ to be isomorphisms. By Lemma 3 E • 5, $f_x \upharpoonright B_{< f(y)}$ is an isomorphism with an initial segment $A_{<a} \subseteq A_{<x}$ so that $f_x \circ f_y : A_{<y} \to A_{<a}$ is an isomorphism. \dashv

So all that suffices is to show that f is surjective onto some initial segment. f is an isomorphism between A and im f. Taking b_0 the least element of $B \setminus \text{im } f$, we get that $B_{< b_0} \subseteq \text{im } f$ by minimality of b_0 . To show that im $f \subseteq B_{< b_0}$, suppose $f(a_0) >_B b_0$ so that there is an isomorphism $g : B_{< f(a_0)} \to A_{< a_0}$. Thus $g \upharpoonright B_{< b_0}$ is an isomorphism between $\mathbf{B}_{< b_0}$ and $\mathbf{A}_{< a}$ for some $a \in B$ by Lemma $3 \mathbf{E} \cdot 5$. But then $b_0 = f(a)$ contradicts that $b_0 \notin \text{im } f$. Hence im $f \subseteq B_{< b_0}$, and so we have equality, and thus f is an isomorphism.

Stated more loosely, for any two well-orders, either they are isomorphic, or one is isomorphic to an initial segment of the other. As a corollary of this, the ordinals exhaust all of the well-orderings in V.

— 3E•7. Corollary -

For every well-order A, there is a unique ordinal α such that $A \cong \langle \alpha, \in \rangle$.

Of course, Ord is well-ordered by \in , but Ord \notin V by Burali–Forti Paradox (3 A • 11), so this isn't an issue: every quantifier ranges over sets and we're only considering structures in V while (Ord, \in) \notin V.

- **3E**•8. **Definition** Let **A** be a well-order. The *order-type* of **A** is the unique ordinal α with $\mathbf{A} \cong \langle \alpha, \in \rangle$.

More than just getting a unique order-type, we also get that the isomorphism is unique.

— 3E•9. Result

Let **A** be a well-order, and $f : A \to \alpha$ and $g : A \to \alpha$ isomorphisms. Therefore, f = g.

Proof .:.

Assume not, and let $a \in A$ be $<_A$ -minimal such that $f(a) \neq g(a)$. For the sake of definiteness, assume f(a) < g(a). Since g is an isomorphism, there is some $b \in A$ where g(b) = f(a) < g(a). In other words, $\mathbf{V} \models "g(b) < g(a)"$ so that as an embedding, $\mathbf{A} \models "b <_A a"$ so by minimality of a, f(a) = g(b) = f(b), contradicting that f is injective.

There are, of course, other questions one can ask of well-orders in the context of model theory, like when two ordinals are elementarily equivalent under membership, for example. But for now, we will only make use of the fact that well-orders are isomorphic to ordinals.

Section 4. Other Well-founded Relations

Recall Axiom $(2 E \cdot 4)$, the axiom of foundation. To further motivate why this axiom should be true, we will show the following result, which holds even in the absence of foundation. In essence, the result says that all well-founded models of set theory in V are isomorphic to transitive sets. So the axiom of foundation in some sense takes the converse to be true: all transitive sets are well-founded.

– 4 • 1. Theorem (The Mostowski Collapse)

Let $\mathbf{A} = \langle A, \langle A \rangle$ (in V) be well-founded such that \mathbf{A} satisfies the axiom of extensionality. Therefore $\mathbf{A} \cong \langle T, \epsilon \rangle$ for a unique transitive set T.

Although we can prove the theorem outright at this point, to get a better perspective on what is going on with the proof, we will introduce a useful idea: rank. Although all well-orders are isomorphic to ordinals, well-founded, extensional structures are not in general. But they can still make use of ordinals according to *chains*, which are then well-ordered. Really, this just means indexing the *levels* of the structure like with a tree.

The most fundamental idea behind rank functions is given by Transfinite Recursion $(3 \text{ C} \cdot 2)$, and so often we want the process to stop at some ordinal. The following lemma, a consequence of the axiom of replacement, will be useful in doing this. Note that the lemma further reinforces the idea that some collections are simply "too big" to be sets. In essence, we will use this to say that there can't be Ord-many levels of a well-founded set.

— 4•2. Lemma

Let A be a set. Therefore there is no surjection $f : A \rightarrow \text{Ord.}$

Proof .:.

Otherwise, the formula $\varphi(x, y, f)$ given by $\langle x, y \rangle \in f$ defines a function on A. By replacement, f''A = Ord exists (i.e. is an element of V), contradicting Burali–Forti Paradox (3 A • 11).

The general idea of a rank function is given below.

— 4•3. Lemma —

Let $A = \langle A, \langle A \rangle$ be well-founded. Therefore there is a unique function $f : A \to \text{Ord such that } f(a)$ is 0 if a is $\langle A \text{-minimal}, \text{ and otherwise } f(a) = \sup\{f(b) + 1 : b <_A a\}$.

Proof .:.

Uniqueness is immediate: for f, g two such functions and $a <_A$ -minimal where $f(a) \neq g(a)$, we have that $f(a) = \sup\{f(b) + 1 : b <_A a\}$. By minimality of a, this supremum is $\sup\{g(b) + 1 : b <_A a\} = g(a)$, which means g(a) is f(a), a contradiction.

We construct such an f by transfinite recursion. Firstly, as **A** is well-founded, define by transfinite recursion

$$X_0 = \emptyset$$

$$X_{\alpha+1} = \left\{ a \in A : a \text{ is } <_A \text{ -minimal in } A \setminus \bigcup_{\beta \le \alpha} X_\beta \right\}$$

$$X_{\gamma} = \emptyset, \text{ for } \gamma \text{ a limit.}$$

If $X_{\alpha+1}$ is ever empty, then we stop, and so $X_{\alpha} = A$. Then we define $f : A \to \text{Ord}$ by taking f(x) to be the least (and only) α such that $x \in X_{\alpha+1}$. By Lemma 4 • 2, this process stops at some $\alpha \in \text{Ord}$ so that $f \in V$.

Note that $x, y \in X_{\alpha}$ implies x and y are $<_A$ -incomparable: $x \not\leq_A y$ and $y \not\leq_A x$ (otherwise, they wouldn't be minimal). Hence f(x) = f(y) implies x and y are $<_A$ -incomparable.

Moreover, the contrapositive then tells us that if $x <_A y$, then $f(x) \neq f(y)$, and in fact f(x) < f(y), as otherwise f(y) < f(x) implies $x <_A y$ is not actually $<_A$ -minimal in $A \setminus \bigcup_{\beta < f(y)} X_\beta$, because $x \in A \setminus \bigcup_{\beta < f(x)} X_\beta \subseteq A \setminus \bigcup_{\beta < f(y)} X_\beta$. Therefore, $f(a) \ge \sup\{f(b) + 1 : b <_A a\}$ for all $a \in A$.

Now if $f(a) > \beta = \sup\{f(b) + 1 : b <_A a\}$, then by the definition of the X_{α} s, *a* wasn't minimal in $A \setminus \bigcup_{\gamma < \beta + 1} X_{\gamma}$, meaning that there is some $b \in A \setminus \bigcup_{\gamma < \beta + 1} X_{\gamma}$ with $b <_A a$. Taking a $<_A$ -minimal such *b* yields that $f(b) = \beta + 1$, contradicting the definition of β .

The point of having a rank function is to proceed by induction on the levels. Indeed, the proof above just defines the function f by induction on the levels of A. So if we can prove something for the elements inductively by level, then we can prove it for the whole well-founded set. So we have the following definition. By uniqueness, we are justified in using "the" rank function, and defining the following as aspects of the structure alone, independent of any choice of rank function.

For well-founded $\mathbf{A} = \langle A, \langle A \rangle$, the *rank function* on \mathbf{A} is the function rank : $A \rightarrow$ Ord such that

• $\operatorname{rank}(a) = 0$ if a is $<_A$ -minimal; and

• $\operatorname{rank}(a) = \sup \{\operatorname{rank}(b) + 1 : b <_A a\}$ for a not $<_A$ -minimal.

A structure $\mathbf{A} = \langle A, \langle A \rangle$ is *extensional* iff it satisfies the axiom of extensionality:

 $\{z \in A : z \ R \ x\} = \{z \in A : z \ R \ y\}$ implies x = y.

For A an extensional, well-founded structure, we can use the rank function to define the following.

- the *levels* of A are the sets $lvl_{\alpha}(A) = \{a \in A : rank(a) = \alpha\}$ for all $\alpha \in Ord$.
- the *height* or *length* of A is $ht(A) = \sup\{rank(a) + 1 : a \in A\} = im rank.$

We include the "+1" in the definition of height (and rank) to ensure that every element has a smaller rank than the height (or rank of the element we're considering). So the empty relation has height 0, and the set with one element has height 1 while the single element has rank 0. Note that for A a set, Lemma $4 \cdot 2$ implies that the height of A is an ordinal, and not just Ord itself. Note some other immediate facts.

- 4 · 5. Result

- Let $\mathbf{A} = \langle A, \langle A \rangle$ be well-founded with rank function, rank. Therefore, the following hold.
 - 1. If $a <_A b$, then rank $(a) < \operatorname{rank}(b)$.

2. If $a, b \in A$ are *comparable*—i.e. $a <_A b$ or $b <_A a$ —then rank $(a) < \operatorname{rank}(b)$ iff $a <_A b$.

Proof .:.

- 1. Clearly $a <_A b$ implies rank $(b) > \sup\{\operatorname{rank}(x) : x <_A b\} \ge \operatorname{rank}(a)$ by definition of rank.
- 2. If a and b are comparable, then either $a <_A b$ (in which case rank(a) < rank(b) implies $a <_A b$ by (1)),
 - or $b <_A a$ (in which case rank(a) < rank(b) implies $b <_A a$ vacuously by (1)).

Note that we cannot ensure in general that rank(a) < rank(b) implies $a <_A b$, since, for example, taking $<_A = \{(0, 1), (2, 3)\}$ yields a well-founded relation with rank(2) = 0, rank(1) = 1, but $2 \not<_A 1$. But this concept of rank is what allows us to *collapse* a well-founded, extensional set to a transitive set. We cannot do with with the above example, because it does not satisfy extensionality. It is extensionality that ensures we can *uniquely* describe elements by talking about their predecessors.

- Proof of The Mostowski Collapse (4 • 1) ...

As **A** satisfies extensionality, there is only one $<_A$ -minimal element, a_\emptyset . This is because any other $a \neq a_\emptyset$ must then have $\operatorname{pred}_{<_A}(a) \neq \operatorname{pred}_{<_A}(a_\emptyset) = \emptyset$. Hence there is some element of $\operatorname{pred}_{<_A}(a)$, which means a isn't minimal.

Proceed by recursion on the levels of A to define an isomorphism. Since there is only one $<_A$ -minimal element a_{\emptyset} , define $f_0(a_{\emptyset}) = \emptyset$. At limit stage γ define $f_{\gamma} = \bigcup_{\alpha < \gamma} f_{\alpha}$. At successor stage $\alpha + 1$, consider $lvl_{\alpha+1}(A)$.

Define $f_{\alpha+1}$ by

$$f_{\alpha+1}(x) = \begin{cases} f_{\alpha}(x) & \text{if } x \in \text{dom}(f_{\alpha}) \\ \{f_{\alpha}(y) : y <_{A} x\} & \text{if } x \in \text{lvl}_{\alpha+1}(\mathbf{A}). \end{cases}$$

This process stops at ht(A). Note that this process is well-defined: inductively, dom $(f_{\alpha}) = \bigcup_{\beta \leq \alpha} \text{lvl}_{\beta}(A)$, and if $y <_A x \in \text{lvl}_{\alpha+1}(A)$, then rank $(y) < \text{rank}(x) = \alpha$ so that y is in the domain of f_{α} . Taking $f = \bigcup_{\alpha < \text{ht}(A)} f_{\alpha}$, it follows that $f(x) = \{f(y) : y <_A x\}$ for all $x \in A$.

Note that T = im f is transitive: if $x \in f(a) \in T$, then x = f(b) for some $b <_A a$, and thus $x = f(b) \in T$. So it suffices to show that f is an isomorphism between **A** and $\langle T, \in \rangle$.

Surjectivity of $f : A \to T$ is immediate. For injectivity, let $a \in A$ be $<_A$ -minimal where f(a) = f(b) for some b. Let $f(x) \in f(b)$ for some $x <_A b$ so that $f(x) \in f(a)$ and thus f(x) = f(y) for some $y <_A a$. By minimality of a, y = x and therefore $x <_A a$. The same idea shows that if $x <_A a$ then $x <_A b$, and thus a = b by extensionality.

Now if $a <_A b$ then $f(a) \in \{f(x) : x <_A b\} = f(b)$. Similarly, suppose $f(a) \in f(b)$. Thus f(a) = f(x) for some $x <_A b$. By injectivity, a = x and thus $a <_A b$.

To see that T is unique, suppose $g : A \to D$ is an isomorphism with D transitive. Let $a \in A$ be of least rank such that $f(a) \neq g(a)$. Note that by extensionality and the inductive hypothesis, $f(a) = \{f(x) : x <_A a\} = \{g(x) : x <_A a\} = g(a)$, a contradiction.

So again, The Mostowski Collapse (4 • 1) should highlight the importance of transitive sets, as they allow us to consider any sort of well-founded, extensional^{xiv} relation. This also motivates the axiom of foundation, which says that membership is well-founded. We will not accept foundation as an axiom just yet, though.

- **4**•6. Definition Let $\mathbf{A} = \langle A, \langle A \rangle$ be well-founded and extensional. The *mostowski collapsing map* of \mathbf{A} is an isomorphism $\pi : A \to T \subseteq V$ defined by recursion on rank: for every $a \in A$, $\pi(a) = \{\pi(b) : b <_A a\}$. The *transitive collapse* of \mathbf{A} is then $\langle \operatorname{im} \pi, \in \rangle$.

The proof of The Mostowski Collapse (4 \cdot 1) shows that π is well-defined, unique, and is in fact an isomorphism.

Note that there is a slightly more general version of The Mostowski Collapse $(4 \cdot 1)$: we don't require that $A \in V$, but instead that at least pred_{<A} $(a) \in V$ for each $a \in A$. For example, V satisfies this, as pred_{\in} $(x) = x \in V$ for each $x \in V$. The proof remains the same, as we never needed ht(A) to be an ordinal: it could be Ord itself, as with V. The point of this generalization is just in case we have a well-founded, partially ordered structure that is not a set. Then we can collapse it down to a transitive class (not necessarily a set) under membership. For now, we will have no use of this generality, but it will be incredibly important later, as we will collapse down various collections into "inner models".

To be slightly more precise than the previous paragraph, for A and R classes, if $\text{pred}_{R}(x)$ is a set for each $x \in A$, then we can define the mostowski collapse as in Definition 4 • 6 as a class, and so yield the image T as a transitive class, which is still isomorphic under membership to A under R.

§4A. Powerset and the cumulative hierarchy

As a consequence of the axiom of foundation, we have the following *iterative* characterization of V in the sense that all collections are formed from things that already exist. In this sense, starting with $V_1 = \{\emptyset\}$, we can take the set of collections of elements in V_1 , which is $V_2 = \{\emptyset, \{\emptyset\}\}$. Then we can take the set of all collections of elements in this: $V_3 = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\{\emptyset\}\}, \{\{\emptyset\}\}\}, \{\{\emptyset\}\}\}, \{\{\emptyset\}\}\}, \{\{\emptyset\}\}, \{\{\emptyset\}\}\}, and so on. More precisely, by Lemma 4 • 3, there is a rank function on V. But what exactly is this rank function? By uniqueness, we just need to give an example of one. A first stab at this would be at stage <math>\alpha$ to define the $\alpha + 1$ st level by $\{y : y \subseteq |v|_{\alpha}(V)\}$. This seems finte, but it's not particularly useful, as it's unclear

xiv in the sense of satisfying the axiom of extensionality

that this results in a set. So in doing defining the rank function, we will introduce another axiom, saying that these levels exist: we can continue to define V_{α} for all α .

- **4A**•**1**. Definition (Axiom)
(*Powerset*) The poweset
$$\mathcal{P}(x) = \{y : y \subseteq x\}$$
 exists: $\forall x \exists P \ \forall y \ (y \in P \leftrightarrow y \subseteq x)$.

Note that although comprehension allows us to say that all sorts of subsets of x exist, without the powerset axiom, we cannot in general form the set of all of these at once. But once we know we can collect these together, we get some immediate properties.

• $x \in \mathcal{P}(x), \emptyset \in \mathcal{P}(x);$

- if x is transitive, $x \subseteq \mathcal{P}(x)$;
- if $x \subseteq y$, then $\mathcal{P}(x) \subseteq \mathcal{P}(y)$;
- $\mathfrak{P}(x) \cap \mathfrak{P}(y) = \mathfrak{P}(x \cap y);$
- $\mathcal{P}(x) \cup \mathcal{P}(y) \subseteq \mathcal{P}(x \cup y).$

Now consider the following collection. Regardless of whether foundation holds, we can still define it in V.

- 4A·2. Definition Define the *cumulative hierarchy* to be the collection WF = $\bigcup_{\alpha \in Ord} V_{\alpha}$ given by transfinite recursion: $V_0 = \emptyset$, $V_{\alpha+1} = \mathcal{P}(V_{\alpha})$, and $V_{\gamma} = \bigcup_{\alpha < \gamma} V_{\gamma}$, for γ a limit ordinal.

Note that WF is not a set, since we will have a surjection from WF onto the ordinals. To see this, consider the following easy to show facts.

— 4 A • 3. Result -

For every $\alpha \in \text{Ord}$, and x,

- 1. If x is transitive, $\mathcal{P}(x)$ is transitive.
- 2. WF is transitive, in particular $V_{\alpha} \subseteq V_{\beta}$ for $\alpha < \beta$.
- 3. if $x \in V_{\alpha}$, then $\{x\} \in V_{\alpha+1}$.
- 4. V_{α} is closed under (finite) unions, intersections, and complements.
- 5. $\alpha \in V_{\alpha+1}$ for each $\alpha \in Ord$, hence $Ord \subseteq WF$.
- 6. For each $x \in WF$, the least α with $x \in V_{\alpha}$ is a successor ordinal.
- 7. (WF, \in) is well-founded with rank function rank(x) as the least α with $x \in V_{\alpha+1}$.
- 8. $x \in WF$ iff $x \subseteq WF$.

Proof .:.

- 1. Suppose x is transitive, and let $z \in y \in \mathcal{P}(x)$, i.e. $z \in y \subseteq x$. Therefore $z \in x$ so by transitivity, $z \subseteq x$, and thus $z \in \mathcal{P}(x)$.
- 2. Proceed by induction on α to show that V_{α} is transitive. For $\alpha = 0$, this is immediate. For $\alpha + 1$, use (1) and the inductive hypothesis. For γ a limit, if $y \in x \in \bigcup_{\alpha < \gamma} V_{\alpha}$, then $y \in x \in V_{\alpha}$ for some $\alpha < \gamma$, in which case $y \in V_{\alpha}$ by the inductive hypothesis, and thus $y \in V_{\gamma}$. Hence every V_{α} is transitive, and for the same reason as with the limit ordinal, WF = $\bigcup_{\alpha \in Ord} V_{\alpha}$ is transitive too.
- 3. This is clear, as $\{x\} \subseteq V_{\alpha}$, and thus $\{x\} \in \mathcal{O}(V_{\alpha}) = V_{\alpha+1}$.
- For any two subsets x, y ⊆ V_β for some β < α, x ∩ y, x ∪ y, and x \ y are all still subsets of V_β, and hence are elements of V_{β+1} ⊆ V_α.
- 5. For $\alpha = 0$, clearly $V_1 = \mathcal{P}(\emptyset) = \{\emptyset\}$ has $0 \in V_1$. For the successor $\alpha + 1$, $\alpha \in V_{\alpha+1}$ so that $\{\alpha\} \in V_{\alpha+2}$ and as transitive sets, using (4), $\alpha + 1 = \alpha \cup \{\alpha\} \in V_{\alpha+2}$. For limit $\gamma, \beta \in V_{\beta+1}$ for all $\beta < \gamma$. As $\beta < \gamma$ implies $\beta + 1 < \gamma$, this means $\beta \in V_{\gamma}$ for every $\beta < \gamma$. Therefore $\gamma \subseteq V_{\gamma}$ and so $\gamma \in V_{\gamma+1}$.
- 6. Let $x \in WF$ be in V_{α} for α least. If α is a limit ordinal, then clearly $x \in \bigcup_{\beta < \alpha} V_{\beta}$ implies $x \in V_{\beta}$ for some $\beta < \alpha$, a contradiction. Also, $x \notin V_0 = \emptyset$, hence α must be a successor.
- 7. Write f for this function. Note that if $x \in y$ then clearly $f(x) \leq f(y)$, as $x \in y \in V_{f(y)+1}$ implies

 $x \in V_{f(y)+1}$. To see that $f(x) \neq f(y), y \in V_{f(y)+1}$ implies $y \subseteq V_{f(y)}$ and hence $x \in V_{f(y)}$, implying that $f(y) \ge f(x) + 1 > f(x)$. So $x \in y$ implies f(x) < f(y).

Now suppose $X \subseteq WF$. If there is no \in -minimal element, then f''X has no \in -minimal element, contradicting the well-foundedness of the ordinals.

To see that this function f is really a rank function, we need to show that $f(x) = \beta = \sup\{f(y) + 1 : y \in x\}$. So clearly, the above argument gives that $f(x) \geq \beta$. And clearly, $y \in V_{f(y)+1}$ for each $y \in x$ implies $x \subseteq \bigcup_{y \in x} V_{f(y)+1} = V_{\beta}$, and hence $x \in V_{\beta+1}$ shows that $f(x) \leq \beta$. Therefore $f(x) = \beta$, and f is a rank function.

8. If $x \in WF$, then $x \subseteq WF$ by transitivity. For the other direction, if $x \subseteq WF$, then $x \subseteq V_{\alpha}$ for $\alpha =$ $\sup\{\operatorname{rank}(y) : y \in x\}$. Hence $x \in V_{\alpha+1} \subseteq WF$. \dashv

We can prove more about the class WF, in particular, that it consists of well-founded transitive sets. To do this, with the added technology of the natural numbers, we have the following definition. Note that we can still make this definition in the absence of foundation.

- **4A**•**4**. Definition
Let x be a set. Define trcl(x), the *transitive closure* of x to be
$$\bigcup_{n \in \omega} \bigcup^n x$$
, where \bigcup^n is defined by recursion on ω :
 $\bigcup^0 x = x, \bigcup^{n+1} x = \bigcup(\bigcup^n x).$

Hence every set is contained in its transitive closure. Of course, the transitive closure trcl(x) is indeed transitive, since $y \in \text{trcl}(x)$ implies $y \in \bigcup^n x$ and hence $y \subseteq \bigcup^{n+1} x \subseteq \text{trcl}(x)$. The key reason that this should be a motivation for the axiom of foundation, is that we only every need to "go down" ω many times. Foundation will tell us that we only need to go down $< \omega$ -many times, although the number of times may be arbitrarily high. Let's first prove some quick results about the transitive closure.

- 4 A • 5. Result -

1 . . 1

- For every x,
 - 1. trcl(x) is transitive, and is the \subseteq -minimal transitive set containing x: if $x \subseteq T$ where T is transitive, then $\operatorname{trcl}(x) \subseteq T$.
 - 2. If x is transitive, then trcl(x) = x.
 - 3. If $x \in y$, then trcl(x) \subseteq trcl(y) (assuming foundation, trcl(x) \subseteq trcl(y)).
 - 4. $\operatorname{trcl}(x) = x \cup \bigcup_{y \in x} \operatorname{trcl}(y)$.

Proof .:.

- 1. If T is transitive and $x \subseteq T$, then clearly $\bigcup^0 x \subseteq T$. And inductively, $\bigcup^n x \subseteq T$ implies $\bigcup^{n+1} x \subseteq T$ by transitivity. Hence $\bigcup_{n \in \omega} \bigcup^n x \subseteq T$, meaning trcl $(x) \subseteq T$.
- 2. If x is transitive, then $trcl(x) \subseteq x$ by (1), and since clearly $x \subseteq trcl(x)$, we have equality.
- 3. If $x \in y \subseteq trcl(y)$ so that $x \subseteq trcl(y)$. Using (1), we again have that $trcl(x) \subseteq trcl(y)$. Assuming foundation, $x \notin \operatorname{trcl}(x)$ (otherwise we would have a finite loop), but $x \in \operatorname{trcl}(y)$.
- 4. Note that $T = x \cup \bigcup_{y \in x} \operatorname{trcl}(y)$ is transitive, since $y \in x$ implies $y \subseteq \operatorname{trcl}(y) = T$, and if $y \in T \setminus x$, then clearly y is in a transitive set, and hence is a subset of T. Therefore, by (1), $trcl(x) \subseteq T$. But $T \subseteq trcl(x)$, since $x \subseteq trcl(x)$, and (3) implies $trcl(y) \subseteq trcl(x)$ for each $y \in x$.

With this, we have the following, demonstrating why the notation "WF" is used.

```
- 4A•6. Theorem
```

```
Let x be a transitive set. Therefore x \in WF iff (x, \in) is well-founded.
```

Proof .:.

Suppose $x \in WF$. Therefore (x, \in) is well-founded, just by the fact that WF is well-founded and transitive:

 $x \subseteq$ WF. So suppose $\langle x, \in \rangle$ is well-founded.

Note that $x \subseteq WF$. To see this, otherwise $A = x \setminus WF$ has a \in -minimal element $a \in A$. Thus $a \subseteq WF$ so that $a \in WF$ by Result $4 \land 4 \land 3$. But then $x \subseteq WF$ yields $x \in WF$ by the same reasoning. \dashv

So if all of this was motivation, let us give the actual result.

- $4A \cdot 7$. Theorem - The axiom of foundation implies V = WF.

Proof .:.

Assuming the axiom of foundation, for each $x \in V$, trcl(x) is a transitive set where $\langle trcl(x), \in \rangle$ is well-founded (just by virtue of V being well-founded). Hence $trcl(x) \in WF$. But then $x \subseteq WF$, and so $x \in WF$. Hence every element of the universe is an element of WF, and so the two are equal.

With all of that out of the way, we will now finally accept the axiom of foundation as a part of the axioms of set theory. The rank function on WF = V is incredibly useful, as it allows us to proof properties of V through induction on rank. The cumulative hierarchy also gives a nice, stratified picture of the un0iverse, as seen below.



4A•8. Figure: The set theoretic universe

The well-foundedness of the universe also gives that any model embedde in V is then well-founded as well. This is just because any infinite decreasing sequence $A \models "a_{n+1} \in a_n$ " for $\{a_n : n \in \omega\} \subseteq A$ implies $V \models "f(a_{n+1}) \in f(a_n)$ " for each $n \in \omega$, where $f : A \to V$ is an embedding. Now this relies on a separate, stronger characterization of well-foundedness than first-order logic alone is able to give. So we present the following meta-theoretic result.

Let $\mathbf{A} = \langle A, \in^{\mathbf{A}} \rangle$ be a structure. Consider the following propositions:

- 1. A is well-founded.
- 2. There are no infinite \in^{A} -decreasing sequences of elements of *A*.
- 3. A satisfies the axiom of foundation.

Therefore (1) implies (2) and (3), but (3) doesn't imply (2) and thus doesn't imply (1) either.

Proof .:.

To see that (1) implies (2), note that any infinite \in^{A} -decreasing sequence of elements of A is a function from some ordinal α to A. Restricting to ω yields the sequence $\langle a_n : n \in \omega \rangle$ still \in^{A} -decreasing, which gives the set $\{a_n : n \in \omega\} \in V$ with no \in^{A} -minimal element. Hence A isn't well-founded.

To see that (1) implies (3), if A doesn't satisfy the axiom of foundation, then for some $a \in A$, $A \models "\forall x \in a \exists y \in a (y \in x)"$. Hence the set $\{x \in A : A \models "x \in a"\} \in V$ has no \in^A -minimal predecessor. Therefore A isn't well-founded.

To see that (3) doesn't imply (1) nor (2), we use compactness to give a model where (3) holds, but (2)—and

thus (1)—fails. In particular, consider ordinal ω in V: $\mathbf{N} = \langle \omega, \in \rangle$. As we saw before, the theory of this model can be "misinterpreted" to give an ill-founded model. Clearly N satisfies the axiom of foundation, because N is well-founded. Therefore, we can consider the theory of N:

Th(N) = { φ a FOL(\in)-sentence : N $\models \varphi$ }

Now consider the additional constant symbols $\{c_n : n \in \omega\}$. Intuitively, each c_n should count "backwards". Formalizing this, let T be the theory $\text{Th}(\mathbf{N}) \cup \{"c_{n+1} \in c_n" : n \in \omega\}$. Note that T has a model by compactness: for each finite subset $\Delta \subseteq T$, there is some largest $N \in \omega$ where c_N occurs in Δ (because Δ is finite). Therefore the model \mathbf{N}' interpretting c_0 as $N \in \omega$, and c_1 as N - 1 and so on—meaning $c_n^{\mathbf{N}} = N - n$ for all $n \leq N$ —has $\mathbf{N}' \models "c_{n+1} \in c_n"$, and $\mathbf{N}' \models \text{Th}(\mathbf{N})$, because we haven't changed any of the structure, just given names to some elements. So \mathbf{N}' is a model of Δ . Hence every finite subset of T has a model, and so T has a model \mathbf{A} . Thus \mathbf{A} satisfies the axiom of foundation in $\text{Th}(\mathbf{N}) \subseteq T$, but \mathbf{A} also has the infinite $\in^{\mathbf{A}}$ -decreasing sequence $\langle c_n^{\mathbf{A}} : n \in \omega \rangle$ in \mathbf{V} . Therefore (3) holds, but neither (1) nor (2) holds for \mathbf{A} .

We will later see that (2) is actually equivalent to (1), but this requires the axiom of choice in the form of König's theorem on trees.

Let us now think about the ranks of sets, and how they can be computed. Recall that the rank of a set x is the least $\alpha \in Ord$ such that $x \in V_{\alpha+1}$. The reason for the "+1" is that V_{α} for α a limit is never the least such that a set appears in it: $V_{\alpha} = \bigcup_{\beta < \alpha} V_{\beta}$. So defining it in this way allows us to say that there is always a set of rank α for $\alpha \in Ord$. Another, easier to remember definition is that the rank of x is the least α with $x \subseteq V_{\alpha}$.

- **4A**•10. Result - For every set x and y,

- if $y \subseteq x$ then rank $(y) \leq \operatorname{rank}(x)$.
- $\operatorname{rank}(\operatorname{trcl}(x)) = \operatorname{rank}(x);$
- $rank({x}) = rank(x) + 1;$
- $\operatorname{rank}(\mathfrak{P}(x)) = \operatorname{rank}(x) + 1$; and
- $\operatorname{rank}(x \cup y) = \max(\operatorname{rank}(x), \operatorname{rank}(y)).$
- $\operatorname{rank}(x) = x$ for $x \in \operatorname{Ord}$.

Proof .:.

- The least α such that $x \subseteq V_{\alpha}$ thus also has $y \subseteq V_{\alpha}$.
- For $x \subseteq V_{\alpha}$ with α least, $\bigcup x \subseteq V_{\alpha}$ by transitivity. Hence inductively, $trcl(x) = \bigcup_{n \in \omega} \bigcup^{n} x \subseteq V_{\alpha}$ and thus $trcl(x) \in \mathcal{P}(V_{\alpha}) = V_{\alpha+1}$. This establishes that $rank(trcl(x)) \leq rank(x)$. Since $x \subseteq trcl(x)$ (1) implies the other inequality and thus the two are equal.
- By Result 4 A 3, rank is a rank function, and thus $rank({x}) = rank(x) + 1$.
- Since {x} ⊆ 𝒫(x), it follows by (1) and (3) that rank(𝒫(x)) ≥ rank(x) + 1. For the other direction, note that y ⊆ x ⊆ V_α implies y ∈ V_{α+1} and thus 𝒫(x) ⊆ V_{α+1} so that rank(𝒫(x)) ≤ rank(x) + 1. Hence the two are equal.
- Let rank $(x) < \operatorname{rank}(y) = \alpha$. Therefore $x, y \subseteq V_{\alpha}$ and thus $x \cup y \subseteq V_{\alpha}$, implying rank $(x \cup y) \le \alpha$. Since $y \subseteq x \cup y$, (1) implies the other inequality.
- Proceed by induction on α . For $\alpha = 0$, this is clear. For $\alpha + 1$,

 $\operatorname{rank}(\alpha \cup \{\alpha\}) = \max(\operatorname{rank}(\alpha), \operatorname{rank}(\{\alpha\})) = \operatorname{rank}(\alpha) + 1 = \alpha + 1$

by (3) and the inductive hypothesis. For limit α , as a rank function, $\operatorname{rank}(\alpha) = \sup_{\beta < \alpha} (\operatorname{rank}(\beta) + 1) = \sup_{\beta < \alpha} (\beta + 1) = \alpha$.

At this point, calculating ranks might seem completely worthless, but they help to understand just how the universe is built up, and at what stages certain sets come into play. For now, we don't have much use for it, but later on, the levels of the cumulative hierarchy (and other hierarchies) will play a big role in understanding their larger structure through the use of reflection properties—properties of the larger structure holding in smaller parts. For example, just by calculating ranks, one can see that for limit α , V_{α} is closed under the powerset operation as well as taking unions, pairs, cartesian products, and so on. In this sense, which we will make precise later, the levels of the cumulative hierarchy

OTHER WELL-FOUNDED RELATIONS

model a great portion of set theory.

Section 5. Ordinals and Cardinality

It is nearly impossible to have a discussion about set theory that doesn't eventually devolve into a discussion about cardinals. What are cardinals? What is cardinality? These are things that need to be addressed, but to address them, we need a better understanding of ordinals.

§5A. Ordinal arithmetic

Recall that we can add 1 to ordinals: $\alpha + 1 = \alpha \cup \{\alpha\}$. Using Transfinite Recursion (3 C • 2), we can also define addition between ordinals in general. The motivating picture is that $\alpha + \beta$ is just the order of α placed before the order of β . In particular, we could define $\alpha + \beta$ to be the unique ordinal corresponding to this well-order via Corollary 3 E • 7. But instead, we have the following definition.

 $-5A \cdot 1. \text{ Definition} \qquad - 5A \cdot 1. \text{ Definition} \qquad - 5A \cdot 1. \text{ Definition} \qquad - 5A \cdot 1. \text{ Definite} + : \text{ Ord } \times \text{ Ord } \rightarrow \text{ Ord as follows: for each } \alpha \in \text{ Ord,}$ $\cdot \alpha + 0 = \alpha;$ $\cdot \alpha + (\beta + 1) = (\alpha + \beta) + 1;$ $\cdot \alpha + \gamma = \sup_{\beta < \gamma} \alpha + \beta \text{ for } \gamma \text{ a limit.}$ $\text{Define } \cdot : \text{ Ord } \times \text{ Ord } \rightarrow \text{ Ord as follows: for each } \alpha \in \text{ Ord,}$ $\cdot \alpha \cdot 0 = 0;$ $\cdot \alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + \alpha;$ $\cdot \alpha \cdot \gamma = \sup_{\beta < \gamma} \alpha \cdot \beta \text{ for } \gamma \text{ a limit.}$

Really, we've defined the class $+_{\alpha}$: Ord \rightarrow Ord for each $\alpha \in$ Ord by Transfinite Recursion (3 C • 2), and so have + defined as the class { $\langle \langle \alpha, \beta \rangle, \gamma \rangle : +_{\alpha}(\beta) = \gamma$ }. So the more formally-minded can be put at ease by knowing that these classes are well-defined^{xv}

Note that these definitions (restricted to ω) then coincide with the definitions of addition and multiplication on the natural numbers. In particular, given these definitions, that 0 isn't a successor, and Induction on ω (3 B • 1), all of the axioms of peano arithmetic are satisfied by ω under these interpretations. Formally, this means the following.

— 5A•2. Theorem

- 5A•3. Lemma

 $ZFC \vdash PA^{\omega}$, where PA^{ω} is the set of axioms of peano arithmetic with all quantifiers restricted to ω , and $+, \cdot, 0, 1$ replaced by the defining FOL(\in)-formulas. In particular, $ZFC \vdash Con(PA)$ by Completeness (1 D • 1).

We haven't quite made precise what all of this means (which we will get to in the next couple sections), but the idea is just that ZFC will show things like the commutativity of + and \cdot on ω . But unlike normal addition, we don't have the same sort of cancellation laws *for general ordinals*, and in fact, commutativity does not hold in general.

For each $\alpha, \beta, \gamma \in \text{Ord}$, if $\beta < \gamma$, then $\alpha + \beta < \alpha + \gamma$. However, it's possible that $\beta + \alpha = \gamma + \alpha$.

Proof .:.

The example to the second sentence can be given easily: take $\alpha = \omega$ with $\beta = 1 < 2 = \gamma$. As a limit ordinal, $\beta + \alpha = \sup_{n \in \omega} 1 + n = \omega = \sup_{n \in \omega} 2 + n = \omega$.

^{xv}And of course, the same idea applies to multiplication.

To show the first, proceed by induction on γ . For $\gamma = 0$, this is immediate. For $\gamma + 1$, $\beta < \gamma + 1$ implies $\alpha + \beta \le \alpha + \gamma$ by the inductive hypothesis. By definition, this is strictly less than $(\alpha + \gamma) + 1 = \alpha + (\gamma + 1)$. For γ a limit, $\beta < \gamma$ implies $\alpha + \beta < (\alpha + \beta) + 1 = \alpha + (\beta + 1) \le \alpha + \gamma$.

Let $\alpha, \beta, \gamma \in \text{Ord.}$ Therefore $\alpha + \beta = \alpha + \gamma$ iff $\beta = \gamma$.

Proof .:.

One direction is immediate. If $\beta \neq \gamma$, say $\beta < \gamma$, then $\alpha + \beta < \alpha + \gamma$, and so the two are unequal. \dashv

We can characterize $\alpha + \beta$ as a copy of α followed by a copy of β . This is formalized by a long definition, but the idea is to produce a copy of α disjoint from a copy of β by considering $\alpha \times \{0\}$ and $\beta \times \{1\}$ instead. We use the map $x \mapsto \langle x, 0 \rangle$ to define the isomorphic order on $\alpha \times \{0\}$, and similarly for $\beta \times \{1\}$. This means $\langle x, 0 \rangle < \langle y, 0 \rangle$ iff x < y, and similarly with a 1 in place of a 0, so we are justified in calling them "copies". We put these two orders together just by saying every element of the copy of β is above all elements of the copy of α . So this is how we formalize this "copy of α followed by a copy of β ". The characterization is then easy, although very formal.

 $-5A \cdot 5$. Theorem

For each $\alpha, \beta \in \text{Ord}, \alpha + \beta$ is the order-type	e of <	$A_{\alpha,\beta}$ on $A_{\alpha,\beta} = (\alpha \times \{0\}) \cup (\beta \times \{1\})$ given by
$\langle \gamma,n angle <_A \langle \delta,m angle$	iff	$(n = 0 \land m = 1) \lor (n = m \land \gamma < \delta).$

Proof .:.

Note that for each $\gamma < \beta$; $A_{\alpha,\gamma} = A_{\alpha,\beta} \setminus \{\langle \delta, 1 \rangle : \gamma \leq \delta < \beta\}$. Similarly, the uniform definition of the ordering yields that $A_{\alpha,\gamma}$ is an $\langle A_{\alpha,\beta} - initial$ segment of $A_{\alpha,\beta}$, and in fact the order $\langle A_{\alpha,\gamma} \rangle$ is equal to $\langle A_{\alpha,\beta+1} \cap (A_{\alpha,\beta} \times A_{\alpha,\beta})$.

Proceed by induction on β . For $\beta = 0$, this is immediate: $A_{\alpha,\beta} = \alpha \times \{0\}$ and $\langle A_{\alpha,\beta} \rangle$ is the same as the order on $\alpha = \alpha + 0 = \alpha + \beta$.

For $\beta + 1$, Note that $\langle \beta, 1 \rangle$ is $\langle A_{\alpha,\beta+1}$ -maximal. If we consider $A_{\alpha,\beta+1} \setminus \{\langle \beta, 1 \rangle\}$, we get $A_{\alpha,\beta}$. By the inductive hypothesis, and the idea above, it follows that

 $\left(\operatorname{pred}_{<_{A_{\alpha,\beta+1}}}(\langle \beta,1\rangle),<_{A_{\alpha,\beta+1}}\right) = \langle A_{\alpha,\beta},<_{A_{\alpha,\beta}}\rangle \cong \alpha+\beta.$

Hence adding on a single element at the end yields $\alpha + (\beta + 1) = (\alpha + \beta) + 1 \cong \langle A_{\alpha,\beta+1}, \langle A_{\alpha,\beta+1} \rangle$.

For limit β , it should be obvious that $A_{\alpha,\beta} = \bigcup_{\gamma < \beta} A_{\alpha,\gamma}$ and $\langle_{A_{\alpha,\beta}} = \bigcup_{\gamma < \beta} \langle_{A_{\alpha,\gamma}}$. The inductive hypothesis tells us that the order-type of $\mathbf{A}_{\alpha,\beta} = \langle A_{\alpha,\beta}, \langle_{A_{\alpha,\beta}} \rangle$, say τ , is at least $\sup_{\gamma < \beta} \alpha + \gamma = \alpha + \beta$. Moreover, if $\tau > \alpha + \beta$, then there must be some initial segment of τ and thus of $\mathbf{A}_{\alpha,\beta}$ which has order-type $\alpha + \beta$, which contradicts that each initial segment has order-type $\alpha + \gamma$ for $\gamma < \beta$.

We have similar sorts of properties for multiplication. The characterization for ordinal multiplication in that $\alpha \cdot \beta$ is the order-type of β copies of α : $\omega \cdot 4 = \omega + \omega + \omega + \omega$, for example.

– 5A•6. Lemma

For each $\alpha, \beta, \gamma \in \text{Ord}$, if $\alpha > 0$, then $\beta < \gamma$, then $\alpha \cdot \beta < \alpha \cdot \gamma$. However, it's possible that $\beta \cdot \alpha = \gamma \cdot \alpha$.

Proof .:.

Again, the example to the second sentence can be given easily: $2 \cdot \omega = \sup_{n \in \omega} 2 \cdot n = \omega = 1 \cdot \omega$ although 1 < 2.

Proceed by induction on γ . For $\gamma = 0$, the statement is vacuously true. For $\gamma + 1$, by the inductive hypothesis, $\alpha \cdot \beta \leq \alpha \cdot \gamma = \alpha \cdot \gamma + 0 < \alpha \cdot \gamma + \alpha = \alpha \cdot (\gamma + 1)$. For γ a limit, we easily have $\beta + 1 < \gamma$ and hence $\alpha \cdot \beta < \alpha \cdot \beta + \alpha = \alpha \cdot (\beta + 1) \leq \alpha \cdot \gamma$. - **5A**•**7**. Corollary For each $\alpha, \beta, \gamma \in \text{Ord}$, if $\alpha > 0$, then $\alpha \cdot \beta = \alpha \cdot \gamma$ iff $\beta = \gamma$.

Proof .:.

One direction is immediate. Now if $\beta \neq \gamma$, say $\beta < \gamma$, then $\alpha \cdot \beta < \alpha \cdot \gamma$, and so the two are unequal. \dashv

Implicit in the restriction that $\alpha > 0$ in Lemma 5 A • 6 is that this doesn't work for $\alpha = 0$. This is, of course, true, since $\alpha \cdot \beta = 0$ for all β when $\alpha = 0$. Although Definition 5 A • 1 only states $\beta \cdot 0 = 0$ for all β , we can inductively show $0 \cdot \beta = 0$ easily: $\beta = 0$ is immediate, and since $0 \cdot (\beta + 1) = 0 \cdot \beta + 0 = 0 + 0 = 0$, it holds at successors, and so trivially at limits.

The characterization of $\alpha \cdot \beta$ as β copies of α , like addition before, relies on a very formal construction to make these "copies" precise. We do this as before by tagging each copy of α : the γ th copy of α is $\alpha \times \{\gamma\}$. Hence we're ordering $\alpha \times \beta$. We ensure the γ_0 th copy of α is completely before the γ_1 th copy of α whenever $\gamma_0 < \gamma_1 < \beta$ by a complicated definition. But once one understands the construction, the idea is easy.

- 5A•8. Theorem

Let $\alpha, \beta \in \text{Ord.}$ Therefore $\alpha \cdot \beta$ is order-type of $<_{\alpha \times \beta}$ on $\alpha \times \beta$ defined by $\langle \alpha_0, \beta_0 \rangle <_{\alpha \times \beta} \langle \alpha_1, \beta_1 \rangle$ iff $(\beta_0 < \beta_1) \lor (\beta_0 = \beta_1 \land \alpha_0 < \alpha_1)$.

Proof .:.

Note that for each $\gamma < \beta$, $\alpha \times \gamma = (\alpha \times \beta) \setminus (\alpha \times (\beta \setminus \gamma))$. Similarly, the uniform definition of the ordering yields that $\alpha \times \gamma$ and its order form an $<_{\alpha \times \beta}$ -initial segment of $\alpha \times \beta$. In fact, $<_{\alpha \times \gamma} = <_{\alpha \times \beta} \cap ((\alpha \times \gamma) \times (\alpha \times \gamma))$.

Proceed by induction on β . For $\beta = 0$, this is immediate, as both $\alpha \cdot \beta$ and $\alpha \times \beta$ are \emptyset . For $\beta + 1$, by the inductive hypothesis, $\alpha \times (\beta + 1)$ is just the order on $\alpha \times \beta$ followed by the normal order on $\alpha \times \{\beta\}$, which is isomorphic to $\alpha \cdot \beta$ followed by α (using Theorem 5 A • 5 for the formal details). Hence this is just $\alpha \cdot \beta + \alpha = \alpha \cdot (\beta + 1)$. The limit case follows similarly as before.

We can continue to define further ordinal operations. In particular, ordinal exponentiation. This will be the last one we develop, as it is hardly every used, but it does give a good picture of the ordinals and how we can describe them.

5A•9. Definition
Define ordinal exponentiation as follows: for each α ∈ Ord,
α⁰ = 1;
α^{β+1} = (α^β) ⋅ α;
α^γ = sup_{β<γ} a^β for γ a limit.

There is another characterization of ordinal exponentiation in terms of functions with finite support, but it is almost never used in practice, and is instead left to the exercises. But the point is that ordinal exponentiation allows us to express more and more ordinals. In particular, we have the following picture of ordinals, beginning with the natural numbers, and building from there using our operations.

 $\overset{\circ}{0} \begin{array}{c} \overset{\circ}{1} \begin{array}{c} \overset{\circ}{2} \end{array} \\ \overset{\circ}{\omega} \begin{array}{c} \omega \\ \omega + 1 \end{array} \\ \overset{\circ}{\omega} + \omega \\ \overset{\circ}{\omega} + \omega \\ \overset{\circ}{\omega} + \omega \\ \overset{\circ}{\omega} \overset{\circ}{\omega} \overset{\circ}{\omega} \\ \overset{\circ}{\omega} \overset{\circ}{\omega} \\ \overset{\circ}{\omega} \\ \overset{$

And this picture, of course, never ends: we can continue to add and multiply ordinals to get larger and larger ordinals like $\omega^{\omega^{\omega}}$, $\omega^{\omega^{\omega^{\omega}}}$, and so on. Actually, taking the supremum of these exponentials— ω raised to ω *n*-times for $n < \omega$ —yields a truly gargantuan ordinal called ε_0 that satisfies $\omega^{\varepsilon_0} = \varepsilon_0$.

Now I would like to raise the question, which ordinals are important? Obviously, this isn't something inherent to the ordinals themselves but instead how we view them. But the question is still one that warrants an answer, given that the

ordinals are the canonical well-orders. Are there any other ordinals that are "canonical" in a sense? The answer turns out to be yes. We will take two approaches to answer this question: one the easier route working in V, and a harder route where we deprive ourselves of an important axiom to show that certain things exist or hold in general.

§5B. Cardinals with choice

This picture of the ordinals is useful as it provides a clear idea of "counting" in set-theoretic terms: we proceed lining up the elements of a given set with ordinals just as a child (or adult) might count something by lining it up with their fingers, associating each finger with a number.

The ordinals play the role of the fingers when counting. The issue is that it doesn't follow from the other axioms that every set can be counted in this way. To motivate the axiom of choice, which we need to demonstrate this, consider the following argument.

For some ordinal $\alpha < \beta$, consider the set $X_{\alpha} \neq \emptyset$. Since each X_{α} is non-empty, consider some $x_{\alpha} \in X_{\alpha}$. Thus $\{x_{\alpha} : \alpha < \beta\}$ exists. This is equivalent to the axiom of choice. Although we can ensure each X_{α} has an element, our finite notion of proof can't ensure give these x_{α} s all at once if there are infinitely many X_{α} s.

To further motivate the idea, consider the following definition, extending a previous one.

5B•1. Definition Let A and B be sets. Write $A \leq_{\text{size}} B$ iff there is an injection $f : A \rightarrow B$. Write $A \geq_{\text{size}} B$ iff there is a surjection $f : A \rightarrow B$. Write $A =_{\text{size}} B$ iff there is a bijection $f : A \rightarrow B$.

It should be clear that $A \leq_{\text{size}} B$ reflects the notion that A has fewer (or as many) elements than B, because any such injective $f : A \to B$ is really just a bijection $f : A \to \text{im}(f)$ where $\text{im}(f) \subseteq B$. Given that A and im(f) have the same size, and $\text{im}(f) \subseteq B$, it makes sense to say that A is no bigger than B.

Similarly, it should be clear that $A \ge_{\text{size}} B$ reflects the notion that A has more (or as many) elements than B, since a surjection *covers* all of B with the transformed elements of A (and many elements of A might be forced to go to the same element of B just to fit inside).

It should also be intuitive that $A \leq_{\text{size}} B$ and $A \geq_{\text{size}} B$ implies $A =_{\text{size}} B$. Proving this with what we know thus far, however, is quite difficult, being impossible. So consider the following axiom that allows us to show that this is true.

— 5B•2. Definition (Axiom)

(*Choice*) for any family of non-empty family of disjoint sets F, there is a set C which has chosen one element from each $z \in F$:

 $\forall F \ (\emptyset \notin F \land \forall x, y \in F \ (x \cap y = \emptyset) \to \exists C \ \forall x \in F \ !y \ (y \in x \cap C).$

We call such a set *C* a *choice set*.^{xvi} Really this axiom is just due to the fact that all of our proofs and formulas are finite. In the real world, each $x \in F$ is non-empty, so there is an element $a_x \in x$. So then we can consider $C = \{a_x : x \in F\}$ as a perfectly good set by replacement. The issue is that the finite nature of proofs and formulas cannot incorporate all of this in a finite number of formulas: it requires a potentially infinite number of existential instantiations. But once we have this axiom, we can show that $A \leq_{size} B$ and $A \geq_{size} B$ implies $A =_{size} B$. With this, we have the following.

- **5B**•3. Result
Moreover, for
$$f : B \to A$$
 a surjection, there is an injection $f' : A \to B$ such that $f(f'(a)) = a$ for all $a \in A$.

Proof .:.

For each $a \in A$, consider the set $f^{-1}(a) = \{b \in B : f(b) = a\}$. As f is surjective, $f^{-1}(a)$ is non-empty for each $a \in A$. Hence the family, which exists by replacement, $\{f^{-1}(a) : a \in A\} = F$ is a family of non-empty sets. Let C then be as in the axiom of choice: for each $f^{-1}(a)$, there is exactly one $b \in C \cap f^{-1}(a)$. Now

xviYou will occasionally see "choice function" as well, since this defines the function taking $x \in F \setminus \{\emptyset\}$ to the unique element of $x \cap C$.

consider the function

 $f' = \{ \langle a, b \rangle \in A \times B : a \in A \land b \in C \cap f^{-1}(a) \}.$ This is an injection $a \neq a' \in A$ requires $f^{-1}(a) \cap f^{-1}(a') = \emptyset$ (any common element *b* would need to have a = f(b) = a'). Moreover, f' is defined on all of *A*, since $C \cap f^{-1}(a)$ is has an element for each *a*; and f(f'(a)) = a because $f'(a) \in f^{-1}(a)$ so that f(f'(a)) = a.

This implies the otherwise intuitive fact below.

- **5B**•**4**. Corollary For all sets A and B, $A \leq_{\text{size}} B$ iff $B \geq_{\text{size}} A$.

Proof .:.

If $B \ge_{\text{size}} A$, then Result $5 B \cdot 3$ tells us that $A \le_{\text{size}} B$. Cleraly if $A \le_{\text{size}} B$, as witnessed by the injection $f : A \to B$, then for any fixed $a_0 \in A$, we get a surjection $g : B \to A$ defined by

$$g(b) = \begin{cases} f^{-1}(b) & \text{if } b \in \text{im } f \\ a_0 & \text{otherwise.} \end{cases}$$

This is a surjection, because $A = \text{dom } f = \text{im } f^{-1} \subseteq \text{im } g \subseteq A$ implies im g = A.

 \dashv

One of the important consequences of choice is that it allows us to count.

- **5B**•**5**. Theorem For each set *A*, there is an ordinal α such that $A =_{\text{size}} \alpha$.

Proof .:.

We construct a bijection by transfinite recursion, using the axiom of choice just once. In particular, we define a sequence of approximations to a bijection $f : \alpha \to A$ where $f \upharpoonright_{\beta} = f_{\beta} : \beta \to A$ such that $f_{\beta} \subseteq f_{\gamma}$ for $\beta < \gamma$. The way to understand the process is just that we use the axiom of choice to choose the "next" element from A. Starting from \emptyset and taking unions at limit stages, this defines the whole process.

Formally, we consider $\mathcal{O}(A) \setminus \{\emptyset\}$. Because this isn't necessarily a family of disjoint sets, consider $P'(A) = \{x \times \{x\} : x \in \mathcal{O}(A) \setminus \{\emptyset\}\}$, tagging each element with names for each subset it appears in. Thus each subset $X \subseteq A$ can be identified with $X \times \{X\} = \{\langle y, X \rangle : y \in X\}$. This P'(A) is a family of non-empty, disjoint sets, and thus there is a set C as in the axiom of choice. Note that this defines a choice function $C : \mathcal{O}(A) \setminus \{\emptyset\} \to A$ by taking C(X) to be the unique y where $\langle y, X \rangle \in (X \times \{X\}) \cap C$. Using this C, we can define our sequence of f_{α} s. In particular, $f_0 = \emptyset$ is an injection, and for γ a limit, define $f_{\gamma} = \bigcup_{\beta < \gamma} f_{\beta}$. For the successor case, suppose $f_{\alpha} : \alpha \to A$ has been defined. If $A = \operatorname{im} f_{\alpha}$, we let $f = f_{\alpha}$ and are done. Otherwise, let $f_{\alpha+1}(\alpha) = C(A \setminus \operatorname{im} f_{\alpha})$.

Note that this process has to stop at some point, because otherwise there is a surjection $g : A \to \text{Ord}$ defined by taking g(a) to be the least ordinal $\alpha \in \text{Ord}$ where $a \in \text{im } f_{\alpha}$, or else g(a) = 0. Replacement implies im g = Ord is a set, contradicting Burali–Forti Paradox (3 A • 11).

So once im $f_{\alpha} = A$, define $f = f_{\alpha}$. Consider the following easy to see facts about f:

- when f_{α} is defined, dom $(f_{\alpha}) = \alpha$.
- In particular, dom(f) is an ordinal, and f_{α} is defined iff $\alpha \leq \text{dom}(f)$.
- $f_{\alpha} \subseteq f_{\beta}$ for all $\alpha < \beta \leq \operatorname{dom}(f)$.
- Hence $f_{\alpha} = f \upharpoonright \alpha$ and so im $f_{\alpha} = f "\alpha$.

By construction $f(\alpha) \in A \setminus f''\alpha$. In particular, f is injective, since for $\alpha < \beta$, $f(\beta) \in A \setminus f''\beta$, yet $f(\alpha) \in f''\beta$. Since f is injective by construction, f is thus a bijection between an ordinal and A.

Note that this ordinal is not necessarily unique. For example, $A = \omega + 1$ has the same size as ω , because we can send $\omega \mapsto 0$ and for $n \in \omega$, we can send $n \mapsto n + 1$. This is clearly surjective onto ω , and it's injective too. So really, just

reordering the elements allows us to see that the two have the same size regardless of order^{xvii}. The notion of couting given by the ordinals is incredibly important, and leads to the next idea of size: cardinality, being the smallest ordinal of the same size.

 $-5B\cdot 6$. Definition -

Let A be a set. Define the *cardinality* of A, written |A|, to be the least ordinal α such that $A =_{\text{size}} \alpha$. An ordinal κ is a *cardinal* iff $\kappa = |\kappa|$.

Hence $A =_{\text{size}} B$ is equivalent to |A| = |B|. So in particular, $\omega + 1$ is not a cardinal. We have a number of other examples of cardinals: the finite numbers and ω , for instance. To show this, note the following easy to see facts about cardinality.

For sets A and B, writing $A <_{\text{size}} B$ for $A \leq_{\text{size}} B$ while $A \neq_{\text{size}} B$, 1. $A \leq_{\text{size}} B$ iff $|A| \leq |B|$. 2. $A \leq_{\text{size}} B$ iff $B \geq_{\text{size}} A$ (from Corollary 5 B • 4). 3. $A \leq_{\text{size}} B$ and $A \geq_{\text{size}} B$ implies $A =_{\text{size}} B$. 4. $A =_{\text{size}} B$, $A <_{\text{size}} B$, or $B <_{\text{size}} A$. 5. For $\alpha \leq \beta \in \text{Ord}$, $|\alpha| \leq |\beta|$.

Proof .:.

1. Let $f : A \to B$ be injective. Let $c_A : A \to |A|$ and $c_B : B \to |B|$ be bijections. Define the function $g : |A| \to |B|$ by taking $g(\alpha)$ to be the least β such that $\beta \notin (c_B \circ f \circ c_A^{-1})^{"}\alpha$. Note that g is therefore order preserving and hence is an embedding from |A| to |B|. If g is bijective, then it is an isomorphism and hence $|A| =_{\text{size}} |B|$, giving that |A| = |B|. Otherwise, by Lemma 3 E • 6, |A| is then isomorphic to an initial segment of |B|, and as a cardinal, |A| must be this initial segment, meaning |A| < |B|.

For the other direction, if $|A| \leq |B|$, then bijections $c_A : A \to |A|$ and $c_B : B \to |B|$ yield the injection $c_B^{-1} \circ c_A : A \to B$.

- 3. This is immediate from (1) and (2): $A \leq_{\text{size}} B$ implies $|A| \leq |B|$. $A \geq_{\text{size}} B$ is equivalent to $A \geq_{\text{size}} B$ which is just saying $|A| \geq |B|$, and therefore |A| = |B|. Using bijections $c_A : A \rightarrow |A|$ and $c_B : B \rightarrow |B| = |A|$ yields the bijection $c_B^{-1} \circ c_A : A \rightarrow B$ telling us that $A =_{\text{size}} B$.
- 4. This follows from the same relation happening for ordinals.
- 5. Clearly $\alpha \leq_{\text{size}} \beta$ since the identity function id $\uparrow \alpha = \{\langle x, x \rangle \in \alpha \times \alpha : x \in \alpha\}$ is an injection from α to β . So by (1), $|\alpha| \leq |\beta|$.

So this notion of counting gives some very nice properties regarding size, most of which should be expected, and allows us to write $|A| \ge |B|$ instead of $A \ge_{size} B$ and so forth. So we will abandon the "size" inequalities until we develop the theory of cardinals without the axiom of choice. Beyond the above results, we also get the following famous principle.

- 5B.8. Corollary (The Pigeonhole Principle) For all sets A and B, suppose |A| < |B|. Therefore, if $f : B \to A$, then f is not injective. Moreover, any $f : A \to B$ is not surjective.

Proof .:.

If $f : B \to A$ is injective, then $B \leq_{\text{size}} A$ and hence $|B| \leq |A|$, contradicting that |A| < |B|. Similarly, if $f : A \to B$ is surjective, then $|A| \geq |B|$, again a contradiction. \dashv

Now let's get on to proving what the cardinals are. Examples of non-cardinals are abundant. For example, $\omega + \omega$ can

^{xvii}Clearly ω and $\omega + 1$ are not *isomorphic* as orders, but disregarding order, they have the same size.

be put in bijection with ω since we can rename the first copy of ω with even numbers, and the second copy of ω with the odd numbers. It will be a goal to show that there exist larger cardinals than ω , since even $\omega \cdot \omega$ can be shown to have cardinality ω . Firstly, we have that every natural number is a cardinal number.

- **5B**•**9**. **Result** Let $n \in \omega$. Therefore *n* is a cardinal.

Proof .:.

Proceed by induction on *n*. For n = 0 this is immediate: a bijection $f : 0 \to m$ will have $f \subseteq 0 \times m = \emptyset$ so that $f = \emptyset$ and thus $0 = \emptyset = \text{im } f = m$.

For n + 1, it suffices to show that |n + 1| > n by (5) of Result 5 B • 7. So suppose $f : n + 1 \rightarrow n$ is a bijection. Consider f"n which then has size n. But $f"n = n \setminus \{f(n)\}$. Now we show that this is impossible. If n = 0 or n = 1, this is clearly impossible, because n = 0 has $f = \emptyset$, and n = 1 has $1 \setminus \{f(1)\} = \emptyset$, which requires that $f \upharpoonright 1 : 1 \rightarrow \emptyset$ is a bijection.

For $n = n^* + 1$ where $n^* \ge 1$, there is a clear bijection between n^* and $n \setminus \{f(n)\}$, as we will show. Explicitly, take $g : n^* \to n \setminus \{f(n)\}$ where

$$g(k) = \begin{cases} k & \text{if } k < f(n) \\ k+1 & \text{if } k \ge f(n). \end{cases}$$

This is a bijection. Clearly it's injective, so it suffices to show surjectivity. To see this, any $k \in n \setminus \{f(n)\}$ has $k \neq f(n)$ and $k \leq n^*$. If $k < f(n) \leq n^*$ then we obviously have $g(k) = k \in \text{im } g$. If $f(n) < k \leq n^*$, then k > 0 and hence there is some $k^* \in \omega$ where $k = k^* + 1$ (this is where we use the fact that ω is the least limit ordinal) and this satisfies $f(n) \leq k^*$. Hence $g(k^*) = k^* + 1 = k$. So g is surjective, meaning g is a bijection between $n \setminus \{f(n)\}$ and n^* . Since $f \upharpoonright n : n \to n \setminus \{f(n)\}$ is a bijection, we have a bijection between n and $n^* < n$, contradicting the inductive hypothesis. Therefore no such f can exist.

We also have that ω is a cardinal.

- **5B**•10. Result — The supremum of cardinals is a cardinal. In particular $\omega = \sup_{n \in \omega} n$ is a cardinal.

Proof .:.

Let X be a set of cardinals with $\chi = \sup X$. Clearly if X has a maximal element, then χ is this, and so $\chi \in X$ is a cardinal. So suppose X has no maximal element. If $|\chi| < \chi$, then there is some cardinal $\kappa \in X$ with $|\chi| \le \kappa$. But since there is some larger cardinal $\lambda \in X$ with then $\chi \ge \lambda > \kappa$, it follows by (5) from Result 5 B • 7 that $|\chi| \ge \lambda > \kappa \ge |\chi|$, a contradiction. Therefore $|\chi| \ge \chi$. We always have by definition of cardinality that $|\chi| \le \chi$, and so $|\chi| = \chi$.

Hence we have a dichotomy between ω and the smaller sets, which has already been talked about. Formally, we have the following.

- **5B**•**11**. **Definition** A set *A* is finite iff $|A| < \omega$. A set is infinite iff $|A| \ge \omega$.

So this gives us limit cardinals like ω . But what comes after ω ? Certainly there are no infinite cardinals that are successor ordinals.

- 5B•12. Result

Let κ be an infinite cardinal. Therefore κ is not a successor ordinal.

Proof .:.

Let $\alpha + 1$ be a successor ordinal. Write $|\alpha| = \lambda$. Consider a bijection $b : \alpha \to \lambda$. Now consider the bijection

defined by

$$f(\xi) = \begin{cases} b(\xi+1) & \text{if } \xi \in \omega \\ b(0) & \text{if } \xi = \alpha \\ b(\xi) & \text{otherwise.} \end{cases}$$

This has $f: \alpha + 1 \to \lambda \le \alpha$ as a bijection, meaning $|\alpha + 1| \ne \alpha + 1$.

But are there any cardinals larger than ω ? The answer to this question is an emphatic yes. In fact, there are just as many cardinals as there are ordinals. And consistently, there are just as many ordinals as there are sets. To generate these cardinals, consider the following theorem, often considered the result that gave birth to the field of set theory, and inspired Russell's Paradox (2 • 6).

5B•13. Theorem (Cantor's Theorem) Let X be a set. Therefore $|X| < |\mathcal{P}(X)|$.

Proof .:.

Let $f: X \to \mathcal{P}(X)$. Consider the set $A = \{x \in X : x \notin f(x)\}$. This is a definable subset of X, and so clearly $A \in \mathcal{P}(X)$. If f were surjective, then A = f(a) for some $a \in A$. So we can ask whether $a \in A$ or not. If $a \in A$, then it meets the definition: $a \notin f(a) = A$, which is a contradiction. Hence $a \notin A$. But this means a doesn't meet the definition: $a \in f(a) = A$. Again, we have a contradiction, and so $A \in \mathcal{P}(X) \setminus \text{im}(f)$.

Therefore $|\omega| < |\mathcal{P}(\omega)|$, and thus there are larger cardinals than ω . In fact, this theorem gives that there is *no* largest cardinal, since any cardinal κ has $\mathcal{P}(\kappa) > \kappa$. With this information under our belt, consider the following definition.

Define by transfinite recursion the infinite cardinals. $\aleph_0 = \omega$ $\aleph_{\alpha+1}$ = the least cardinal greater than \aleph_{α} $\aleph_{\gamma} = \sup \aleph_{\beta}$, for γ a limit.

 $\beta < \gamma$

Although the two are the same as sets, when referring to \aleph_{α} as an ordinal rather than a cardinal, write ω_{α} .

So this allows us to consider truly large sets: \aleph_2 , \aleph_{ω_1} , $\aleph_{\omega_{\omega}}$, and so on.

— 5B•15. Corollary

5B•14. Definition

The sequence of $n < \omega$ and \aleph_{α} s exhausts all of the cardinals and cardinalities.

Proof .:.

Proceed by induction on α where $\alpha = |\alpha|$. Clearly if $\alpha < \omega$ then we're done. Otherwise, consider $X = \{\beta < \alpha : |\beta| = \beta\}$. This is a set of cardinals, and its supremum λ is then a cardinal by Result 5 B • 10. Note that inductively each $\beta \in X$ has $\beta = \aleph_{\gamma}$ for some γ . In particular, for $\delta = \sup\{\gamma + 1 : \aleph_{\gamma} \in X\}$, we have that $X = \{\aleph_{\gamma} : \gamma < \delta\}$ and thus sup $X = \aleph_{\delta}$. Because $\alpha \ge \sup X$, either $\alpha = \sup X = \aleph_{\delta}$, or $\alpha > \sup X$, and is thus the least cardinal greater than \aleph_{δ} , meaning $\alpha = \aleph_{\delta+1}$.

But the definition of the alephs raises the question that allowed us to even consider larger cardinals: what is $|\mathcal{P}(\omega)|$? Where on the long line of alephs is this? Note that the above tells us that $|\mathcal{P}(\omega)| \ge \aleph_1$, but it's not clear whether this equality holds or not. The statement that $|\mathcal{P}(\omega)| = \aleph_1$ is often referred to as *the continuum hypothesis* or CH. Many set theorists have—often very complicated—reasons for thinking that CH is false and instead that $|\mathcal{P}(\omega)| = \aleph_2$.^{xviii} We will return to this question after investigating what cardinality looks like in a world without choice.

^{xviii}Most of this will not be covered in this book, but for those interested, a search for PFA will lead one in the right direction. Be warned, however, that the proof that $ZFC \vdash PFA \rightarrow |\mathcal{P}(\omega)| = \aleph_2$ is incredibly long, dealing with complicated set theoretic postulates independent of the other axioms, and full of a technical method called "forcing".

§ 5 C. Cardinality without choice

In the world of choice, the equivalence relation of $=_{size}$ has canonical representatives in the form of ordinals called *cardinals*, and so every set can be compared in size. In particular, \leq_{size} , \geq_{size} and $=_{size}$ are all just different parts of a single linear order: modulo $=_{size}$, \leq_{size} (the existence of an injection) is one direction and \geq_{size} (the existence of a surjection) is the reverse direction.

In general, $=_{size}$ is still an equivalence relation, and \leq_{size} is still a partial order modulo $=_{size}$, but it's not necessarily the case that it's linear, nor that $A \leq_{size} B$ is equivalent to $B \geq_{size} A$. How, then, do we define cardinality? How do we choose canonical representatives for the equivalences classes of $=_{size}$? The issue is that we can't, and so in a choiceless context, we don't even try to define representatives of the $=_{size}$ equivalence classes in general. We can still do this for ordinals, yielding the same notion of what a cardinal is, but this is only because the ordinals have a canonical order on them. Without choice, there isn't always a well-order on sets.

— 5C•1. Lemma

The axiom of choice (AC) is equivalent to the statement that every set has a well-order.

Proof .:.

Let X be an arbitrary set, and suppose AC holds. By Theorem 5 B • 5, there is a bijection $f : X \to \alpha$ for some $\alpha \in \text{Ord.}$ Hence the order $W = \{\langle x, y \rangle \in X \times X : f(x) < f(y)\}$ induced by f makes f an isomorphism between $\langle X, W \rangle$ and $\langle \alpha, \in \rangle$, meaning W is a well-order.

Now suppose every set can be well-ordered. Let F be an arbitrary set of disjoint, non-empty sets. Consider $X = \bigcup F$. This has a well-order W. Hence each $x \in F$ has a W-least element, called $a_x \in x$. Moreover, $a_x \notin y$ for each $y \in F \setminus \{x\}$ since the elements of F are pairwise disjoint. Therefore, the set $C = \{a_x : x \in F\}$ works as a choice set for F.

5C·2. Corollary -

AC holds for families of sets of ordinals. Hence all parts of Result 5 B \cdot 7 holds for $A, B \in$ Ord.

Ostensibly, as with choice, for arbitrary X we can take the \in -least element of { $\alpha \in \text{Ord} : \alpha =_{\text{size}} X$ } and thus arrive at a cardinality for X as before. The issue is that it's not clear this set is non-empty, and in fact, if choice fails then this will be empty for some X as Lemma 5 C • 1 shows.

— 5C•3. Definition

Let X be a set. The *choiceless-cardinality* of X is the equivalence class $[X]_{size} = \{A : A =_{size} X\}$. A *cardinal* is still an ordinal κ such that κ is <-minimal in Ord $\cap [\kappa]_{size}$.

Note that $[X]_{size}$ will be a class rather than a set. With these concepts, we still have the following results about \leq_{size} . Namely, that \leq_{size} is antisymmetric modulo $=_{size}$. A similar result was shown with choice: Result 5 B • 7 (3), where $A \leq_{size} B$ and $A \geq_{size} B$ implies $A =_{size} B$. But this was done by comparing cardinality rather than defining a bijection outright.

— 5C·4. Theorem (Cantor-Bernstein) ——

Let A and B be sets. Suppose $A \leq_{\text{size}} B$ and $B \leq_{\text{size}} A$. Therefore $A =_{\text{size}} B$.

Proof .:.

Let $\mathbb{A} : A \to B$ and $\mathbb{B} : B \to A$ be injections, witnessing the hypothesis. We will categorize the elements of *B* in the following way. Call elements $b \in B \setminus im(\mathbb{A})$ starting points.

For each starting point $b_0 \in B$, we can then identify the path it takes by going to A via \mathbb{B} , then then back to B via \mathbb{A} . Write $(\mathbb{A} \circ \mathbb{B})^n$ for $(\mathbb{A} \circ \mathbb{B}) \circ (\mathbb{A} \circ \mathbb{B}) \circ \cdots \circ (\mathbb{A} \circ \mathbb{B})$, meaning $\mathbb{A} \circ \mathbb{B}$ composed *n*-times for each $n \in \mathbb{N}$. So a point $b \in B$ is on the path of b_0 iff $b = (\mathbb{A} \circ \mathbb{B})^n (b_0)$ for some $n \in \mathbb{N}$, possibly 0. Now we define $f : A \to B$

via replacement by

$$f(a) = \begin{cases} \mathbb{B}^{-1}(a) & \text{if } \mathbb{A}(a) \text{ is on the path of a starting point,} \\ \mathbb{A}(a) & \text{otherwise.} \end{cases}$$

This makes sense as \mathbb{B} is injective: \mathbb{B}^{-1} is a function. To see this, if $\langle a, b \rangle, \langle a, b' \rangle \in \mathbb{B}^{-1}$ for $b \neq b'$, then $\mathbb{B}(b) = \mathbb{B}(b') = a$ contradicts injectivity. Hence f is a function defined on all of A, and clearly im $f \subseteq B$.

So it suffices to show that f is injective, and surjective.

— Claim 1 –

f is injective.

Proof .:.

Suppose f(a) = f(a') for $a \neq a'$. If both $\mathbb{A}(a)$ and $\mathbb{A}(a')$ are on the path of a starting point, then $f(a) = \mathbb{B}^{-1}(a) = f(a') = \mathbb{B}^{-1}(a')$. This contradicts that \mathbb{B} is a function: $\langle a, b \rangle, \langle a', b \rangle \in \mathbb{B}^{-1}$ implies $\mathbb{B}(b)$ is both a and a'. So this case can't happen. Similarly, if neither $\mathbb{A}(a)$ nor $\mathbb{A}(a')$ is on the path of a starting point, then $f(a) = \mathbb{A}(a) = f(a') = \mathbb{A}(a')$ contradicts the injectivity of \mathbb{A} .

So suppose for the sake of definiteness that A(a) is on the path of a starting point, but A(a') isn't. Note that f(a) is then on the path of a starting point, because $A \circ B(f(a)) = A \circ B(B^{-1}(a)) = A(a)$ on the path of a starting point. A(a) of course is not itself a starting point, since it's in the image of A, but f(a) might be. Anyway, f(a) being on a path means that f(a) = f(a') = A(a') is too, a contradiction.

All that remains to be shown is that f is surjective. To do this, let $b \in B$. If b is on the path of a starting point, $a = \mathbb{B}(b)$ yields f(a) = b. If b is not on the path of a starting point, then certainly b itself is not a starting point, meaning $b \in \text{im}(\mathbb{A})$. So taking such an a with $\mathbb{A}(a) = b$ yields that \mathbb{A} isn't on the path of a starting point, and thus $f(a) = \mathbb{A}(a) = b$. Thus f is surjective, and so a bijection.

As detailed above, it's tempting for each X to define $\{\alpha \in \text{Ord} : \alpha =_{\text{size}} X\}$, and then take |X| to be the <-least element of this class. Although we can't do this because the class might be empty, we still at least have the following result, showing that the ordinals can still overtake any set in the \leq_{size} -ordering.

– 5C•5. Theorem (Hartogg's Number)

Let *X* be a set. Therefore there is a cardinal κ such that $\kappa \not\leq_{\text{size}} X$.

Proof .:.

Consider the approximations to a well-order of X. In particular, consider the set

 $\mathcal{W} = \{W \in \mathcal{O}(X \times X) : W \text{ is a well-order of } \operatorname{dom}(W) \cup \operatorname{ran}(W).\}$

Now by Corollary 3 E • 7, we can consider the set of the corresponding order-types.

 $\mathcal{O} = \{ \alpha \in \operatorname{Ord} : \exists W \in \mathcal{W}(\langle \alpha, \epsilon \rangle \cong \langle \operatorname{dom}(W) \cup \operatorname{ran}(W), W \rangle) \}.$

Note that \mathcal{O} must be an ordinal, since it is transitive: $\beta < \alpha \in \mathcal{O}$ has that the well-order $W \in \mathcal{W}$ with order-type α can be restricted to an initial segment with order-type β and thus $\beta \in \mathcal{O}$. So it suffices to show that $\mathcal{O} \not\leq_{\text{size}} X$.

Suppose $f : \mathcal{O} \to X$ is an injection. Therefore the order $W = \{\langle f(\alpha), f(\beta) \rangle \in X \times X : \alpha < \beta\}$ yields a well-order of a subset of X that is isomorphic to \mathcal{O} . In particular, $W \in W$ and $\mathcal{O} \in \mathcal{O}$, contradicting that the ordinals are well-founded.

If choice holds, the cardinal described above is just any cardinal greater than |X|. But without choice, it's not clear that $X =_{size} \alpha$ for any $\alpha < \kappa$, as this would guarantee by Cantor–Bernstein (5 C • 4) that X can be well-ordered, and thus any family $F \subseteq \mathcal{P}(X)$ would have a choice set just by selecting the least-elements in the non-empty sets. So if every X has a cardinality, then we always get choice sets, and thus the axiom of choice holds.

Hartogg's Number (5 C • 5) is especially useful in confirming the other properties of the cardinals that we know from Subsection 5 B. There, ω_1 was shown to exist from a well-order of $\mathcal{P}(\omega)$. But without choice, it's possible for $\mathcal{P}(\omega)$

to have no well-order. How then do we show that there are larger cardinalities? We use Hartogg's Number (5 C • 5). Note that we still have the usual properties of \leq_{size} due to choice holding on the ordinals by Corollary 5 C • 2.

```
For each cardinal \kappa \in \text{Ord}, there is a cardinal \lambda > \kappa.
```

Hence without choice we can still define $\aleph_1, \aleph_2, \dots, \aleph_{\omega}$, and so on. It's just that not every set needs to be in bijection with one of these.

§ 5 D. cofinality and cardinal arithmetic

We now return to the world of choice, although often it is unnecessary for this subsection. There will be times when it is needed, but mostly this is just in requiring that functions from κ to λ can be well-ordered for κ and λ ordinals.

As introduced before, there are operations defined on ordinal numbers: addition, multiplication, and exponentiation, for example. We have similar operations on cardinals, although they do not obey the same rules. It will happen that everything becomes either dramatically simpler, or else impossible to know. We begin with some notable properties of cardinals. We begin with addition.

- **5D**•1. **Definition** Let X and Y be sets. Write $X \sqcup Y$ for $(X \times \{0\}) \cup (Y \times \{1\})$ the disjoint union. Let κ and λ be cardinals. Define $\kappa + \lambda$ to be the cardinality of $\kappa \sqcup \lambda$. Define $\kappa \cdot \lambda$ to be the cardinality of $\kappa \times \lambda$.

Note that the cardinality of the ordinal addition $\kappa + \lambda$ is the cardinal addition $\kappa + \lambda$. To make this more apparent what is meant, $|\omega + \omega_1| = \aleph_0 + \aleph_1$ for example. Unlike with ordinal addition, where $\omega + 1 \neq \omega$, both of these cardinal operations simplify to just being the maximum of the two cardinals. First we have some immediate properties about these operations, just following from the existence of easy to find injections or surjections. Below, for κ , λ , and θ cardinals:

- Cardinal addition and multiplication are commutative.
- $\kappa < \lambda$ implies $\theta \cdot \kappa \le \theta \cdot \lambda$ and $\theta + \kappa \le \theta + \lambda$ (possibly with equality, as we shall see).
- $\kappa + 0 = \kappa$, and $\kappa \cdot 0 = 0$.
- $\kappa + 1 = \kappa$, and $\kappa \cdot 1 = \kappa$.
- $\kappa + \lambda \leq \kappa \cdot \lambda$ when $\lambda \neq 0$.
- $\alpha \leq \aleph_{\alpha}$ (possibly with equality, as we shall see).

Trivially, however, these facts won't be important to know, since we will get that $\kappa + \lambda = \kappa \cdot \lambda = \max(\kappa, \lambda)$. To show this, we first consider the case where $\kappa = \lambda$.

- **5D**•**2**. Lemma — Let κ be an infinite cardinal. Therefore $\kappa \cdot \kappa = \kappa + \kappa = \kappa$.

```
Proof .:.
```

Clearly $\kappa \leq \kappa + \kappa \leq \kappa \cdot \kappa$ so it suffices to show $\kappa \cdot \kappa \leq \kappa$. We consider the following ordering on Ord × Ord. Write $\langle \alpha_0, \beta_0 \rangle \prec \langle \alpha_1, \beta_1 \rangle$ iff 1. $\max(\alpha_0, \beta_0) < \max(\alpha_1, \beta_1)$; or else 2. $\alpha_0 < \alpha_1$; or else 3. $\beta_0 < \beta_1$. In essence, we have a lexicographic order where things in the square $\gamma \times \gamma$ always precede things in the square $\delta \times \delta$ for $\gamma < \delta$. As a result, this means we follow the edges of increasingly bigger squares. Claim 1 \prec is a (class) well-order of Ord × Ord. Proof .:.

That \prec is a linear order should be easy to see from the definition: transitivity follows from from progressing through the cases each time, and the other requirements follow from < being a linear order on Ord. To show that \prec is well-founded, let X be a set of pairs of ordinals. Consider the set $Y = \{\max(\alpha, \beta) : \langle \alpha, \beta \rangle \in X\}$. This has a \in -least element $\alpha_0 \in Y$ so consider the class $\{\langle \alpha, \beta \rangle \in X : \max(\alpha, \beta) = \alpha_0\}$. Now we similarly choose the \in -least first entry in this set, and of those entries with the same max and same first entry, we conside the \in -least second entry. This gives a \prec -least element of X just by definition of \prec .

Proceed by induction on γ to show $\aleph_{\gamma} \cdot \aleph_{\gamma} = \aleph_{\gamma}$ by showing $\prec_{\gamma} = \prec \cap ((\aleph_{\gamma} \times \aleph_{\gamma}) \times (\aleph_{\gamma} \times \aleph_{\gamma}))$ has order-type \aleph_{γ} . Because we prioritize smaller squares, for each $\langle \alpha, \beta \rangle \in \text{Ord} \times \text{Ord}$, $\text{pred}_{\prec}(\langle \alpha, \beta \rangle)$ is a set, and in particular, it has order-type at most (using ordinal multiplication) $\max(\alpha, \beta) \cdot \max(\alpha, \beta)$. So for $\alpha, \beta < \aleph_{\gamma}$, the inductive hypothesis tells us that this $\text{pred}_{\prec}(\langle \alpha, \beta \rangle)$ has cardinality $|\max(\alpha, \beta)| \cdot |\max(\alpha, \beta)| = |\max(\alpha, \beta)| < \aleph_{\gamma}$, and thus the order-type of this initial segment is an ordinal strictly less than \aleph_{γ} . Thus every initial segment of \prec_{γ} has order-type strictly less than \aleph_{γ} , and therefore the order-type of \prec_{γ} is at most \aleph_{γ} . Since clearly the order-type is at least \aleph_{γ} (consider $\{\langle \alpha, 0 \rangle : \alpha < \aleph_{\gamma}\}$, still well-ordered by \prec and isomorphic to $\langle \aleph_{\gamma}, \langle \rangle$, and use Lemma 3 $E \cdot 6$), we have equality and thus $|\aleph_{\gamma} \times \aleph_{\gamma}| = \aleph_{\gamma}$.

We can then conclude that $\kappa \cdot \lambda = \kappa + \lambda = \max(\kappa, \lambda)$ for infinite cardinals κ and λ .

 $\begin{array}{c} \textbf{5D-3. Corollary} \\ \hline \text{Let } \kappa < \lambda \text{ be cardinals with } \lambda \text{ infinite. Therefore } \kappa \cdot \lambda = \kappa + \lambda = \lambda. \end{array} \\ \hline Proof ::. \\ \hline \lambda = 1 \cdot \lambda \leq \kappa \cdot \lambda \leq \lambda \cdot \lambda = \lambda \text{ by Lemma 5 D • 2, and similarly for addition.} \end{array}$

Now we will discuss some aspects of cardinal arithmetic that are more complicated in the sense that it's impossible to write down precisely which \aleph_{α} the answer is. But there are still interesting results we can give.

- **5D**•**4**. **Definition** Let *A* and *B* be sets. Define ${}^{A}B = \{f \in \mathcal{P}(A \times B) : f \text{ is a function from } A \text{ to } B\}$. For κ and λ cardinals, define $\kappa^{\lambda} = |{}^{\lambda}\kappa|$. We often write $\kappa^{<\lambda}$ for $\sup_{\theta < \lambda} \kappa^{\theta}$.

We will interchangeably write $X^{<\omega} = \bigcup_{n < \omega} X^n$ or ${}^{<\omega}X = \bigcup_{n < \omega} {}^nX$, which can be identified in a similar way. In particular, rather than the *n*-tuple $\langle x_0, \dots, x_{n-1} \rangle \in X^n$, we can think of this as instead a function specifying the value of the *k*th entry: $f : n \to X$ where $f(k) = x_k$ for k < n. Similarly for a function $f : n \to X$, we can think of this as a tuple that merely lists out the values of $f : \langle f(0), \dots, f(n-1) \rangle$. So mixing up the two only slightly changes the implementation of the same idea: finite lists of elements from *X*.

Now obviously we get the following facts about exponentiation: for all cardinals κ , λ , $\theta \in$ Ord, and all sets A,

- $^{\emptyset}A = \{\emptyset\}, {}^{1}A = 1 \times A.$
- $\kappa < \lambda$ implies $\theta^{\kappa} < \theta^{\lambda}$;
- $\kappa < \lambda$ implies $\kappa^{\theta} \leq \lambda^{\theta}$;
- $\kappa^0 = 1, \kappa^1 = \kappa.$
- $\kappa^2 = \kappa \cdot \kappa$, and so successively, $\kappa^n = \kappa$ for each $n \in \omega$, if κ is infinite. Therefore $\kappa^{<\omega} = \kappa$.

The notation of exponentiation makes sense for this operation because if $f : A \sqcup B \to C$, meaning $f \in {}^{A \sqcup B}C$, then we can view f according to how it acts on A and how it acts on B. In particular, every function in ${}^{A \sqcup B}C$ can be viewed as a pair of functions in ${}^{A}C \times {}^{B}C$, and vice versa (because we're taking the disjoint union). Moreover, the idea of evaluation just gives that ${}^{A}({}^{B}C)$ is effectively the same as ${}^{A \times B}C$ in that every function $f : A \to {}^{B}C$ can be uniquely identified with the map $g : A \times B \to C$ where g(a, b) = f(a)(b). Hence we get the following facts about cardinal exponentiation. - **5D** • **5**. **Result** Let κ , λ , and θ be cardinals. Therefore, 1. $\theta^{\kappa+\lambda} = \theta^{\kappa} \cdot \theta^{\lambda}$. 2. $(\theta^{\kappa})^{\lambda} = \theta^{\kappa \cdot \lambda}$.

Moreover, these concepts also allow us to view the powerset as exponentiation.

- **5D**•6. Result Let X be a set. Therefore $|\mathcal{P}(X)| = |^X 2| = 2^{|X|}$.

Proof .:.

Each subset corresponds to its characteristic function: $A \subseteq X$ yields $\chi_A : X \to 2$ where $\chi_A(x) = 1$ iff $x \in A$ and otherwise $\chi_A(x) = 0$. Hence $\chi_A^{-1}(1) = A$ for all $A \subseteq X$. In particular, if $\chi_A = \chi_B$ then they both have the same preimage of 1 and so A = B. Similarly, every function $f : X \to 2$ yields a unique subset of X just by the preimage of 1: $A_f = \{x \in X : f(x) = 1\} = f^{-1}(1)$ and so $\chi_{A_f} = f$. Hence the map $F : \mathcal{O}(X) \to X^2$ where $A \mapsto \chi_A$ is a bijection. Therefore $|\mathcal{O}(X)| = |X^2|$ which is just $2^{|X|}$ by definition.

In particular, $|\mathcal{O}(\aleph_0)| = 2^{\aleph_0} > \aleph_0$ by Cantor's Theorem (5 B • 13), which more generally gives the following.

- **5D**•**7**. **Corollary** - Let κ be a cardinal. Therefore $2^{\kappa} > \kappa$.

Note that $\kappa^{\kappa} = 2^{\kappa}$ for infinite κ , since

 $2^{\kappa} \leq \kappa^{\kappa} \leq (2^{\kappa})^{\kappa} = 2^{\kappa \cdot \kappa} = 2^{\kappa}.$

Another proof that $2^{\kappa} > \kappa$ follows from a very useful theorem. First, note that we can generalize exponentiation to other products, and we generalize multiplication to other sums.

- **5D**•8. **Definition** Let *I* be a set, and let $\{\kappa_i : i \in I\}$ be a set of ordinals. The cardinal sum $\sum_{i \in I} \kappa_i$ is the cardinality of the union $\bigcup_{i \in I} \kappa_i \times \{i\}$. The cardinal product $\prod_{i \in I} \kappa_i$ is the cardinality of the cartesian product $\prod_{i \in I} \kappa_i$.

Obviously we have $\sum_{i \in I} \kappa_i \leq \prod_{i \in I} \kappa_i$ just by looking at the map sending $\langle \alpha, i \rangle \in \bigcup_{i \in I} \kappa_i \times \{i\}$ to the function in the cartesian product $\prod_{i \in I} \kappa_i$ where $i \mapsto \alpha$ and $j \mapsto 0$ for every $j \in I$ with $j \neq i$. We also have the following easy to confirm properties.

- $I \subseteq J$ with $\{\kappa_j : j \in J\}$ a set of cardinals implies $\sum_{i \in I} \kappa_i \leq \sum_{i \in J} \kappa_i$; and
- if in addition, $\emptyset \notin J$, $\prod_{i \in I} \kappa_i \leq \prod_{j \in J} \kappa_j$.
- $\kappa_i \leq \theta_i$ implies $\sum_{i \in I} \kappa_i \leq \sum_{i \in I} \theta_i$, and similarly for products.
- $\sum_{i \in I} 1 = |I|$ and $\prod_{i \in I} 2 = 2^{|I|}$; and more generally,
- $\sum_{i \in I} \kappa = |I| \cdot \kappa$ and $\prod_{i \in I} \kappa = \kappa^{|I|}$.

Mostly we will look at sums as given by partitions: if we can cover a set, then the cardinality is given by how many pieces we need, and how big the pieces are.

– 5 D•9. Result -

Let X be a set, and $P \subseteq \mathcal{O}(X)$ a partition of X such that |P| is infinite. Therefore $|X| = \sum_{Y \in P} |Y| = |P| \cdot \sup_{Y \in P} |Y|$.

Proof .:.

Since X can be written as the disjoint union $X = \bigcup_{Y \in P} Y$, it's clear that $\{|Y| : Y \in P\}$ is a set of cardinals, and $\bigcup_{Y \in P} |Y| \times \{Y\}$ is in bijection with X, just by sending $\langle \alpha, Y \rangle$ to the $f_Y(\alpha)$ where $f_Y : |Y| \to Y$ is a bijection. As a result, $|X| = \sum_{Y \in P} |Y|$.

This is the same as $|P| \cdot \sup_{Y \in P} |Y|$, since $|P| = \sum_{Y \in P} 1 \le \sum_{Y \in P} |Y| \le \sum_{Y \in P} |P| = |P| \cdot |P| = |P|$. \dashv

Infinite sums in general work like this.

— 5 D•10. Corollary -

Let *I* be a set, and $\{\kappa_i : i \in I\}$ a set of cardinals. Therefore $\sum_{i \in I} \kappa_i = |I| \cdot \sup_{i \in I} \kappa_i$.

A less trivial theorem is the following.xix giving an alternative proof of Cantor's Theorem (5 B • 13).

👝 5 D • 11. Theorem (König's Theorem) -

Let *I* be a set (used only as an index), and let $\{\kappa_i : i \in I\}$ and $\{\theta_i : i \in I\}$ be two sets of cardinals. Suppose $\kappa_i < \theta_i$ for all $i \in I$. Therefore $\sum_{i \in I} \kappa_i < \prod_{i \in I} \theta_i$.

Proof .:.

Without loss of generality, instead consider the situation where we have pairwise disjoint families

$$\{\mathbf{K}_i : i \in I\}, \qquad |\mathbf{K}_i| = \kappa_i$$

$$\{\Theta_i : i \in I\}, \qquad |\Theta_i| = \theta_i.$$

For example, $K_i = \kappa_i \times \{i\}$ and $\Theta_i = \theta_i \times \{i\}$ works. Let $K = \bigcup_{i \in I} K_i$, and Θ equal the cartesian product $\prod_{i \in I} \Theta_i$ (i.e. the set of all functions f from I with $f(i) \in \Theta_i$). It's clear that there's an injection from K to Θ , just because $\kappa_i < \theta_i$: send $\langle \alpha, i \rangle \in K$ to the map $f \in \Theta$ defined by, for $j \in I$,

$$f(j) = \begin{cases} \langle \alpha, j \rangle & \text{if } j = i \\ \langle \alpha + 1, j \rangle & \text{if } j \neq i. \end{cases}$$

Now suppose we had a surjection $F : K \to \Theta$. We will diagonalize out of this using evaluation maps: for $x \in I$ and $f \in \Theta$, $\iota_x(f) = f(x)$. Let $F_i = (\iota_i \circ F) \upharpoonright K_i$, a function from K_i to $\bigcup_{i \in I} \Theta_i$.

Since $|\Theta_i| > |K_i|$, as a function from K_i , F_i can never cover Θ_i . So let $g(i) \in \Theta_i \setminus im(F_i)$ for each $i \in I$. The resulting function g cannot be in the image of F. To see this, if we let $k \in K$ be such that F(k) = g, then we know $k \in K_i$ for exactly one $i \in I$. Hence $k \in dom(F_i)$ and so $g(i) \in \Theta_i \setminus \{F_i(k)\}$ by construction. Yet $F_i(k) = (\iota_i \circ F)(k) = F(k)(i) = g(i)$, a contradiction.

– 5D·12. Corollary –

Let κ be a cardinal. Therefore $2^{\kappa} > \kappa$.

Proof .:.

Since 2 > 1, $\kappa = \sum_{i \in \kappa} 1 < \prod_{i \in \kappa} 2 = 2^{\kappa}$ by König's Theorem (5 D • 11).

 \dashv

This raises the question, how much more can we know about 2^{κ} , and λ^{κ} more generally? Because cardinal exponentiation grows in both arguments, we at least know that $2^{\kappa} = \kappa^{\kappa}$ since $2^{\kappa} \leq (2^{\kappa})^{\kappa} = 2^{\kappa \cdot \kappa} = 2^{\kappa}$ by Result 5 D • 5 and Lemma 5 D • 2. For now, the question we will address is just when does 2^{λ} cross from being below κ to being above? To give the best possible answer we can give, we need a new concept.

— 5D•13. Definition -

A *poset* is a structure $\mathbf{A} = \langle A, \langle A \rangle$ where $\langle A \rangle$ is a partial order over A. Let $\mathbf{A} = \langle A, \langle A \rangle$ be a poset. A subset $X \subseteq A$ is *cofinal* in \mathbf{A} iff $\forall a \in A \exists x \in X \ (a = x \lor a \land A x)$. For α an ordinal, the *cofinality* of α , written $cof(\alpha)$, is the least order-type of a cofinal subset of α .

If α is an ordinal, we say that $X \subseteq \alpha$ has order-type β for the more formal statement that $\langle X, \in \rangle$ has order-type $\langle \beta, \in \rangle$. For linear orders, being cofinal is the same as being unbounded. So $cof(\alpha)$ is also the least order-type of an unbounded subset of α . One may expect that if $\alpha > \beta > cof(\alpha)$, then there is a cofinal subset of α with order-type β . But this may not be true, paradoxically. The main reason is that after using $cof(\alpha)$ many elements of β , we might run out of room to place the other elements of β while preserving the order. First, note the following easy examples: for α an ordinal,

- $cof(\alpha + 1) = 1$, as witnessed by $\{\alpha\} \subseteq \alpha + 1$.
- $\operatorname{cof}(\aleph_{\omega}) = \omega$, as witnessed by $\{\aleph_n : n < \omega\}$.
- $cof(\aleph_{\alpha}) = cof(\alpha)$ for α a limit, by the same reason.
- $\operatorname{cof}(\aleph_0) = \omega$.

xix named after Kőnig Gyula who often published under the pseudonym "Julius König".

• $cof(\alpha) \le \alpha$ for each ordinal α .

An arguably easier way to characterize cofinality is with functions.

```
5D-14. Definition 
An ordinal \lambda is cofinal in \alpha \in \text{Ord iff there is an increasing function } f : <math>\lambda \to \alpha such that im f is cofinal in \alpha.
```

This is an alternative way to characterize it in the following sense.

```
– 5 D•15. Result
```

Let α be an ordinal. There is a subset of α of order-type β iff β is cofinal in α .

Proof .:.

Obviously if β is cofinal in α , then there is a subset of α of order-type β : im f where $f : \beta \to \alpha$ is increasing. So suppose $X \subseteq \alpha$ has order-type β . Thus there is a function $f : \beta \to X$ which is an isomorphism and thus order preserving, and increasing in particular. It follows that f witnesses that β is cofinal in α . \dashv

Let's investigate what kinds of ordinals can be cofinalities. Note that being unbounded in an ordinal isn't unique: for α an ordinal, obviously both $cof(\alpha)$ and α itself have unbounded sequences in α . For a less trivial example, $\omega + \omega + \omega$ has $\{\omega + \omega + n : n \in \omega\}$ as a subset with order-type ω , $\{\omega + \alpha + m : \alpha \leq \omega \land m \in \omega\}$ as a subset with order-type $\omega + \omega$, and both are unbounded in $\omega + \omega + \omega$.

Nevertheless, we do get a kind of uniqueness in the following sense.

5D•16. Lemma Let β be cofinal in α . Therefore $cof(\beta) = cof(\alpha)$

Proof .:.

Enumerate $X = \{x_{\xi} : \xi < \beta\}$, and let $Y = \{y_{\xi} : \xi < \operatorname{cof}(\alpha)\}$ be cofinal by definition of $\operatorname{cof}(\alpha)$. For each $y \in Y \subseteq \alpha$, as X is cofinal in α , there is some $x \in X$ with y < x. So for $y_{\xi} \in Y$, let $x'_{\xi} \in X$ be the least element of X such that $y_{\xi} < x'_{\xi}$. Hence $\{x'_{\xi} : \xi < \operatorname{cof}(\alpha)\}$ is a subset of X that is cofinal with order-type $\operatorname{cof}(\alpha)$. Since $\langle X, \langle \rangle \cong \langle \beta, \langle \rangle$, taking the relevant transformation of the x'_{ξ} 's yields then that $\operatorname{cof}(\beta) \le \operatorname{cof}(\alpha)$.

But any cofinal subset of β of order-type $cof(\beta)$ yields a cofinal subset of X of order-type $cof(\beta)$, and thus a cofinal subset of α of order-type $cof(\beta)$. So by minimality of α , $cof(\alpha) \le cof(\beta)$. Therefore $cof(\beta) = cof(\alpha)$. \dashv

As a result, cofinalities are their own cofinality.

- **5D**•17. Corollary Let α be an ordinal. Therefore $cof(cof(\alpha)) = cof(\alpha)$.

More than just this, it turns out that they will be cardinals.

- **5D**•18. Theorem - Let α be an ordinal. Therefore $cof(\alpha)$ is a cardinal.

Proof .:.

Let $\kappa = |\operatorname{cof}(\alpha)|$ with $b: \kappa \to \operatorname{cof}(\alpha)$ a bijection. For each $\xi < \kappa$, define $f(\xi)$ to be the least element of $\operatorname{cof}(\alpha)$ larger than $\max(\sup_{\gamma < \xi} b(\gamma), \sup_{\gamma < \xi} f(\gamma))$.

This is well defined, since the max will always be less than $cof(\alpha)$. To see this, otherwise, If either supremum (take f for definiteness) has $\sup_{\gamma < \xi} f(\gamma) = cof(\alpha)$, then $\{f(\gamma) : \gamma < \xi\}$ is a cofinal subset of $cof(\alpha)$ with order-type ξ so that $cof(\alpha) = cof(\xi) \le \xi < \kappa \le cof(\alpha)$, which is a contradiction. Therefore $f(\xi)$ is always defined for $\xi < \kappa$.

By definition, f is increasing. Moreover, im f is cofinal in $cof(\alpha)$, since b is a bijection: each $\zeta < cof(\alpha)$ has $b(\gamma) = \zeta$ for some $\gamma < \kappa$ so that $f(\gamma + 1) > b(\gamma) = \zeta$. Because im f has order-type κ , $cof(\alpha) = cof(\kappa) \le cof(\alpha)$

 \dashv

 $\kappa \leq \operatorname{cof}(\alpha)$. Hence $\kappa = \operatorname{cof}(\alpha)$ is a cardinal.

Hence being a cofinality is a property of cardinals. We also introduce some notation.

- **5D**•**19**. **Definition** Let κ be a cardinal. κ is *regular* iff $cof(\kappa) = \kappa$. κ is *singular* iff $cof(\kappa) < \kappa$. For κ a cardinal, κ^+ is the least cardinal greater than κ .

Note that regular cardinals appear all over the place, as do singular cardinals. In particular, all successor cardinals are regular.

- **5D**•20. Result Let κ be a cardinal. Therefore κ^+ is regular.

Proof .:.

Let $X \subseteq \kappa^+$ be cofinal with order-type $\alpha = cof(\kappa^+) < \kappa^+$. By Theorem 5 D • 18, $\alpha \le \kappa$. Note that X actually forms a partition of κ^+ by looking at the spaces between elements of X. For now, write

 $[\beta, \alpha) = \{ \xi \le \alpha : \beta \le \xi < \alpha \} = \alpha \setminus \beta text.$

Define $\alpha \approx \beta$ iff $X \cap [\beta, \alpha) = \emptyset$ and $X \cap [\alpha, \beta) = \emptyset$, meaning $\alpha \approx \beta$ iff (for $\alpha < \beta$) $\alpha \notin X$ and there are no elements of X strictly between α and β . Note that this is an equivalence relation: it's clearly symmetric and reflexive. \approx is transitive since if $\alpha \approx \beta \approx \gamma$ with $\alpha < \gamma$, then one of the following holds:

- $\beta < \alpha < \gamma$, in which case $[\alpha, \gamma) \cap X \subseteq [\beta, \gamma) \cap X = \emptyset$;
- $\alpha < \gamma < \beta$, in which case $[\alpha, \gamma) \cap X \subseteq [\alpha, \beta) \cap X = \emptyset$;
- $\alpha < \beta < \gamma$, in which case $[\alpha, \gamma) = [\alpha, \beta) \cup [\alpha, \gamma)$ so that the intersection with X is $\emptyset \cup \emptyset = \emptyset$.

Note that each equivalence class of \approx has size at most κ , since a class C is bounded by an element of X as it's cofinal: $C \subseteq \sup C + 1 < \sup X = \kappa^+$.

Since the number of equivalence classes is $|\kappa_{\alpha}^+| = |X| = |\alpha| = \alpha$, it follows that as the partition covers κ^+ ,

$$\kappa^+ \leq \sum_{C \in \kappa^+_{/\infty}} |C| \leq \sum_{C \in \kappa^+_{/\infty}} \kappa = |\kappa^+_{/\infty}| \cdot \kappa = \alpha \cdot \kappa = \max(\alpha, \kappa) < \kappa^+,$$

a contradiction. Thus $\alpha = cof(\kappa^+) \ge \kappa^+$, and so we have equality.

 \dashv

Where does all of this talk of regularity get us? Recall that we started this rabbit hole with a question: for which λ is $2^{\lambda} > \kappa$? It turns out that the answer to this question is unknowable in the sense that different models of set theory will give different answers. But, we do know at least the following, which is often also referred to as "König's theorem".

— 5D•21. Theorem (König's Cofinality Theorem) –

Let κ be a cardinal. Therefore $\kappa < \kappa^{\operatorname{cof}(\kappa)}$. Moreover, $\kappa < \operatorname{cof}(2^{\kappa})$.

Proof .:.

Let $X = \{x_{\alpha} : \alpha < cof(\kappa)\}$ be an increasing enumeration of a cofinal subset of κ . By König's Theorem (5 D • 11) and Corollary 5 D • 10, we get that

$$\kappa = \sup_{\alpha < \operatorname{cof}(\kappa)} x_{\alpha} = \sum_{\alpha < \operatorname{cof}(\kappa)} x_{\alpha} < \prod_{\alpha < \operatorname{cof}(\kappa)} \kappa = \kappa^{\operatorname{cof}(\kappa)}$$

Moreover, if we instead choose a κ -length increasing enumeration $Y = \{y_{\alpha} : \alpha < \kappa\} \subseteq 2^{\kappa}$, we get that

$$\sup_{\alpha < \kappa} y_{\alpha} \leq \sum_{\alpha < \kappa} y_{\alpha} < \prod_{\alpha < \kappa} 2^{\kappa} = (2^{\kappa})^{\kappa} = 2^{\kappa \cdot \kappa} = 2^{\kappa}.$$

Hence *Y* isn't cofinal in 2^{κ} , and therefore $cof(2^{\kappa}) > \kappa$.

The concept of cofinality also is the source of many other results about regular cardinals, especially successor cardinals.

 \neg

- **5D**•**22**. Lemma Let α be an ordinal and $X \subseteq \alpha$. If $|X| < cof(\alpha)$ then sup $X < \alpha$.

Proof .:.

If sup $X = \alpha$ then for β the order-type of X, noting that then $\beta \leq |X|$, we have by Lemma 5D•16 that $cof(\alpha) = cof(\beta) \leq |X| < cof(\alpha)$, a contradiction.

– 5 D•23. Corollary –

Let κ be a regular cardinal. Suppose $2^{<\kappa} = \kappa$. Therefore $\kappa^{<\kappa} = \kappa$.

Proof .:.

 $\kappa^{<\kappa}$ is the cardinality of $\bigcup_{\beta \in \kappa} {}^{\beta}\kappa$. Because functions from β to κ are bounded in κ by Lemma 5 D • 22, each such function is a function from β to some $\alpha < \kappa$. In particular, $\bigcup_{\beta \in \kappa} {}^{\beta}\kappa = \bigcup_{\alpha,\beta \in \kappa} {}^{\beta}\alpha$. Thus

$$\kappa^{<\kappa} \le \kappa \cdot \sup_{\alpha,\beta} |\alpha|^{|\beta|} \le \kappa \cdot \sup_{\alpha,\beta \in \kappa} 2^{|\alpha| \cdot |\beta|} = \kappa \cdot 2^{<\kappa} = \kappa \dashv$$

The hypothesis that $2^{<\kappa} = \kappa$ is an odd one, and seems very strong. This is really a statement similar to the continuum hypothesis, a very useful hypothesis regardless of its truth value.

§5E. The continuum hypothesis

We know that 2^{\aleph_0} is some cardinal, and thus is \aleph_{α} for some α . Cantor's Theorem (5 B • 13) tells us that $2^{\aleph_0} \ge \aleph_1$. König's Cofinality Theorem (5 D • 21) tells us that $cof(2^{\aleph_0}) \ge \aleph_1$ as well. But this is really all we can know.

- 5E•1. Definition CH is the statement that $2^{\aleph_0} = \aleph_1$.

Without choice, it's not clear that $\mathcal{P}(\omega)$ has an ordinal cardinality. So there are a number of formulations of CH that are equivalent under choice, and we must be careful which we choose if we are in a choiceless context.

Now we introduce a term that is so essential to much of mathematics, it's a wonder we have gotten so far without its introduction.

- **5E**•2. Definition Let X be a set. X is *countable* iff $|X| \le \aleph_0$.

- **5E**•**3**. Result CH is equivalent to the statement CH': for every $X \subseteq \mathcal{P}(\omega)$, either X is countable, or $X =_{\text{size}} \mathcal{P}(\omega)$.

Proof .:.

If CH is true, every $X \subseteq \mathcal{P}(\omega)$ has $|X| \leq \aleph_1$ and therefore $|X| < \aleph_1$ or $|X| \leq \aleph_0$. If CH fails, then $2^{\aleph_0} > \aleph_1$. So the bijection $b : \mathcal{P}(\omega) \to 2^{\aleph_0}$ yields a preimage b^{-1} " \aleph_1 of size \aleph_1 that is a subset of $\mathcal{P}(\omega)$. Hence CH' fails. \dashv

CH' is in essence an equivalent formulation of CH, but it is often more appropriate of a formulation, because we can ask if it holds in restricted contexts. Really, CH is a statement about well-orders while CH' is a statement more about subsets of $\mathcal{P}(\omega)$. As such, we can ask which families $X \subseteq \mathcal{P}(\mathcal{P}(\omega))$ have an analogous version of CH' hold of them. We will see later that CH' holds of closed subsets of \mathbb{R} , for instance: every closed subset is either countable or of size 2^{\aleph_0} .

First, we note that \mathbb{R} has size 2^{\aleph_0} so that the analogous version of CH' makes sense for subsets of $\mathcal{O}(\mathbb{R})$ rather than just $\mathcal{O}(\mathcal{O}(\omega))$. To be formal, this requires a specific construction of the real numbers, which is not done here. Instead, we rely on a more informal knowledge of real numbers as decimal expansions.

 $-5E \cdot 4. \text{ Theorem } |\mathbb{R}| = 2^{\aleph_0}.$

Proof .:.

It suffices to show that $|\mathbb{R}| = \aleph_0^{\aleph_0}$ since this is equal to 2^{\aleph_0} . Each real $r \in \mathbb{R}$ can be identified with a decimal expansion: $r = r_0.r_1r_2r_3\cdots$, meaning an $\omega + 1$ -length sequence in ω , where $r_0 \in \omega$ and $r_n \in 10$ for n > 0. The number of such sequences is $\aleph_0 \cdot 10^{\aleph_0}$, and so there are that many real numbers. But

$$2^{\aleph_0} \leq 10^{\aleph_0} \leq \aleph_0^{\aleph_0} \leq (2^{\aleph_0})^{\aleph_0} = 2^{\aleph_0},$$

and thus $|\mathbb{R}| = \aleph_0 \cdot 2^{\aleph_0} = 2^{\aleph_0}$.

Η

Really this just says that $\mathbb{R} =_{\text{size}} \mathcal{P}(\omega)$ which is clearly 2^{\aleph_0} by the above argument. In a choiceless context, we still get that $\mathbb{R} =_{\text{size}} \mathcal{P}(\omega) =_{\text{size}} \omega^2$, but it's not clear that this has an ordinal cardinality: that it can be well-ordered.

One has to have a little care about the decimal expansion in the proof of Theorem $5 \ge 4$ to ensure that it is unique, for example, $1.000 \cdots = 0.999 \cdots^{xx}$ But this can be done just by specifying that each decimal expansion should end in an infinite sequence of 0s if it has one that ends in 9s.

- 5E.5. Definition GCH is the statement that for all cardinals κ , $2^{\kappa} = \kappa^+$.

A much stronger statement than CH, GCH imposes a kind of regularity property on cardinal exponentiation.

- 5E•6. Theorem -

GCH implies that for any cardinals $\lambda > 1, \kappa > 0$ with at least one infinite,

• $\lambda \leq \kappa$ implies $\lambda^{\kappa} = \kappa^+$;

- $\operatorname{cof}(\lambda) \leq \kappa < \lambda$ implies $\lambda^{\kappa} = \lambda^+$;
- $\kappa < \operatorname{cof}(\lambda)$ implies $\lambda^{\kappa} = \lambda$

Proof .:.

- We have $\kappa^+ \leq \lambda^{\kappa}$ by Cantor's Theorem (5 B 13). Simple combinatorics yields $\lambda^{\kappa} \leq \kappa^{\kappa} \leq (2^{\kappa})^{\kappa} = 2^{\kappa} = \kappa^+$.
- We have $\lambda < \lambda^{\kappa}$ by König's Cofinality Theorem (5 D 21) and thus $\lambda^{+} \leq \lambda^{\kappa} \leq \lambda^{\lambda} = 2^{\lambda} = \lambda^{+}$.
- Proceed by induction on λ . The first place we can have $\kappa < \operatorname{cof}(\lambda)$ with at least one of the two infinite is for $\lambda = \aleph_0$, which is clear: $\lambda^n = \lambda$ for any $n < \omega$. For $\lambda > \aleph_0$, since $\kappa < \operatorname{cof}(\lambda)$, any function from κ to λ is bounded, and therefore $\lambda^{\kappa} = \sup_{\alpha < \lambda} |^{\kappa} \alpha|$. By the previous two results and induction, each $|^{\kappa} \alpha| \le \max(\alpha, \kappa)^+ \le \lambda$ and therefore $\lambda^{\kappa} \le \lambda \le \lambda^{\kappa}$.

This can be summed up with the following figure where λ and $cof(\lambda)$ are fixed and the exponentiation λ^{κ} is calculated for each interval κ is in.



5 E • 7. Figure: Calculating λ^{κ} under GCH

In general, there's very little that can be said about cardinal exponentiation, so GCH seems to dramatically simplify the situation. We will see later that this combinatorial property is very *useful* if perhaps unlikely.

^{xx}To see this, note that $r_0.r_1r_2\cdots$ is formally just $\sum_{n\in\omega}r_n\cdot 10^{-n}$ and 0.99999 \cdots is then equal to

$$\lim_{N \to \omega} \sum_{n=1}^{N} \frac{9}{10^n} = \lim_{N \to \omega} \frac{10^N - 1}{10^N} = \lim_{N \to \omega} 1 - \frac{1}{10^N} = 1.$$

With this section, we have introduced all of the axioms of what is commonly referred to as set theory.^{xxi} The whole collection of axioms (as well as their actual first-order formulas) are written at the beginning of the document.

^{xxi}There are two notions of "set theory": one is just "set theory" in the sense of "the axioms of sets"; and the other is the field of study with the same name. Often I will use "set theory" as a more informal way of writing ZFC or some sufficiently large fragment of it.

Section 6. Another Look at Model Theory

Some have described the field of set theory as being more about the model theory of set theory. Regardless of opinion about this, it does note of a relationship between the two. With the ideas of cardinality at our disposal, we may investigate further some properties of first-order logic. Then we will look more precisely at how these theorems interact with ideas surrounding set theory.

§6A. Further into first-order logic and model theory

The first result we will consider is the idea of a model *generated* by a set, and formulas. There are two or three versions of this theorem. The first two versions are certainly useful for logic, and have the most applications outside of logic, especially algebra, in detailing what is first-order expressible. The third version is the most useful for our purposes, and implies the other two. First we introduce a definition.

— 6A•1. Definition ——

Let **A**, and **B** be $FOL(\sigma)$ -models.

A is a submodel of B, written $A \subseteq B$, iff the interpretations of A are the same in B, but restricted to being functions and relations over A.

A is an elementary submodel of B, written $A \leq B$ or $A \leq_{\sigma} B$ if the signature is unclear, iff $A \subseteq B$, and for all FOLp(σ)-formulas with parameters in $A \cap B = A$, we have $A \models \varphi$ iff $B \models \varphi$.

It should be clear that being an elementary submodel implies being a submodel just by looking at the atomic $FOLp(\sigma)$ -formulas. But being a submodel does not entail being elementary. For example, the order of the real numbers on the unit interval $\langle (0, 1), \leq \rangle$ is the same as for the closed unit interval $\langle [0, 1], \leq \rangle$ so that they are submodels: $\langle (0, 1), \leq \rangle \subseteq \langle [0, 1], \leq \rangle$. But $\langle [0, 1], < \rangle \models "\exists x \forall y (y \leq x)"$ while $\langle (0, 1), \leq \rangle \not\models "\exists x \forall y (y \leq x)"$: $\langle [0, 1], \leq \rangle$ has a maximal element whereas $\langle (0, 1), \leq \rangle$ does not. In essence, being an elementary submodel is the strongest amount of agreement two models can have on first-order formulas. So note the following properties of elementary submodels: for all $FOL(\sigma)$ -models **A**, **B**, and **C**;

• $A \preccurlyeq A$.

- $A \leq B \leq A$ iff A = B (since $A \subseteq B \subseteq A$, and they interpret the signature the same way).
- $A \preccurlyeq B \preccurlyeq C$ implies $A \preccurlyeq C$.
- $A \preccurlyeq C$ and $B \preccurlyeq C$ implies $A \preccurlyeq B \leftrightarrow A \subseteq B$.

The next theorem, one of the versions of the Löwenheim–Skolem theorem, then tells us that we can *generate* elementary submodels using arbitrary subsets of the original model we start with.

– 6 A • 2. Theorem (Taking a Skolem Hull) –

Let A be an infinite $FOL(\sigma)$ -model, and $X \subseteq A$. Therefore there is a model $Hull^A(X)$ called the *skolem hull* of X, such that

1. $X \subset \operatorname{Hull}^{\mathsf{A}}(X) \subset A;$

- 2. $|\operatorname{Hull}^{\mathsf{A}}(X)| \leq |X| \cdot |\sigma| \cdot \aleph_0;$
- 3. Hull^A(X) \preccurlyeq A.

Most of the time, $Hull^{A}(X)$ will not be unique. So despite the calling it *the* skolem hull, we really are interested in *any* model with the properties (1)–(3).^{xxii}

^{xxii}And sometimes we also care about some nice properties given from the construction of such models. In particular, the elements of the hull are generated from F0Lp-formulas, and so we can get representations of these elements by F0L-formulas and parameters. This is sometimes used in the literature, but will not be used explicitly here.

 \neg

To prove this result, we essentially do a careful proof of Completeness $(1 D \cdot 1)$, building up a model from X by closing under the functions of σ and whatever witnesses existential statements need from A. So the following combinatorial result will be useful in showing that we do not add too many elements in building up the skolem hull.

6A•3. Lemma Let X be a set. Let f be a function with $X \subseteq \text{dom } f$. Therefore the *closure* of X under f—meaning the \subseteq -least set Y with $X \subseteq Y$ and $f''Y \subseteq Y$ —has size at most $|X| \cdot \aleph_0$.

Proof .:.

Write $X_0 = X$, and define $X_{n+1} = X_n \cup f'' X_n$. Let $Y = \bigcup_{n \in \omega} X_n$. Note that for each $x \in Y$, $f(x) \in X_{n+1}$ where $x \in X_n$. Hence $f(x) \in Y$. Thus Y is closed under f. Moreover, for each $n \in \omega$, $|X_{n+1}| \le |X_n| + |X_n| =$ $2|X_n|$ because $|f''X_n| \leq |X_n|$. Therefore, inductively, $|X_n| \leq \aleph_0 \cdot |X|$ for each $n \in \omega$. Therefore the union Y has $|X| \leq |Y| \leq \aleph_0 \cdot \aleph_0 \cdot |X| = \aleph_0 \cdot |X|$. Regardless of whether Y is the \subseteq -least set containing X, any $Z \subseteq Y$ which is the real closure of X has $|Z| < |X| \cdot \aleph_0$.

As a result, we can close under entire sets of functions as well, and still we can bound the size of the resulting set.

- 6A•4. Corollary

Let X be a set. Let σ be a set of functions with $X \subseteq \text{dom } f$ for each $f \in \sigma$. Therefore the *closure* of X under σ —meaning the \subseteq -least set Y with $X \subseteq Y$ and $f''Y \subseteq Y$ for each $f \in \sigma$ —has size at most $|X| \cdot |\sigma| \cdot \aleph_0$.

Proof .:.

As before, write $X_0 = X$, and define $X_{n+1} = X_n \cup \bigcup_{f \in \sigma}$ (the closure of X_n under f). Thus by Lemma 6 A • 3, $|X_{n+1}| \leq |X_n| + |X_n| \cdot |\sigma| \cdot \aleph_0 = |X_n| \cdot |\sigma| \cdot \aleph_0$ For each $n \in \omega$. So inductively, it follows that $|X_n| \le |X| \cdot |\sigma|^n \cdot \aleph_0 = |X| \cdot |\sigma| \cdot \aleph_0$. Taking the union $Y = \bigcup_{n \in \omega} X_n$ yields that Y is closed under each $f \in \sigma$ as in Lemma 6 A • 3, and moreover, $|Y| \le |X| \cdot |\sigma| \cdot \aleph_0^2$.

Therefore, when we build up the skolem hull, we aren't adding too many elements to X. Note that in the following proof of Taking a Skolem Hull (6 A • 2), indirectly confirm that we have an elementary submodel by the idea of skolem functions: functions which map existential statements to elements that witness them. This allows us to see that the agreement between A and $Hull^{A}(X)$ includes existential statements. The propositional connectives are practically free, and so by induction on formulas, this implies the hull is an elementary submodel.

- Proof of Taking a Skolem Hull $(6A \cdot 2)$...

For each existential FOL(σ)-formula $\psi(\vec{x})$ being $\exists v \varphi(v, \vec{x})$, add the function symbol f_{ψ} (with arity being the length of \vec{x}) to the signature. Thus we now consider the signature

 $\sigma' = \sigma \cup \{ f_{\psi} : \psi \text{ is an existential FOL}(\sigma) \text{-formula} \}.$

We interpret the functions f_{ψ} in the model A by the axiom of choice: for $\psi(\vec{x})$ being " $\exists v \, \varphi(v, \vec{x})$ ", if A \models " $\exists v \varphi(v, \vec{x})$ ", choose $f_{\psi}^{\mathsf{A}}(\vec{x}) \in A$ such that $\mathsf{A} \models "\varphi(f_{\psi}^{\mathsf{A}}(\vec{x}), \vec{x})$ ". Obviously, if $\mathsf{A} \not\models "\exists v \varphi(v, \vec{x})$ ", then we can set $f_{\psi}^{A}(\vec{x})$ to be any particular, fixed element of A that we want (this is only done to ensure that f_{ψ}^{A} is indeed a function defined over all of A). Hence we can consider the $FOL(\sigma')$ -model A' with these new interpretations, noting that we have only added interpretations: we still have $X \subseteq A' = A$, for instance.

With this, by Corollary 6 A • 4, we can consider the closure of X under the functions of σ' , yielding Hull^A(X). This clearly has $X \subseteq \text{Hull}^{A}(X) \subseteq A$, meaning (1) holds. Moreover, by Corollary $6 \land 4$, $|\text{Hull}^{A}(X)| \leq |X| \cdot |\sigma| \cdot |\aleph_{0}|$, meaning (2) holds.

Now we take the model Hull^A(X) to have the same function and relation interpretations as A, but restricted to Hull^A(X). To show (3), suppose $w_0, \dots, w_n \in \text{Hull}^A(X)$. We proceed by induction on the FOL(σ)-formula $\varphi(\vec{x})$ to show that $\operatorname{Hull}^{\mathsf{A}}(X) \models "\varphi(\vec{w})"$ iff $\mathsf{A} \models "\varphi(\vec{w})"$.

- For $\varphi(\vec{x})$ atomic, the result is immediate by definition.
- For $\varphi(\vec{x})$ being $\neg \psi(\vec{x})$, the inductive hypothesis clearly give the result.
- For $\varphi(\vec{x})$ being $\psi(\vec{x}) \wedge \chi(\vec{x})$, $\mathbf{A} \models "\psi(\vec{w}) \wedge \chi(\vec{w})$ " iff it models each individually. By the inductive hypothesis, this is equivalent to $\mathsf{Hull}^{\mathsf{A}}(X)$ modeling each individually, meaning $\mathsf{Hull}^{\mathsf{A}}(X) \models "\varphi(\vec{w})$ ".
- For φ(x) being ∃v ψ(v, x), A ⊨ "∃v ψ(v, w)" iff A ⊨ "ψ(f_φ^A(w), w)". Since Hull^A(X) is closed under these skolem functions, by the inductive hypothesis, this is equivalent to Hull^A(X) ⊨ "ψ(f_φ^A(w), w)", iff Hull^A(X) ⊨ "∃v ψ(v, w)".

Hence by induction on FOL(σ)-formulas, it follows that Hull^A(X) \leq A, and thus (1)–(3) hold.

Н

This method of taking skolem hulls, which are effectively the models first-order generated by the set X we look at, is incredibly powerful, and yields the next result, which has some counter intuitive consequences.

- 6A • 5. Theorem (Löwenheim-Skolem) Let T be a FOL(σ)-theory with an infinite model. Therefore, for every cardinal $\kappa \ge |\sigma| \cdot \aleph_0$, there is a model $\mathbf{M} \models T$ with $|M| = \kappa$.

Proof .:.

Let $\mathbf{A} \models T$ be an infinite model. For the downward version, $\operatorname{suppose} |\sigma| \cdot \aleph_0 \le \kappa \le |A|$, and let $X \subseteq A$ be a subset of size κ . Therefore, by Taking a Skolem Hull (6 A • 2), there is a hull $\operatorname{Hull}^{\mathsf{A}}(X) \preccurlyeq A$ with $X \subseteq \operatorname{Hull}^{\mathsf{A}}(X)$ meaning $\kappa \le |\operatorname{Hull}^{\mathsf{A}}(X)|$. Moreover, $|\operatorname{Hull}^{\mathsf{A}}(X)| \le \kappa \cdot |\sigma| \cdot \aleph_0 = \kappa$ so that the hull has size κ . By elementarity, each sentence $\varphi \in T$ has $\mathbf{A} \models \varphi$ iff $\operatorname{Hull}^{\mathsf{A}}(X) \models \varphi$. Therefore $\operatorname{Hull}^{\mathsf{A}}(X) \models T$, and so the hull works.

For the upward version, we use compactness: let $\kappa > |A|$. Consider the expanded signature $\sigma' = \sigma \cup \kappa$, where each ordinal $< \kappa$ is a constant symbol (devoid of its meaning as an ordinal). Now consider the theory $T' = T \cup \{``\alpha \neq \beta``: \alpha \neq \beta < \kappa\}$. Any FOL (σ') -model $\mathbf{B} \models T'$ has $\mathbf{B} \models T$ and must have at least κ many elements: the interpretations of the ordinal symbols which are all different. T' has a model by Compactness (1 D • 2): every finite subset $\Delta \subseteq T$ is modeled by an expansion \mathbf{A}' of \mathbf{A} by interpreting the *finitely many* ordinalsymbols in Δ as just different elements of A. Since A has infinitely many elements, we can always do this. Therefore $\mathbf{A}' \models \Delta$, and so T' has a model, which is then of size $\lambda \geq \kappa \geq |\sigma| \cdot \aleph_0$. Using the downward statement, it follows that we have a model $\mathbf{M} \models T'$ of size exactly κ .

Some immediate consequences of this are that if ZFC is consistent, then there is a *countable* model in addition to a model of size \aleph_1 , and models of every cardinality. One might be very confused about this, since supposedly ω and thus $\mathcal{P}(\omega) >_{\text{size}} \aleph_0$ should be in the model, but this isn't necessarily true: the model will contain fewer subsets than in the real world, as not every subset of ω can be described by formulas, and Löwenheim–Skolem (6 A • 5) really only deals with what is minimally required of the formulas of FOL. In essence, the model won't realize its $\mathcal{P}(\omega)$ is small, because it doesn't conain the necessary bijection between its interpretation of $\mathcal{P}(\omega)$ and \aleph_0 .

Now often in model theory, one deals with chains of elementary submodels. So it's nice to have the following theorem.

- 6A · 6. Theorem (Tarski-Vaught Theorem) Let A_{α} be a FOL(σ)-model for each $\alpha < \gamma \in$ Ord with γ a limit ordinal. Suppose $A_{\alpha} \preccurlyeq A_{\beta}$ for all $\alpha < \beta < \gamma$. Therefore there is a model $\bigcup_{\alpha < \gamma} A_{\alpha}$ where $A_{\alpha} \preccurlyeq \bigcup_{\beta < \gamma} A_{\beta}$ for all $\alpha < \gamma$.

Proof .:.

The "direct limit" $\bigcup_{\alpha \in \gamma} A_{\alpha}$ is just given by the union of the corresponding models: the universe the union of the universes, the relations are the unions of the relations, and the functions are the unions of the functions. The constants are necessarily the constants as interpreted by A_0 .

Using this, any FOLp(σ)-formula φ with parameters in A_{α} that is atomic clearly has $A_{\alpha} \models \varphi$ iff $\bigcup_{\beta < \gamma} A_{\beta} \models \varphi$. Similarly, the propositional connectives follow easily. For the existential case, it should be clear that $A_{\alpha} \models \varphi$. " $\exists x \ \varphi(x)$ " implies $\mathbf{A}_{\alpha} \models$ " $\varphi(a)$ " for some $a \in A_{\alpha}$ and thus by the inductive hypothesis, $\bigcup_{\beta < \gamma} \mathbf{A}_{\beta} \models$ " $\varphi(a)$ " and thus $\bigcup_{\beta < \gamma} \mathbf{A}_{\beta} \models$ " $\exists x \ \varphi(x)$ ". For the reverse direction, since $\bigcup_{\beta < \gamma} \mathbf{A}_{\beta} \models$ " $\varphi(a)$ " for some $a \in \bigcup_{\beta < \gamma} A_{\beta}$, we have that $a \in A_{\beta}$ for some $\beta < \gamma$. Therefore by the inductive hypothesis, $\mathbf{A}_{\beta} \models$ " $\varphi(a)$ " and thus $\mathbf{A}_{\beta} \models$ " $\exists x \ \varphi(x)$ ". But by elementarity, it follows that $\mathbf{A}_{\alpha} \models$ " $\exists x \ \varphi(x)$ ".

Often we don't want to consider an elementary submodel directly, but instead a model which *maps* to an elementary submodel by way of an embedding.

- $6A \cdot 7$. Definition Let A and B be FOL(σ)-models. For $f : A \to B$ an injective map, the structure f "A is the structure with universe f"A and with interpretations of σ given by f applied to the interpretations in A. $f : A \to B$ is an *embedding* (A is *embedded* in B) iff f "A \subseteq B. $f : A \to B$ is an *elementary embedding* (A is *elementarily embedded* in B) iff f "A \leq B.

An alternative characterization of being an elementary embedding would be that for every FOL(σ)-formula $\varphi(\vec{x})$ and \vec{a} members of A, $\mathbf{A} \models "\varphi(\vec{a})$ " iff $\mathbf{B} \models "\varphi(f(\vec{a}))$ ". This characterization is arguably a better way of thinking about it. Similarly, f is an embedding iff $\mathbf{A} \models "R(\vec{a})$ " iff $\mathbf{B} \models "R(f(\vec{a}))$ " for every relation R and \vec{a} in A, and similarly for functions: $\mathbf{A} \models "F(\vec{a}) = a_0$ " iff $\mathbf{B} \models "F(f(\vec{a})) = f(a_0)$ ".

But with this added concept, we also can generalize the union model of Tarski–Vaught Theorem ($6 A \cdot 6$) to the direct limit proper. The direct limit is essentially a least upper bound with respect to embedability. The figure below represents the general idea: any **M** with embeddings following the diagram is "larger" than the direct limit in that the direct limit embeds in **M**.



6A•8. Figure: The direct limit embeddings

And here the diagram commutes.

— 6A•9. Definition —

Let \mathcal{A} be a set of FOL(σ)-models and \mathcal{F} be a set of (elementary) embeddings between models of \mathcal{A} . $\langle \mathcal{A}, \mathcal{F} \rangle$ is called a *directed system of (elementary) embeddings* iff

• for each $A, B \in A$, there is at most one $f : A \to B$ in \mathcal{F} , denoted $f_{A,B}$ with $f_{A,A} = id \upharpoonright A$;

• for each $A, B \in \mathcal{A}$ there is some $C \in \mathcal{A}$ with $f_{A,C}, f_{B,C} \in \mathcal{F}$; and

• if $f_{A,B}$, $f_{B,C} \in \mathcal{F}$, then there is an (elementary) embedding $f_{A,C} \in \mathcal{F}$ with $f_{B,C} \circ f_{A,B} = f_{A,C}$.

For $\langle \mathcal{A}, \mathcal{F} \rangle$ a directed system of embeddings, the *direct limit* is the FOL(σ)-model **dir lim** $_{\mathcal{F}}$ \mathcal{A} such that

- 1. there is an embedding $f_{A,\infty}: A \to \dim_{\mathscr{F}} A$ such that $f_{B,\infty} \circ f_{A,B} = f_{A,\infty}$ whenever $f_{A,B}$ exists; and
- 2. for every model M satisfying (1) in place of dir $\lim_{\mathcal{F}} \mathcal{A}$, there is an embedding f: dir $\lim_{\mathcal{F}} \mathcal{A} \to M$ such that $f \circ f_{A,\infty} = f_{A,M}$ for all $A \in \mathcal{A}$.

The general idea behind directed systems of embeddings is that we can continually embedd things in a "larger" model, and we can do so in a way where the embeddings work well together. The idea behind the direct limit is that there should be an upper bound to this: one where everything in \mathcal{A} embedds into it (in a way that works nicely with), and it's the "least" such model. So the direct limit should be thought of as a least upper bound on a poset^{xxiii} $\langle \mathcal{A}, R \rangle$ where

xxiii The directed set doesn't necessarily form a poset, since the existence of embeddings need not be antisymmetric, but it will at least be reflexive

A *R* **B** iff $f_{A,B} \in \mathcal{F}$, although the least upper bound isn't necessarily in \mathcal{A} .

Of course, we should confirm that every directed system of embeddings has a direct limit, which mostly just amounts to checking that a certain construction works. The general idea behind the construction is that we just take the disjoint copies of all the models, and then take the union as in Tarski–Vaught Theorem ($6 A \cdot 6$). From there, we mod out by "eventual equivalence" when transformed by elements of \mathcal{F} .

– 6 A•10. Result

Let $\langle \mathcal{A}, \mathcal{F} \rangle$ be a directed system of embeddings between FOL(σ)-models. Therefore the direct limit **dir** $\lim_{\mathcal{F}} \mathcal{A} = \mathbf{D}$ exists, and is isomorphic to the disjoint union of \mathcal{A} modulo eventual equivalence through the embeddings of \mathcal{F} .

Proof .:.

Without loss of generality, each $A \cap B = \emptyset$ for $A, B \in A$ just by replacing A with the isomorphic model A' with universe $A \times \{A\}$, tagging each element of the universe with the model it comes from. This ensures that the union $\bigcup A$ is a disjoint union. This union is defined as in Tarski–Vaught Theorem ($6A \cdot 6$): the universe is the disjoint union $\bigcup_{A \in A} A$, and the relations and functions are the disjoint unions of the corresponding relations and functions. This union model does not interpret the constant symbols of σ . To form a FOL(σ)-model **D**, consider the relation on $\bigcup_{A \in A} A$ defined by (for $x \in A$ and $y \in B$) $x \approx y$ iff there is some **C** where $f_{A,C}(x) = f_{B,C}(y)$, meaning that x and y are eventually equal in the embeddings.

— Claim 1

 \approx , eventual equivalence through the embeddings of \mathcal{F} , is an equivalence relation on $\bigcup_{A \in \mathcal{A}} A$.

Proof .:.

Clearly \approx is reflexive (take $f_{A,A}$ = id to witness this) and symmetric (by symmetry of =). \approx is transitive, since if $f_{A,B}(x) = f_{A',B}(y)$ and $f_{A',C}(y) \approx f_{A'',C}(z)$, then for some **M** with $f_{B,M}$, $f_{C,M} \in \mathcal{F}$, it follows by injectivity and the embeddings working well together that

$$f_{\mathbf{A},\mathbf{M}}(x) = f_{\mathbf{B},\mathbf{M}} \circ f_{\mathbf{A},\mathbf{B}}(x)$$

= $f_{\mathbf{B},\mathbf{M}} \circ f_{\mathbf{A}',\mathbf{B}}(y)$
= $f_{\mathbf{A}',\mathbf{M}}(y)$
= $f_{\mathbf{C},\mathbf{M}} \circ f_{\mathbf{A}',\mathbf{C}}(y)$
= $f_{\mathbf{C},\mathbf{M}} \circ f_{\mathbf{A}'',\mathbf{C}}(z) = f_{\mathbf{A}'',\mathbf{M}}(z).$

Hence \approx is transitive, and so an equivalence relation.

Now consider the model **D** with universe $(\bigcup_{A \in \mathcal{A}} A)_{\approx}$ —the equivalence classes of \approx —and corresponding relation and function interpretations as per Result 2 C • 11: $R^{D}([x_{0}]_{\approx}, \cdots, [x_{n}]_{\approx})$ iff for $x_{i} \in A_{i}$ and $\mathbf{B} \in \mathcal{A}$ such that $f_{A_{i}, \mathbf{B}} \in \mathcal{F}$ for each $i \leq n$, $R^{B}(f_{A_{i}, \mathbf{B}}(x_{0}) \cdots, f_{A_{n}, \mathbf{B}}(x_{n}))$. As the $f \in \mathcal{F}$ are embeddings and so respect R, this will be well-defined. We do the same process for the functions of σ . Note that the constant symbols work out nicely after modding out by \approx : each constant symbol c of σ is interpreted as $[c^{A}]_{\approx}$ for any $\mathbf{A} \in \mathcal{A}$. As embeddings, the constant symbols are mapped to the corresponding constant symbols, and thus eventual equivalence always holds between c^{A} and c^{B} . This completes the construction of **D**. Now we must show that **D** is the direct limit.

Firstly, note that each $\mathbf{A} \in \mathcal{A}$ has an embedding $f_{\mathbf{A},\mathbf{D}} : A \to D$ defined by $a \mapsto [a]_{/}$. This is an embedding, because for $a_0, \dots, a_n \in A$, $R^{\mathsf{D}}([a_0]_{\approx}, \dots, [a_n]_{\approx})$ by definition is equivalen to $R^{\mathsf{B}}(f_{\mathsf{A},\mathsf{B}}(a_0), \dots, f_{\mathsf{A},\mathsf{B}}(a_n))$ for some B with $f_{\mathsf{A},\mathsf{B}} \in \mathcal{A}$. In particular, for $\mathsf{B} = \mathsf{A}$, this is just $R^{\mathsf{A}}(a_0, \dots, a_n)$. The same idea applies for functions and constants to show that $f_{\mathsf{A},\mathsf{B}}$ is an embedding. Moreover, this embedding plays nicely with the $f \in \mathcal{F}$, since eventual equivalence yields $f_{\mathsf{A},\mathsf{B}}(a) \approx a$ so that $f_{\mathsf{B},\mathsf{D}} \circ f_{\mathsf{A},\mathsf{B}}(a) = [f_{\mathsf{A},\mathsf{B}}(a)]_{\approx} = [a]_{\approx} = f_{\mathsf{A},\mathsf{D}}(a)$. So (1) holds of Definition $6 \mathsf{A} \cdot 9$.

To see that **D** is the least such model—that (2) holds of Definition $6 \land 9$ —suppose **M** has the same property. Let

 \dashv

and transitive.

 $f_{\mathsf{D},\mathsf{M}}$ be defined by, for $x \in A$ and $\mathsf{A} \in \mathcal{A}$, $f_{\mathsf{D},\mathsf{M}}([x]_{\approx}) = f_{\mathsf{A},\mathsf{M}}(x)$. This is well defined since if $x \approx f_{\mathsf{B},\mathsf{A}}(x)$ then $f_{\mathsf{D},\mathsf{M}}(f_{\mathsf{B},\mathsf{A}}(x)) = f_{\mathsf{B},\mathsf{M}} \circ f_{\mathsf{A},\mathsf{B}}(x) = f_{\mathsf{A},\mathsf{M}}(x) = f_{\mathsf{D},\mathsf{M}}(x)$

for any $\mathbf{B} \in \mathcal{A}$ with $f_{\mathbf{A},\mathbf{B}} \in \mathcal{F}$. In fact, $f_{\mathbf{D},\mathbf{M}}$ will be injective since $[x]_{\approx} \neq [y]_{\approx}$ implies the transformations of x and y by $f \in \mathcal{F}$ are always different so that applying the embedding $f_{\mathbf{A},\mathbf{M}}$ where \mathbf{A} contains transformations of both x and y, the transformations are still different in \mathbf{M} . The reverse direction holds in the same way. To see that $f_{\mathbf{D},\mathbf{M}}$ respects the relations, functions, and constant symbols of σ , suppose c is a constant symbol of σ . $f_{\mathbf{D},\mathbf{M}}(c^{\mathbf{D}}) = f_{\mathbf{A},\mathbf{M}}(c^{\mathbf{A}}) = c^{\mathbf{M}}$ for any $\mathbf{A} \in \mathcal{A}$ as $f_{\mathbf{A},\mathbf{M}}$ is an embedding and $c^{\mathbf{D}} = [c^{\mathbf{A}}]_{\approx}$. For the relation R, if $R^{\mathbf{D}}([x_0]_{\approx}, \cdots, [x_n]_{\approx})$, then $R^{\mathbf{A}}$ holds of the transformations of the \vec{x} where \mathbf{A} contains all of these transformed elements. But then applying $f_{\mathbf{A},\mathbf{M}}$ yields that the relation holds of the $f_{\mathbf{A},\mathbf{M}}$ transformations of the \vec{x} so that the relation holds of the $f_{\mathbf{D},\mathbf{M}}$ transformations. The reverse direction is the same, and the argument for functions proceeds similarly. Therefore $f_{\mathbf{D},\mathbf{M}}$ is an embedding, and so \mathbf{D} is the direct limit.

Thus Tarski–Vaught Theorem (6 A • 6) can be reformulated as saying that if we have a chain of elementary embeddings, then each is elementarily embedded in the direct limit. So to generalize this, we have the following result, whose proof is precisely the same as Tarski–Vaught Theorem (6 A • 6), although translated through the elementary embeddings of \mathcal{F} instead of the elements themselves.

− 6A • 11. Corollary Let $\langle \mathcal{A}, \mathcal{F} \rangle$ be a directed system of elementary embeddings. Therefore A is elementarily embedded in dir lim_{\mathcal{F}} \mathcal{A} for each A $\in \mathcal{A}$.

The point of all of this talk about elementary embeddings will become clear in the next chapter. But it is an important idea if we want to learn about V, as the first-order truths of V are then reflected in any model it elementarily embeds into. So-called large cardinals often state the existence of elementary embeddings from V into another model, and so commonly uses the techniques of this subsection.

§6B. Logic within set theory

One might worry that, since the above ideas depend on set theory, although the meta-theoretic ZFC can prove the above results about first-order logic the formal^{xxiv} ZFC can't. Readers worried about this can put their minds at ease. But although objects in ZFC are hereditarily sets, we can still code non-set things like formulas using sets.

Rather than give a tedious account of the syntax of first-order logic, and an even more tedious account of how to formalize this, we merely give an impression on how these things are formalized in ZFC.

The *logical symbols of formal first-order logic* is the set ω , consisting of the codes for logical symbols $(\wedge)^{2} = 0, \quad (\neg)^{2} = 1, \quad (\exists)^{2} = 2, \quad ((2 = 3, -1))^{2} = 4, \quad ((2 = 3, -1))^{2} = 4$

and variables ' v_n ' = n + 6 for $n < \omega$.

For A a set, a *variable assignment* for A is a function $f : \{ v_n : n \in \omega \} \to A$, i.e. a function $f : \omega \setminus 6 \to A$. For σ a set of relations, functions, and constants, a σ -formula is a $\varphi \in (\omega \sqcup \sigma)^{<\omega}$ obeying the usual syntax rules. For σ a set of relations, functions, and constants, a σ -proof is a finite sequence of σ -formulas that obeys the usual syntax rules for proofs.

Once we have the syntax of first-order logic in ZFC, we can start to address the satisfaction relation. This is done by induction on formulas. Firstly, we have a couple definitions that allows us to more precisely see why we can do this in ZFC: the relations are well-founded.

— 6 B • 2. Definition

Suppose \leq_n is a linear order on B_n for each $n \in \omega$. Then the *length-prioritized lexicographic ordering* $\leq_{\text{lex}} \subseteq \bigcup_{N \in \omega} \prod_{n=0}^{N} B_n$ is the order defined by, for $f : n \to \bigcup_{k \in \omega} B_k$ and $g : m \to \bigcup_{k \in \omega} B_k$ where $n, m \in \omega$, $f \leq_{\text{lex}} g \leftrightarrow f = g \lor |f| < |g| \lor (|f| = |g| \land \text{the least } k \in \omega \text{ with } f(k) \neq g(k) \text{ has } f(k) <_k g(k)).$

xxiv"Formal" here means relating to formulas, i.e. "syntactic", rather than the opposite of casual.

This is really just the dictionary order on $\langle \omega B \rangle$ where each component potentially has its own ordering. This is best understood when each $\langle n \rangle$ is the same ordering on $B = B_n$. In particular, working with triplets, $\langle n \rangle = \langle n$

•
$$a \prec a'$$
; or

- a = a' and $b \prec b'$; or
- a = a' and b = b' and $c \prec c'$.

Sequences of different lengths are compared in the same way, but the longer one comes after in the order. To save space, we would also write the above as the more intelligible conditions:

- $a \prec a'$; or else
- $b \prec b'$; or else
- $c \prec c'$.

Definition $6 \text{ B} \cdot 2$ just generalizes this to larger product sequences with more relations. What's important for us is when this is a well-ordering.

− 6B•3. Lemma Suppose each $\leq_n \subseteq B_n \times B_n$ is a well-order of B_n . Therefore \leq_{lex} is a well-order of $\bigcup_{N \in \omega} \prod_{n=0}^N B_n$.

Proof .:.

It should be clear that \leq_{lex} is a linear order of $B^{<\omega}$: transitivity follows since each $<_n$ is transitive. Totality clearly holds since any two distinct sequences differ somewhere, and since each $<_n$ is total, wherever they differ is ordered. Clearly anti-symmetry holds by anti-symmetry of each $<_n$. So \leq_{lex} is clearly linear, and all that suffices is to show well-foundedness.

Let $\langle f_n : n \in \omega \rangle$ be \langle_{lex} -decreasing. Therefore $\langle \text{dom}(f_n) : n \in \omega \rangle$ is non-increasing. So without loss of generality, we can assume each $\text{dom}(f_n) < k$ for some $k \in \omega$. For each m < k consider $\langle f_n(m) : n < \omega \land m \in \text{dom}(f_n) \rangle$. If each of these is finite or eventually stabilizes, then eventually f_{n+1} is an initial segment of f_n . If this were the case, then the only way for $\langle f_n : n \in \omega \rangle$ to be \langle_{lex} -decreasing is for their lengths to be decreasing, contradicting the well-foundedness of ω . Thus for some m, $\langle f_n(m) : n < \omega \land \text{dom}(f_n) \rangle$ is infinite and doesn't stabilize. Take the least $m \in \omega$ for which this happens. Therefore, eventually, $f_{n+1}(m) <_m f_n(m)$, contradicting the well-foundedness of $<_m$.

The point of having $<_{\text{lex}}$ prioritize length is to ensure that inductive hypotheses hold for subformulas: for ψ a subformula of φ , $\psi \leq_{\text{lex}} \varphi$. Hence, we can proceed by induction on formulas.

— 6 B•4. Corollary

For any signature σ well-ordered by $<_{\sigma}$, the σ -formulas are well-ordered by $<_{\text{lex}}$.

Proof .:.

Strictly speaking, the order is on $\omega \cup \sigma$ where were merely place all elements of ω before σ . In other words, for α the order-type of $\langle \sigma, <_{\sigma} \rangle$, the order on $\omega \sqcup \sigma$ is given by $\omega + \alpha$. Hence $<_{\text{lex}}$ well-orders $(\omega \sqcup \sigma)^{<\omega}$, which contain all of the σ -formulas.

For set structures, V can define the satisfaction relation by induction on formulas (the property of being a subformula is well-founded, as formulas are certain finite sequences of an alphabet). In fact, we can define this relation *uniformly*.

— 6 B • 5. Definition

For A a set, and σ a signature, an *interpretation* of σ in A is a map ς with $\sigma = \text{dom}(\varsigma)$ where for R an *n*-placed relation, $\varsigma(R) = R^{\mathsf{A}} \subseteq A^n$, and similarly for functions and constants.

- 6B•6. Theorem

Let σ be a signature, A be a set, v a variable assignment for A, ς an interpretation of σ in A, and x a $\sigma \sqcup A$ -formula coding the real-world formula $\psi(\vec{y})$. Therefore, there is a FOL(\in)-formula "models($\sigma, A, \varsigma, v, x$)" such that

 $\langle A, \varsigma \rangle \vDash "\psi(v(\vec{y}))"$ iff $\mathbf{V} \vDash "models(\sigma, A, \varsigma, v, x)"$.

ANOTHER LOOK AT MODEL THEORY

Proof .:.

In particular, models(σ , A, ς , v, x) iff

- x is a $\sigma \sqcup A$ -formula;
- ς is a set of relations, functions, and constants over *A*;
- v is a variable assignment, a function mapping variables to elements of A;
- there is a function $f_v: (A \cup \sigma)^{<\omega} \to \{0, 1\}$ such that for any $z \in (A \cup \sigma)^{<\omega}$,
 - z is of the form " $R(\vec{y})$ " where $R \in \sigma$ is a relation symbol, and $v(\vec{y}) \in \varsigma(R) \leftrightarrow f_v(z) = 1$,
 - z is of the form " $y_0 = y_1$ " and $v(y_0) = v(y_1) \leftrightarrow f_v(z) = 1$,
 - z is of the form " $\psi \wedge \theta$ " and $f_v(z) = 1$ iff both $f_v("\psi") = 1$ and $f_v("\theta") = 1$.
 - z is of the form " $\neg \psi$ " and $f("\psi") = 1$ iff f(z) = 0, and
 - z is of the form " $\forall t \ \psi(\vec{y}, t)$ " and f(z) = 1 iff any function h_w obeying these rules for all $z' \leq_{\text{lex}} z$ where w is any variable assignment w for A with $v \setminus \{\langle a, v(a) \rangle\} \subseteq w$ —has $h_w("\psi(a, t)") = 1$; and
- $f_v(x') = 1$ where x' has every free variable y in x replaced by v(y).

The above is really only a partial proof, since it only applies to signatures with no function symbols. But this isn't an issue with the result, it just makes the defining formula even longer to have an auxiliary function interpreting terms through the variable assignment.

Doing this then allows us to confirm by the same sort of proofs before that Completeness $(1 D \cdot 1)$, Compactness $(1 D \cdot 2)$, and so forth hold. But the important thing about Theorem $6B \cdot 6$ is that ZFC has the ability to understand when something is true in a given set model. We will often use Theorem $6B \cdot 6$ without stating so, because the idea of a formula being true of a (set) structure is so widely used. Of course, we may not have access to classes since they aren't objects in the universe.

§ 6 C. Common applications to set theory

For now, our main application will be with respect to Taking a Skolem Hull ($6 \land 2$) and elementarity. The great thing about taking skolem hulls of transitive sets is that we end up with well-founded sets, and thus can collapse them.

Let $\mathbf{A} = \langle A, R \rangle$ be well-founded. Therefore any $\mathbf{B} = \langle B, R' \rangle$ embedded in \mathbf{A} is also well-founded.

Proof .:.

Let $f : B \to A$ be an embedding and let $X \subseteq B$ be arbitrary. Since **A** is well-founded, $f''X \subseteq A$ has an R-minimal element $a \in f''X$. Thus for every $y \in f''X$, $\neg y R a$. As an embedding, $\neg(f^{-1}(y) R' f^{-1}(a))$ for each $y \in f''X$, meaning $\neg(x R' f^{-1}(a))$ for each $x \in X$. Therefore $f^{-1}(a)$ is R'-minimal. Thus **B** is also well-founded.

- 6C•2. Corollary -

Let T be a transitive set and $X \subseteq T$. Therefore $\operatorname{Hull}^{\langle T, \epsilon \rangle}(X)$ is well-founded, and is isomorphic to the transitive collapse $\operatorname{cHull}^{\langle T, \epsilon \rangle}(X)$, which is then elementarily embedded in $\langle T, \epsilon \rangle$. Moreover, if X is transitive, X is left uncollapsed: the collapsing map $\pi : \operatorname{Hull}^{\langle T, \epsilon \rangle}(X) \to \operatorname{cHull}^{\langle T, \epsilon \rangle}(X)$ has $\pi \upharpoonright X = \operatorname{id} \upharpoonright X$.

Proof .:.

Write T' for cHull^{(T, \in)}(X). By Taking a Skolem Hull (6 A • 2), Hull^{(T, \in)} $(X) \leq \langle T, \in \rangle$ so that the hull is wellfounded. By elementarity, the hull satisfies the axiom of extensionality. By The Mostowski Collapse (4 • 1), the hull is isomorphic to the transitive $\langle T', \in \rangle$ by the map inductively defined by $\pi(x) = \{\pi(a) : Hull^{<math>(T, \in)$} $(X) \models$

 \neg
" $a \in x$ "}. Note that as a substructure of V, for $a, x \in H$, $\operatorname{Hull}^{(T, \in)}(X) \models$ " $a \in x$ " iff $a \in x$. Moreover, $\operatorname{Hull}^{(T, \in)}(X) \models$ " $a \in x$ " implies $a \in H$ just by virtue of the semantics. Therefore $\operatorname{Hull}^{(T, \in)}(X) \models$ " $a \in x$ " iff $a \in x \cap H$ and thus $\pi(x)$ is equal to $\{\pi(a) : a \in x \cap H\}$. In particular, if X is transitive, the inductive hypothesis tells us that $\pi(x)$ for $x \in X$ is equal to $\{\pi(a) : a \in x \cap H\} = \{a : a \in x \cap H\} = x \cap H$. Since X is transitive, $x \subseteq X \subseteq H$ so that $x \cap H = x$. Therefore $\pi(x) = x$ and so $\pi \upharpoonright X = \operatorname{id} \upharpoonright X$ by induction on rank. \dashv

Note that the use of "collapse" especially makes sense here, because every $\pi(x) \in \text{cHull}^{\mathsf{T}}(X)$ has $\operatorname{rank}(\pi(x)) \leq \operatorname{rank}(x)$. Of course, strict inequality requires that $x \not\subseteq \operatorname{Hull}^{\mathsf{T}}(X)$. Using Tarski–Vaught Theorem (6 A • 6) and direct limits in general, we can build up skolem hulls to have less and less collapsed while still being relatively small.

In particular, if we take the hull that includes all of an ordinal, we get a model that contains all of the ordinals below it. Using the elementary chains, this allows us to conclude the following, showing we can get ordinals in our uncollapsed model before collapsing.

- 6C·3. Corollary

Let T be a transitive set with $\kappa \in T$ an uncountable, regular cardinal and $X \subseteq T$ of size $\langle \kappa$. Therefore, there is an elementary $\mathbf{H} \preccurlyeq \langle T, \epsilon \rangle$ with $H \cap$ Ord an ordinal, $|H| < \kappa$, and $X \subseteq H$.

Proof .:.

Take the skolem hull $H_0 = \text{Hull}^{(T,\epsilon)}(X)$. This may not have $H_0 \cap \text{Ord}$ as an ordinal although it will satisfy that $H_0 \leq \langle T, \epsilon \rangle$ and $|H_0| \leq \aleph_0 \cdot 1 \cdot |X| < \kappa$. For H_n already defined, if $H_n \cap \text{Ord}$ is an ordinal, then stop the process, and take $H = H_n$. Otherwise let $H_{n+1} = \text{Hull}^{(T,\epsilon)}(H_n \cup \sup(H_n \cap \text{Ord}))$. As a regular cardinal, $\sup(H_n \cap \text{Ord}) < \kappa$ because inductively $|H_n| < \kappa$, which also tells us that $|H_{n+1}| < \kappa$. Define H_{ω} to be the direct limit of the H_n s for $n < \omega$ as in Tarski–Vaught Theorem (6 A • 6).

Note that $H_{\omega} \leq \langle T, \in \rangle$ with $X \subseteq H_{\omega}$ and $|H_{\omega}| \leq \aleph_0 \cdot \sup_{n \in \omega} |H_n|$. As each $|H_n| < \kappa$ and κ has cofinality $\kappa > \omega$, it follows that $\sup_{n \in \omega} |H_n| < \kappa$ and thus $|H_{\omega}| < \kappa$. To see that $H_{\omega} \cap \text{Ord}$ is an ordinal, it suffices to show that $H_{\omega} \cap \text{Ord}$ is transitive. For $\beta \in H_{\omega} \cap \text{Ord}$, it follows that $\beta \in H_n \cap \text{Ord}$ for some $n < \omega$. Thus $\beta \subseteq H_{n+1}$ and so $\beta \subseteq H_{\omega} \cap \text{Ord}$.

Theorems and ideas like this will play a big role in what we can do with small models of fragments of set theory as well as inner models (which haven't been defined yet). To proceed further in this direction, we will need to consider the FOLp agreement between **V** and other transitive sets in the next section.

Section 7. Absoluteness

Absoluteness in some sense refers to how correct our definitions of concepts are. Formally, a definition is absolute between two models if the two agree on what the definition applies to. For example, $x \in y$ is absolute between any two transitive models containing x and y: $\mathbf{A} \models "x \in y"$ iff $\mathbf{B} \models "x \in y"$ since they both interpret membership the same way.

— 7•1. Definition -

Let $\varphi(\vec{x})$ be a FOL-formula. Let A and B be models. We say that φ is *absolute* between A and B iff $A \models \varphi(\vec{a})$ iff $B \models \varphi(\vec{a})$ whenever \vec{a} are parameters belonging to both A and B.

– 7•2. Corollary

If **A** and **B** model some theory T, and if $T \vdash \forall \vec{x} \ (\varphi \leftrightarrow \psi)$, then φ is absolute between **A** and **B** iff ψ is.

This general definition isn't much to work with. It does, however, tell us that many of our set-theoretic conceptions are not absolute between models of set theory. For example, the argument in Result $4 \land 9$ shows that well-foundedness isn't absolute. A similar argument shows that even a set being infinite isn't even absolute between models of set theory: consider any particular infinite set $A \in V$ and the signature $\sigma \cup A \cup \{A\}$ where we have a constant symbol ffor A and every element of A. Then consider the theory $T = \mathsf{ZFC} + \{a \in A \land A \text{ is finite}^n : a \in A\}$. We can always interpret the symbol 'A' as some finite subset of A, and in particular, for any finite subtheory of T, the set of all constants appearing in the subtheory. This shows that each finite subset of T is satisfiable, and thus that T has a model where A is finite.

So these ideas mean first-order logic on its own doesn't tell us much about the deeper structure of V. They only tell us is that asking whether something is absolute in general, without any further restrictions, is not a good question to ask. So for the most part, we will restrict our view to models which are transitive. And in doing so, we also can refine this notion a bit. Note that we are assuming, as structures, that the models are non-empty.

— 7·3. Definition

For $A \subseteq V$ transitive, write **A** for (A, \in) , and call **A** a *transitive model*. Let **A** \subseteq **B** be two transitive models, and let $\varphi(\vec{x})$ be a FOL-formula.

- φ is downward-absolute between A and B iff $B \models \varphi(\vec{a})$ implies $A \models \varphi(\vec{a})$ whenever \vec{a} are in $A \cap B = A$.
- φ is upward-absolute between A and B iff $A \models \varphi(\vec{a})$ implies $B \models \varphi(\vec{a})$ whenever \vec{a} are in A.
- φ is *downward-absolute* iff φ is downward-absolute between all transitive models and submodels.
- φ is *upward-absolute* iff φ is upward-absolute between all transitive models and submodels.
- φ is *absolute* iff φ is absolute between all transitive models.

Equivalently, φ is absolute iff φ is absolute between V and its transitive submodels. Because membership for transitive models is always the same, we get that " $x \in y$ " is absolute. " $x = \emptyset$ " is absolute: for A transitive with $x \in A$, x is non-empty iff there is a $y \in x \subseteq A$. By absoluteness of " $y \in x$ " and transitivity, this is equivalent to $A \models$ " $\exists y (y \in x)$ ", which just says $A \models$ " $x \neq \emptyset$ ". Hence " $x \neq \emptyset$ " (and thus " $x = \emptyset$ ") is absolute.

The above idea should indicate that absoluteness can often be proven in a kind of inductive way, beginning with simple formulas like " $x \in y$ " and working with increasingly more complex formulas. We can prove a great number of absoluteness results by studying this kind of complexity, which is mostly just due to the number of quantifiers. But because we're working with transitive models, *bounded* quantifiers do not increase complexity: a bounded quantifier $Qx \in X$ ranges over the same elements (namely the elements of X) in V as in the transitive model A, because both properly understand what it means to be a member of X. And this is really the best understanding of transitivity: properly understanding membership. The following hierarchy of formulas is called the *Lévy hierarchy* after Azriel Lévy (χ_{I} (χ_{I} (χ_{I} (χ_{I} (χ_{I} (χ_{I})).

— 7•4. Definition

Let φ be a FOL-formula. A *bounded quantifier* is a quantifier of the form " $\exists x \in X$ " or " $\forall x \in X$ " for some x and X, being short-hand for " $\exists x (x \in X \land \cdots)$ " and " $\forall x (x \in X \rightarrow \cdots)$ " respectively.

- φ is Σ_0 (and Π_0) iff all quantifiers occurring in φ are bounded.
- φ is Σ_{n+1} iff φ is of the form $\exists x \ \psi$ where ψ is Π_n .
- φ is Π_n iff φ is of the form $\neg \psi$ where ψ is Σ_n .

Note that the X in " $\exists x \in X$ " is a variable rather than a parameter. We also get a variant hierarchy where we allow parameters. In particular, something is $\Sigma_n(A)$ iff it satisfies the same definition, but allows parameters in A. If we allow parameters, we get more formulas. For example, we could bound quantifiers by elements of A, allowing more $\Sigma_0(A)$ -formulas than the standard Σ_0 -formulas.^{xxv}

This is the first hierarchy of formulas we will encounter, although we will encounter many more later revolving around \mathbb{R} and \mathbb{N} . The fact that transitive sets understand bounded quantifiers (when they contain the parameters), tells us that Σ_0 -formulas are absolute. Note that we can formalize this by first noting that truth in transitive classes can be known by **V**, just by bounding quantifiers.

- 7 • 5. Definition

Let C be a (FOLp-definable) transitive class and φ a FOLp-formula with parameters in C. The formula φ^{C} is the formula where each quantifier " $\exists x$ " and " $\forall x$ " is replaced by " $\exists x \in C$ " and " $\forall x \in C$ ", respectively.

Alternatively, we can define φ^{C} by induction on φ :

 $\begin{array}{lll} "(x = y)^{C}" & \text{is} & "x = y"; \\ "(x \in y)^{C}" & \text{is} & "x \in y"; \\ "(\neg \varphi)^{C}" & \text{is} & "\neg \varphi^{C}"; \\ "(\varphi \wedge \psi)^{C}" & \text{is} & "(\varphi^{C} \wedge \psi^{C})"; \\ "(\exists x \varphi)^{C}" & \text{is} & "\exists x \in C \varphi"; \text{and} \\ "(\forall x \varphi)^{C}" & \text{is} & "\forall x \in C \varphi". \end{array}$

More explicitly, since membership in C is really a formula, for C is defined by ψ , then " $(\exists x \phi)^{C}$ " is " $\exists x (\psi(x) \land \phi^{C})$ ". The same idea applies to C a set with $\psi(x)$ just being " $x \in C$ ". Note that if we already have a bounded quantifier, the restriction of " $\exists x \in X$ " to C then gives " $\exists x \in X \cap C$ ", and similarly " $\forall x \in X$ " maps to " $\forall x \in X \cap C$ ". Therefore, we can recast truth about φ in **C** as truth of φ^{C} in **V**. Note that we already knew how to do with with sets by Theorem 6 B • 6, but not classes in general.

– 7•6. Lemma

Let C be a transitive class and φ a FOLp-formula with parameters in C. Therefore $(C, \in) = C \vDash \varphi$ iff $V \vDash \varphi^C$.

Proof .:.

Proceed by induction on formula complexity. As a transitive class, $\mathbf{C} \models "x \in y$ " iff $\mathbf{V} \models "x \in y$ " so that the result holds if φ is atomic. The sentential connectives follow easily from the inductive hypothesis. So suppose φ is of the form $\exists x \ \psi$.

- If $\mathbf{C} \models ``\exists x \psi$ '', then $\mathbf{C} \models ``\psi(c)$ '' for some $c \in \mathbf{C}$ and thus inductively—since $\psi^{\mathbf{C}}(c)$ is $(\psi(c))^{\mathbf{C}}$ $\mathbf{V} \models ``\psi^{\mathbf{C}}(c)$ ''. Thus $\mathbf{V} \models ``\exists x \in \mathbf{C} \psi^{\mathbf{C}}$ '' which is just to say that $\mathbf{V} \models ``\phi^{\mathbf{C}}$ ''.
- Conversely, if $\mathbf{V} \models ``\exists x \in C \psi^C$ '', then for some $c \in C$, $\mathbf{V} \models ``\psi^C(c)$ '', which inductively says $\mathbf{C} \models ``\psi(c)$ '' and thus $\mathbf{C} \models \varphi$.

^{xxv}Note that, a priori, not every formula can be placed in the Lévy hierarchy as we've stated it here. Other sources will do away with this issue by allowing *blocks* of quantifiers rather than single ones. This is avoided here both to show the importance of the background assumptions, and to show how to get around the issue. In particular, " $\exists x \exists y (x \in y)$ " can't be placed in the hierarchy. It is only by assuming some additional set theory that this formula is *equivalent* to one in the Lévy hierarchy.

More precisely, every formula is equivalent to one in prenex normal form: all quantifiers appear at the beginning of the formula. Assuming some basic set theory, each block of quantifiers of the form $Qx_0 Qx_1 \cdots Qx_n \varphi$ can instead be written as $Qx(x = \langle x_0, \cdots, x_n \rangle \land \varphi')$ where φ' replaces each x_i with the defined notion of being the *i*th entry in x, something which can be said using only bounded quantifiers. This allows us to show each formula is equivalent to one in the Lévy hierarchy. In this sense, we say that φ is Π_n (or Σ_n) iff φ is equivalent to a formula which is Π_n (or Σ_n). In doing so, however, we need to specify the theory they are equivalent under.

So just by rewriting the definition, we get the following, alternative characterization of absoluteness which is a statement just about V. Note that this statement isn't something that V can evaluate though: what transitive classes exist depends on V, but V cannot quantify over all transitive classes.

— 7•7. Corollary

Let φ be a FOL-formula. Therefore φ is absolute iff $\mathbf{V} \models "\varphi \Leftrightarrow \varphi^{C}$ " for each transitive class C.

§7A. Easy absoluteness results

Important examples of absolute formulas include all of the Σ_0 -formulas of the Lévy hierarchy.

— 7A•1. Result

Let φ be a Σ_0 -formula. Therefore φ is absolute.

Proof .:.

Proceed by structural induction on φ . Since all quantifiers in φ are bounded, and $\forall x \in X \ \psi$ is equivalent to $\neg \exists x \in X \neg \psi$, we only need to consider the sentential operations and the bounded quantifier $\exists x \in X \ \psi$. The sentential operations are immediate by induction. So it suffices to consider bounded quantification: suppose φ is $\exists x \in X \ \psi$. Let **A** be an arbitrary, transitive model with $X \in A$ so that $X \subseteq A$. Inductively, for all $a \in A, \ \psi^A(a)$ holds iff $\psi(a)$ holds.

- Since $X \subseteq A$, if there is an $a \in X$ such that $\psi(a)$ holds then there is an $a \in A$ such that $a \in X$ and $\psi^A(a)$ holds, i.e. φ^A holds.
- Conversely, if φ^A holds, then there is some $a \in X \cap A = X$ such that $\psi^A(a)$ holds. Inductively, this means $\psi(a)$ holds and thus $\exists x \in X \psi$.

More generally, this says that if φ is absolute between transitive **A** and **B**, then " $\exists x \in X \varphi$ " is absolute between them as well when X is in both. And of course, boolean combinations^{xxvi} of absolute formulas are absolute as well. This is stated as follows with the same proof as Result 7 A • 1.

- 7A•2. Result

Let A be a transitive model. Therefore the set of FOLp-formulas absolute between A and V is closed under bounded quantification, conjunctions, and negations.

- 7 A • 3. Corollary

The following axioms are absolute, because they are true in all non-empty, transitive models:

- the axiom of extensionality,
- the axiom of the empty set, and
- the axiom of foundation.

Proof .:.

Since all of these are true in V, the only way for these to fail to be absolute is if they are false in some transitive model. So we will show this does't happen. All of these can be shown through careful analysis of the forms of the axioms.

- Extensionality says that for every x and y, " $x = y \leftrightarrow \forall v \in x \ (x \in y) \land \forall v \in y \ (v \in x)$ " holds. The property of extensionality holding at x, y is Σ_0 and thus absolute. So if it holds for all $x, y \in V$, then it holds for all $x, y \in C$ for any class C. Therefore extensionality holds in C, and is thus absolute.
- By (3) of Corollary 2 E 5, the universe has the empty set in it. By the argument just after Definition 7 3,
 "x = Ø" is absolute and thus the axiom of the empty set is satisfied and thus absolute: "∃x (x = Ø)" is absolute.
- The axiom of foundation says that for every $x \neq \emptyset$, $\exists y \in x \forall z \in y (z \notin x)$. Since " $x \neq \emptyset$ " is absolute and the other part is Σ_0 , foundation holding for x is absolute, meaning for every $x \in C$, **C** believes foundation

xxvimeaning formulas built up from the sentential connectives starting from some given formulas

 \neg

holds for x, because it holds in V. Therefore C must satisfy foundation.

Result 7 A • 1 gives a great number of absoluteness results. For the most part, we will not give the completely formal definitions that show these are Σ_0 . Instead, like with much of first-order logic, we will resort to giving impressions and instructions that allow one to carefully check that they are. The following are absolute all because they are defined by Σ_0 -formulas.

- *x* being an (un-ordered) pair: everything in *x* is either some $y \in x$ or $z \in x$.
- *x* being an ordered pair.
- x being the first-coordinate of an ordered pair y: there is a $z \in y$ such that for every $w \in z, w = x$.
- *x* being a relation.
- x being the domain of R: for every $z \in x$, there is a pair $(z, y) \in R$, and vice versa.
- *x* being the range of *R*.
- x being a function: for every y in the domain of x, there is a unique z in the range of x with $(y, z) \in x$.
- x being the output of a function f with input y, i.e. x = f(y).
- *x* being an injective function.
- *x* being a surjective function.
- x being a subset of y: every $z \in x$ is in y.
- x being transitive: every $z \in x$ is a subset of x.
- x being an ordinal: x is transitive, and \in linearly orders x.

And many, many more concepts are absolute by Result 7 A • 1. As a result of the above absoluteness examples, we have some nice consequences about what it means for transitive sets to model the axioms of set theory. Most of the axioms of set theory state the closure of the universe under certain sets. Pair, for example, says that for every x and y, $\{x, y\}$ exists. Now while the property of *being* an (un-ordered) pair is absolute, this doesn't tell us that the *existence* of un-ordered pairs is absolute. Similarly, being a subset is absolute, but being the powerset isn't, because a model might contain fewer subsets than another: $4 = \{0, 1, 2, 3\}$ thinks $\mathcal{O}(2)$ exists and is $3 = \{0, 1, \{0, 1\}\}$, because every subset of 2 that 4 contains is in 3: $\{1\} \notin 4$ although $\{1\} \subseteq 2$.

But we can have a better picture of what these sorts of defined sets will look like, because their defining formulas are absolute. Really, the following is just another way to state absoluteness.

- **7** A • 4. Result - Let φ be a FOLp-formula absolute between a transitive model M and V. Therefore $\{x : \varphi(x)\}^{M} = \{x : \varphi(x)\} \cap M$.

Proof .:.

 $\mathbf{M} \models "\varphi(x)"$ iff $x \in M$ and $\varphi^M(x)$ holds. By absoluteness, this is equivalent to $x \in M$ and $\varphi(x)$.

As a result, $\mathcal{O}^{M}(X) = \mathcal{O}(X) \cap M$, $\operatorname{Ord}^{M} = \operatorname{Ord} \cap M$, and so on. This idea also gives an understanding of when transitive models satisfy (some of the) axioms of set theory.

— 7 A • 5. Corollary -

Let **A** be a transitive model. Therefore,

• $A \models$ Union iff $x \in A$ implies $\bigcup x \in A$.

- $A \models \text{Pair iff } x, y \in A \text{ implies } \{x, y\} \in A.$
- $A \models \text{Comp iff } \{x \in y : \varphi^A(x)\} \in A \text{ for each } y \in A \text{ and FOLp-formula } \varphi$.
- $A \models P$ iff $x \in A$ implies $\mathcal{P}(x) \cap A \in A$.
- $\mathbf{A} \models \text{Inf if } \omega \in A$.

Proof .:.

For Union, Pair, and Comp, by absoluteness, the only way A can interpret " $x = \bigcup y$ " and " $x = \{y, z\}$ " is the same way V does. The only way A can interpret " $x = \{y \in z : \varphi(y)\}$ " is the way V interprets " $x = \{y \in z : \varphi^A(y)\}$ ".

 \dashv

Hence A being closed under unions, pairing, or comprehension as A interprets it (i.e. satisfying Union, Pair, or Comp) is the same as being closed under unions, pairing, or comprehension as V interprets it.

For P, note that being a subset is absolute. Hence

 $\mathcal{P}^{\mathsf{A}}(x) = \{ y \in A : \mathsf{A} \vDash "y \subseteq x" \} = \{ y \in A : y \subseteq x \} = \mathcal{P}(x) \cap A.$

Hence being closed under powersets (i.e. satisfying P) is the same as being closed under powersets intersected with the universe.

For Inf, clearly if $\omega \in A$, then by the absoluteness results above, A satisfies Inf. The reverse may not hold, since $\omega \cup \{\{1\} + n : n < \omega\} \cup \{\omega \cup \{\{1\} + n : n < \omega\}\},\$

where $x + 1 = x \cup \{x\}$ and $x + n = ((x + 1) + \cdots) + 1$, is a transitive set that models the axiom of infinity, but ω is merely a subset of the universe (and a subset of a set in the universe), not a set inside it. \dashv

Of course, not everything turns out to be absolute, but we can get partial absoluteness for some formulas, as we've used in Corollary $7 \text{ A} \cdot 3$. For example, we have the following easy consequences of Result $7 \text{ A} \cdot 1$.

Proof .:.

Let φ be $\forall x \ \theta$ where θ is Σ_0 . If $\mathbf{V} \models \forall x \ \theta$, then, in particular, $\theta(a)$ holds for every $a \in \mathbf{C}$. By absoluteness, $\theta(a) \leftrightarrow \theta^{\mathbf{C}}(a)$ and thus $\mathbf{C} \models \forall x \ \theta$.

Let ψ be $\exists x \ \theta$ where θ is Σ_0 . Thus $\neg \psi$, being $\forall x \ \neg \theta$ is downward absolute. Taking the contrapositive means that ψ is upward absolute: $\mathbf{V} \models ``\neg \psi \rightarrow \neg \psi^{C}$ '' implies $\mathbf{V} \models ``\psi^{C} \rightarrow \psi$ ''.

Again, more generally, when φ is absolute between $\mathbf{A} \subseteq \mathbf{B}$, then $\exists x \varphi$ is upward absolute between them and $\forall x \varphi$ is downward absolute between them. We cannot ask for stronger than this mere partial absoluteness. For example, the existence of ω —meaning the axiom of infinity—is Σ_1 :

$$"\exists N(\emptyset \in N \land \forall x \in N(x \cup \{x\} \in N))",$$

but the transitive set $\{\emptyset\}$ doesn't have such an N although V does. So upward absoluteness is all we can say about Σ_1 -formulas in general. Similarly, in $\{\emptyset\}$, we have the Π_1 -sentence " $\forall x \ (x = \emptyset)$ " as true although it's false for V. So downward absoluteness is all we can say about Π_1 -formulas in general.

§7B. The Lévy hierarchy and absoluteness with some set theory

As Corollary 7 A • 3 shows, relatively few things will be absolute, especially if they require more axioms of set theory to even state properly. For example, while *being* the union of two sets is absolute, the *existence* of such a set isn't absolute. For example, $\{0, 1, \{1\}\}$ is transitive, but $2 = \{1\} \cup 1$ isn't in the set. So often it will be useful to restrict our attention to transitive models of some fragment of ZFC.

 $- 7B \cdot 1. \text{ Definition}$ Let *T* be a theory, and φ a FOL-formula. We say that φ is Σ_n^T (or Π_n^T) iff $T \vdash "\varphi \leftrightarrow \psi$ " for some Σ_n (or Π_n) formula ψ . We say that φ is Δ_n^T iff φ is both Σ_n^T and Π_n^T .

As a result, Σ_n^{\emptyset} just consists of all formulas logically equivalent to Σ_n -formulas, and similarly for Π_n^{\emptyset} . As a result, the placement of a formula φ isn't unique: if φ is a Σ_n^T -sentence, then " $\forall x \varphi$ "—just adding on a dummy quantifier—is a Π_{n+1}^T -sentence that is logically equivalent to φ .

Note also that if φ is Σ_n^T and $T \subseteq T'$, then φ is $\Sigma_n^{T'}$ as well. Hence absoluteness results for $T \subseteq ZFC$ extend to absoluteness results for ZFC.

— 7 B • 2. Corollary -

 Δ_1^{\emptyset} -formulas are absolute.

Proof .:.

If φ is a Δ_1^{\emptyset} -formula, then φ is Π_1^{\emptyset} and thus downward absolute by Result 7 A • 6; and φ is Σ_1^{\emptyset} and thus upward absolute by Result 7 A • 6.

7B·3. Corollary –

Well-foundedness is absolute between transitive models of Lemma $4 \cdot 3$, e.g. of ZF - P.

Proof .:.

Well-foundedness is downward absolute because a relation R is ill-founded iff the following Σ_1 -formula holds: $\exists x \forall y \in x \exists z \in x (z R y)$. This means well-foundedness is Π_1 , and thus downward absolute.

Upward absoluteness holds as it is Σ_1^T for T such a theory as in the statement of the corollary: it states the existence of a pair: a function and an ordinal which constitute a rank function. By Lemma 4 • 3: if $R \subseteq A \times A$ is well-founded in a model **C** of this, then **C** believes that there is a rank function $f : A \to \text{Ord}^{\mathsf{C}}$. Since the following are Σ_0 and so absolute between transitive models:

- being a function, and being f(x);
- being an ordinal—which implies $Ord^{c} = Ord \cap C$;
- being 0;
- being *R*-minimal—which is Σ_0 as seen by " $\forall y \in A (\neg y R x)$ ";
- being x + 1; and
- being the supremum of a set of ordinals,

it follows that $f : A \to \text{Ord} \cap C$ is still a rank function in V. Hence there can be no infinite *R*-decreasing sequence in *A* without the ranks decreasing and so violating the well-foundedness of the ordinals. Therefore, well-foundedness is upward absolute between such models, and hence absolute between such models. \dashv

We also get that functions and sets defined by transfinite recursion using absolute notions will be absolute.

— 7B•4. Theorem

Suppose "F(x) = y" is absolute between transitive models of ZF - P. Let $G : Ord \rightarrow V$ be defined by transfinite recursion: $G(\beta) = F(G \upharpoonright \beta)$. Therefore "G(x) = y" is absolute between transitive models of ZF - P.

Proof .:.

Applying transfinite recursion in \mathbb{M} , $G^{\mathbb{M}}$ is such that for every ordinal $\alpha \in \operatorname{Ord} \cap M$, $G^{\mathbb{M}}(\alpha) = F^{\mathbb{M}}(G^{\mathbb{M}} \upharpoonright \alpha)$. By transfinite *induction* on α , the inductive hypothesis that $G^{\mathbb{M}} \upharpoonright \alpha = G \upharpoonright \alpha$ and the absoluteness of F implies $G^{\mathbb{M}}(\alpha) = F(G \upharpoonright \alpha) = G(\alpha)$, as desired. Hence $G^{\mathbb{M}} = G \cap M$.

As a result, the rank of a set is absolute between transitive models of ZF - P being defined by transfinite recursion.

Proof .:.

 $\operatorname{rank}(x) = \alpha$ is absolute between such models as a consequence of Lemma 4 • 3 where we take membership to be the well-founded relation. As noted, the proof works just as well for classes as for sets, since it just relies on transfinite induction and recursion.

To show $V_{\alpha}^{\mathsf{M}} = V_{\alpha} \cap M$, just note that we can define $x \in V_{\alpha}$ iff rank $(x) < \alpha$, which is absolute.

 \neg

Some of the most common axioms satisfied are those of basic set theory.

7B•6. Definition

Basic set theory (BST) consists of the following axioms:

- 1. extensionality, empty set, foundation;
- 2. comprehension, pairing, union; and
- 3. the existence of cartesian products: $\forall x \forall y \exists z \forall w \ (w \in z \Leftrightarrow \exists a \in x \exists b \in y \ (w = \langle a, b \rangle))$.

We know from Corollary $7 A \cdot 3$ that (1) is already absolute and satisfied by all transitive sets. So the addition of (2) adds more absoluteness between models we care about. (3) is not an explicit axiom of ZFC, but it does follow from both powerset and replacement. Since we will work in contexts in which either might be missing, we use the weaker result that cartesian products exist.

The reason for this is that under BST, a greater number of things are equivalent, and the Lévy hierarchy will be closed under various operations. What we mean by this is that for T a theory, a formula φ is Σ_n^T iff φ is equivalent over Tto a Σ_n formula. So we mean that Σ_n^{BST} is a much larger class than Σ_n^{\emptyset} . For the most part, we will just need slight weakenings of ZFC, but working in more generality will help later. In particular, Σ_1^{BST} is closed under existential quantification: $\exists y \exists x \psi$ for ψ being Σ_0 is equivalent to $\exists z \exists y \in z \exists x \in z \psi$ by pairing in BST. Similarly, Π_1^{BST} is closed under universal quantification in addition to \lor and \land .

- 7 B • 7. Result

For each $n < \omega$, Σ_n^{BST} is closed under existential quantification, disjunction, and conjunction. Π_n^{BST} is closed under universal quantification, disjunction, and conjunction. Δ_n^{BST} is closed under bounded quantification, disjunction, conjunction, and negation.

To get more than this, we need more set theory. For example, in full ZFC, each Σ_n^{ZFC} is closed under existential quantification and both bounded quantifiers (and similarly for Π_n^{ZFC} and universal quantification). Showing this, however, requires some complicated tricks better suited for the end of the chapter.

But these calculations give some partial absoluteness about cardinality and cofinality.

— 7 B • 8. Result

Being a cardinal is Π_1^{BST} and therefore downward absolute between models of BST. κ being singular is Σ_1^{BST} and therefore upward absolute between models of BST. Hence being regular is Π_1^{BST} and so downward absolute between models of BST.

Proof .:.

 κ being a cardinal is equivalent to being an ordinal (which is Σ_0) and



The calculation above shows that this is equivalent over BST to a Π_1 -formula.

 κ being singular is equivalent to there being an increasing function whose image is cofinal in κ and whose domain is an ordinal less than κ :



It should be noted that any $T \supseteq BST$ also has these absoluteness results because T will prove the equivalences that BST does. But we can do more of these kinds of calculations to get that ω is absolute when ω is in the model.

— 7 B•9. Result

Finiteness is absolute between transitive models of BST.

Proof ...

- 1. $x = y \cup \{y\} = y + 1$ is an absolute relation.
- 2. x being a limit ordinal is equivalent to x being an ordinal (absolute between transitive models) and $\forall y \in x \ (y \cup \{y\} \in x)$, which is absolute by (1).
- 3. x being the least ordinal of a set of ordinals X is absolute. To see this, the least ordinal of X can be defined by a Σ_0 -formula: x is the least ordinal of X iff x is an ordinal and $x \in X$ and $\forall y \in X(y \text{ is an ordinal} \rightarrow x = y \lor x \in y)$, which is Σ_0 .
- 4. x being ω is the same as being the least ordinal in the class of limit ordinals (if there are any).
- 5. x being $n < \omega$ is just defined by iteratively considering (1).

x being finite is just to say that there is some $n \in \omega$ with a bijection $f : x \to n$. This form is Σ_1^{BST} and thus upward absolute between transitive models of BST. For downward absoluteness, suppose $\mathbf{A} \models \text{BST}$ with $x \in A$. If *x* is really finite, then there is some $n < \omega$ where we can then write out that $x = \{x_0, \dots, x_{n-1}\}$ and so define *f* by $f = \{\langle x_0, 0 \rangle, \dots, \langle x_{n-1}, n-1 \rangle\}$, just using a single Σ_0 -formula. Since $n \in A$ by pairing and union, it follows by cartesian products and comprehension that $f \in A$ and thus $f : x \to n$ is a bijection showing *x* is finite in \mathbf{A} .

BST is mostly brought up because almost every model we would like to consider will be a model of it. With stronger theories, like ZF, we get more absoluteness, and learn more about V.

§7C. Toy models for set theory

So far we've investigated the absoluteness between models of various fragments of set theory, but we haven't given many concrete examples of what these models look like. Firstly, the levels of the cumulative hierarchy serve as a nice introduction to models of (fragments of) set theory. Just by their form, we immediately get some axioms holding in them.

− 7C • 1. Result Let $\alpha \in \text{Ord.}$ Therefore $V_{\alpha} = \langle V_{\alpha}, \in \rangle \models \text{BST} - \text{Pair} - \text{``the existence of cartesian products''}$.

Proof .:.

Extensionality, empty set, and foundation all hold by Corollary 7 A • 3, because V_{α} is transitive. It suffices by Corollary 7 A • 5 to show that V_{α} is closed under pairing, unions, and subsets. But just by the rank argument given in Result 4 A • 10, $x, y \in V_{\alpha}$ implies $x \cup y \in V_{\alpha}$, and $y \subseteq x \in V_{\alpha}$ implies $y \in V_{\alpha}$. In particular, $\{z \in x : \varphi^{V_{\alpha}}(z)\} \in V_{\alpha}$.

The issue with pairing and the existence of cartesian products is that they increase rank. Thus for V_{α} to be closed under these, α should be a limit ordinal. If this is the case, then V_{α} models much more than just BST. In particular, V_{α} models almost all of ZFC.

7C·2. Result

Let α be a limit ordinal. Therefore $V_{\alpha} = \langle V_{\alpha}, \in \rangle \models \mathsf{ZFC} - \mathsf{Rep} - \mathsf{Inf}$.

Proof .:.

Recall that $V_{\beta} \subseteq V_{\gamma}$ for $\beta < \gamma$. Hence as a limit ordinal, $V_{\beta+n} \subseteq V_{\alpha}$ for any $\beta < \alpha$ and $n < \omega$. Let $x \in V_{\alpha_x+1}$ and $y \in V_{\alpha_y+1}$ be arbitrary with $\alpha_x, \alpha_y < \alpha$.

- For Pair, x, y ∈ V_{max(αx,αy)+1} and thus {x, y} ∈ V_{max(αx,αy)+2}. As a limit ordinal, αx, αy < α implies max(αx, αy) + 2 < α and thus {x, y} ∈ V_α. So by Corollary 7 A 5, V_α ⊨ Pair.
 - V_{α} satisfies Comp trivially, since any subset $y \subseteq x \in V_{\alpha_x+1}$ (like $y = \{z \in x : \varphi^{V_{\alpha}}(z)\}$) has $y \subseteq x \subseteq V_{\alpha_x}$ and thus $y \in V_{\alpha_x+1}$ by definition of cumulative hierarchy. In fact, $\mathcal{P}(x) \in V_{\alpha_x+2} \subseteq V_{\alpha}$. So since

- $\mathcal{O}^{\mathsf{V}_{\alpha}}(x) = \mathcal{O}(x) \cap \mathcal{V}_{\alpha} = \mathcal{O}(x)$, it follows that $\mathsf{V}_{\alpha} \models \mathsf{P}$ too.
- For Union, if $x \in V_{\alpha_x+1}$, then rank $(x) > \sup\{\operatorname{rank}(y) + 1 : y \in x\}$ Hence $y \subseteq V_{\alpha_x+1}$ for each $y \in x$. Therefore $\bigcup x \subseteq V_{\alpha_x+1}$ and so $\bigcup x \in V_{\alpha_x+2}$. By Corollary $7 \land \bullet 5$, $V_{\alpha} \models$ Union. As a result, the existence of cartesian products holds, since $x \times y \in \mathcal{O}(\mathcal{O}(\mathcal{O}(x \cup y))) \in V_{\alpha}$ for each $x, y \in V_{\alpha}$, and being the cartesian product is absolute.
- For AC, for any non-empty family of non-empty, disjoint sets F ∈ V_α, there is a set C in V that has chosen one element from each set in F. Note that C ∈ P(∪F) ∈ V_α so that C ∈ V_α.

So by Corollary 7 A • 5, since rank(ω) = ω , $V_{\alpha} \models$ Inf whenever $\alpha > \omega$ is a limit. This should be taken to be evidence of the consistency of ZFC; the two independent hurdles to this being the axiom of infinity, and replacement.

- 7C•3. Corollary — Let $\alpha > \omega$ be a limit ordinal. Therefore $V_{\alpha} \models ZFC - Rep$. Moreover, $V_{\omega} \models ZFC - Inf$, and in fact $V_{\omega} \models \neg Inf$.

Proof .:.

 $\omega \in V_{\alpha}$ implies $V_{\alpha} \models ZFC - Rep$ by Result 7 C • 2 and Corollary 7 A • 5. As for V_{ω} , note that every element of V_{ω} is finite: by induction, $V_0 = \emptyset$, and $|V_{n+1}| = 2^{|V_n|+1}$ is finite as well. Hence $V_{\omega} = \bigcup_{n < \omega} V_n$ has no infinite set in it and so $V_{\omega} \models \neg Inf$, as any set following such a definition would be infinite (in V).

To see that $V_{\omega} \vDash \mathsf{Rep}$, suppose φ is a FOLp-formula that defines a function over $D \in V_{\omega}$, which is to say

$$\mathcal{U}_{\omega} \models ``\forall x \in D \exists ! y \varphi(x, y) ".$$

We now wish to show that the image of φ is in V_{ω} . Note that in V, there is then a function $f : D \to V_{\omega}$. As D is finite, there is some finite subset $R \subseteq V_{\omega}$ with $f : D \to R$. As each $r \in R$ has rank $n_r < \omega$ and R is finite, it follows that the rank of R is max $\{n_r + 1 : r \in R\} < \omega$ and thus $R \in V_{\omega}$. Hence $V_{\omega} \models \text{Rep}$.

Of course, by Gödel's incompleteness theorem—assuming that ZFC is consistent—we can't construct from ZFC alone a model of ZFC, as this would imply $ZFC \vdash Con(ZFC)$. But the issues with replacement and the axiom of infinity can be dealt with at the cost of the powerset axiom.

Note that in the proof of Corollary 7 C • 3, the reason why replacement holds in V_{ω} is due to rank being bounded: the domain is small enough, and so the outputs are bounded. If we could ensure that our toy model was "regular" in a similar sense as with V_{ω} , we can ensure replacement holds. To make this idea precise, we have the following definition.

- 7C • 4. Definition Let $\kappa \ge \aleph_0$ be a cardinal. Let H_{κ} be the set of *hereditarily* < κ -sized sets defined by $x \in H_{\kappa}$ iff $|x| < \kappa$ and every $y \in trcl(x)$ has $|y| < \kappa$.

To give a concrete example, $V_{\omega} = H_{\aleph_0}$ by similar reasoning as in Corollary 7 C • 3. There's an alternative characterization of H_{κ} for regular κ . For the most part, we will not be interested in H_{κ} for singular κ , since it will not model as much set theory as with regular cardinals.

− **7C** • **5**. **Result** For $\kappa \ge \aleph_0$ a regular cardinal, $x \in H_{\kappa}$ iff $|\operatorname{trcl}(x)| < \kappa$.

Proof .:.

If $|\operatorname{trcl}(x)| < \kappa$, then clearly each $y \in \operatorname{trcl}(x)$ has $|y| < \kappa$ because the transitive closure is transitive: $y \subseteq \operatorname{trcl}(x)$. So suppose $|x| < \kappa$ and every $y \in \operatorname{trcl}(x)$ has $|y| < \kappa$. Note that $\operatorname{trcl}(x) = \bigcup_{n <} \bigcup^n x$. Clearly $|\bigcup^0 x| = |x| < \kappa$. Inductively, $|\bigcup^n x| < \kappa$ so that $\bigcup^{n+1} x = \bigcup(\bigcup^n x)$ is the union of $< \kappa$ -many sets of size $< \kappa$. As a regular cardinal, it follows that this has size $< \kappa$. Hence $\operatorname{trcl}(x)$, being the union of countably many sets of size $< \kappa$, has size $< \kappa$.

Clearly $H_{\kappa} \subseteq H_{\lambda}$ for $\kappa < \lambda$, and these are all transitive. Now just by it's definition, it's not clear that H_{κ} is a set. But by dealing just with regular cardinals—for each singular cardinal λ , $H_{\lambda} \subseteq H_{\lambda+}$ —we can show that each H_{κ} is a set. The proof of this is non-trivial, and we be the first real use of The Mostowski Collapse (4 • 1).

− **7C**•6. Result For $\kappa \ge \aleph_0$ a regular cardinal, $H_{\kappa} \subseteq V_{\kappa}$ is a set.

Proof .:.

Let $x \in H_{\kappa}$ be arbitrary. Write $T = \operatorname{trcl}(x \cup \{x\})$. Proceed by induction on rank to show every $y \in T$ has rank less than κ . For $y = \emptyset$, this is obvious. For y of rank $\alpha + 1$ with $\alpha < \kappa$, because κ is a limit ordinal, $\alpha + 1 < \kappa$ so y has rank $< \kappa$.

For y of rank γ a limit, $\gamma = \sup\{\operatorname{rank}(z) + 1 : z \in y\}$. Note that $|y| < \kappa$, and each rank $(z) + 1 < \kappa$ for $z \in y$ by induction. In other words, we have a function from $|y| < \kappa$ to κ . This is then bounded in κ , because κ is regular. Hence γ , being at most this bound, is less than κ .

Thus each $y \in T$ has rank $< \kappa$, and in particular, $x \in T$ has rank $< \kappa$. Therefore $H_{\kappa} \subseteq V_{\kappa}$.

Let's now investigate how much set theory H_{κ} will satisfy. We of couse have the basics.

– /C•/. Lemma ———	
Let $\kappa \geq \aleph_0$ be a regular cardinal.	Therefore $\mathbf{H}_{\kappa} = \langle \mathbf{H}_{\kappa}, \in \rangle \models BST + Rep.$

Proof .:.

- For Pair, note that $trcl(\{x, y\}) = trcl(x) \cup trcl(y) \cup \{x, y\}$ by (4) of Result 4A•5. So if $x, y \in H_{\kappa}$, then $|trcl(\{x, y\})| < \kappa + \kappa + \kappa = \kappa$ and therefore $\{x, y\} \in H_{\kappa}$.
- For Union, $trcl(\bigcup x) \subseteq trcl(x)$ so if $x \in H_{\kappa}$, then $trcl(\bigcup x)$ has size $\leq |trcl(x)| < \kappa$ and thus $\bigcup x \in H_{\kappa}$.
- For Comp, since $y \subseteq x$ implies $trcl(y) \subseteq trcl(x)$, it follows that $x \in H_{\kappa}$ implies $y \in H_{\kappa}$. Hence $\mathcal{O}(x) \subseteq H_{\kappa}$. In particular, all definable subsets x are in H_{κ} , and thus $H_{\kappa} \models Comp$.
- For replacement, we argue as with V_{ω} . Suppose φ defines a function from $D \in H_{\kappa}$. We wish to show that the image of φ under D, being R, is in H_{κ} . As each $r \in R$ has $|\operatorname{trcl}(r)| < \kappa$ and $|R| \leq |D| < \kappa$, it follows that $\operatorname{trcl}(R) = R \cup \bigcup_{r \in R} \operatorname{trcl}(r)$ has size $< \kappa$, being the union of $< \kappa$ -many sets each of size $< \kappa$. Therefore $R \in H_{\kappa}$ so that $H_{\kappa} \models \operatorname{Rep}$.
- The existence of cartesian products follows from replacement.

$$\dashv$$

More than just basic set theory, we get all of the axioms, except perhaps for powerset. The issue with powerset is Cantor's Theorem (5 B • 13): the powerset will have a higher cardinality. For example, H_{\aleph_1} , the hereditarily countable sets, will contain ω , but $\mathcal{P}(\omega) \cap H_{\aleph_1} = \mathcal{P}(\omega)$ will not be a set in H_{\aleph_1} because it will be too large: $|\mathcal{P}(\omega)| \ge \aleph_1$.

Now as we've seen, $H_{\aleph_0} = V_{\omega} \models ZFC - Inf$. With uncountable κ , however, we gain Inf at the expense of P.

— 7C•8.	Theorem ·			
Let $\kappa > \aleph_0$	be a regula	r cardinal.	Therefore H_{κ}	$\vDash ZFC - P.$

Proof .:.

As $\operatorname{trcl}(\omega) = \omega < \kappa, \omega \in \mathcal{H}_{\kappa}$ so that $\mathcal{H}_{\kappa} \models \operatorname{Inf.}$ For AC, for any non-empty family $F \in \mathcal{H}_{\kappa}$ of non-empty, disjoint sets, a choice set $C \in \mathcal{V}$ has $C \in \mathcal{P}(\bigcup F) \subseteq \mathcal{H}_{\kappa}$ so that $C \in \mathcal{H}_{\kappa}$. The rest follow from Lemma 7 C • 7. \dashv

As a result, if a regular cardinal $\kappa > \aleph_0$ has $H_{\kappa} = V_{\kappa}$, then $V_{\kappa} \models ZFC$. So the existence of such κ cannot be proven to exist just within ZFC. Such axioms stating the existence of such κ are effectively stronger axioms of infinity, since ZFC – Inf cannot prove the existence of a model of ZFC – Inf although ZFC can. The analogy being that ZFC cannot prove the existence of a model of ZFC + LC (LC standing for "large cardinals") can.

- 7C•9. Definition

A cardinal κ is weakly inaccessible iff κ is regular and a limit cardinal: $\lambda < \kappa$ implies $\lambda^+ < \kappa$. A cardinal κ is strongly inaccessible or just inaccessible iff κ is regular and a strong limit cardinal: $\lambda < \kappa$ implies $2^{\lambda} < \kappa$.

Note that being weakly or strongly inaccessible is downward absolute between models of ZF: being regular is down-

ward absolute, and being a limit cardinal is equivalent to $\{\alpha < \kappa : |\alpha| = \alpha\}$ being unbounded in κ , which is clearly downward absolute. Similarly, being a strong limit is downward absolute between models of ZFC.

— 7C•10. Corollary ——

Let κ be strongly inaccessible. Therefore $V_{\kappa} = H_{\kappa}$ and $V_{\kappa} \models ZFC$.

Proof .:.

We already know that $H_{\kappa} \subseteq V_{\kappa}$ since κ is regular. So it suffices to show that $V_{\kappa} \subseteq H_{\kappa}$. Firstly, note that $|V_{\alpha}| < \kappa$ for $\alpha < \kappa$. This is obvious for $\alpha = 0$. For $\alpha < \kappa$, inductively $|V_{\alpha}| = \lambda < \kappa$ implies $|V_{\alpha+1}| = 2^{\lambda} < \kappa$ as κ is strongly inaccessible. Similarly, if $\gamma < \kappa$ is a limit, $|V_{\gamma}| = \sup_{\alpha < \gamma} |V_{\alpha}|$. Since $\gamma < \kappa$ and inductively each $|V_{\alpha}| < \kappa$ for $\alpha < \gamma$, this must have size $|V_{\gamma}| < \kappa$ since κ is regular.

But for any $x \in V_{\alpha+1}$ with $\alpha < \kappa$, it follows that $trcl(x) \in V_{\alpha+1}$ and thus $|trcl(x)| \le |V_{\alpha+1}| < \kappa$ and therefore $x \in H_{\kappa}$. By Theorem 7 C • 8, all axioms of ZFC except possibly P are satisfied by $V_{\kappa} = H_{\kappa}$. By Result 7 C • 2, P is satisfied too, and thus all axioms of ZFC.

Weakly inaccessible cardinals will have their uses later: $L_{\kappa} \models ZFC$ for weakly inaccessible κ , for example.

§7D. Reflection theorems

The levels of V are able to capture a lot of information about V itself. This idea generalizes to other classes $M \subseteq V$ that have a similar construction as the cumulative hierarchy.

7D•1. Definition
A transitive class M is *stratified* iff there is a (class) function mapping α ∈ Ord to M_α ∈ V such that
M = ∪_{α∈Ord} M_α and M_γ = ∪_{α<γ} M for limit γ;
α ≤ β implies M_α ⊆ M_β;
M_α ∈ M for each α ∈ Ord;
Each M_α is transitive.

Note that, for example V is stratified as witnessed by the cumulative hierarchy. Being stratified entails that M satisfies some weakenings of the axioms of ZFC. In particular, M might not satisfy comprehension. So whereas the following axioms are equivalent to the usual axioms under the theory {Comp, Ext}, they are weaker in its absence.

— 7 D•2. Definition

- (wPair) $\{x, y\} \subseteq z$ for some $z: \forall x \forall y \exists z (x \in z \land y \in z).$
- (wUnion) $\bigcup F \subseteq z$ for some $z: \forall F \exists U \forall v (\exists x (x \in F \land v \in x) \rightarrow v \in U).$
- (wP) for each x, $\mathcal{P}(x) \subseteq z$ for some z: $\forall x \exists P \ \forall v \ (v \subseteq x \to v \in P)$.

• (wRep) the image of a function over a set is contained a set: for each $FOL(\in)$ -formula φ ,

 $\forall w_0 \cdots \forall w_n \; \forall D \; \left(\forall x \in D \; \exists ! y \; \varphi(x, y, \vec{w}) \right) \rightarrow \exists R \; \forall x \in D \; \exists y \in R \; \varphi(x, y, \vec{w}) \right).$

Writing wZFC for ZFC (and similarly wZF for ZF) replacing axioms with these weak versions, we have the following. Note that wZFC is equivalent to ZFC in the sense that wZFC $\vdash \varphi$ for each $\varphi \in$ ZFC and vice versa. But if we remove axioms like comprehension—as we do below—then the resulting theories are not equivalent.

```
- 7 D \cdot 3. Result - If M is stratified, then M \models wZF - Comp - Inf
```

Proof .:.

Extensionality, empty set, and foundation all hold by virtue of M being stratified.

For wPair, if x, y ∈ M, then x ∈ M_α and y ∈ M_β for some ordinals α, β ∈ Ord. Therefore, as M_α, M_β ⊆ M_{max(α,β)}, we have x, y ∈ M_{max(α,β)} ∈ M witnessing the axiom.

- For wUnion, if $x \in M$, then $x \in M_{\alpha}$ for some $\alpha \in Ord$ and as a transitive set with $x \subseteq M_{\alpha}$, $\bigcup x \subseteq trcl(x) \subseteq M_{\alpha}$ and therefore $M_{\alpha} \in M$ witnesses the axiom.
- For wP, suppose x ∈ M. For each y ∈ 𝒫(x) ∩ M, let α_y be the least α with y ∈ M_α. In V, we can thus consider the supremum β = sup_{y∈𝒫(x)∩M} α_y. Hence 𝒫(x) ∩ M ⊆ M_β ∈ M witnesses the axiom.
- For wRep, suppose φ defines a function in **M** on some $D \in M$. For $x \in D$, let α_x be the least α with $y \in M_{\alpha}$, where y is the output of x. By replacement in **V**, the supremum $\beta = \sup_{x \in D} \alpha_x < \text{Ord.}$ But then M_{β} contains the pointwise output of D. Therefore $M_{\beta} \in M$ witnesses the axiom. \dashv

For any stratified M, we get that the levels of M reflect the truth of M itself. To show this, we need some restricted versions of Tarski–Vaught Theorem $(6 A \cdot 6)$.

— 7 D•4. Lemma

Let φ be a FOLp-formula and $M_0 \subseteq M$ be non-empty, transitive classes. Therefore the following are equivalent:

- 1. φ and all of its subformulas are absolute between M₀ and M.
- 2. for each subformula of φ (possibly including φ itself) of the form " $\exists y \ \psi(\vec{x}, y)$ ", for all $\vec{x} \in M^{<\omega}$,

 $\exists y \in M \ \psi^M(\vec{x}, y) \to \exists y \in M_0 \ \psi^M(\vec{x}, y).$

Proof .:.

The same proof for Tarski–Vaught Theorem (6 A • 6) applies to show that (1) implies (2).

So assume (2) holds. Proceeding by induction on subformulas, given a subformula ψ of φ , we can assume each proper subformula of ψ is absolute between M_0 and M. If ψ is atomic or of the form " $\chi \wedge \theta$ " or " $\neg \theta$ " then clearly ψ is absolute between M_0 and M.

So consider the subformula " $\exists y \psi$ ". Note that then ψ^M iff ψ^{M_0} by the inductive hypothesis. Therefore, $\exists y \in M_0 \psi^{M_0}$ implies $\exists y \in M_0 \psi^M$ and therefore $\exists y \in M \psi^M$. By (2), the reverse implications hold. So we know for any \vec{m} in M,

 $\mathsf{M} \vDash ``\exists y \ \psi(\vec{m}, y)'' \quad \text{iff} \quad \exists y \in M \ \psi^{M}(\vec{m}, y) \quad \text{iff} \quad \exists y \in M_0 \ \psi^{M_0}(\vec{m}, y) \dashv$

This allows us to perform an induction on formulas so that when we close under the property of (2), we get absoluteness.

- 7 D • 5. Theorem (The Reflection Principle)

Let M be stratified, and let φ be a FOLp-formula. Therefore, there are arbitrarily large $\alpha \in \text{Ord}$ where φ is absolute between M and M_{α}.

Proof .:.

Proceed by induction on φ to show the variant result that φ and all of its subformulas are absolute between **M** and \mathbf{M}_{α} for arbitrarily large $\alpha \in \text{Ord}$. For φ being "x = y" or " $x \in y$ ", this is obvious, as they are absolute between all transitive models, which **M** and \mathbf{M}_{α} are. Similarly the propositional connectives are obvious assuming the result holds of their subformulas.

Let $\beta \in \text{Ord}$ be arbitrary. For each subformula of φ of the form " $\exists y \psi$ ", we will show there is an $\alpha > \beta$ where $\mathbf{M} \models "\forall \vec{x} \ (\exists y \ \psi(\vec{x}, y) \rightarrow \exists y \in M_{\alpha} \ \psi(\vec{x}, y))$ "

and thus by (2) of Lemma 7 D • 4, conclude that φ and its subformulas are absolute between M and M_{α}.

For each subformula " $\exists y \psi$ " and $\vec{x} \in M$, let $F_{\psi}(\vec{x})$ be the least ordinal $\alpha \in \text{Ord such that } \exists y \in M_{\alpha} \psi^{M}(\vec{x}, y)$ (if there is no such ordinal, set $\alpha = 0$). Such an ordinal $\neq 0$ will exist if $\mathbf{M} \models$ " $\exists y \psi(\vec{x}, y)$ ", since M is stratified. So F_{ψ} points to a level where there is a witness to ψ .

For $\alpha \in Ord$, consider

 $G(\alpha) = \sup\{F_{\psi}(\vec{x}) : \vec{x} \in \mathbf{M}_{\alpha}^{<\omega} \land ``\exists y \psi ``is one of the existential subformulas of \varphi\}.$

This means that in **M**, every input in M_{γ} has its witness to ψ^{M} somewhere in $M_{G(\gamma)}$. So now we just continually

apply G and then union up to get a model closed under this.

Take $\alpha_0 > \beta$ arbitrary. Let $\alpha_{n+1} = \max(G(\alpha_n), \alpha_n + 1)$ and set $\alpha = \sup_{n < \omega} \alpha_n$. Clearly α is a limit ordinal and hence $M_{\alpha} = \bigcup_{n < \omega} M_{\alpha_n}$. But then any $\vec{x} \in M_{\alpha}^{<\omega}$ has $\vec{x} \in M_{\alpha_n}^{<\omega}$ for some $n < \omega$ and therefore if there is a y where $\psi^M(\vec{x}, y)$, there is a y in $M_{\alpha_{n+1}}$. Therefore M_{α} and M satisfy (2) of Lemma 7 D • 4. Hence φ and all of its subformulas are absolute between M_{α} and M, and $\alpha > \beta$.

An alternative proof of The Reflection Principle $(7 D \cdot 5)$ can be given by more combinatorial means^{xxvii}, but this is not done here, since the relevant concepts will not be introduced until Chapter II. Note that The Reflection Principle $(7 D \cdot 5)$ is equivalent to the result holding for finitely many formulas φ , as we can just take the single formula which is the conjunction of the finitely many. In particular, we have the following.

- 7 D • 6. Corollary

ZFC is not finitely axiomatizable: there is no finite set of FOL(\in)-formulas T such that $T \vdash \varphi$ iff ZFC $\vdash \varphi$.

Proof .:.

For each model $\mathbf{M} \models \mathsf{ZFC}$, the hierarchy V_{α}^{M} witnesses that M is stratefied in M . If there were such a finite collection, the conjunction of these finitely many formulas is a formula φ . By The Reflection Principle (7 D • 5) in M , since $\mathsf{M} \models \varphi$, there is some $\mathsf{V}_{\alpha}^{\mathsf{M}} \models \varphi$ and therefore $\mathsf{V}_{\alpha}^{\mathsf{M}} \models \mathsf{ZFC}$. Consider the (according to M) least $\alpha \in \mathrm{Ord}^{\mathsf{M}}$ where $\mathsf{V}_{\alpha}^{\mathsf{M}} \models \varphi$. By the same argument above, by the absoluteness of rank and thus the V_{α} s between transitive submodels of M , there is some $\beta \in \mathsf{V}_{\alpha}^{\mathsf{M}}$ where $\mathsf{V}_{\beta}^{\mathsf{M}} \models \varphi$, contradicting the minimality of α in M .

The above corollary highlights an important idea regarding the relativity of transitivity. In principle, everything we've done thus far has been in an arbitrary model of ZFC, and so the notions of "transitive", "well-founded", and so forth are notions relative to this background model. In particular, for $M \models ZFC$,

- A model N is transitive in M iff $\operatorname{trcl}^{\mathsf{M}}(N) = N$.
- A relation R is well-founded in **M** iff there is no $m \in M$ with $\mathbf{M} \models "\forall x \in m \exists y \in m (y R x)"$.

As we've seen, these can differ between different models of set theory. But the same absoluteness results above hold; it's just that they are restricted to transitive models *of our given model* rather than the more philosophically based notion of V.

Corollary 7 D • 6 also highlights an important distinction between a theorem and a theorem scheme. The Reflection Principle (7 D • 5) is a theorem scheme in that for each φ , we get a different theorem. The Reflection Principle (7 D • 5) is not equivalent to any single formula by the same sort of reasoning as in Corollary 7 D • 6. That said, we're still effectively working in an arbitrary model of ZFC, and so a coded version of The Reflection Principle (7 D • 5) still holds in an arbitrary model of ZFC, it's just that the coded notion of "formula" etc. in a non-standard model may not agree with the actual universe, just like well-foundedness.

In particular, if **M** has $\langle \omega^{\mathsf{M}}, \in^{\mathsf{M}} \rangle \not\cong \langle \omega, \in \rangle$, then **M** will have an $n \in \omega^{\mathsf{M}}$ that **M** thinks is a coded formula, but doesn't correspond to any real-world formula, because *n* isn't even an actual natural number. For a more concrete example, ZFC \nvDash Con(ZFC) implies there are models of ZFC where \neg Con(ZFC). In such a model **M**, the coded proof that ZFC $\vdash "\varphi \land \neg \varphi"$ corresponds to one of these "natural numbers" of **M**, and not the code of an actual proof.

We can actually get a slightly stronger reflection theorem. Note that for $\kappa = \text{Ord}$ —and thus writing M_{κ} for M—this is the same as The Reflection Principle (7 D • 5).

– 7 D•7. Theorem (The Reflection Theorem) -

Let M be stratified. Let $\kappa > \aleph_0$ be a regular cardinal (allowing for $\kappa = \text{Ord}$). Let φ be a FOLp-formula. Therefore, there are arbitrarily large $\alpha < \kappa$ where φ is absolute between M_{κ} and M_{α} .

^{xxvii}In particular, each formula is absolute between M_{α} and M on a *club* of $\alpha \in$ Ord. Since the intersection of two clubs is a club, the propositional connectives are dealt with easily in the induction on formulas. The existential case " $\exists x \psi$ " can be dealt with by considering the map sending parameters to the least β with a witness in M_{β} . Taking the supremum of such β s and then closing the club for ψ under this yields another club that gets the job done, similar to the approach taken above, but avoiding Lemma 7 D • 4 at the cost of giving a background on clubs.

Proof .:.

Сн I §7 D

Proceed by induction on φ . Let $\beta \in$ Ord be arbitrary. As before, we only need to deal with the existential subformulas of φ . For each subformula of φ of the form " $\exists y \psi$ ", we will show there is an α with $\beta < \alpha < \kappa$ where

 $"\forall \vec{x} \in \mathcal{M}_{\kappa} \ (\exists y \in \mathcal{M}_{\kappa} \ \psi^{\mathcal{M}_{\kappa}}(\vec{x}, y) \to \exists y \in \mathcal{M}_{\alpha} \ \psi(\vec{x}, y))".$

For $\vec{x} \in M$, let $F_{\psi}(\vec{x})$ be the least ordinal $\alpha < \kappa$ such that $\exists y \in M_{\alpha} \psi^{M_{\kappa}}(\vec{x}, y)$ (if there is no such ordinal, set $\alpha = 0$). So F_{ψ} points to a level where there is a witness to ψ (if there is one). For $\alpha < \kappa$, consider

 $G(\alpha) = \sup\{F_{\psi}(\vec{x}) : \vec{x} \in \mathbf{M}_{\alpha}^{<\omega} \land \exists y \ \psi$ is one of the existential subformulas of $\varphi\}$.

As before, we just continually apply G and then union up to get a model closed under this.

Take α_0 to be arbitrary such that $\beta < \alpha_0 < \kappa$. Let $\alpha_{n+1} = \max(G(\alpha_n), \alpha_n + 1)$ and set $\alpha = \sup_{n < \omega} \alpha_n$. As $\operatorname{cof}(\kappa) > \omega$ and inductively each $\alpha_n < \kappa$, it follows that $\alpha < \kappa$. Consider $M_{\alpha} = \bigcup_{n < \omega} M_{\alpha_n} \subseteq M_{\kappa}$. As before, M_{α} and M_{κ} satisfy (2) of Lemma 7 D • 4. Hence φ and all of its subformulas are absolute between M_{α} and M_{κ} , and $\kappa > \alpha > \beta$.

The point of this will be to have the ability to take skolem hulls of the levels of a stratified model, and end up with smaller models of the same statements. With The Reflection Principle (7 D • 5), we can't take a skolem hull of M and expect it to be in the model of set theory, since M is a proper class, and not a set. But M_{κ} for $\kappa \in \text{Ord } is$ a set, and so we can take the skolem hull.

A useful result of The Reflection Principle $(7 D \cdot 5)$ with Taking a Skolem Hull $(6 A \cdot 2)$ gives the following.

— 7 D•8. Corollary

Let $\Delta \subseteq \mathsf{ZFC}$ be a finite subset. Therefore there are countable, transitive models of Δ . In fact, for any hereditarily countable $M \in \mathsf{H}_{\aleph_1}$, there is a countable, transitive model $\mathsf{M}' \vDash \Delta$ with $M \subseteq M' \in \mathsf{H}_{\aleph_1}$.

Proof .:.

Let $\operatorname{trcl}(M) \subseteq V_{\beta}$ for some β . Since $M \in H_{\aleph_1}$, $\operatorname{trcl}(M)$ is countable. By The Reflection Principle (7 D • 5), there is some $\alpha > \beta$ where $V_{\alpha} \models \Delta$ iff $V \models \Delta$, meaning $V_{\alpha} \models \Delta$. By Corollary 6 C • 2, cHull^{V_{α}} (trcl(M)) is transitive, countable, contains M, and satisfies Δ .

Section 8. The First Inner Models

We begin with the definition of an inner model. The general picture of an inner model is just a "skinny" version of the background model V, where the class of all ordinals constitutes the "backbone" of the universe.



8.1. Figure: An inner model

8•2. Definition

A class M ⊆ V is an *inner model* iff
M is transitive;
Ord ⊆ M; and
M = ⟨M, ∈⟩ ⊨ ZFC

If we replace ZFC in the last condition with some theory T, we say M is an inner model of T.

So clearly V is an inner model. Moreover, as we've defined things, for any model $W \models ZFC - Found$, $WF^W \subseteq W$ is an inner model of ZFC. Of course, in V, both of these are just V. So these examples are not particularly illuminating for us. The goal of this section is to introduce two more inner models: one of which is very rigid, and one of which is very flexible.

Note that being an inner model is a scheme, and not a singular formula. It's saying that $Ord \subseteq M$ is transitive (a single sentence) and that $\mathbf{M} \models \varphi$ for each $\varphi \in ZFC$ (infinitely many sentences).

§8A. The constructible universe and definability

Recall that the levels of V were defined by iteratively taking the powerset operation. Gödel's definition of the constructible universe, L, does the same, but restricts to subsets which are definable over the previous levels.

- 8A • 1. Definition Let $L = \bigcup_{\alpha < Ord} L_{\alpha}$ where $\alpha \mapsto L_{\alpha}$ is defined by transfinite recursion: $L_0 = \emptyset$, $L_{\gamma} = \bigcup_{\alpha < \gamma} L_{\alpha}$ for γ a limit, and $L_{\alpha+1} = \{x \in \mathcal{P}(L_{\alpha}) : x \text{ is definable over } \langle L_{\alpha}, \in \rangle\}.$ Here, x being definable over L_{α} means that there is a FOLp-formula $\varphi(y)$ where $L_{\alpha} \models "\varphi(y)$ " iff $y \in x$.

Keep in mind that L_{α} is usually not $V_{\alpha}^{L} = V_{\alpha} \cap L$: the levels L_{α} are formed in a fairly different way from the levels $V_{\alpha} \cap L$. It may be that to define a subset x of ω , we need to talk about some really large ordinals, and so while $x \in V_{\omega+1} \cap L$, we might not have $x \in L_{\omega+1}$. This is just a caution to not conflate the notation L_{α} with V_{α}^{L} .

The importance of L to set theory is hard to overstate. There are three main ideas why. Firstly, every transitive model of enough set theory has an interpretation of L, and this interpretation is the same across all transitive models of ZF - P.

Secondly, it's the *only* model with this property, demonstrating a strong minimality condition. In fact, it's defining formula is so rigid that any transitive model elementarily equivalent to one of the L_{α} levels is actually one of the L_{α} levels. Thirdly, as a result of all of this, L is always the smallest inner model. And thus it can be seen as the transitive model "generated" by the theory of ZFC. This is analogous to the situation with arithmetic, where \mathbb{N} is the smallest model of the peano axioms, PA, and so can be thought of as being generated by them.

To confirm all of this, we begin with showing that $L \models ZFC$. First we will show that L is stratified, which gives a great portion of set theory. Note that $L_{\alpha} \in L_{\alpha+1}$ as witnessed by the formula "x = x". Furthermore, because we're allowing parameters, $L_{\alpha} \subseteq L_{\beta}$ for $\alpha \leq \beta$. One might think the only thing needed to confirm that L is stratified is that each L_{α} is transitive. But in fact, we need to ensure that the function taking α to L_{α} is definable: that L is a class.

To do this, we need to understand how to formalize definability within set theory. We know from Theorem $6B \cdot 6$ that we can do a lot of the work based on this in ZF, since we can look at the full powerset, and then restrict to those subsets which have a first-order definition as per Theorem $6B \cdot 6$. But to help us later, it will be useful to work in ZF – P, which requires instead a reliance on the replacement axiom. So rather than rely on Theorem $6B \cdot 6$, we will instead think of closing a given set under operations corresponding to the logical operations.

BA•2. Definition Let A be a set. For $n, m, k \in \omega$, define

• Exists^{*n*}_{*A*}(*D*) = { $\tau \in A^n : \exists x \in A \ (\tau^x \in D)$ };

• Memb^{*n*,*m*,*k*} = { $\tau \in A^k : n, m \in \text{dom}(\tau) \land \tau(n) \in \tau(m)$ };

• Equal_A^{*n,m,k*} = { $\tau \in A^k : n, m \in \text{dom}(\tau) \land \tau(n) = \tau(m)$ }

Define FOL(A) to be the closure of $\{\text{Memb}_A^{n,m}, \text{Equal}_A^{n,m} : n, m \in \omega\}$ under Exists_A^n , intersections, and complements in A^n for $n < \omega$.

Note that $\text{Exists}_{A}^{n}(D)$ corresponds to the existential quantification while $\text{Memb}_{A}^{n,m,k}$ and $\text{Equal}_{A}^{n,m,k}$ correspond to membership and equality. Similarly, intersections correspond to conjunction. Complements correspond to negations. Hence starting with the atomic formulas and closing under existential quantification, conjunction, and relative complement, we get all of the first-order formulas, and this corresponds precisely to looking at their defined sets, closing under these operations. Hence FOL(A) corresponds to the FOL-defined subsets of $A^{<\omega}$.

Note that we can define FOL(A) by finitary recursion, just repeatedly applying the operations to the sets in the previous stage, starting with the first stage of {Memb}_A^{n,m}, Equal_A^{n,m} : n, m \in \omega}. Thus without powerset, FOL(A) exists.

- 8A•3. Definition Let A be a set. For $\sigma \in A^n$ where $n < \omega$, define FOL_{σ}(A) to be the closure of FOL(A) under the operations of Definition 8A•2 and the operation

$$\operatorname{Param}_{A}^{\sigma}(D) = \{ \tau \in A^{n} : \sigma^{\frown} \tau \in D \}.$$

Define FOLp(A) to be $\bigcup_{\sigma \in A^{\leq \omega}} FOL_{\sigma}(A)$, the set of all subsets of $A^{\leq \omega}$ that are FOLp-definable.^{xxviii}

As before, in ZF - P, FOLp(A) exists.

BA•4. Corollary The function $\alpha \mapsto L_{\alpha}$ is definable, and hence L is a class: $x \in L$ iff $\exists \alpha \in Ord(x \in L_{\alpha})$.

Proof .:.

Using Definition 8 A • 3 we can talk about which sets are FOLp-definable over L_{α} . So we can define recursively $x = L_{\alpha}$ iff there exists a function L with dom $(L) = \alpha + 1$ and $L(\alpha) = x$ such that • $L(0) = \emptyset$:

^{xxviii}In principle, we should be a bit careful arguing about FOLp(A) as the FOLp-definable subsets of A, since FOLp(A) is a formally defined concept that may not mesh with the real world notions. For example, in model with a non-standard ω , it's not immediately obvious that all "finite" subsets of A are in FOLp(A), even though all the *actual* finite subsets are. The proof that all the subsets of size $< \omega^{M}$ are in FOLp^(A) for any M is a bit tedious, going through the details of FOLp(A) being defined as a closure of certain operations. It is not particularly difficult, but it is not particularly enlightening, and doesn't serve to help understand the general ideas.

• for every $\beta < \alpha$, $L(\beta + 1) = \{y \in \text{FOLp}(L(\beta)) : y \subseteq L_{\beta}\}.$

• for every limit ordinal $\gamma \leq \alpha$, $L(\gamma) = \bigcup_{\beta < \gamma} L(\beta)$.

Now we can show that L is stratified, and hence get a large portion of ZFC by Result 7 D • 3

8A • 5. Lemma For each $\alpha \in \text{Ord}$, L_{α} is transitive, and hence L is stratified.

Proof .:.

Proceed by induction on α . For $\alpha = 0$, $L_0 = \emptyset$ is obviously transitive. For $\alpha + 1$, each $x \in L_{\alpha+1}$ is a subset of L_{α} . Clearly each $a \in L_{\alpha}$ is in $L_{\alpha+1}$ defined by the F0Lp-formula " $y \in a$ ": $a = \{y \in L_{\alpha} : y \in a\} \in FOLp(L_{\alpha})$. Hence $x \subseteq L_{\alpha} \subseteq L_{\alpha+1}$. The result for limits holds clearly by the inductive hypothesis. \dashv

Thus by Result 7 D • 3, $L \models wZF - Comp - Inf$. To confirm that $L \models Inf$, we will show $Ord \subseteq L$ and thus $\omega \in L$ showing Inf holds by Corollary 7 A • 5.

 $-8A \cdot 6. \quad \text{Lemma} \quad ---- \\ \text{Ord} \subseteq L \text{ so that } L \vDash \text{Inf.}$

Proof .:.

We show by induction that $\alpha \in L_{\alpha+1} \setminus L_{\alpha}$. For $\alpha = 0$, $L_0 = \emptyset$. Since $L_{\alpha} \in L_{\alpha+1}$ for each α , it follows that $0 \in L_{0+1} = L_1 \setminus L_0$. Inductively, $\alpha \in L_{\alpha+1} \setminus L_{\alpha}$. Therefore $\alpha + 1 = \alpha \cup \{\alpha\} \subseteq L_{\alpha+1}$ is F0Lp-definable over $L_{\alpha+1}$ by " $x < \alpha \lor x = \alpha$ ", showing $\alpha + 1 \in L_{\alpha+2}$. Since $\alpha \notin L_{\alpha+1}$, $\alpha + 1 \notin L_{\alpha+1}$ so that $\alpha + 1 \in L_{\alpha+2} \setminus L_{\alpha+1}$. This deals with the successor case.

For the limit case, the inductive hypothesis tells us that $\alpha \subseteq L_{\alpha}$ and that α is the least ordinal not in L_{α} ($\alpha \notin L_{\beta}$ for any $\beta < \alpha$ as otherwise this would imply by transitivity of L_{β} that $\beta \in L_{\beta}$). Therefore α is definable over L_{α} by "x is an ordinal". Hence $\alpha \in L_{\alpha+1}$. Hence each ordinal α is in $L_{\alpha+1} \subseteq L$ so that $Ord \subseteq L$.

Hence we only need to show that comprehension and choice hold in L. To do this, we use Corollary 7 A • 5.

- 8A.7. Theorem - $L \models Comp and therefore L is an inner model of ZF.$

Proof .:.

Let φ be arbitrary. We want to show that for each $A \in L$, $A_{\varphi} = \{x \in A : \varphi^{L}(x)\} \in L$. To see this, note that $A \in L_{\alpha}$ for some $\alpha \in Ord$. Note that there are arbitrarily large $\beta \in Ord$ where φ is absolute between L and L_{α} by The Reflection Principle (7 D • 5). In particular, there is some L_{β} where $A \in L_{\beta}$, and $\forall x (\varphi^{L_{\beta}}(x) \leftrightarrow \varphi^{L}(x))$. As a result, φ defines A_{φ} over L_{β} and thus $A_{\varphi} \in L_{\beta+1} \subseteq L$. Therefore, $L \models Comp$ so that $L \models wZF$. As wZF is equivalent to ZF, $L \models ZF$.

So all that remains is the axiom of choice. Note that all of the above work on L didn't use the axiom of choice. So if we were to start in a universe $W \models ZF + \neg AC$, we would still have $(L^W, \in^W) \models ZFC$. The basic idea behind the proof is to well-order all of the sets in L according to the formulas that defined the sets.^{xxix}

```
— 8A•8. Theorem —
L is an inner model of ZFC.
```

Proof .:.

It suffices to show $L \models AC$. For each $x \in L$, write α_x for the least α where $x \in L_{\alpha_x+1}$. Well-order the (codes of the) FOL(\in)-formulas with \leq_{lex} as in Definition $6B \cdot 2$. For each $x \in F$ let φ_x be the (code of the) \leq_{lex} -

Η

^{xxix}Again, formally, we would do this by ordering them by the lexicographically least sequence of operations that yield the element to be in L. We are merely thinking of this sequence of operations as a formula built up by the corresponding syntactic operations.

least formula which defines x over L_{α_x} for some parameters \vec{w}_x of L_{α_x} . Now define by recursion the order $<_{L_{\alpha}} \subseteq L_{\alpha} \times L_{\alpha}$ by taking $<_{L_0} = \emptyset$, and $<_{L_{\gamma}} = \bigcup_{\alpha < \gamma} <_{\alpha}$ for γ a limit, and $x <_{L_{\alpha+1}} y$ iff

1. $x, y \in L_{\alpha}$ and $x <_{L_{\alpha}} y$; or else

2. $\alpha_x < \alpha_y$; or else

3. " φ_x " <_{lex} " φ_y "; or else

4. $\vec{w}_x <_{\text{lex}} \vec{w}_y$ under the $<_{\text{L}_{\alpha}}$ -order.

It follows by induction and Lemma 6 B • 3 that each $<_{L_{\alpha}}$ is a well-order of L_{α} . In fact, $<_{L_{\alpha}} \in L_{\alpha+\omega}$, because the notions above are all easily definable.

So let $F \in L$ be a non-empty family of non-empty, disjoint sets. Note that $F \subseteq L_{\alpha}$ for some α . Consider

 $C = \{ y \in \bigcup F : \exists x \in F(y \text{ is the } <_{L_{\alpha}} \text{ -least element of } x) \}.$

It follows that C is a choice set for F, and is in L. Thus $L \models AC$.

\dashv

§8B. L as a canonical inner model

As stated before, L has many "canonicity" properties. In particular, it has a strong minimality condition, being contained (up to a given height) in any transitive model of ZF - P. As a result, it's the smallest inner model, and is determined by its theory. We state these three facts as follows. Firstly, we have the absoluteness of L, leading to L being the smallest inner model.

— 8 B • 1. Theorem (Absoluteness of L) –

For any transitive model $M \models ZF - P$, writing L_{Ord} for L;

1. For each $\alpha \in \text{Ord} \cap M$, $L_{\alpha} \subseteq M$.

2. $L^{\mathsf{M}} = L_{\operatorname{Ord} \cap \mathsf{M}}$.

In particular, $L^{L} = L$. In fact, if $Ord \subseteq M$, then all of L is contained in M.

- 8B•2. Corollary (Smallest Inner Model) - $L \subseteq M$ for any inner model M of ZF - P.

Next, since we can write "V = L" as a FOL(\in)-sentence, considering it as an axiom yields the following.

```
- 8B·3. Theorem (Condensation)
Suppose M \models ZF - P + "V = L" where M is transitive. Therefore M = L_{Ord \cap M}.
```

This theorem can be strengthened significantly, although we will prove stronger versions later. In particular, if $\mathbf{M} \preccurlyeq_{\Sigma_1} \mathbf{L}_{\alpha}$ for some $\alpha \in \text{Ord or } \alpha = \text{Ord}$, then $\mathbf{M} \cong \mathbf{L}_{\beta}$ for some $\beta \leq \alpha$. Here " \preccurlyeq_{Σ_1} " refers to being an elementary substructure with respect to Σ_1 -formulas.

To show the above results, we need to show the absoluteness of the construction of L. Firstly, note the following absoluteness result.

- 8B•4. Lemma - "y = FOLp(x)" is absolute between transitive models of ZF – P.

Proof .:.

This follows since the closure of a set under these operations is given by recursion. Given that each of the operations is clearly absolute, it follows that the output of this is absolute by Theorem 7 B \cdot 4.

- Proof of Absoluteness of $L(8B \cdot 1)$...

Proceed by induction on α to show $L_{\alpha}^{\mathsf{M}} = L_{\alpha}$ and thus $L_{\alpha} \subseteq \mathsf{M}$ for $\alpha \in \mathsf{Ord} \cap \mathsf{M}$. Clearly, for $\alpha = 0$, $L_{\alpha} = \emptyset \in \mathsf{M}$. Similarly, by the absoluteness of unions and the inductive hypothesis, for limit $\gamma \in \mathsf{Ord} \cap \mathsf{M}$,

$$\begin{split} L_{\gamma}^{\mathsf{M}} &= \bigcup_{\alpha < \gamma} L_{\alpha}^{\mathsf{M}} = \bigcup_{\alpha < \gamma} L_{\alpha} = L_{\gamma}. \text{ For the successor stage } \alpha + 1, \text{ Lemma 8 B • 4 tells us that } L_{\alpha+1}^{\mathsf{M}} = \text{FOLp}^{\mathsf{M}}(L_{\alpha}^{\mathsf{M}}) = \text{FOLp}(L_{\alpha}) = L_{\alpha+1}. \text{ Hence } L_{\alpha} \in M \text{ for each } \alpha \in \text{Ord} \cap M. \end{split}$$
We have $L^{\mathsf{M}} = \bigcup_{\alpha \in \text{Ord} \cap M} L_{\alpha}^{\mathsf{M}} = \bigcup_{\alpha \in \text{Ord} \cap M} L_{\alpha} = L_{\text{Ord} \cap M} \subseteq M.$ In particular, for M an inner model, $L^{\mathsf{M}} = L \subseteq M.$

This shows the first two canonical properties of L: Absoluteness of L (8 B • 1) and Smallest Inner Model (8 B • 2). To show the third, Condensation (8 B • 3), we first should note that the sentence "V = L" does indeed exist, being defined through FOLp: for every x there is an ordinal α such that some function L defined on $\alpha + 1$ has $x \in L(\alpha)$ and L satisfies the properties as laid out by Corollary 8 A • 4. More succinctly, " $\forall x \exists \alpha \in Ord (x \in L_{\alpha})$ ". The proof of condensation is easy given these first two properties.

- Proof of Condensation $(8B \cdot 3)$...

Since
$$M \vDash "V = L"$$
, $M = V^M = L^M = L_{M \cap Ord}$ by Absoluteness of L (8 B • 1). \dashv

Being such a minimal model allows us to say more about absoluteness.

- 8B.5. Theorem - Suppose φ is upward absolute between inner models of ZF – P. Suppose L $\vDash \varphi$. Therefore φ is absolute between inner models of ZF – P.

Proof .:.

Suppose $M \vDash ZF - P$. Therefore $L^M = L_{Ord \cap M} = L \subseteq M$ has $L \vDash \varphi$. By upward absoluteness, $M \vDash \varphi$.

If absoluteness is generally regarded as $\varphi \leftrightarrow \varphi^M$ being true for all appropriate **M**, the above tells us that this is equivalent to $\varphi^L \leftrightarrow \varphi^M$ for all appropriate **M** \models ZF – P.

§8C. Applications and properties of L

The importance of Condensation $(8 B \cdot 3)$ comes from its use with Taking a Skolem Hull $(6 A \cdot 2)$ in the form of Corollary $6 C \cdot 2$. In particular, since any skolem hull is elementarily equivalent to a level of L, when we collapse it, it becomes a level of L.

If we investigate further the levels of L, we get some quick examples of models of "V = L". Note that the levels of L, although defined similarly, develop differently to the levels of V. In particular, $V_{\alpha} \neq L_{\alpha}$ in general, even if we assume V = L. An easy example of this is that in ZFC, $\mathcal{P}(\omega) \subseteq V_{\omega+1}$, meaning $|V_{\omega+1}| \ge 2^{\aleph_0} > \aleph_0$. But $L_{\omega+1}$ has only countably many new elements, corresponding to the defining formulas and parameters.^{XXX} Hence $|L_{\omega+1}| = \aleph_0$. So the point is that subsets of ω don't appear in $L_{\omega+1}$. In particular, one should not make the mistake of thinking $V_{\alpha}^L = L_{\alpha}$. This is (almost always) false.

- 8C·1. Lemma
Let
$$\alpha \ge \omega$$
. Therefore $|L_{\alpha}| = |\alpha|$.

Proof .:.

Proceed by induction on α . For $\alpha = \omega$, this is clear as L_{ω} is the countable union of sets, each of which is countable by induction: $L_0 = \emptyset$ is clearly countable, and $L_{n+1} \subseteq \mathcal{O}(L_n)$ which is also finite for $n < \omega$. For $\alpha + 1$, $L_{\alpha+1}$ is the closure of L_{α} under countably many operations and is thus $|L_{\alpha+1}| \leq |L_{\alpha}| \cdot \aleph_0$. Since clearly $\omega \subseteq L_{\alpha}$ for $\alpha \geq \omega$ and $L_{\alpha} \subseteq L_{\alpha+1}$, it follows that the reverse inequality holds and in fact $|L_{\alpha+1}| = |L_{\alpha}| = |\alpha| = |\alpha + 1|$.

xxx More formally, it's the closure of L_{ω} under countably many operations, and hence adds only countably many elements.

For limit
$$\gamma$$
, $|L_{\gamma}| = \left| \bigcup_{\alpha < \gamma} L_{\alpha} \right| \le |\gamma| \cdot \sup_{\alpha < \gamma} |L_{\gamma}| = |\gamma| \cdot \sup_{\alpha < \gamma} |\alpha| = |\gamma|.$

This allows us to more precisely understand what the levels of L look like.

 $- 8C \cdot 2. \quad \text{Result} \quad - \\ \text{Let } \kappa > \aleph_0 \text{ be a regular cardinal. Therefore } \mathsf{L}_{\kappa} \vDash \mathsf{ZFC} - \mathsf{P} + "\mathsf{V} = \mathsf{L}".$

Proof .:.

Once we show $L_{\kappa} \models ZF - P$, by absoluteness, $L^{L_{\kappa}} = L_{\kappa} = V^{L_{\kappa}}$ so that $L_{\kappa} \models "V = L$ ". So let $x, y \in L_{\kappa}$ be arbitrary. Thus $x \in L_{\alpha}$ and $y \in L_{\beta}$ for some $\alpha, \beta < \kappa$.

- For Pair, assume without loss of generality that $\alpha < \beta$. Thus $x, y \in L_{\beta}$ and so $\{x, y\} \in L_{\beta+1}$.
- For Union, $\bigcup x \subseteq \operatorname{trcl}(x) \subseteq L_{\alpha}$ and thus $\bigcup x \in L_{\alpha+1}$ as it is easily definable.
- For Comp, use The Reflection Theorem (7 D 7). In particular, the same proof as Theorem 8 A 7 applies to show $L_{\kappa} \models$ Comp: for each φ , there are arbitrarily large $\gamma < \kappa$ (e.g. $\gamma > \alpha$ where $x \in L_{\alpha}$) where φ is absolute between L_{γ} and L_{κ} . Therefore in $L_{\gamma+1}$, the set defined by comprehension, $\{z \in x : \varphi^{L_{\kappa}}\} = \{z \in x : \varphi^{L_{\gamma}}\} \in L_{\gamma+1} \subseteq L_{\kappa}$.
- For wRep, suppose D ∈ L_κ, φ is a F0Lp-formula, and that L_κ ⊨ "∀x ∈ D ∃! y φ(x, y)". We need to find an R ∈ L_κ such that every x ∈ D has a y ∈ R with φ^{L_κ}(x, y). Say D ∈ L_α so that |D| ≤ |L_α| = |α| < κ. Thus the L-ranks of the image of D should be bounded in L_κ. Explicitly, consider the function f ∈ L defined by φ^{L_κ} on D. Note that as f : L_κ → L_κ, im f ⊆ L_κ. Let γ = sup{γ_x + 1 : f(x) ∈ L_{γx+1}}. Because κ is regular and |D| ≤ |L_α| = |α| < κ, we have that γ < κ. Hence im f ⊆ L_γ and thus L_γ witnesses the axiom of wRep for φ and D. Comprehension then gives Rep.
- For AC, the definition of $<_{L_{\kappa}}$ in Theorem 8 A 8 yields a definable well-order of all of L_{κ} . Hence for any non-empty family *F* of non-empty, disjoint sets in L_{κ} , $F \subseteq L_{\alpha}$ for some $\alpha < \kappa$ so that the $<_{L_{\alpha}}$ -least (i.e. the $<_{L_{\kappa}}$ -least) element of each $x \in F$ yields a choice set just as in Theorem 8 A 8.

– 8C•3. Corollary –

If $L \models "\kappa > \omega$ is a cardinal", then $L_{\kappa} \models ZFC - P + "V = L$ ".

Proof .:.

If $\mathsf{ZFC} \vdash ``\kappa > \aleph_0$ is regular $\rightarrow \varphi^{L_{\kappa}}$ " for each φ of $\mathsf{ZF} - \mathsf{P} + ``V = L$ ", then in particular, $\mathsf{ZFC} + ``V = L$ " proves this. But then ZFC proves each relativization to L, i.e. if $``(\kappa > \aleph_0 \text{ is regular})^{L}$ " then $(\varphi^{L_{\kappa}})^{L}$ which is just $\varphi^{L_{\kappa}} \dashv$

One of the more important corollaries of Result 8 C \cdot 2 and Condensation (8 B \cdot 3) is what happens when we take skolem hulls.

- 8C·4. Corollary

Let $\kappa > \aleph_0$ be a regular cardinal. Let $X \subseteq L_{\kappa}$. Therefore the collapsed skolem hull $\operatorname{cHull}^{L_{\kappa}}(X) = L_{\alpha}$ for some $\alpha < \kappa$. Moreover, if X is transitive, then $X \subseteq \operatorname{cHull}^{L_{\kappa}}(X)$.

Proof .:.

By The Mostowski Collapse $(4 \cdot 1)$, the collapsed hull models "V = L":

 $\operatorname{cHull}^{\mathsf{L}_{\kappa}}(X) \cong \operatorname{Hull}^{\mathsf{L}_{\kappa}}(X) \preccurlyeq \mathsf{L}_{\kappa} \vDash \operatorname{ZFC} - \mathsf{P} + "\mathsf{V} = \mathsf{L}".$

Hence by Condensation (8 B • 3), cHull^{L_k}(X) = L_{α} for some α . As Hull^{L_k}(X) \subseteq L_k, α can be calculated as $\alpha = \text{Ord} \cap \text{cHull}^{L_{\kappa}}(X) \leq \kappa$.

Thus it suffices to show $X \subseteq \text{cHull}^{L_{\kappa}}(X)$ when X is transitive. To do this, we show that the collapsing map fixes X. Let π : Hull^{L_{κ}} $(X) \rightarrow \text{cHull}^{L_{\kappa}}(X)$ be the collapsing isomorphism, defined inductively by $\pi(x) = {\pi(y) : y \in x}$. We show that $\pi(x) = x$ for each $x \in X$. Suppose not. Let $x \in X$ be the ϵ -least element of X where $\pi(x) \neq x$. Thus $\pi(x) = {\pi(y) : y \in x}$. As $x \subseteq X$, it follows by minimality that each $y \in x$ has $\pi(y) = y$

and hence $\pi(x) = \{y : y \in x\} = x$. \dashv

Now so far, we've been investigating and developing this theory for seemingly no reason. But an important application of this Corollary 8 C • 4 gives the relative consistency of lots of combinatorial principles. For now, we just show that the *generalized* continuum hypothesis (GCH) holds: $2^{\kappa} = \kappa^+$ for infinite cardinals κ . Recall from Cantor's Theorem (5 B • 13) (and Result 5 D • 6) that we only know $2^{\kappa} \ge \kappa^+$. From the method of forcing (which hasn't been introduced here), 2^{κ} can consistently be *any* cardinal of cofinality > κ . So L thinks 2^{κ} is as small as it can possibly be all of the time.

The general idea behind the proof is that all of the subsets of κ appear by stage $L_{\kappa+1}$. Recall that although $\kappa \in L_{\kappa+1}$, not every subset of κ in L may appear at stage $L_{\kappa+2}$, unlike V where $\kappa \in V_{\kappa+1}$ but $\mathcal{O}(\kappa) \in V_{\kappa+2}$.

- 8C•5. Theorem L \models GCH, meaning L \models " $\forall \kappa (|\kappa| = \kappa \ge \aleph_0 \rightarrow 2^{\kappa} = \kappa^+)$ ".

Proof .:.

Argue in a model of "V = L" to suppress so many superscripts of L. Let $\kappa \ge \aleph_0$ be a cardinal, and let $x \in \mathcal{P}(\kappa)$ be arbitrary so that $x \in L_{\alpha}$ for some $\alpha \in \text{Ord}$. Let θ be a regular cardinal larger than $\max(\kappa, \alpha)$ (for example, $\theta = \max(\kappa^+, |\alpha|^+)$ works, but we just need it to be regular and sufficiently large). Therefore $L_{\theta} \models \mathsf{ZF}-\mathsf{P}+\mathsf{"V}=\mathsf{L"}$.

Let $H = \text{cHull}^{L_{\theta}}(\{x\} \cup \kappa)$ so that $H = L_{\alpha}$ for some $\alpha < \theta$. As $|H| \le \aleph_0 \cdot |\kappa \cup \{x\}| = \kappa$, it follows by Lemma 8 C • 1 that $\alpha \le \kappa^+$. Note also that $\kappa \cup \{x\}$ is transitive, so that $\kappa \cup \{x\} \subseteq H$ by Corollary 8 C • 4. In particular, $x \in L_{\alpha} \subseteq L_{\kappa^+}$. As $x \in \mathcal{P}(\kappa)$ was arbitrary, $\mathcal{P}(\kappa) \subseteq L_{\kappa^+}$ and therefore $2^{\kappa} \le |L_{\kappa^+}| = \kappa^+$. By Cantor's Theorem (5 B • 13), $\kappa^+ \le 2^{\kappa}$ and thus we have equality.

Note that this shows there is no hope of proving the consistency of \neg GCH from ZFC with our current methods: trying to define an inner model with this true in it. Any attempts to define a class C by a formula φ to show ZFC \vdash ZFC^C $+ \neg$ GCH^C would also need to have ZFC $+ "V = L" \vdash ZFC^{C} + \neg$ GCH^C. But as the smallest inner model, any model $\mathbf{M} \models$ ZFC + "V = L" has by absoluteness of L,

$$\mathsf{L}^{\mathsf{M}} = \mathsf{L}^{\mathsf{C}^{\mathsf{M}}} \subseteq \mathsf{C}^{\mathsf{M}} \subseteq \mathsf{M} = \mathsf{L}^{\mathsf{M}}$$

and thus would have $C = L^M \models \neg GCH$, a contradiction.

The regularity property of GCH in L is also manifested in another regularity property in the levels of L, as suggested by the fact that both H_{κ}^{L} and L_{κ} model ZFC – P.

- 8 C • 6. Result Let $\kappa > \aleph_0$ be a regular cardinal. Therefore $L_{\kappa} = H_{\kappa}^{L}$.

Proof .:.

Suppose $x \in L_{\kappa}$. It suffices to show that $\mathbf{L} \models ``|\operatorname{trcl}(x)| < \kappa$ ''. As a limit ordinal, $x \in L_{\alpha}$ for some $\alpha < \kappa$ and thus $\operatorname{trcl}(x) \subseteq L_{\alpha}$ by transitivity. Without loss of generality, we can assume $\alpha \ge \omega$ so by Lemma 8 C • 1, $|\operatorname{trcl}(x)| \le |L_{\alpha}| = |\alpha| < \kappa$ and therefore $x \in H^{\mathsf{L}}_{\kappa}$ so that $L_{\kappa} \subseteq H^{\mathsf{L}}_{\kappa}$.

So now we consider $x \in H_{\kappa}^{\mathsf{L}}$. Note that $\operatorname{trcl}(x) \cup \{x\} \in \mathsf{L}_{\alpha}$ for some $\alpha \in \operatorname{Ord}$. Let $\theta > \alpha$ be a regular cardinal (which is then regular in L). Consider the skolem hull $H = \operatorname{cHull}^{\mathsf{L}_{\theta}}(\operatorname{trcl}(x) \cup \{x\})$ which then has size $|H| \leq \aleph_0 \cdot |\operatorname{trcl}(x) \cup \{x\}| < \kappa$. By Condensation (8 B • 3), $H = \mathsf{L}_{\beta}$ for some β . In fact, since $\beta = \operatorname{Ord} \cap H$ and $|H| < \kappa$, $\beta < \kappa$. Moreover, Corollary 8 C • 4 implies that, as a transitive set, $\operatorname{trcl}(x) \cup \{x\} \subseteq H$ is left uncollapsed. Hence $\operatorname{trcl}(x) \cup \{x\} \in \mathsf{L}_{\beta} \subseteq \mathsf{L}_{\kappa}$. In particular, $x \in \mathsf{L}_{\kappa}$ and thus $\mathsf{H}_{\kappa}^{\mathsf{L}} = \mathsf{L}_{\kappa}$.

In combination with Corollary 7 C • 10, this shows that weakly inaccessible cardinals yield set models of ZFC.

- 8C•7. Theorem - Let κ be weakly inaccessible. Therefore $L_{\kappa} \vDash$ ZFC.

 \neg

Proof .:.

Any weakly inaccessible cardinal is regular so that $L_{\kappa} = H_{\kappa}^{L}$. By GCH in L, κ being weakly inaccessible is the same as κ being strongly inaccessible. Therefore, $L \models "\varphi^{H_{\kappa}}"$ and so $\varphi^{L_{\kappa}}$ for each φ of ZFC.

- 8C•8. Corollary

If κ is (strongly or weakly) inaccessible, then there is a countable, transitive model of ZFC.

Proof .:.

 L_{κ} is a model of ZFC so that cHull^{L_{κ}} (\emptyset) \models ZFC and is countable.

Note that The Reflection Principle (7 D • 5) tells us that every finite fragment of ZFC has a countable transitive model, but it's not provable in ZFC that there is a countable, transitive model of the entirety of ZFC, as this would imply $ZFC \vdash Con(ZFC)$, contradicting Gödel's incompleteness theorems.

As a final note about L for this section, the ordering $<_{L} = \bigcup_{\alpha \in Ord} <_{L_{\alpha}}$ described in Theorem 8 A • 8 is definable, and is a well-order of all of L. So L \models AC just because there of a much stronger principle holding: the existence of a *global* well-order. This is stronger than mere AC, which says there is a well-order of each individual set, but perhaps not of the entire universe, a proper class. Next, we will investigate what happens when there is a definable global well-order in general. This yields another inner model HOD. Of course, it's consistent that "HOD = L" holds, since any definable inner model contained L must be L itself by Absoluteness of L (8 B • 1): L = L^{HODL} \subseteq HOD^L \subseteq L.

§ 8 D. Hereditarily ordinal definable sets

Whereas L is a very rigid inner model by Condensation (8 B • 3), the next inner model we will consider will be very flexible. It is so flexible, in fact, that V = HOD might be false in HOD, which is to say HOD^{HOD} might not be HOD. Consistently, for any countable, transitive $W \models ZFC$, there is a U with W an inner model of U such that $U \models "W = HOD"$. So no matter where we start, it's consistent we're starting from HOD of a larger model. We begin—as with L—with the closure under definability.

Recall that we needed the clumsy closure definition of definability from Definition 8 A • 2 because we wanted to work with models of ZF - P to ensure the absoluteness of L, and in particular to get Corollary 8 C • 4. For HOD, we have no such interest, because HOD is so flexible, even under full ZFC. So we will use P with the ostensibly more complicated formula from Theorem 6 B • 6, defining what it means to have $A \models \varphi$ for A a set.^{xxxi}

— 8D•1. Definition -

A set x is *ordinal definable* or OD iff there is an $\alpha \in \text{Ord such that for some (coded) formula } \varphi \in (\omega \sqcup \{\in\})^{<\omega}$ and parameters $\vec{\beta} \in \text{Ord} \cap V_{\alpha}, V_{\alpha} \models \forall y \ (\varphi(\vec{\beta}, y) \leftrightarrow x = y)$ ".

The reason for taking a level of V is just to make the definition more concrete: using The Reflection Principle (7 D • 5), x is ordinal definable iff there is a formula with only ordinal parameters such that x is defined by this formula.^{xxxii} And of course, every set which is ordinal definable over V_{α} (using parameters $\vec{\beta}$) is ordinal definable over V using the ordinals $\vec{\beta}$ and now with α .

— 8 D•2. Corollary -

For $\varphi(\vec{w}, x)$ a FOL(\in)-formula, ZF $\vdash \forall \vec{\alpha} \in \text{Ord } \forall x \ (\forall y \ (x = y \leftrightarrow \varphi(\vec{\alpha}, y)) \rightarrow x \in \text{OD}).$

^{xxxi}Note there that for finite signatures, φ can be regarded as a sequence of natural numbers where each number corresponds to a symbol. This makes the formal definition of satisfaction and definability more complicated, but dramatically simplifies the presentation. In particular, when quantifying over "formulas" of any given model, we're quantifying over elements of ω where the sequence $\langle n_0, \dots, n_m \rangle \in \omega^{<\omega}$ is encoded by the number $2^{n_0+1}3^{n_0+1} \cdots p_m^{n_m+1} \in \omega$ where p_m is the *m*th prime number of ω starting with $p_0 = 2$.

^{xxxii}Of course, as with other issues about definability, we should be slightly careful about this. If a model of ZFC misinterprets ω , then it will have formulas that are not actual (coded) formulas. The issue is, as before, trying to identify the real world formulas with the formulas of the model. All real world formulas yield formulas of the model, but there may be formulas of the model that are not actual formulas.

 \neg

Note further some immediate consequences of this definition.

- Ord \subseteq OD;
- $x, y \in OD$ implies $\{x, y\} \in OD$;
- $x \in OD$ implies $\mathcal{P}(x) \in OD$.
- $V_{\alpha} \in OD$ for each $\alpha \in Ord$, as it is definable from α .

Now, unfortunately, OD might not be transitive (otherwise $V = \bigcup_{\alpha \in Ord} V_{\alpha} \subseteq OD \subseteq V$ implies that every set is ordinal definable). To counteract this issue, we bring in the concept of begin *hereditarily* ordinal definable. Then we can confirm the other axioms of ZFC, as the resulting class will be transitive, allowing us to use results from Section 7.

- 8 D · 3. Definition The class HOD of *hereditarily ordinal definable sets* consists of all sets x such that $x \in OD$ and $trcl(x) \subseteq OD$.

One consequence of this is that if $x \subseteq \text{HOD}$ and $x \in \text{OD}$ then $x \in \text{HOD}$.

— 8D•4. Corollary — HOD is a transitive class.

Note how this reflects the idea of "hereditarily" as in Definition $7 C \cdot 4$. There are two basic ideas about HOD that we care about:

- 1. **HOD** \models ZFC; and
- 2. M = HOD iff there is a definable, global well-order of M over M.

First we define what a global well-order is.

— 8 D • 5. Definition

A *class well-ordering (definable in* **V**) of a class C is a class $\leq \subseteq C \times C$ defined over **V** by some FOLp(\in)-formula φ such that every non-empty $X \in C$ has a \leq -least element.

A *definable*, global well-order of V is a class well-order (definable in V) of V.

Clearly if a class C has a well-order (definable in V) \preccurlyeq , then V can *uniformly* get choice sets for sets in C just by considering the \preccurlyeq -least elements in whatever non-empty family of non-empty, disjoint sets. Note that HOD has such a well-order defined according to the defining formula and ordinal parameters used to define its members.

– 8 D•6. Lemma

There is a (definable in V) class well-ordering of OD. Moreover, all initial segments of this well-ordering are sets.

Proof .:.

Consider the *Gödel ordering* of $<_{\text{Ord}}<_{\omega}$ on finite sequences of ordinals where $\langle \alpha_1, \cdots, \alpha_n \rangle <_{\text{Ord}}<_{\omega} \langle \beta_1, \cdots, \beta_m \rangle$ for $n, m \in \omega$ iff

• $\max(\vec{\alpha}) < \max(\vec{\beta})$; or else

- *n* < *m*; or else
- $\vec{\alpha} <_{\text{lex}} \vec{\beta}$.

It should be clear that $<_{\text{Ord}}<\omega$ is well-founded, and seeing that it's linear isn't difficult. Hence $<_{\text{Ord}}<\omega$ is a well-order. It should also be clear that for any particular $\vec{\alpha}$, all $<_{\text{Ord}}<\omega$ -predecessors of $\vec{\alpha}$ are contained in $(\max(\vec{\alpha}) + 1)^{<\omega}$, and thus the initial segments of $<_{\text{Ord}}<\omega$ are sets.

As a result, define the ordering \leq_{OD} on OD by taking $x \leq_{OD} y$ iff

- the least α where x is OD in V_{α} is less than the least for y; or else
- the $<_{Ord} < \omega$ -least set of parameters used to define x in this $V_{\alpha} <_{Ord} < \omega$ -precedes those for y; or else
- the formula φ used to define x with those parameters in $V_{\alpha} <_{lex}$ -precedes that formula for y.

It should be clear that this yields a well-ordering of OD.

As HOD \subseteq OD, we have the following.

— 8 D • 7. Corollary

HOD has a well-order definable over V and thus $HOD \models AC$.

Proof .:.

Using $<_{\text{OD}}$, if $F \in \text{HOD}$ is a non-empty family of non-empty, disjoint sets, then let F be defined by ψ in V_{α} with ordinal parameters $\vec{\beta}$. In V, we can then consider the $<_{\text{OD}}$ -least elements of $x \in F : z \in C$ iff $\exists x \exists F' (x \in F' \land \psi^{V_{\alpha}}(\vec{\beta}, F') \land z \in x$ is $<_{\text{OD}}$ -least). Note that then $C \in \text{HOD}$, which shows that $\text{HOD} \models \text{AC}$. \dashv

Now we can confirm HOD \models ZFC. Note that HOD is, of course, non-empty as Ord \subseteq HOD. The majority of the proof of this comes down to coming up with a definition to show that whatever set we're interested in x has $x \in$ OD and $x \subseteq$ HOD, implying that $x \in$ HOD because trcl $(x \cup \{x\}) \in$ OD.

- 8D \cdot 8. Theorem - HOD is an inner model.

Proof .:.

Let $x, y \in \text{HOD}$ as witnessed by ψ_x and ψ_y with parameters $\vec{\alpha}_x$ and $\vec{\alpha}_y$ in V_{γ_x} and V_{γ_y} respectively: $x \in V_{\gamma_x}$ is such that $V_{\gamma_x} \models \forall z \ (\varphi(\vec{\alpha}_x, z) \leftrightarrow x = z))$, and similarly for y.

- As usual, extensionality, empty set, and foundation follow from HOD being a (non-empty) transitive class.
- For Pair, note that $\{x, y\}$ is OD defined by $z \in \{x, y\}$ iff $\mathbf{V}_{\gamma_x} \models "\psi_x(\vec{\alpha}_x, z)" \lor \mathbf{V}_{\gamma_y} \models "\psi_y(\vec{\alpha}_y, z)"$. This formula has parameters $\vec{\alpha}_x$, $\vec{\alpha}_y$, γ_x , and γ_y so that $\{x, y\} \in \text{OD}$ and therefore $\{x, y\} \in \text{HOD}$ as $\operatorname{trcl}(\{x, y\}) = \operatorname{trcl}(x) \cup \operatorname{trcl}(y) \cup \{x, y\} \subseteq \text{HOD}$.
- For Union, the union $\bigcup x$ is defined by $z \in \bigcup x$ iff $\mathbf{V}_{\gamma_x} \models \exists y \exists x'(\psi_x(\vec{\alpha}_x, x') \land y \in x' \land z \in y))$. Thus $\bigcup x$ is OD. As $\operatorname{trcl}(\bigcup x) \subseteq \operatorname{trcl}(x) \subseteq \operatorname{HOD}$, it follows that $\bigcup x \in \operatorname{HOD}$.
- For Comp, let $\varphi(\vec{w}, z)$ be arbitrary. We want to show $x_{\varphi} = \{z \in x : \varphi^{\text{HOD}}(\vec{t}, x, z)\} \in \text{HOD}$ for each $\vec{t} \in \text{HOD}^{<\omega}$. Note that x_{φ} can be defined in **V** by

 $v = x_{\varphi} \quad \text{iff} \quad \forall z \left(z \in v \leftrightarrow \exists \vec{w} \exists x' \left(\vec{w} = \vec{t} \land \mathbf{V}_{\gamma_x} \vDash "\psi(\vec{\alpha}, x') " \land \varphi^{\text{HOD}}(\vec{w}, x', z) \right) \right),$

where by $\vec{w} = \vec{t}$ we really mean the conjunction of " w_i is the unique element satisfying the defining formula for t_i with the corresponding parameters in the corresponding level of V" for each i. As a result, x_{φ} is OD and so $\text{trcl}(x_{\varphi}) \subseteq \text{trcl}(x) \subseteq \text{HOD}$ implies $x_{\varphi} \in \text{HOD}$.

For P and Rep, first note that $V_{\gamma_x+1} \cap HOD \in HOD$. To see this, because $V_{\gamma_x+1} \cap HOD \subseteq HOD$, it suffices to show $V_{\gamma_x+1} \cap HOD \in OD$. Every element z of $V_{\gamma_x+1} \cap HOD$ is such that $trcl(\{z\}) \subseteq OD$. So let β be sufficiently large such that every element of V_{γ_x+1} is ordinal definable over V_{β} . Thus for all $z, z \in V_{\gamma_x+1} \cap HOD$ iff

$$x \subseteq V_{\gamma_x} \land \forall s \in \operatorname{trcl}(\{z\}) \exists \varphi \exists \vec{w} \in \operatorname{Ord}^{<\omega} \forall t \ (t \in s \leftrightarrow \mathsf{V}_{\beta} \models "\varphi(t, \vec{w})")$$

In other words, $z \in V_{\gamma_x+1} \cap HOD$ iff $z \subseteq V_{\gamma_x}$ and every $s \in trcl(\{z\})$ is ordinal definable over V_{β} . The only parameters in the above definition are $\gamma_x, \beta \in Ord$, and so $V_{\gamma_x+1} \cap HOD \in OD$ and hence in HOD.

- For P, note that $\mathcal{P}(x) \cap \text{HOD} \subseteq V_{\gamma_x+1} \cap \text{HOD} \in \text{HOD}$. This implies wP. By Comp, P holds.
- For Rep, let ψ define in HOD a function over x ∈ HOD. Let α ∈ Ord be such V_α contains the range of the function defined by ψ^{HOD} over x. Since V_α ∩ HOD ∈ HOD, this witnesses wRep. Therefore by Comp, Rep holds.

Therefore, as L is the smallest inner model, we have that L is its own HOD.

— 8 D • 9. Corollary ——

 $L = HOD^{L}$ and thus $L \models "L = HOD"$.

The issue with asking questions of HOD in general is due to its highly non-constructive nature: a subset of something (relatively) small like ω might need parameters in V_{α} for extremely large α to be defined, for example. In this sense, HOD requires knowing about all of the sets of V. And this is the general idea why HOD^{HOD} might not be HOD. More precisely, because the $V_{\alpha}s$ are not absolute between inner models, being ordinal definable is not absolute.

So far we have shown HOD \models ZFC, and V = HOD implies V has a definable, global well-order, because V = HOD is clearly equivalent to V = OD, which provably has a definable, global well-order. Our final goal for this section is then to show the converse: if V has a definable, global well-order, then V = HOD.

```
- 8D·10. Theorem
```

Let φ be a FOL(\in)-formula. Suppose \preccurlyeq , defined by $x \preccurlyeq y$ iff $\varphi(x, y)$, is a global well-order of V. Therefore V = HOD.

Proof .:.

By Lemma 4 • 3, there is a FOL-definable (class) function $f : V \to Ord$ where f(x) is the rank of x in $\langle V, \preccurlyeq \rangle$. But then x is definable from the ordinal f(x). In particular, if f(x) = y is defined by $\psi(x, y)$, then z = x iff $\psi(z, \alpha)$ for $\alpha = f(x)$. Hence V = OD and therefore $trcl(x \cup \{x\}) \subseteq OD$ automatically. Thus V = HOD. \dashv

- $8D \cdot 11$. Corollary - V = HOD iff there is a FOLp-definable, global well-order.

Proof .:.

One direction was just proven in Theorem 8 D \cdot 10. If V = HOD, then since V has a class well order of HOD = V definable in V, there is a global well-order.

So how is it possible for HOD^{HOD} to not be HOD? Ostensibly, HOD has a definable, global well-order by virtue of the one from V and thus $HOD \models "V = HOD"$, i.e. $HOD = HOD^{HOD}$. The issue is that this definable well-order is not absolute, because it depends on the levels of V rather than the levels of HOD. Again, because HOD^{HOD} has lost the information about the other sets in V, we don't know that the definition for the global well-order still yields a global well-order. We only have $HOD = HOD^{HOD}$ if this (or some other) global well order is definable over HOD, not V.

Section 9. Variants of the Axioms

There are many propositions ostensibly stronger than other axioms, but turn out to be equivalent to the rest of ZFC. We detail some of these proposals, as well as some variant axioms that are actual weakenings of other axioms.

§9A. Equivalents of the axiom of choice

Recall the official definition of AC.

- 9A • 1. Definition (Axiom) (AC) for any family of non-empty family of disjoint sets F, there is a set C which has chosen one element from each $z \in F$:

 $\forall F \ (\emptyset \notin F \land \forall x, y \in F \ (x \cap y = \emptyset) \to \exists C \ \forall x \in F \ \exists ! y \ (y \in x \cap C).$

We will give three equivalent (over ZF) formulations of AC. Recall that a *chain* for a poset (A, \preccurlyeq) is just a \preccurlyeq -linearly order subset of A.

— 9A•2. Definition (Axiom) –

(*Zorn's Lemma*, AC_{Zorn}) For every (non-empty) poset $\langle A, \preccurlyeq \rangle$, if every chain is bounded in A, then A has a \preccurlyeq -maximal element.

 (AC_{Prod}) If F is a non-empty set of non-empty sets, then $\prod_{x \in F} x$ is non-empty.

(AC_{Card}) Every set is bijective with an ordinal.

 $(AC_{WellOrd})$ Every set has a well-order.

Note that by Lemma 5 C • 1, $AC_{WellOrd}$ is equivalent to AC over models of ZF. We will use this extensively to show the equivalences $AC \leftrightarrow AC_{Zorn} \leftrightarrow AC_{Card} \leftrightarrow AC_{WellOrd}$. First we have Zorn's lemma.

− 9A•3. Theorem – ZF \vdash "AC \leftrightarrow AC_{Zorn}".

Proof .:.

• (AC_{Zorn} \rightarrow AC) This implication holds in BST. For F a non-empty family of non-empty, disjoint sets, by comprehension, union, and powerset, consider the set

 $C = \left\{ y \in \mathcal{O}\left(\bigcup F\right) : \forall x \in F \ (y \cap x = \emptyset \lor \exists ! z \in x \ (z \in y)) \right\}$

of approximations to a choice set for F. Note that the poset $\langle C, \subseteq \rangle$ exists by the existence of cartesian products. Note that for any chain $c \subseteq C$, $\bigcup c$ yields another chain so that $\bigcup c \in C$ and c is bounded by $\bigcup c$. Therefore if AC_{Zorn} holds, then there is a \subseteq -maximal element X of C. But then for any $x \in F$, there must be some $z \in x$ with $z \in X$, as otherwise (by pairing and union) $X \cup \{z\}$ would contradict maximality. Therefore X must be a choice set.

- $(AC_{WellOrd} \rightarrow AC_{Zorn})$ Let $\langle A, \preccurlyeq \rangle$ be a (non-empty) poset such that every chain is bounded in A. By $AC_{WellOrd}$, let $<_A$ be a well-order of A. Define by transfinite recursion an injective function f from ordinals into A.
 - Let f(0) be the $<_A$ -least element of A.
 - Let $f(\alpha + 1)$ to be the \prec_A -least element of $\{a \in A : f(\alpha) \prec_A a\}$.
 - For limit γ , note that $f'' \gamma \subseteq A$ is a chain. Hence we can take $f(\gamma)$ to be the $<_A$ -least element that bounds $f'' \gamma$.

If A has no \leq -maximal element, then $f(\alpha)$ can be defined for all α . Note, however, that f is injective by transitivity. This contradicts replacement given that A is a set.

Arguably the easiest equivalence to prove is that AC \leftrightarrow AC_{Prod}. Note that any $f \in \prod_{x \in F} x$ is referred to as a *choice* function in that $f(x) \in x$ for each $x \in F$.

9A·4. Theorem BST \vdash "AC_{Prod} \rightarrow AC" and with P or Rep, "AC \rightarrow AC_{Prod}".

Proof .:.

Suppose AC_{Prod} holds. Let F be a non-empty set of disjoint, non-empty sets. Let $f \in \prod_{x \in F} x$ be a choice function. Therefore im f is a choice set so that AC holds.

Suppose AC holds. Let *F* be an aribitrary non-empty set of non-empty sets. Consider $F' = \{\{x\} \times x : x \in F\}$ which exists by the existence of cartesian products, and either replacement or powerset with comprehension. Note that F' is a non-empty family of now disjoint (by considering the first component), non-empty sets. Therefore a choice set *C* exists by AC. Now every element of *C* is of the form $\langle x, a \rangle$ for $x \in F$ and $a \in x$. Moreover, for each $x \in F$ there is exactly one *a* such that $\langle x, a \rangle \in C$. Hence $C \in \prod_{x \in F} x$ is a choice function.

The last equivalence is that $AC \leftrightarrow AC_{Card}$. We have already proven that $ZF \vdash "AC \rightarrow AC_{Card}$ " with Theorem 5B•5. So it suffices to show $AC_{Card} \rightarrow AC_{WellOrd}$.

Proof .:.

Let X be arbitrary. By AC_{Card} , there is a bijection $f : X \to \alpha$ for some ordinal $\alpha \in Ord$. Therefore $\{\langle x, y \rangle : f(x) \in f(y)\} \subseteq X \times X$ is a well-order of X.

Therefore, over ZF, AC_{Card}, AC_{WellOrd}, AC_{Zorn}, AC_{Prod}, and AC are all equivalent.

§9B. Weakenings of the axiom of choice

There are various weakenings of AC that suppose choice only holds in certain contexts. These variant axioms are often useful in models of determinacy, which is incompatible with full choice, but sill compatible with weakenings. In the end, we will have the following implications with these principles defined throughout the subsection:

Global Choice, aka "V = HOD" \rightarrow AC \rightarrow DC \rightarrow Countable Choice \rightarrow Finite Choice,

with $ZF \vdash$ Finite Choice, but $ZF \nvDash$ Countable Choice, and all of the above implications are strict for ZF.

If AC is saying that we can make choices infinitely often, there are then variations of this that weaken how many choices we can make, or when we can make them. In particular, the first way we can weaken AC is by restricting outselves to making just finitely many choices.

```
— 9B•1. Definition — The axiom of finite choice says that for any finite family of non-empty, disjoint sets, there is a choice set for the family.
```

This weakening is so weak, that it is provable.

```
- 9B•2. Theorem ZF \vdash finite choice.
```

Proof .:.

Suppose *F* is a finite family of non-empty, disjoint sets, meaning there is a bijection $f : |F| \to F$ where $|F| < \omega$. Proceed by induction on |F| to show that there is a choice function for *F*. For |F| = 1, this is easy: $F = \{x\}$ for some $x \neq \emptyset$. As $x \neq \emptyset$, by basic properties of first-order logic, there is some $y \in x$ where then $\{y\}$ is a choice VARIANTS OF THE AXIOMS

For |F| = n + 1, let $x \in F$ be arbitrary. Inductively, $F \setminus \{x\}$ has a choice set C. Now take $y \in x$ and let $C' = C \cup \{y\}$ to get a choice set for F. By induction, all such F have a choice set. \dashv

So really finite choice is just a consequence of applying existential instantiation from FOL-proofs finitely many times. The reason AC is needed, is because proofs are finite, and we can't apply existential instantiation infinitely many times with a finite proof.

The second weakening of AC is that we can make *countably* many choices, which is then clearly stronger than finite choice, and is in fact independent of ZF: ZF doesn't prove it, nor disprove it.

- 9B•3. Definition The axiom of *countable choice* says that for any non-empty, countable family (meaning there is an injection from it into ω) of disjoint, non-empty sets, there is a choice set for the family.

One important consequence of countable choice is Kőnig's theorem on trees, which requires some version of choice. First we introduce the concept of a tree, which is incredibly important in set theory in that it is a slight generalization of ordinals.

- 9B•4. Definition

A *tree* is a poset (A, \leq) such that for every $a \in A$, $\operatorname{pred}_{\leq}(a)$ is well-ordered by \leq .

A tree is *finitely splitting* iff there are finitely many least elements, and for every node $a \in A$ there are at most finitely many direct \leq -successors to a.

A *branch* is a \subseteq -maximal, \leq -linearly ordered subset of *A*.

In particular, if a tree has height $n < \omega$, then the tree is finite and so finite choice yields a branch with height n. But is there an infinite branch if the tree is infinite (of height ω)? AC and König's theorem in particular state that this is true.

– 9 B • 5. Theorem (Kőnig's Lemma on Trees)

Let $\mathbf{T} = \langle T, \leq_T \rangle$ be an finitely splitting tree of height ω . Therefore

ZF + "countable choice" \vdash "there is an infinite branch of T".

Proof .:.

First we will show that T must be countable.

– Claim 1 ––––––

There is an injection $f: T \to \omega$.

Proof .:.

It should be clear by induction that for each $n < \omega$, the *n*th level of **T** is finite. For each $n < \omega$, consider the set B_n of bijections from $|lvl_n(\mathsf{T})| < \omega$ to $|vl_n(\mathsf{T})$ (which is a set because it's a subset of ${}^{<\omega}lvl_n(\mathsf{T})$). Note that the family $\{B_n : n < \omega\}$ of these bijections is then countable, nonempty, and each B_n is also non-empty and disjoint as the levels are disjoint. So by countable choice, there is a choice *C* set for this family.

So let $f_n \in C$ be $f_n : \operatorname{lvl}_n(\mathsf{T}) \to |\operatorname{lvl}_n(\mathsf{T})|$ for each $n < \omega$. Let $f : T \to \omega$ be defined by $f(\tau) = 2^n \cdot 3^{f_n(\tau)}$, for $\tau \in \operatorname{lvl}_n(\mathsf{T})$. Since each element is only in one level, and for each level $n < \omega$, f_n is a bijection, it follows that f is an injection.

For each $\tau \in T$, let S_{τ} consist of all $\sigma \in T$ such that

- σ is a direct successor to τ (i.e. $\tau <_T \sigma$ and there are no $\rho \in T$ with $\tau <_T \rho <_T \sigma$); and
- there are infinitely many \leq_T -successors to σ (i.e. the set of $\rho \in T$ with $\sigma \leq_T \rho$ is infinite).

Note that S_{τ} might be empty for some τ . But clearly, for some $\tau_0 \in T$ being \leq_T -minimal, S_{τ_0} is non-empty (otherwise T will be the union of finitely many finite sets and thus be finite). Therefore $F = \{S_{\tau} : \tau \in T \land S_{\tau} \neq \emptyset\}$ is a countable, non-empty family of non-empty, disjoint sets. Thus by countable choice, there is a choice set C for F.

Now we proceed by recursion to give an infinite path. Let τ_0 be a \leq_T -least element of T with infinitely many successors. By the same reasoning as above one of the direct successors also must have infinitely many successors: $S_{\tau_0} \neq \emptyset$ (meaning $S_{\tau_0} \in F$). For τ_n already defined with $S_{\tau_n} \in F$, take τ_{n+1} to be the unique element of $C \cap S_{\tau_n}$. By the same reasoning before, we must have $S_{\tau_{n+1}} \in F$. Therefore the sequence defined by recursion, $\langle \tau_n : n \in \omega \rangle$, yields an infinite branch of T.

Let's consider an alternative way to state countable choice. This generalizes to the axiom of dependent choice.

- 9B•6. Definition The axiom of *dependent choice* (DC) says that for $R \subseteq X \times X$, if $\forall x \in X \exists y \in X (x R y)$ then there is a sequence $\langle x_n : n \in \omega \rangle$ such that $x_n R x_{n+1}$ for all $n \in \omega$.

Just from this definition, it's not immediately clear that DC implies countable choice, but we can show this without much effort.

- 9B•7. Theorem \longrightarrow ZF \vdash "DC \rightarrow countable choice".

Proof .:.

Assume DC and suppose F is a non-empty, countable set of non-empty, disjoint sets, as witnessed by an injection $f: F \to \omega$. Without loss of generality, F is infinite as finite choice follows from ZF alone. Consider the relation $R \subseteq (\bigcup F) \times (\bigcup F)$ defined as follows. For each $x \in \bigcup F$ let $F_x \in F$ be the unique element of F that has x as an element. Thus $x \in F_x$ for each $x \in \bigcup F$. Define x R y iff

1. $f(F_x) < f(F_y)$; and

2. there is no $X \in F$ where $f(F_x) < f(X) < f(F_y)$.

Since F is infinite, im f is unbounded in ω , meaning that for each $x \in \bigcup F$, there is some $y \in \bigcup F$ where $f(F_x) < f(F_y)$. Therefore by DC, there is a sequence $\langle x_n \in \bigcup F : n \in \omega \rangle$ where $f(F_{x_n}) < f(F_{x_{n+1}})$ for all $n \in \omega$. Without loss of generality (just by finite choice in ZF to add in finitely many entries in the sequence) we can assume min $\{f(F_{x_n}) : n \in \omega\} = \min\{f(X) : X \in F\}$. Let $C = \{x_n : n < \omega\}$. This is a choice set for $\{F_{x_n} : n < \omega\}$.

So now we show that $\{F_{x_n} : n < \omega\} = F$. Since clearly $\{F_{x_n} : n < \omega\} \subseteq F$, suppose $X \in F \setminus \{F_{x_n} : n < \omega\}$. Note that f(X) = n for some $n < \omega$ where then $n \ge \min\{f(F_{x_m}) : m < \omega\}$. Clearly if equality holds, then $X = F_{x_m}$ for some $m < \omega$ by injectivity of f. Thus, as $\{f(F_{x_m}) : m < \omega\}$ is unbounded in ω , there is some $m < \omega$ where $f(F_{x_m}) < f(X) < f(F_{x_{m+1}})$. But then $\neg(x_m R x_{m+1})$, a contradiction. Therefore there can be no such X and thus $\{F_{x_n} : n < \omega\} = F$, meaning C is a choice set for F.

The reverse does not hold, implying that DC is strictly stronger than countable choice. DC is used for a great number of theorems and basic results in analysis, particularly in the use of sequences. For instance DC will show the equivalence between continuity (in a general topological sense for \mathbb{R}) and sequential continuity. Indeed, in a very vague sense, much of analysis can be carried out in DC, or at least DC relativized to \mathbb{R} . Given this, DC is really the first serious weakening of AC used for mathematics.

To show the power of DC over countable choice, we have the following theorem. Countable choice only gave Kőnig's Lemma on Trees (9 B • 5): for a finitely branching tree of height ω , there is an infinite path. The main reason why we needed finitely many branches is to ensure the resulting tree was countable. DC does not need this restriction, only that there are no finite braches (meaning branches with a finite length). In fact, the following consequence of DC is equivalent to DC over ZF.

- 9B•8. Theorem

Therefore $\mathsf{ZF} \vdash \mathsf{``DC} \leftrightarrow \mathsf{every}$ tree of height ω has a branch''.

Proof .:.

Assume DC and let $\mathbf{T} = \langle T, \leq_T \rangle$ be a tree of height ω . If there is some $\tau \in T$ with no $\sigma \in T$ where $\tau <_T \sigma$, then

VARIANTS OF THE AXIOMS

T has a finite branch, which can be seen just by considering $\operatorname{pred}_{<_T}(\tau)$. So suppose *T* has no finite branches. Thus for each $\tau \in T$, there is some $\sigma \in T$ with $\tau <_T \sigma$. Hence by dependent choice, there is a sequence of $\langle \tau_n \in T : n < \omega \rangle$ where $\tau_n < \tau_{n+1}$ for all $n < \omega$. Closing $\{\tau_n : n < \omega\}$ under $<_T$ -predecessors then yields an infinite branch.

Assume every tree of height ω has a branch. Let $R \subseteq X \times X$ be such that $\forall x \in X \exists y \in X (x R y)$. Consider the tree of finite sequences

$$T = \{ f : n+1 \to X : n < \omega \land \forall m < n \ (f(m) \ R \ f(m+1)) \},\$$

ordered by \leq_T where $\tau <_T \sigma$ iff $\tau \subsetneq \sigma$. Note that *T* has no finite branches, since each $f : n + 1 \to X$ has some $y \in X$ where $f(n) \ R \ y$ so that $f' = f \cup \{\langle n + 1, y \rangle\}$ extends *f*. Hence *T* has height ω and therefore has an *infinite* branch $f : \omega \to T$ where $f(n) \subsetneq f(n + 1)$. Therefore, $\langle x_n : n < \omega \rangle$ defined by $x_n = f(m)(n)$ for any (and all) m > n (since the domains are strictly increasing as *m* increases, this is just to make sure $n \in \text{dom}(f(m))$) yields that DC holds for *R*.

Since the existence of finite branches will be a part of ZF, DC is really equivalent to the existence of branches when there are no finite branches. So the theorem above (and König's Lemma on Trees $(9B \cdot 5)$) is really talking about the existence of branches when they are forced to be infinite. AC says that every tree (of ordinal height) has branches, whereas these various weakenings say that only certain trees have branches:

- finite choice says finite trees have branches;
- countable choice says countable trees of height ω have branches;
- DC says all trees of height ω have branches;
- AC says all set trees have branches; and
- the extension of AC *global choice*—formally understood as "V = HOD" through Corollary 8 D 11—says that trees of height Ord have branches.

And we can prove almost all of these to be equivalences. The odd one out is countable choice, which seems to be strictly stronger than countable trees of height ω having branches, and instead seems to be equivalent to when we can also take, in some sense, finite approximations to choice sets.

— 9 B • 9. Theorem

- 1. $\mathsf{ZF} \vdash$ "finite choice \leftrightarrow finite trees have branches".
- 2. $\mathsf{ZF} \vdash$ "countable choice \Leftrightarrow countable trees of height ω have branches \land every countable $F \neq \emptyset$ of nonempty, disjoint sets has a C' where $\forall x \in F$ ($|C' \cap x| < \aleph_0$)".
- 3. $\mathsf{ZF} \vdash \mathsf{``DC} \leftrightarrow \mathsf{trees} \mathsf{ of height } \omega \mathsf{ have branches''}.$
- 4. $ZF \vdash$ "AC \leftrightarrow trees of height < Ord have branches".
- 5. $W \models ZF$ has $W \models "V = HOD"$ iff class trees of W of height $\leq Ord^{W}$ have (W-definable) class branches.

Proof .:.

- 1. Since both are proven from ZF, it follows that they are equivalent over ZF.
- 3. This is Theorem 9B 8.
- 2. We've proven that countable choice implies countable trees of height ω have branches by Kőnig's Lemma on Trees (9 B 5), and obviously an actual choice set C for F yields the second statement. So suppose countable trees of height ω have branches, and we can take choice sets modulo finite subsets. To show countable choice, let $F \neq \emptyset$ be an arbitrary countable set of non-empty, disjoint sets. Let C' be such that $C' \cap x$ is finite for each $x \in F$. Consider the following tree of refinements on C. First, let $f : \omega \to F$ be a bijection. Define

 $T = \left\{ c \subseteq C' : |c| < \omega \land \forall n < |c| \exists ! x \ (x \in c \cap f(n)) \right\},\$

with $c \leq_T d$ iff $c \subseteq d$. This is a tree $\mathbf{T} = \langle T, \leq_T \rangle$ with height ω and since $C' \cap f(n)$ is finite for each n, \mathbf{T} is finitely splitting. Hence T is countable, and thus there is a branch $C \subseteq T$. Note that this branch is itself a subset of C', and for each $n < \omega$, $\exists ! x \ (x \in C \cap f(n))$. Hence C is a choice set for F.

4. Let T = ⟨T, ≤_T⟩ be a tree. Suppose AC holds. Therefore, by Theorem 9A•3, Zorn's lemma holds. So consider the non-empty poset ⟨A, ≼⟩ where A consists of chains of T and a ≤ b iff a ⊆ b. Note that every chain of A is bounded in A since a chain c ⊆ A has ∪ c as a chain of T and a ∈ c implies a ≤ ∪ c. Thus A has a ≤-maximal element C, which is then a branch of T.

Now suppose that all set trees have branches. We will show that AC_C holds. Let X be an arbitrary set. Consider the tree where

 $T = \{c : \alpha \to X : \alpha \in \text{Ord} \land c \text{ is injective}\},\$

where $c \leq_T c'$ iff $c \subseteq c'$. Note that $<_T$ is a well-founded relation, since any infinite $<_T$ -decreasing sequence yields a corresponding decreasing sequence of ordinals according to the domains of the entries in the sequence: $c_{n+1} <_T c_n$ implies dom $(c_{n+1}) < \text{dom}(c_n)$. Moreover, \leq_T linearly order predecessors so that $\mathbf{T} = \langle T, \leq_T \rangle$ is a tree. By Hartogg's Number (5 C • 5), T is a set, being a subset of ${}^{\kappa}X$ where $\kappa \in$ Ord has $\kappa \not\leq_{\text{size}} X$. Therefore there is a branch $C \subseteq T$. Note that then $f = \bigcup C$ is an injective function from some ordinal $\alpha < \kappa$ to X. Moreover, f must be surjective, as any $x \in X \setminus \text{im } f$ has $f <_T f' = f \cup \{\langle \alpha, x \rangle\}$, contradicting the maximality of f. Therefore $f : \alpha \to X$ is a bijection, showing AC_C holds.

5. Suppose V = HOD holds, and let \leq be the definable well-order of V. Let T $\neq \emptyset$ and the tree order \leq_T be classes. Define by transfinite recursion a (possibly class) branch of T. Let t_0 be a minimal element of T. For t_{α} defined for $\alpha < \beta$, if there is no t with $t_{\alpha} <_T t$, then the closure of $\{t_{\alpha} : \alpha < \beta\}$ under \leq_T -predecessors is a branch of T. Otherwise, take t_{β} to be the \leq -least such t. The class $\{t \in T : \exists \alpha \in \text{Ord} (t \leq_T t_{\alpha})\}$ is then a branch of T.

Suppose all trees of height Ord have (class) branches. This means that there is a FOLp-definable branch of every tree of height Ord. We will define a class well-ordering of V from this. Consider the class T of functions f from ordinals (to V) such that

i. *f* is injective; and

ii. If $x \in V_{\alpha} \cap \text{im } f$, then $V_{\beta} \subseteq \text{im } f$ for $\beta < \alpha$.

This is ordered by $f \leq_T g$ iff $f \subseteq g$. As before, \leq_T is well-founded, and linear on collections of predecessors. Thus T and \leq_T form a tree so that there is a FOLp-definable branch $C \subseteq T$.

 \bigcup C is then a class function from Ord (to V). Moreover, \bigcup C is injective. To see that \bigcup C is surjective, note $x \in V$ with rank $(x) = \alpha$ implies $x \in V_{\alpha+1}$. Since $|V_{\alpha+1}| < Ord$ and f is injective, im f cannot be contained in $V_{\alpha+1}$. Hence there is some $y \in V_{\gamma} \cap \text{ im } f$ for $\gamma > \alpha + 1$. But then by (ii), (a sufficiently large initial segment of) f has $V_{\alpha+1} \subseteq \text{ im } f$. Hence $x \in \text{ im } f$ so that $f : Ord \to V$ is a class bijection. This yields the class well-order of V by $x \leq y$ iff $f^{-1}(x) < f^{-1}(y)$.

§9C. The axiom scheme of collection

We begin with Scott's trick, a clever way of restricting formulas which ostensibly define proper classes to sets.

- 9C·1. Theorem (Scott's Trick) Let $\approx \subseteq C \times C$ be a definable equivalence relation on a proper class C. Therefore, for each $c \in C$ we can define the equivalence class of c under \approx to be the *set*

 $[c]_{\approx} = \{ d \in \mathbb{C} : c \approx d \land \operatorname{rank}(d) \text{ is the least such rank} \}.$

Proof .:.

For each $c \in C$, the class {rank $(d) \in Ord : d \approx c$ } has a least element, α . Hence { $d \in V_{\alpha+1} : c \approx d$ } is a set. Moreover, for any $d \approx c$, it follows that $[d]_{\approx} = [c]_{\approx}$.

Scott's trick really just says that we can still make sense of equivalence classes for what might ordinarily be proper classes. For example, this allows us to define equivalence classes of $=_{size}$ in ZF + \neg AC that are sets. The idea can be

slightly generalized to other relations.

— 9C·2. Theorem

Let $\varphi(x, y)$ be a FOLp-formula. Suppose for every $x \in D$ there is some y where $\varphi(x, y)$. Therefore there is a set R where for every $x \in D$ there is a $y \in R$ with $\varphi(x, y)$ (and for every $y \in R$ there is a $x \in D$ with $\varphi(x, y)$).

Proof .:.

As with Scott's Trick (9 C • 1), for each $d \in D$, let [d] be the set $\{y : \varphi(d, y) \land \operatorname{rank}(y) \text{ is least}\}$. Therefore, by replacement, $R = \bigcup_{d \in D} [d]$ yields a set witnessing the result.

Notice that the above theorem is really a strengthening of replacement. Whereas replacement requires φ to define a function, the above shows that we can weaken this requirement to just being a relation.

- 9C·3. Definition The axiom (scheme) of collection (Coll) consists of formulas of the form $\forall w_0 \cdots \forall w_n \forall D (\forall x \in D \exists y \varphi(x, y, \vec{w})) \rightarrow \exists R \forall x \in D \exists y \in R \varphi(x, y, \vec{w})),$ where φ is a FOL(\in)-formula.

So just by examining the form of this, this is stronger than replacement, although Theorem 9C • 2 shows they are equivalent over ZF. Note, however, that for Scott's trick and this idea to hold, we required powerset to ensure that $V_{\alpha} = \{x : \operatorname{rank}(x) < \alpha\}$ is a set for each α . In fact, under ZF – P, Coll is *strictly* stronger than Rep, although the proof of this is quite complicated.

Given this, in the absence of powerset, we often will work with collection instead of replacement.

```
- 9C·4. Definition \overline{ZF^-} is the theory ZF - P + Coll. Similarly ZFC^- is ZF^- + AC.
```

Note that we have encountered many toy models of (fragments of) set theory that model ZF^- rather than merely ZF - P. In particular, $H_{\kappa} \models ZFC^-$ for regular κ , showing $L_{\kappa} \models ZFC^-$ for κ regular in L.

− 9 C • 5. Result Let $\kappa > \aleph_0$ be a regular cardinal. Therefore $H_{\kappa} \models$ Coll and thus $H_{\kappa} \models$ ZFC⁻.

Proof .:.

Suppose φ defines a relation over $D \in H_{\kappa}$. By choice and collection in V, there is an $R \subseteq H_{\kappa}$ such that for each $d \in D$ there is some unique $r \in R$ with $\varphi(d, r)$, and for each $r \in R$ there is some (possibly multiple) $d \in D$ with $\varphi(d, r)$. Hence there is a surjection from D onto R. Since then $|R| \leq |D| < \kappa$ with $R \subseteq H_{\kappa}$, it follows by regularity of κ that $R \in H_{\kappa}$.

§9D. Refinements of collection and comprehension

Because comprehension, collection, and replacement are all schemes—meaning that for each appropriate φ we get a new axiom—there are various refinements of these that restrict what kinds of formulas are allowed.

- 9D•1. Definition

For $n \in \mathbb{N}$, Σ_n -Comp refers to the axiom scheme of comprehension restricted to Σ_n -formulas. And similar definitions hold for Σ_n -Coll, and Σ_n -Rep.

The benefit of having these refinements is being able to have "enough" comprehension, or "enough" collection to play around with. Although a structure might not satisfy full collection or full comprehension, often it will at least satisfy Σ_0 -collection or Σ_0 -comprehension. This is particularly important in the fine structure theory of L and L[E] as explored in later chapters.

Similar to these refinements on the axiom schemes, we also have refinements on elementarity. For example,

- 9D•2. Definition Let σ be a signature. Let A and B be FOL(σ)-structures. Let $n \in \mathbb{N}$. A function $j : A \to B$ is a Σ_n -embedding iff for all Σ_n -FOL-formulas φ (defined similarly as with Definition 7•4) $A \models "\varphi(a_0, \dots, a_m)"$ iff $B \models "\varphi(j(a_0), \dots, j(a_m))"$. In this case, we may also write $j : A \to \Sigma_n B$. If j = id so that $A \subseteq B$ is a submodel, we also write $A \preccurlyeq_{\Sigma_n} B$.

Thus for transitive classes of set theory, a Σ_0 -embedding is just an embedding while a Σ_{ω} -embedding (meaning an embedding which is Σ_n for each $n < \omega$) is an embedding with full elementarity. And this is where the refinements of collection and comprehension become useful: if **B** satisfies some fragment of comprehension, and there is a sufficiently elementary embedding from **A** into **B**, then **A** will also satisfy some fragment of comprehension.

— The Axioms of **ZFC**

1. (Extensionality, Ext) two sets are equal whenever they have the same members:

 $\forall x \; \forall y \; (x = y \leftrightarrow \forall v \; (v \in x \leftrightarrow v \in y)).$

- 2. (Empty set) there is a set \emptyset with no members: $\exists z \forall x (x \notin z)$.
- 3. (Comprehension, Comp) for each x, and for each $FOL(\in)$ -formula $\varphi(v, \vec{w}), \{v \in x : \varphi(v, \vec{w})\}$ exists:

 $\forall w_0 \cdots \forall w_n \ \forall x \ \exists z \ \forall v \ (v \in z \leftrightarrow v \in x \land \varphi(v, \vec{w})).$

- 4. (Pairing, Pair) for any two sets x and y, the pair $\{x, y\}$ exists: $\forall x \forall y \exists z \forall v (v \in z \leftrightarrow (v = x \lor v = y))$.
- 5. (Union, Union) for any family of sets F, there is a set containing the elements of all of those sets:

 $\forall F \exists U \ \forall v \ (v \in U \leftrightarrow \exists x (x \in F \land v \in x)).$

6. (Foundation, Found) for each x, there is a \in -minimal element of x, meaning a member $y \in x$ with no $z \in y$ being in x:

$$\forall x \exists y (y \in x \land \forall z (z \in y \to z \notin x)).$$

- 7. (Infinity, Inf) an infinite set exists: $\exists N(\emptyset \in N \land \forall x (x \in N \to x \cup \{x\} \in N))$.
- 8. (Replacement, Rep) the image of a function over a set is a set: for each $FOL(\in)$ -formula φ ,

 $\forall w_0 \cdots \forall w_n \forall D \left(\forall x (x \in D \to \exists ! y \varphi(x, y, \vec{w})) \to \exists R(y \in R \leftrightarrow \exists x (x \in D \land \varphi(x, y, \vec{w}))) \right).$

- 9. (Powerset, P) for each x, $\mathcal{O}(x)$ exists: $\forall x \exists P \forall v \ (v \in P \leftrightarrow \forall y \ (y \in v \rightarrow y \in x))$.
- 10. (Choice, AC) for any family of non-empty family of non-empty, disjoint sets F, there is a set C which has chosen one element from each $z \in F$:

$$\forall F \ (\emptyset \notin F \land \forall x, y \in F \ (x \cap y = \emptyset) \to \exists C \ \forall x \in F \ \exists ! y \ (y \in x \cap C).$$

Variant Axioms and Axiom Systems –

- 1. (Weak pairing, wPair) for any two x, y, there is a z with $x, y \in z$.
- 2. (Weak union, wUnion) for any family F, there is a z with $\forall x \in F \ (x \subseteq z)$.
- 3. (Weak replacement, wRep) the image of a function over a set is contained in a set.
- 4. (Weak powerset, wP) for any x, there is a set containing all subsets of x.
- 5. (Collection, Coll) there is a range for a relation with over a given domain: for each $FOL(\in)$ -formula φ ,

 $\forall w_0 \cdots \forall w_n \forall D (\forall x \in D \exists y \varphi(x, y, \vec{w})) \rightarrow \exists R \forall x \in D \exists y \in R \varphi(x, y, \vec{w})).$

- 6. (Σ_n -Comprehension, Σ_n -Comp) for each x, and for each Σ_n -formula $\varphi(v, \vec{w}), \{v \in x : \varphi(v, \vec{w})\}$ exists.
- 7. (Σ_n -Collection, Σ_n -Coll) Coll holds for Σ_n -formulas.
- 8. (Dependent choice, DC) for $R \subseteq X \times X$, if $\forall x \in X \exists y \in X (x R y)$ then there is a sequence $\langle x_n : n \in \omega \rangle$ such that $x_n R x_{n+1}$ for all $n \in \omega$.
- 9. For every $x, y, x \times y$ exists.

With these axioms, we have the following theories:

- BST consists of (1)–(6) plus (ix).
- wZF consists of (1), (2), (3), (6), (7), and (i)–(iv). wZFC also adds (10).
- ZF⁻ consists of (1)–(8) plus (v). ZFC⁻ also adds (10).
- $ZF = ZF^- + P$ consists of (1)–(9). ZFC also adds (10).

Сн I §10 В

Section 10. Exercises for Transitivity

§10 A. Easier Exercises

- **10 Ex1. Exercise:** Prove or disprove: $\mathcal{O}(X \setminus Y)$ can be equal to $\mathcal{O}(X) \setminus \mathcal{O}(Y)$.
- **10 Ex2.** Exercise: Suppose R is an equivalence relation over Y. Let $f : X \to Y$ and define $S = \{\langle x, x' \rangle : f(x) \ R \ f(x')\}$. Show S is an equivalence relation over X.
- **10 · Ex3. Exercise:** Verify the following equalities for all sets X, Y, and Z:
 - $X \times (Y \cap Z) = (X \times Y) \cap (X \times Z);$
 - $X \times (Y \cup Z) = (X \times Y) \cup (X \times Z)$; and
 - $X \times (Y \setminus Z) = (X \times Y) \setminus (X \times Z).$
- **10. Ex4. Exercise:** Prove ordinal addition, multiplication, and exponentiation is equal to cardinal addition, multiplication, and exponentiation for ordinals (equivalently cardinals) less than ω .
- **10 Ex5. Exercise:** Let X be an infinite set. Show $|^{<\omega}X| = |X|$.
- **10 · Ex6.** Exercise: Suppose R and R^{-1} well-order some set X. Show $|X| < \omega$.
- **10 · Ex7. Exercise:** Show there are arbitrarily large cardinals κ such that $\kappa^{\aleph_0} = \kappa$.
- **10 Ex8. Exercise:** Show there are arbitrarily large cardinals κ such that $\kappa^{\aleph_0} > \kappa$.
- **10 · Ex9. Exercise:** Show there are arbitrarily large ordinals α such that $\aleph_{\alpha} = \alpha$.
- **10 · Ex10.** Exercise: Suppose κ is weakly inaccessible. Show $\aleph_{\kappa} = \kappa$.

For $\alpha \in \text{Ord}$, define by transfinite recursion the sequence of \beth_{α} s where $\beth_0 = \aleph_0$, $\beth_{\alpha+1} = 2^{\beth_{\alpha}}$, and $\beth_{\gamma} = \sup_{\alpha < \gamma} \beth_{\alpha}$ for γ a limit.

- **10**•Ex11. Exercise: Show there are arbitrarily large α such that $\beth_{\alpha} = \alpha$.
- **10 Ex12.** Exercise: Show there are arbitrarily large α such that $\aleph_{\alpha} = \beth_{\alpha}$.
- **10 Ex13. Exercise:** Show GCH is equivalent to " $\forall \alpha \in \text{Ord} (\aleph_{\alpha} = \beth_{\alpha})$ ".
- **10 · Ex14.** Exercise: Show κ is weakly inaccessible iff $\aleph_{\kappa} = \kappa$ and κ is regular. Show κ is strongly inaccessible iff $\beth_{\kappa} = \kappa$ and κ is regular.
- **10 Ex15. Exercise:** Show x is an ordinal iff x is transitive and every $y \in x$ is transitive.
- **10 Ex16**. **Exercise**: Let $[\kappa]^{\lambda}$ denote the λ -sized subsets of κ . Show $|[\kappa]^{\lambda}| = \kappa^{\lambda}$.
- **10 · Ex17. Exercise:** For each cardinal $\kappa \geq \aleph_0$, show that the number of bijections $f : \kappa \to \kappa$ is 2^{κ} .
- 10 Ex18. Exercise: Show $\sup_{N < \omega} \prod_{n \in N} n = \aleph_0$ but $\prod_{n \in \omega} n = 2^{\aleph_0}$.
§10B. Medium Exercises

- **10**•Ex19. Exercise: Suppose there are inacessible cardinals. We know $V_{\kappa} \models \mathsf{ZFC}$ if κ is inaccessible. But show that the least α where $V_{\alpha} \models \mathsf{ZFC}$ is *not* inaccessible.
- 10 Ex20. Exercise: Show $V_{\kappa} = H_{\kappa}$ iff $\kappa = \beth_{\kappa}$.
- **10 · Ex21.** Exercise: or $\kappa \geq \aleph_0$, show $|H_{\kappa}| = 2^{<\kappa}$.
- **10**•Ex22. Exercise: Show using Konig's Lemma on Trees (9 B 5) that a relation R is well-founded iff there are no infinite, decreasing sequences of elements of dom(R) \cup ran(R).
- 10 Ex23. Exercise: Show that every formula is equivalent under ZF⁻ to a formula in the Lévy-hierarchy.
- **10**•Ex24. Exercise: Show every ordinal α can be represented as the finite sum $\omega^{\beta_0} + \omega^{\beta_1} + \cdots + \omega^{\beta_n}$ where $\beta_{k+1} \leq \beta_k$ for all k < n (and here, exponentiation refers to *ordinal* exponentiation).
- **10 · Ex25. Exercise:** Let $M \subseteq V$ be a class. Suppose $V \models "\forall x \ (x \subseteq M \rightarrow x \in M)$ ". Show V = M.
- **10 · Ex26.** Exercise: Let κ be strongly inaccessible. Show the following are absolute between V_{κ} and V:
 - "y = P(x)";
 - " $y = \omega_{\alpha}$ ";
 - " $y = V_{\alpha}$ ";
 - " $y = cof(\alpha)$ "; and
 - "*α* is strongly inaccessible".
- **10 · Ex27.** Exercise: Suppose X is an (actually) infinite set and that ZFC is consistent. Show there is a model $W \models ZFC$ with $X \subseteq Y \in W$ where $W \models "Y$ is finite".
- **10 Ex28. Exercise:** Consider the game between two players—I and II—alternating turns where on their *n*th turns, I plays some $\alpha_n < \omega_1$ and then II plays some $\beta_n < \omega_1$. After ω turns, the game ends, and we say II wins iff the set $\{\alpha_n : n < \omega\} \cup \{\beta_n : n < \omega\}$ is an ordinal less than ω_1 , and otherwise I wins. Show that II has a winning strategy, meaning II can always win regardless of what I plays.

§10C. Harder Exercises

- **10 · Ex29. Exercise:** Suppose $M \models ZFC$ and $N \models ZF$ are two inner models. Show M = N iff M and N have the same sets of ordinals, meaning $\mathcal{P}(Ord) \cap M = \mathcal{P}(Ord) \cap N$.
- **10 Ex30.** Exercise: A κ -tree is a tree $\mathbf{T} = \langle T, \leq_T \rangle$ of height κ such that $|\text{lvl}_{\alpha}(\mathbf{T})| < \kappa$ for each $\alpha < \kappa$. Thus Kőnig's Lemma on Trees (9 B 5) says that every ω -tree has a cofinal branch. This does not hold for ω_1 . Show there is an \aleph_1 -tree with no cofinal branch. (*Hint*: consider finite, injective functions from \aleph_0 to \aleph_1 and then thin out the tree to ensure the levels are small).
- **10 · Ex31.** Exercise: Show under ZF that AC is equivalent to " $\forall \alpha \in \text{Ord} (\mathcal{P}(\alpha) \text{ has a well-order})$ ".
- **10 · Ex32.** Exercise: Let $\alpha < \omega_1$. Show there is an order-preserving function $f : \alpha \to \mathbb{Q}$ where \mathbb{Q} is the set of rational numbers, and the orders are the usual orders on ordinals and \mathbb{Q} .
- 10 · Ex33. Exercise:
 - Let A be a set and suppose $F : \mathcal{P}(A) \to \mathcal{P}(A)$ is such that $X \subseteq Y \to F(X) \subseteq F(Y)$ for all $X, Y \in \mathcal{P}(A)$. From ZF, show that there is a $Z \in \mathcal{P}(A)$ such that F(Z) = Z.
 - Use the above to find an alternative proof of Cantor-Bernstein $(5 C \cdot 4)$ from ZF.

EXERCISES FOR TRANSITIVITY

Сн I §10

Chapter II. Filters, Embeddings, and Extenders*

Section 11. Ultrafilters and Logic

Ultrafilters can be seen in a variety of places around mathematics, especially within set theory. For our purposes, ultrafilters give rise to ultraproducts, and certain ultraproducts result in inner models. The existence and structure of these inner models give rise to deep results about the original universe we start in, and they present important connections to large cardinal assumptions and consistency strength.

§11 A. Filters

The notion of a filter over a set makes precise the notion of largeness as well as "almost every". Its association with measure also leads to saying a set is "measure overe" to mean that it is in the filter, alluding to measuring subsets of $[0,1] \subseteq \mathbb{R}$ similar to probability. This way of referring to sets in a filter $F \subseteq \mathcal{P}(X)$ is motivated by the idea that if $x \in \mathcal{P}(X)$ is "large" and $x \subseteq y \in \mathcal{P}(X)$, then y is "large" too. This leads to the following definition.

- 11A • 1. Definition Let $A \neq \emptyset$ be a set. A *filter* over A is a non-empty subset $F \subsetneq \mathcal{P}(A)$ such that the following hold: for all $x, y \in \mathcal{P}(A)$,

1. If $x \in F$ and $x \subseteq y$, then $y \in F$; and

2. If $x, y \in F$, then $x \cap y \in F$.

An *ultrafilter* over *A* is a \subseteq -maximal filter $U \subseteq \mathcal{P}(A)$.

Other references will often require $A \in F \subseteq \mathcal{P}(A)$ and $\emptyset \notin F$, but these are implied by (1) and that $F \neq \emptyset$ is a *proper* subset $F \subsetneq \mathcal{P}(A)$. Without this requirement, we'd have trivial filters like all of $\mathcal{P}(A)$, or just \emptyset . We wouldn't want to allow such sets to be filters, because it would muck with the definition of ultrafilters.

To help grasp the concept a bit more, we have some relatively easy examples of filters.

— 11 A • 2. Example —

- 1. Let $A \neq \emptyset$ be any set with $a \in A$. Therefore $\{x \in \mathcal{P}(A) : a \in x\}$ is a filter, and in fact an ultrafilter.
- 2. Let $A \neq \emptyset$ be any set with $\emptyset \neq x \subsetneq A$. Therefore $\{y \in \mathcal{P}(A) : x \subseteq y\}$ is a filter, but not an ultrafilter unless x is a singleton.
- 3. Suppose A is infinite. Therefore $\{x \in \mathcal{P}(A) : A \setminus x \text{ is finite}\}$ is a filter, but not an ultrafilter.
- 4. Let κ be an uncountable, regular cardinal. Call $x \subseteq \kappa$ a *club* iff $\sup x = \kappa$, and for all bounded $y \subseteq x$, $\sup y \in x$. Therefore $\{x \in \mathcal{O}(\kappa) : x \text{ contains a club of } \kappa\}$ is a filter called the *club filter*, but it is not an ultrafilter.

The first is in effect the most trivial kind of filter, and it is something we will try to avoid. Note that we can come up with all sorts of filters. First we just start with a family of non-pairwise-disjoint sets X, and then we close under intersections and supersets. This yields a filter containing X. So this is the process by which we constructⁱ ultrafilters: just keep adding a set or its complement until we can't anymore.

For now, let's try to generate more examples of filters. To do this, we need to be somewhat careful. The general idea is to simply close a given set under finite intersections, and then add all supersets. The issue with this is that we need

ⁱChoice is needed to show the existence of (nonprincipal) ultrafilters, so it's not exactly an explicit construction.

ULTRAFILTERS AND LOGIC

— 11A•3. Definition -

A set X has the *finite intersection property* iff for all finite subsets $\{x_0, \dots, x_n\} \subseteq X$, $\bigcap_{i \leq n} x_i \neq \emptyset$.

– 11 A•4. Result

Let $A \neq \emptyset$ be a set, and let $X \subseteq \mathcal{P}(A)$ have the finite intersection property. Therefore there is a filter $F \supseteq X$.

Proof .:.

Consider the closure Y of X under pairwise intersections. By the finite intersection property, $\emptyset \notin Y$. Now define $F = \{x \in \mathcal{O}(A) : \exists y \in Y(y \subseteq x)\}$. As $\emptyset \notin Y, \emptyset \notin F$ and hence $F \subsetneq \mathcal{O}(A)$. F is clearly closed under supersets and pairwise intersection because Y is. Hence F is a filter with $X \subseteq Y \subseteq F$.

The filter given in the proof is generated by X not just in the sense that the construction is given by X, but also in the sense that it is the \subseteq -minimal filter containing X. Now the question becomes how to generate an *ultrafilter*. Without AC, the situation is a bit odd and differentⁱⁱ, but in our case, *every* filter can be extended to an ultrafilter. The proof of this can be easily shown through Zorn's lemma: consider the set of filters containing F, and for each \subseteq -chain, just take the union to get another filter, and end up with a \subseteq -maximal filter $U \supseteq F$.

The characterization of ultrafilters just as maximal filters is useful to prove their existenceⁱⁱⁱ, but for the most part, it doesn't help one understand properties of ultrafilters. A much more useful characterization is the following.

- 11A.5. Result
Let
$$U \subseteq \mathcal{P}(A)$$
 be a filter. Therefore U is an ultrafilter iff for all $x \in \mathcal{P}(A)$, either $x \in U$ or $A \setminus x \in U$.

Proof .:.

Suppose U contains every subset of A or its complement, but there is some other filter with $U \subsetneq F \subsetneq \mathcal{P}(A)$. Take $x \in F \setminus U$ and note that we must have $A \setminus x \in U \subseteq F$. Since F is a filter, $\emptyset = x \cap (A \setminus x) \in F$ which implies $F = \mathcal{P}(A)$, a contradiction.

Now suppose U is an ultrafilter. Let $x \subseteq A$ be such that $x, A \setminus x \notin U$. Consider the set $X = \{u \setminus x : u \in U\}$ which contains $A \setminus x$, for example. Note $\emptyset \notin X$ since otherwise $u \setminus x = \emptyset$ for some $u \in U$, meaning $u \subseteq x \in U$. Therefore X has the finite intersection property because U does. So let F be the filter generated by X: $F = \{y \subseteq X : \exists z \in F (z \subseteq y)\}$. This contains U, contradicting that U is maximal: $x \in F \setminus U$.

We will only be interested in ultrafilters over infinite sets, since the only ultrafilters over finite sets are principal: for U an ultrafilter over $N \in \omega$, each $m \in N$ has some $X_m \in U$ with $m \notin X_m$. Therefore, intersecting these finitely many sets yields $\bigcap_{m < N} X_m = \emptyset \in U$, contradicting that U is a filter.

§11 B. Background on clubs

Without loss of generality, we will consider filters on infinite cardinals. This somewhat simplifies the notation and situation, but it is a primer for later ideas which work directly with concepts related to cardinals. Example 11 A \cdot 2 (4) already contains an example of how working cardinals can give additional information. This example also includes the notion of a *club*, being a closed and unbounded subset. These sets have further properties that can be connected to ultrafilters later. First we repeat a definition.

ⁱⁱThings being odd is usually the case without AC. In particular, the existence of non-principal ultrafilters cannot be proven without AC.

iii and it's the only definition that works if you define filters for posets in general rather than just for the poset $\langle \Phi(A), \subseteq \rangle$

11B-1. Definition Let κ be a cardinal with $cof(\kappa) > \omega$. A subset $x \subseteq \kappa$ is *club* in κ or *a club* iff $x \cup \{\kappa\}$ is closed under supremum of subsets and sup $x = \kappa$. Let $\{x_{\alpha} : \alpha < \lambda\}$ be a family of sets indexed by $\lambda \in Ord$. The *diagonal intersection* of this family is

 $\mathop{\triangle}_{\alpha < \kappa} x_{\alpha} := \{ \alpha < \kappa : \alpha \in \bigcap_{\beta < \alpha} x_{\alpha} \}.$

The diagonal intersection is important because usually we work with regular cardinals, and the set of clubs is closed under diagonal intersections of length κ if κ is regular. This cannot be strengthened to full intersections, however. To see this, for each $\alpha < \kappa$, take the club $C_{\alpha} = \{\beta < \kappa : \alpha < \beta\}$. This gives that $\bigcap_{\alpha < \kappa} C_{\alpha} = \emptyset$, and in fact $\bigcap_{\alpha < \operatorname{cof}(\kappa)} C_{\gamma\alpha} = \emptyset$ for any confinal sequence $\langle \gamma_{\alpha} < \kappa : \alpha < \operatorname{cof}(\kappa) \rangle$. The diagonal intersection, however, will still be a club, and in fact will be κ itself: for every $\alpha < \kappa$, $\alpha \in \bigcap_{\beta < \alpha} C_{\beta}$.

— 11 B • 2. Result —

Let
$$\kappa$$
 be a cardinal with $cof(\kappa) > \omega$. Let $\{C_{\alpha} : \alpha < \kappa\}$ be a collection of clubs. Therefore

- 1. For each $\lambda < \operatorname{cof}(\kappa)$, $\bigcap_{\alpha < \lambda} C_{\alpha}$ is a club.
- 2. If κ is regular, $\Delta_{\alpha < \kappa} C_{\alpha}$ is a club.

Proof .:.

1. Let $\lambda < \operatorname{cof}(\kappa)$ be given. First we will show that $\bigcap_{\alpha < \lambda} C_{\alpha}$ is unbounded. So let $\gamma < \kappa$ be arbitrary. Choose an increasing sequence of x_{α} s such that each $x_{\alpha} \in C_{\alpha}$ and $\gamma < x_0$. Now we have a sequence $\langle x_{\alpha} : \alpha < \lambda \rangle = \langle x_{0+\alpha} : \alpha < \lambda \rangle$. Since $\lambda < \operatorname{cof}(\kappa)$, this is bounded by some $x_{\lambda+0} > \sup_{\alpha < \lambda} x_{0+\alpha}$, and again choose an increasing sequence as before: $x_{\lambda+\alpha} \in C_{\alpha}$. In the end, we get an interlaced, increasing sequence $X = \langle x_{\lambda\cdot n+\alpha} : n < \omega \land \alpha < \lambda \rangle$ where $x_{\lambda\cdot n+\alpha} \in C_{\alpha}$ for each $\alpha < \lambda$. Notice that as the sequence was interlaced and increasing, each C_{α} slice for $\alpha < \lambda$ has the same supremum:

 $\sup(X \cap C_{\alpha}) = \sup\{x_{\lambda \cdot n + \alpha} : n \in \omega\} = \sup X.$

Hence this supremum is in $\bigcap_{\alpha < \lambda} C_{\alpha}$, and is bigger than γ . Thus the intersection is unbounded.

To see that $\bigcap_{\alpha < \lambda} C_{\alpha}$ is closed, any bounded subset $Y \subseteq \bigcap_{\alpha < \lambda} C_{\alpha}$ has $Y \subseteq C_{\alpha}$ for each $\alpha < \lambda$. Yet sup $Y \in C_{\alpha}$ by the hypothesis, for each $\alpha < \lambda$, implying that sup $Y \in \bigcap_{\alpha < \lambda} C_{\alpha}$ as desired.

2. We will again show that $\triangle_{\alpha < \kappa} C_{\alpha}$ is unbounded, as closure is the easier of the two. Let $\gamma < \kappa$ be arbitrary. Choose an increasing sequence $\langle x_n > \gamma : n < \omega \rangle$ with $x_0 \in C_0 \setminus \gamma$ and $x_{n+1} \in \bigcap_{\alpha < x_n} C_{\alpha}$. This can be done since each $\bigcap_{\alpha < x_n} C_{\alpha}$ is club by (1) and $\kappa = \operatorname{cof}(\kappa)$ is regular. Now write $X = \{x_n : n \in \omega\}$ with $x = \sup X$.

To see that $x \in \Delta_{\alpha < \kappa} C_{\alpha}$, we just need to see that $x \in \bigcap_{\alpha < \kappa} C_{\alpha}$. For each $\alpha < x, \alpha \leq x_m$ for some *m*, which means the tail of *X* is contained in C_{α} :

$$\{x_n : n > m\} \subseteq \bigcap_{\beta < x_m} C_\beta \subseteq C_\alpha$$

Therefore $x = \sup X \in C_{\alpha}$ and hence $x \in \bigcap_{\alpha < x} C_{\alpha}$.

To see that $\triangle_{\alpha < \kappa} C_{\alpha}$ is closed, let $X \subseteq \gamma$ be a bounded subset of it with $x = \sup X$. Note that for any $\alpha < \kappa$, we have $X \setminus \alpha \subseteq \bigcap_{\beta < \alpha} C_{\beta}$. In particular, for $\alpha < x$, the tail of X is a subset of C_{α} and hence $x = \sup X \in C_{\alpha}$. Therefore $x \in \bigcap_{\alpha < x} C_{\alpha}$ and so $x \in \triangle_{\alpha < \kappa} C_{\alpha}$.

The importance of the diagonal intersection is primarily for the purpose of Fodor's lemma, which motivates an important property for filters. Fodor's lemma talks about stationary sets: sets which intersect every club set, but which might not be clubs themselves.

```
- 11 B • 3. Definition
Let \kappa be a cardinal with cof(\kappa) > \omega. A subset X \subseteq \kappa is stationary iff C \cap X \neq \emptyset for every club C \subseteq \kappa.
```

The existence of stationary sets is easy to see just from the fact that every club set is stationary: κ itself is trivially a stationary subset of κ . The existence of stationary, co-stationary subsets—i.e. stationary subsets that do not contain a club—can be shown through direct example. Since $\kappa > \aleph_0$, we can consider $S_{\omega}^{\kappa} = \{\alpha < \kappa : cof(\alpha) = \omega\}$. It's clear that S_{ω}^{κ} is stationary, since each club contains a sequence of length ω , whose supremum is then in S_{ω}^{κ} since $\kappa > \aleph_0$ is

CH II §11 C

regular. More generally, the set S_{λ}^{κ} of ordinals with cofinality $\lambda < cof(\kappa)$ will be stationary whenever $\lambda = cof(\lambda) < \kappa$ for precisely the same reason as with ω .

- 11B•4. Definition Let $X \subseteq$ Ord and let $f : X \rightarrow$ Ord. f is *regressive* iff $f(\alpha) < \alpha$ for all $\alpha \in X$.

— 11 B•5. Lemma (Fodor's Lemma)

Let κ be a regular, uncountable cardinal. Let $S \subseteq \kappa$ be stationary, and let $f : S \to \kappa$ be regressive. Therefore f is constant on a stationary set: f^{-1} "{ β } is stationary for some $\beta \in S$.

Proof .:.

Otherwise for each $\beta < \kappa$, let $C_{\beta} \cap f^{-1}$ "{ β } = \emptyset with C_{β} a club. Consider $\Delta_{\beta < \kappa} C_{\beta}$, which is a club by Result 11 B • 2, and hence $S \cap \Delta_{\beta < \kappa} C_{\beta} \neq \emptyset$. Taking $\alpha \in S \cap \Delta_{\beta < \kappa} C_{\beta}$ requires that $f(\alpha) < \alpha$. But note that f^{-1} "{ β } $\cap C_{\beta} \subseteq f^{-1}$ "{ β } $\cap \bigcap_{\gamma < \alpha} C_{\gamma} = \emptyset$

for each $\beta < \alpha$. In particular, for $\beta = f(\alpha) < \alpha$, $\alpha \in f^{-1}$ "{ $f(\alpha)$ } has $\alpha \notin \bigcap_{\gamma < \alpha} C_{\gamma}$, contradicting that $\alpha \in \Delta_{\beta < \kappa} C_{\beta}$. Hence there must be some β with f^{-1} "{ β } stationary, meaning that f is constant on a stationary set.

Such a result is extremely useful for combinatorial parts of set theory, being used to prove statements like the generalized Δ -system lemma, tremendously useful in methods of forcing. Stated in terms of filters, any ultrafilter extending the club filter will necessarily contain only stationary sets, and thus will abide by Fodor's Lemma (11 B • 5). This is a nice property of ultrafilters for various reasons, as will be covered later.

§11 C. Logic and filters

Now as stated above, filters and ultrafilters give a notion of "size" or "largeness" to subsets, but they also then give a notion of "how often" something is in a given subset. In this way, as with a probability measure, ultrafilters give a notion of how often something is true. To make this connection a little more apparent, the following notation will be used extensively.

11C•1. Definition Let F be a filter over a set K. Let $\varphi(x, \vec{w})$ be a FOL(\in)-formula. Write " $\forall^* x \ \varphi(x, \vec{w})$ " to say that $\{x \in K : \varphi(x, \vec{w})\} \in F$. We write $\exists^* x \ \varphi(x, \vec{w})$ to say that $K \setminus \{x \in K : \varphi(x, \vec{w})\} \notin F$.

 \forall^* should be read as "for almost every", and \exists^* doesn't have a standard phrase, but one can read it as "there is a positive set", analogous to measure over the real numbers as if to say it's not measure 0. If we need to specify the ultrafilter, we write \forall_U^* for "for *U*-almost every". In everyday language, words and phrases like "almost every", "by-and-large", and "many" come into play to gloss over details. These words are usually vague or ambiguous, but the notion of an ultrafilter makes them precise in a way that is consistent with ordinary usage. Moreover, the new quantifiers have their own sort of logic to them based just on Definition 11 A • 1. This new vocabulary dramatically simplifies some proofs, and is overall a better way of thinking about ultrafilters, as well as their properties. Definitions that may seem unmotivated or hard to understand can become more intuitive and natural with the new logical framework.

It's useful to present some easy results about how this quantifier interacts with the other connectives of first order logic. Note that the two properties of Definition $11 \text{ A} \cdot 1$ can be restated as

- 1. If $\forall^* \alpha \varphi$ and $\forall \alpha (\varphi \rightarrow \psi)$, then $\forall^* \alpha \psi$; and
- 2. If $\forall^* \alpha \ \varphi$ and $\forall^* \alpha \ \psi$, then $\forall^* \alpha \ (\varphi \land \psi)$.

This isn't difficult to see if we allow parameters, since $\varphi(\alpha, x)$ might just be $\alpha \in x$ and $\psi(\alpha, y)$ might just be $\alpha \in y$. Regardless, these immediately give the following. As a result of the results to follow, (1) above can be weakened so that if $\forall^* \alpha \varphi$ and $\forall^* \alpha (\varphi \to \psi)$, then $\forall^* \alpha \psi$.

— 11 C • 2. Result

Let U be a filter over set K. Let φ and ψ be FOLp (\in) -formulas. Therefore

1. $\forall^* x \ \varphi \rightarrow \exists^* x \ \varphi$. The two are equivalent for U an ultrafilter.

- 2. $\neg \forall^* x \neg \varphi \leftrightarrow \exists^* x \varphi$.
- 3. $(\forall^* x \ \varphi \land \forall^* x \ \psi) \leftrightarrow \forall^* x \ (\varphi \land \psi);$
- 4. $(\exists^* x \ \varphi \lor \exists^* x \ \psi) \leftrightarrow \exists^* x \ (\varphi \lor \psi);$
- 5. $\exists y \forall^* x \varphi$ implies $\forall^* x \exists y \varphi$;
- 6. $\forall^* x \forall y \varphi$ implies $\forall y \forall^* x \varphi$; and
- 7. $\forall x \ \varphi$ implies $\forall^* x \ \varphi$, which implies $\exists^* x \ \varphi$, which implies $\exists x \ \varphi$.

Proof .:.

- Suppose {x ∈ K : φ(x)} ∈ U. If ¬∃*x φ then K \ {x ∈ K : φ(x)} ∈ U. By closure under intersections, this would imply Ø ∈ U, a contradiction. For the other direction, suppose ∃*x φ, meaning K \ {x ∈ K : φ(x)} ∉ U. By Result 11 A 5, this means the complement {x ∈ K : φ(x)} ∈ U, meaning ∀*x φ.
- 2. We have that $\neg \forall^* x \neg \varphi$ iff $\{x \in K : \neg \varphi(x)\} = K \setminus \{x \in K : \varphi(x)\} \notin U$ iff $\exists^* x \varphi$.
- 3. The ' \leftarrow ' direction is immediate since filters are closed under supersets: $\forall x \ (\varphi \land \psi \rightarrow \varphi)$ and $\forall^* x \ (\varphi \land \psi)$ implies $\forall^* x \ \varphi$ and similarly for ψ . For the ' \rightarrow ' direction, use that filters are closed under intersections.
- 4. This is (3) used with the fact that $\neg(\varphi \land \psi)$ is equivalent to $\neg \varphi \lor \neg \psi$).
- 5. This is just from basic first-order logic: if there is a y such that $\{x \in K : \varphi(x, y)\} \in U$ then as a superset, $\{x \in K : \exists y \ \varphi(x, y)\} \in U$.
- 6. If $\forall^* x \forall y \varphi$, then $\{x \in K : \forall y \varphi(x, y)\} \in U$. This set is contained in $\{x \in K : \varphi(x, y)\}$ for any given y, meaning that for every $y, \{x \in K : \varphi(x, y)\} \in U$. Hence $\forall y \forall^* x \varphi$.
- 7. These implications are is clear: K ∈ U so ∀x φ implies {x ∈ K : φ(x)} = K ∈ U. The second implication is from (1). The third implication follows from K ∈ U: if ∃*x φ then K \ {x ∈ K : φ(x)} ∉ U. But ¬∃x φ would imply K = K \ {x ∈ K : φ(x)} ∉ U, a contradiction. ⊢

The weakness of (5) and (6) cannot be improved in general, since

- $\forall^* \alpha \exists x \ (x = \alpha)$ doesn't satisfy $\exists x \forall^* \alpha \ (x = \alpha)$ unless U is principal; and
- Often $\forall \beta \ \forall^* \alpha \ (\alpha > \beta)$ —i.e. almost everything is bigger than any particular β —but we likely won't have $\forall^* \alpha \ \forall \beta \ (\alpha > \beta)$ —i.e. almost every α is bigger than *everything*.

To find specific examples where this happens, we need to consider some particular properties of ultrafilters. Note that when considering ultrafilters, (1) implies that we don't need the notation of \exists^* . But the two are distinct for filters. For example, the *cofinite* subsets of ω , $F = \{x \subseteq \omega : |\omega \setminus x| < \aleph_0\}$, has \exists^*n (*n* is even), but $\{n \in \omega : n \text{ is even}\} \notin F$ so $\neg \forall^*n$ (*n* is even). By Result 11 C • 2 (2), this would imply \exists^*n ($\neg n$ is even). So very often, if *X* if *F*-positive, so too is the complement of *X*. In fact we get the following stating that this always happens whenever *X* lies strictly between being measure 0 and measure 1.

— 11 C • 3. Result

For any FOLp(\in)-formula φ and filter U over a set K, if $\neg \forall^* x \varphi$, then either $\forall^* x \neg \varphi$, or else both $\exists^* x \neg \varphi$ and $\exists^* x \varphi$.

Proof .:.

Suppose $\neg \forall^* x \varphi$ and $\neg \forall^* x \neg \varphi$. From Result 11 C • 2 (2), we get $\exists^* x \neg \varphi$ and $\exists^* x \varphi$, as desired.

\neg

§11 D. Ultrafilter properties

With all of this logical notation at our disposal, we can more easily state some definitions.

```
- 11D.1. Definition 
Let \kappa be a cardinal, and let U be a filter over \kappa.
```

- U is uniform iff $|x| = \kappa$ for every $x \in U$.
- *U* is *unbounded* iff for every $\beta < \kappa, \forall^* \alpha \ (\alpha > \beta)$.
- U is normal iff for every f such that $\forall^* \alpha$ ($f(\alpha) < \alpha$) there is some $\beta < \kappa$ with $\forall^* \alpha$ ($f(\alpha) = \beta$).

ULTRAFILTERS AND LOGIC

• *U* is η -complete iff for $\gamma < \eta$ and formulas φ_{ξ} for $\xi < \gamma$, $\bigwedge_{\xi < \gamma} (\forall^* \alpha \ \varphi_{\xi})$ iff $\forall^* \alpha \ (\bigwedge_{\xi < \gamma} \varphi_{\xi})$. Equivalently, *U* is η -complete iff for $\gamma < \eta$, $\{X_{\alpha} : \alpha < \gamma\} \subseteq U$ implies $\bigcap_{\alpha < \gamma} X_{\alpha} \in U$.

The club filter over a regular, uncountable cardinal will have all of these properties, for example, although it isn't an ultrafilter.

— 11 D•2. Example ————

Let κ be an uncountable, regular cardinal. Let Club_{κ} be the club filter over κ (the filter generated by closed, unbounded [i.e. club] subsets of κ). Therefore Club_{κ} is uniform, unbounded, normal, and κ -complete.

Proof .:.

Obviously all club sets are unbounded, and so by regularity of κ , each club has cardinality κ . So Club_{κ} is uniform. That Club_{κ} is normal is just Fodor's lemma, since each element of Club_{κ} contains a club and is thus stationary. κ completeness follows from facts about club sets (closure of the intersection is immediate, and for unboundedness, interlace a λ -length sequence of members in the $\lambda < \kappa$ clubs and take the supremum which is in all the clubs). \dashv

Many of the properites of Definition 11 D \cdot 1 are connected as shown below for ultrafilters over a cardinal κ . Implications are denoted with arrows (a dashed arrow denotes the implication merely of the *existence* of an ultrafilter with both properties). If κ is regular, unboundedness is equivalent to uniformity, which is otherwise stronger.



11 D•3. Figure: Properties of non-principal ultrafilters over $\kappa \geq \aleph_0$

Arguably the most difficult of the properites in Figure 11 D • 3 to achieve is normality, which isn't directly implied by any combination of the other properties. However the strongest two properties here are clearly κ -completeness and normality. This combination is important enough to get its own name.

— 11 D•4. Definition

Let κ be an uncountable cardinal with U a non-principal ultrafilter over κ . We say that U is a *measure* iff U is κ -complete and normal.

We call this a measure as motivated from the fact that the function

$$\mu(X) := \begin{cases} 1 & \text{if } X \in U \\ 0 & \text{if } X \notin U \end{cases}$$

is a κ -additive, two-valued, probability measure over κ . Other authors often drop the requirement of normality in the definition of a measure (which makes sense with this motivation), but then always work with normal measures.

What this definition tells us is that the dashed arrowof Figure 11 D•3 means that if κ has a κ -complete, non-principal ultrafilter, then it also has a measure. Another notable property of κ -completeness is that it implies that κ is regular. To see this, since it implies unboundedness, if $\langle \gamma_{\beta} : \beta < cof(\kappa) \rangle$ is unbounded with $cof(\kappa) < \kappa$, then the infinitary conjunction $\bigwedge_{\beta < cof(\kappa)} \forall^* \alpha \ (\alpha > \gamma_{\beta})$ implies $\forall^* \alpha \ (\alpha > \gamma_{\beta} \text{ for all } \beta < cof(\kappa))$, which contradicts that the sequence of $\gamma_{\beta}s$ is unbounded in κ .

The two difficult arrows in Figure 11 D \cdot 3 are that κ -complete, non-principal ultrafilters yield measures, and that unbounded, normal, non-principal ultrafilters are measures. The second of these is easier.

— 11 D • 5. Result

Let κ be an uncountable cardinal with U an unbounded, normal ultrafilter over κ . Therefore U is κ -complete, and hence a measure.

Proof .:.

Let $\{X_{\alpha} : \alpha < \lambda\} \in \mathcal{P}(U)$. If $\bigcap_{\alpha < \lambda} X_{\alpha} \notin U$, we may assume without loss of generality that $\bigcap_{\alpha < \lambda} X_{\alpha} = \emptyset$. So define $f : \kappa \to \lambda$ to be such that $f(\alpha)$ is the least $\xi < \lambda$ with $\alpha \notin X_{\xi}$. As $\lambda < \kappa$, $f''\kappa$ is bounded in κ and thus (as U is unbounded),

 $\forall^* \alpha \ (\alpha > \sup(f"\kappa)) \text{ implies } \forall^* \alpha \ (\alpha > f(\alpha)).$

So f is regressive on a set in U. By normality, there is then a $\beta < \kappa$ with $\forall^* \alpha$ $(f(\alpha) = \beta)$. But this means $\forall^* \alpha \ (\alpha \notin X_\beta)$, contradicting that $X_\beta \in U$.

Normality is kind of a weird definition, but its usefulness will become more apparent as we investigate ultrapowers and elementary embeddings. Really, one should think of the dashed arrow of Figure 11 D \cdot 3 as being a property of κ rather than of the ultrafilters. We could still prove now, without reference to later material, the dashed arrow: we can get normal ultrafilters through *possibly different* κ -complete ultrafilters. However, the proof of this with our current understanding is not the best proof as it is fairly technical without additional concepts. But with later material, the idea becomes much more natural.

It should be noted that the existence of measures isn't provable just from ZFC. The reason for this is that any κ which admits such a measure, called a *measurable cardinal*, will be quite large. We will see later that they will be inaccessible, for example, and thus can't be shown to exist just from ZFC. But they will be much more and much larger than mere inaccessibles. To further investigate measurable cardinals, it is useful to take a look at ultrapowers and elementary embeddings.

Section 12. Ultrapowers and Elementary Embeddings

At this point, we can get to our first true application of filters. Ultraproducts are a model-theoretic notion which serve two purposes. Firstly, they are a sort of average of the starting models: what's true in the ultraproduct is what's "almost always" true in the models. Secondly, they are a way to enlarge the universe: the ultraproduct using just one universe yields an elementary embedding. The usefulness for set theory comes when the ultraproduct is well-founded so that it may be collapsed down into an inner model.

In model theory, there is a general concept of a *reduced product* where you kind of "average out" a set of models over a filter. We will not be so concerned with general reduced products, since we will be focused on ultrafilters. The idea is that your objects are now sequences of elements in these models, and a statement $\varphi(\vec{f})$ is true iff for almost every α , $\varphi(\vec{f}(\alpha))$ is true in the corresponding model. So in particular, a sentence is true iff it is true in "most" of the models. The result is an ultraproduct instead of a mere reduced product. Even still, we will not be concerned with ultraproducts in general, but ultraproducts where the models we're "averaging" are all the same model.

— 12•1. Definition -

Let σ be a signature. Let **A** be an FOL(σ)-model. Let U be an ultrafilter over a set K.

- For $f, g: K \to A$, say $f \approx g$ iff $\forall^* x \ (f(x) = g(x))$. The *ultrapower* of **A** by U is the structure **Ult**(**A**, U)
 - with universe $[f]_U = \{g : f \approx g\};$
 - σ -relation interpretations $R^{\text{Ult}(A,U)}([\vec{f}])$ iff $\forall^* x \ R^{A}(\vec{f}(x))$; and
 - σ -function interpretations $F^{\text{Ult}(A,U)}([\vec{f}]) = [g]$ iff $\forall^* x \ (F^{\text{A}}(\vec{f}(x)) = g(x))$.

One should check that these are actually well-defined, but this is easy given that U is an ultrafilter and filters are closed under finite intersections.^{iv}

What's happening here is that the choice of what is true at the level of atomic formulas is left up to the ultrafilter: what happens often enough in the factors happens in the ultraproduct. This goes through to all levels of FOLp-formula complexity, as shown in the following indispensable theorem known as Łoś's Theorem. The theorem fully characterizes FOLp-truth in ultrapowers based on FOLp-truth in the factors. As the name "Łoś" is Polish, it is pronounced ['wse]. Also note that the statement of the theorem here uses only one parameter x in $\varphi(x)$, but the result actually allows for arbitrarily many: $\varphi(\vec{x})$. This would clutter notation for the proof, which essentially the same either way.

- 12·2. Theorem (Łoś's Theorem)

- Let σ be a signature.
- Let **A** be an FOL(σ)-model.
- Let U be an ultrafilter over K, and write Ult for Ult(A, U).
- Let $\varphi(x)$ be an FOL(σ)-formula, and let $[f] \in$ Ult be a parameter.

Therefore **Ult** \models " $\varphi([f])$ " iff $\forall_{II}^* x (\mathbf{A} \models$ " $\varphi(f(x))$ ").

Proof ...

Write Ult for Ult(A, U). This is a proof by structural induction on φ . For atomic formulas, this is just by definition. Sentential connectives \wedge and \neg follow easily as well, since U is an ultrafilter:

 $\begin{aligned} \mathsf{U}\mathsf{lt} \vDash ``\neg \varphi([f])'' & \text{iff} \quad \mathsf{U}\mathsf{lt} \nvDash ``\varphi([f])'' & \text{iff} \quad \neg \forall^* x \ (\mathsf{A} \vDash ``\varphi(f(x))'') \\ & \text{iff} \quad \forall^* x \ (\mathsf{A} \nvDash ``\varphi(f(x))'') & \text{iff} \quad \forall^* x \ (\mathsf{A} \vDash ``\neg \varphi(f(x))''). \end{aligned}$ $\begin{aligned} \mathsf{U}\mathsf{lt} \vDash ``\varphi([f]) \land \psi([f])'' & \text{iff} \quad \mathsf{U}\mathsf{lt} \vDash ``\varphi([f])'' & \text{and} \ \mathsf{U}\mathsf{lt} \vDash ``\psi([f])'' \end{aligned}$

^{iv}Also, technically each $[f]_U$ defined here is a proper class, but this is unimportant by Scott's Trick (9 C • 1).

Сн II §12 А

iff
$$\forall^* x (\mathbf{A} \vDash ``\varphi(f(x))'') \land \forall^* x (\mathbf{A} \vDash ``\psi(f(x))'')$$

iff $\forall^* x (\mathbf{A} \vDash ``\varphi(f(x)) \land \psi(f(x))'').$

For existential quantification, suppose Ult \models " $\exists y \varphi(y, [f])$ ". Thus there is some $[g] \in$ Ult where Ult \models " $\varphi([g], [f])$ ". By (a modified version of) the inductive hypothesis, this happens iff $\forall^* x \ (\mathbf{A} \models "\varphi(g(x), f(x))")$, and so clearly it follows that $\forall^* x \ (\mathbf{A} \models "\exists y \varphi(y, f(x))")$.

For the other direction, we need AC: for each $x \in K$ such that $\mathbf{A} \models \exists y \varphi(y, f(x))$, let g(x) witness this. Otherwise, let g(x) be any arbitrary element of A_x . The resulting function $g = \{\langle x, g(x) \rangle : x \in K\}$ witnesses that **Ult** $\models \varphi([g], [f])$ and thus that **Ult** $\models \exists y \varphi(y, [f])$.

— 12·3. Corollary
 Any model is elementarily equivalent to any of its ultrapowers.

§12 A. Elementary embeddings

We have actually a much stronger correspondence between truth in **A** and its ultrapowers, but to talk further about this relation, we need the concept of an elementary embedding. The issue is that the two models of universes composed of fundamentally different things, and so we can't just compare them outright. Instead, we translate by a function.

- 12A • 1. Definition Let σ be a signature. Let **A** and **B** be FOL(σ)-models. Let $f : A \to B$ be a function. f is an FOL(σ)-elementary embedding iff for all FOL(σ)-formulas φ and parameters $a_0, \dots, a_n \in A$, $\mathbf{A} \models "\varphi(a_0, \dots, a_n)"$ iff $\mathbf{B} \models "\varphi(f(a_0), \dots, f(a_n))"$.

Any elementary embedding will be an embedding just by considering the atomic formulas: $x \in^{A} y$ iff $f(x) \in^{B} f(y)$. It should be obvious from this that any elementary embedding is injective: for $x, y \in A$,

 $x \neq y$ iff $\mathbf{A} \models "x \neq y$ " iff $\mathbf{B} \models "f(x) \neq f(y)$ " iff $f(x) \neq f(y)$.

Elementary embeddings aren't necessarily surjective, however, meaning that they are stronger than a mere embedding, but weaker than a full isomorphism.

Elementary embeddings are crucial to the understanding of ultrapowers and inner models. Some of the basic facts are not recorded because they are seen to be obvious. To better familiarize the reader with some of these basic facts, the following extensively used results will be given explicit proofs.

— 12 A • 2. Lemma

Let $j : V \to M$ be FOL(\in)-elementary (usually, we just write "elementary"). Therefore the following hold.

- 1. If $\varphi(x)$ is a FOL(\in)-formula that is absolute between V and M, then $\varphi(x)$ iff $\varphi(j(x))$. Hence if x is defined by an absolute formula—meaning y = x iff $\varphi(y)$ —then j(x) = x.
- 2. If f is a function, then j(f) is a function, and j(f(x)) = j(f)(j(x)).
- 3. If f, g are functions, then $j(f \circ g) = j(f) \circ j(g)$.

Proof .:.

- 1. By elementarity, $\varphi(x)$ iff $\mathbf{M} \models "\varphi(j(x))$ ". By absoluteness, this happens iff $\varphi(j(x))$. Now if x is defined by φ —i.e. x = y iff $\varphi(y)$ —then x = j(x).
- Being an ordered pair is definable by a formula absolute between transitive models. So being a set of ordered pairs, x ∈ f implies x is an ordered pair. Elementarity then gives that every x ∈ j(f) has that x is an ordered pair. Moreover, f is a function iff ∀x (∃y ⟨x, y⟩ ∈ f → ∃!y ⟨x, y⟩ ∈ f), which is absolute between transitive models. By elementarity,

$$\mathsf{M} \vDash ``\forall x (\exists y \langle x, y \rangle \in j(f) \to \exists ! y \langle x, y \rangle \in j(f))"$$

By absoluteness, this holds in **V** so that j(f) is then a function.

Let $x \in \text{dom}(f)$ be arbitrary. f(x) is the unique y such that $\langle x, y \rangle \in f$. Hence j(f(x)) is the unique y such that $\langle j(x), y \rangle \in j(f)$. Hence j(f) is a function, and it obeys j(f)(j(x)) = j(f(x)) whenever

 $x \in \text{dom}(f)$.

3. Following easily from (1), for all x and z, $\langle x, z \rangle \in f \circ g$ iff there is some $y \in f$ where $\langle x, y \rangle \in g$ and $\langle y, z \rangle \in f$. By elementarity, for all $x, z, \langle x, z \rangle \in j(f \circ g)$ iff $\exists y (\langle x, y \rangle \in j(g) \land \langle y, z \rangle \in j(f))$, meaning $j(f \circ g) = j(f) \circ j(g)$.

Often arguments like these will be written in shorthand, so it's important to know what will be moved by j or won't be moved by j. For example, consider $\gamma = \{\alpha : \alpha < \gamma\}$. Note that $j(\gamma)$ will generally not be $\{j(\alpha) : j(\alpha) < j(\gamma)\}$: α is a dummy variable that plays no role here, but γ is still a parameter. So $j(\gamma) = \{\alpha : \alpha < j(\gamma)\}$ (as expected). Similarly, $j(\{x : f(x) = \alpha\}) = \{x : j(f)(x) = j(\alpha)\}$, again, x is just a dummy variable, but f and α aren't.

Before heading too deep into this, however, we need to think about how we regard these elementary embeddings. As functions from perhaps the entire universe of sets, they will not be sets. And a priori, there's no reason to think they need to be definable. To counter this issue, we will work in a relatively simple class theory, like NBG – GC + AC—von Neumann–Bernays–Gödel class theory with choice for sets. In other words, we have V being the usual set-theoretic universe adjoining predicates for classes of V (whatever those happen to be, but at least including the definable classes). One can show that this is a conservative extension of ZFC, meaning that no new theorems with quantifiers ranging over sets are proven by NBG.

Obviously there is a kind of trivial elementary embedding from V into an inner model: the identity map id : $V \rightarrow V$. This map isn't exactly interesting, however, and so we will be interested with maps that actually move sets. It turns out that if an elementary embedding moves a set, then it moves an ordinal.

Let $j : V \to M$ be elementary. Let $\alpha \in Ord$. Therefore $j \upharpoonright \alpha = id \upharpoonright \alpha$ iff $j \upharpoonright V_{\alpha} = id \upharpoonright V_{\alpha}$.

Proof .:.

Obviously $j \upharpoonright V_{\alpha} = id \upharpoonright V_{\alpha}$ implies $j \upharpoonright \alpha = id \upharpoonright \alpha$ since $\alpha \subseteq V_{\alpha}$.

For the other direction, in essence, the rank of $x \in V_{\alpha}$ is still preserved. For $\alpha = \emptyset$ and α a limit, the result clearly holds. For the successor case, we assume $j \upharpoonright \alpha = id \upharpoonright \alpha$ and that $j(\alpha) = \alpha$. Let $x \in V_{\alpha+1}$. Hence $x \subseteq V_{\alpha}$ so by elementarity, $j(x) \subseteq V_{j(\alpha)}^{\mathsf{M}}$. Now for any fixed $y \in V_{\alpha}$, $y \in j(x)$ inductively is equivalent to $y = j(y) \in j(x)$. So by elementarity, this is equivalent to $y \in x$. Thus j(x) = x, as desired.

Note that we need *j* to be traditional here for us to conclude $V_{\alpha} \subseteq M^{v}$.

- 12A·4. Corollary If $j : V \to M$ elementary, then the least α with $j(\alpha) \neq \alpha$ is also the least rank of a set moved by j.

This motivates the following definition of a critical point, below which j is just the identity, and which is moved by j.

- 12A.5. Definition $j : V \to M$ be elementary and $j \neq id$. An ordinal α is a *critical point* of *j*—denoted cp(*j*)—iff α is the least ordinal where $\alpha \neq j(\alpha)$.

If $\alpha < \kappa = cp(j)$ then by elementarity $\alpha = j(\alpha) < j(\kappa)$ so that $j(\kappa) > \kappa$. So the first ordinal moved is always moved up. This implies that nontrivial elementary embeddings will never be surjective: no ordinal α with between $j(\kappa)$ and κ (more precisely, $\kappa \le \alpha < j(\kappa)$) is in the image of j.

Because we will so often be working with elementary embeddings into inner models, we will use the following nonstandard terminology to describe this.

— 12A•6. Definition

Let $A, B \models ZFC$ be class models. A function $j : A \rightarrow B$ is *traditional* iff $j \neq id, j : A \rightarrow B$ is an elementary embedding, and B is an inner model of A (i.e. A thinks B is transitive).

^vThe issue is that for $j : M \to N$, we only get that $V_{\alpha}^{\mathsf{M}} \subseteq V_{\alpha}^{\mathsf{N}}$ if $j \upharpoonright \alpha = \mathrm{id} \upharpoonright \alpha$. But we need not have $V_{\alpha}^{\mathsf{N}} \subseteq V_{\alpha}^{\mathsf{M}}$.

On the topic of the identity embedding, there is a kind of ceiling to how close $M \subseteq V$ can be to V when $j : V \to M$ is elementary. The following theorem tells us that in particular, M cannot be V. This is important for ruling out the existence of *reinhardt cardinals*—the critical points of elementary embeddings from V into itself, rather than merely an inner model. The proof of this theorem can only be given after we introduce the concept of measurable cardinals.

− 12A•7. Theorem (Kunen's Inconsistency Theorem) There are no traditional $j : V \rightarrow V$. More precisely, suppose $j : V \rightarrow M$ is traditional and a class. Therefore $M \neq V$.

Partial Proof .:.

Here is one proof in a restricted context where j is class definable without parameters (in other, weak *class* theories, this still holds true, like NBG). Let j be FOL(\in)-defined by φ in the sense that j(x) = y iff $\varphi(x, y)$ then we can define $cp(j) = \kappa$ by the least α such that $\neg \varphi(\alpha, \alpha)$. Clearly φ is absolute between V and itself, so by Lemma 12 A • 2 (1), $j(\kappa) = \kappa$, contradicting that $\kappa = cp(j)$.

A nice property of elementary *j* from V into classes of V is that they will preserve $\mathcal{P}(cp(j))$. In general, if $j : N \to M$ is elementary between two transitive classes, there's no guarantee that the powerset is preserved, and we'd only get $\mathcal{P}(cp(j)) \cap N \subseteq \mathcal{P}(cp(j)) \cap M$.

- 12A • 8. Result Let $j : V \to M$ be traditional with $cp(j) = \kappa$. Therefore $\mathcal{P}(\kappa)^M = \mathcal{P}(\kappa)$. In fact, $V_{\kappa+1} \subseteq M$ so that $x \subseteq V_{\kappa}$ is just $x = j(x) \cap V_{\kappa}$.

Proof .:.

As M is transitive, $\mathcal{P}(\kappa)^{\mathsf{M}} = \mathcal{P}(\kappa) \cap \mathsf{M}$. Let $x \in \mathcal{P}(\kappa) \cap V$. For every α , since $x \subseteq \kappa$,

 $\alpha \in x$ iff $j(\alpha) \in j(x) \land \alpha = j(\alpha) < \kappa$,

Hence $x = j(x) \cap \kappa \in \mathcal{P}(\kappa) \cap M$, and thus $\mathcal{P}(\kappa) \subseteq \mathcal{P}(\kappa) \cap M$. And since $M \subseteq V$, $\mathcal{P}(\kappa) \cap M \subseteq \mathcal{P}(\kappa) \cap V$.

By Result 12 A • 3, looking at $j \upharpoonright V_{\alpha+1}$, $j(V_{\alpha}) = V_{\alpha}$ for each $\alpha < \kappa$ so that

$$\mathbf{V}_{\kappa}^{\mathsf{M}} = \bigcup_{\alpha < \kappa} \mathbf{V}_{\alpha}^{\mathsf{M}} = \bigcup_{\alpha < \kappa} \mathbf{V}_{\alpha} = \mathbf{V}_{\kappa}.$$

Thus $V_{\kappa} \in M$. Now consider $x \subseteq V_{\kappa}$. Since $j \upharpoonright V_{\kappa} = id \upharpoonright V_{\kappa}$ again follows from Result 12 A • 3, we have by elementarity that $y \in j(x) \cap V_{\kappa}$ iff $y \in x$, which means that $j(x) \cap V_{\kappa} = x \in M$ and thus $V_{\kappa+1} \subseteq M$. \dashv

Hence the "strength" of a non-trivial, elementary embedding $j : V \to M$ is at least cp(j) + 1 in the sense that we always have $V_{cp(j)+1} \subseteq M$. It may be possible^{vi} for j to have a larger strength, but to do this, we would need *extenders* rather than mere ultrafilters (this also motivates the notion of a *strong* cardinal). Now while something like $V_{cp(j)+2}$ -strength is out of reach at the moment for measures, it seems rather innocuous. The following shows, however, that this gives agreement on $H_{cp(j)++}$, and this generalizes. To introduce some notation, for a cardinal κ , let $\kappa^{+\alpha}$ be the α th cardinal larger than κ . Explicitly, $\kappa^{+0} = \kappa$, $\kappa^{+\alpha+1} = (\kappa^{+\alpha})^+$, and $\kappa^{+\gamma} = \sup_{\alpha < \gamma} \kappa^{+\alpha}$ for limit $\gamma \in Ord$.

– 12 A•9. Result

Let $j : V \to M$ be traditional with $cp(j) = \kappa$ inaccessible^{vii}. Let $\alpha \in Ord$ and suppose $V_{\kappa+\alpha} \subseteq M$. Therefore, if $\kappa^{+\alpha}$ is regular, $H_{\kappa+\alpha} = H_{\kappa+\alpha}^{M}$.

Proof ...

We immediately have by upward absoluteness that $H_{\kappa^+\alpha}^{\mathsf{M}} \subseteq H_{\kappa^+\alpha}$. To show the other inclusion, proceed by induction on α firstly to show that $\kappa^{+\alpha}$ can be coded as a subset of $\mathfrak{P}^{\alpha}(\kappa) \subseteq V_{\kappa+\alpha}$, defined recursively by $\mathfrak{P}^{0}(\kappa) = \kappa, \mathfrak{P}^{\alpha+1}(\kappa) = \mathfrak{O}(\mathfrak{P}^{\alpha}(\kappa))$, and $\mathfrak{O}^{\gamma}(\kappa) = \bigcup_{\alpha < \gamma} \mathfrak{P}^{\alpha}(\kappa)$ for limit γ .

^{vi}This consistency of this can't be proven in ZFC alone, as such embeddings yield the existence of certain large cardinals, which in turn imply the consistency of ZFC.

^{vii}We will prove later that cp(j) is always inaccessible in V for $j : V \to M$ traditional.

Claim 1 For each $\beta \leq \alpha$, there is an injection $f_{\beta} : \kappa^{+\beta} \to \mathbb{P}^{\beta}(\kappa) \subseteq V_{\kappa+\beta}$ such that $f_{\beta} \in M$. In particular, $\kappa^{+\beta} = (\kappa^{+\beta})^{\mathsf{M}}$.

Proof .:.

Note that for $\beta < \kappa$, $\mathcal{O}^{\beta}(\kappa)^{\mathsf{M}} = \mathcal{O}^{\beta}(\kappa)^{\mathsf{V}}$. This clearly holds for $\beta = 0$. Inductively, let $f_{\gamma} : \kappa^{+\gamma} \to \mathcal{O}^{\gamma}(\kappa)$ be injective for $\gamma < \beta$ with $f_{\gamma} \in \mathsf{M}$ and $\kappa^{+\gamma} = (\kappa^{+\gamma})^{\mathsf{M}}$. Without loss of generality, since $|\mathcal{O}^{\gamma}(\kappa) \setminus \mathcal{O}^{\xi}(\kappa)| = |\mathcal{O}^{\gamma}(\kappa)|$ whenever $\xi < \gamma$, we might as well assume $\operatorname{im}(f_{\gamma}) \cap \operatorname{im}(f_{\xi}) = \emptyset$ whenever $\xi \neq \gamma$.

For successor $\beta + 1$, note that $(\kappa^{+\beta+1})^{\mathsf{M}} = ((\kappa^{+\beta})^+)^{\mathsf{M}}$ which inductively injects into $|\mathcal{O}^{\beta}(\kappa)|^+ \leq \mathcal{O}^{\beta+1}(\kappa)$. This yields an injection $f_{\beta+1} : (\kappa^{+\beta+1})^{\mathsf{M}} \to \mathcal{O}^{\beta}(\kappa)$ in M. So it suffices to show $(\kappa^{+\beta+1})^{\mathsf{M}} = \kappa^{+\beta+1}$, which can be shown just by proving $(\kappa^{+\beta+1})^{\mathsf{M}} \leq \kappa^{+\beta+1}$ since $\mathsf{M} \subseteq \mathsf{V}$. Any $\alpha < (\kappa^{+\beta})^+$ has a bijection $f : \alpha \to \kappa^{+\beta}$. This gives an well-order $R \subseteq \kappa^{+\beta} \times \kappa^{+\beta}$ of order-type α just by $R = \{\langle f(\xi), f(\zeta) \rangle : \xi < \zeta < \alpha\}$. Now using the definable coding of pairs of ordinals from Lemma 5 D • 2, we can code R as a subset of $\kappa^{+\beta}$ and hence a subset of $\mathcal{O}^{\beta}(\kappa)$: $A = f_{\beta}^{"}(\text{code "}R) \in \mathcal{O}^{\beta+1}(\kappa) \subseteq \mathcal{O}^{\alpha}(\kappa)$ so that $A \in \mathsf{M}$. Since the coding is definable and $f_{\beta} \in M$, we get that $R \in M$ of order-type α . Using a transitive collapse in M gives the bijection $f : \alpha \to \kappa^{+\beta}$. Hence $\alpha < (\kappa^{+\beta+1})^{\mathsf{M}}$. This completes the successor step.

For limit β , the following definition in **M** inductively agrees with the one in **V** to show that

$$f_{\beta}(\xi) = \begin{cases} \xi & \text{if } \xi < \kappa \\ f_{\gamma+1}(\xi) & \text{if } \kappa^{+\gamma} \le \xi < \kappa^{+\gamma+1} \text{ for some } \gamma < \beta \end{cases}$$

is an injection from $\kappa^{+\beta}$ to $\mathcal{O}^{\beta}(\kappa)$, and completes this initial induction.

Now we show that any element of $H_{\kappa^{+\alpha}}$ can be coded as a subset of $V_{\kappa^{+\alpha}}$ by coding it as a subset of $\kappa^{+\alpha}$ and hence as a subset of $\mathcal{O}^{\alpha}(\kappa) \subseteq V_{\kappa^{+\alpha}}$. For $\alpha = 0$, since κ is inaccessible, $H_{\kappa} = V_{\kappa}$ by Corollary 7 C • 10, and so by Result 12 A • 8, $H_{\kappa} = H_{\kappa}^{M}$.

Now suppose $\alpha > 0$ is such that $\kappa^{+\alpha}$ is regular. Let $x \in H_{\kappa^{+\alpha}}$ so that $|\operatorname{trcl}(\{x\})| < \kappa^{+\alpha}$ by the regularity of $\kappa^{+\alpha}$. Let $f : \operatorname{trcl}(\{x\}) \to \kappa^{+\beta}$ be injective for some $\beta < \alpha$. Now using the definable coding of pairs of ordinals, we can consider coding $(\operatorname{trcl}(x), \in)$ by a subset of $\kappa^{+\beta}$:

$$A = \{ \operatorname{code}(\langle f(a), f(b) \rangle) : a \in b \in \operatorname{trcl}(x) \} \subseteq \kappa^{+\beta}.$$

This subset contains all the information of x. Using $f_{\beta} : \kappa^{+\beta} \to \mathbb{O}^{\beta}(\kappa)$ from Claim 1, we get that $f_{\beta} "A \subseteq \mathbb{O}^{\beta}(\kappa)$ and hence $f_{\beta} "A \in \mathbb{O}^{\beta+1}(\kappa) \subseteq V_{\beta+1}$. Because $\beta + 1 \leq \alpha$, $V_{\beta+1}^{\mathsf{M}} = V_{\beta+1}$. Hence $f_{\beta} "A \in \mathsf{M}$. Since $f_{\beta} \in \mathsf{M}$, we can decode to get that A and hence, by a transitive collapse, trcl($\{x\}$) $\in \mathsf{M}$ so that $x \in \mathsf{M}$. By Claim 1, M and V agree on the cardinality of trcl($\{x\}$) and the calculation of $\kappa^{+\beta}$ yielding $x \in \mathrm{H}_{\kappa+\beta}^{\mathsf{M}}$.

These types of codings are very common when dealing with similar models of set theory, and it's important to be able to mimic them. But now that we have thought about elementary embeddings in general, let us return to the notion of an ultrapower.

§ 12 B. Characterizing ultrapowers

With all of this talk about elementary embeddings, we should perhaps note that we always have an elementary embedding from a model into its ultrapower.

- 12B•1. Theorem

Let σ be a signature. Let A be an FOL (σ) -model, and let U be an ultrafilter over a set K. Therefore A is elementarily embedded in Ult(A, U) by $x \mapsto [const_x]_U$, where $const_x : K \to A$ is the constant x map.

Сн II §12 В

-

Proof .:.

By Łoś's Theorem (12 • 2), Ult(A, U) $\models "\varphi([const_x])"$ iff for almost every $k \in K$, $A \models "\varphi(const_x(k))"$ (i.e. $A \models "\varphi(x)"$) which is just to say that the following set is in U:

$$\{k \in K : \mathbf{A} \vDash "\varphi(x)"\} = \begin{cases} K & \text{if } \mathbf{A} \vDash "\varphi(x)" \\ \emptyset & \text{otherwise.} \end{cases}$$

As an ultrafilter over $K, \emptyset \notin U$ and $K \in U$ so that $Ult(A, U) \models "\varphi([const_x])"$ iff $A \models "\varphi(x)"$.

Note that for a proper class like V, each equivalence class $[f] \in Ult(V, U)$ will be a proper class as well.^{viii} This can be rectified if we just consider $[f] = \{g : g \approx f \land rank(g) \text{ is minimal}\}$. Doing this, we get the usual equivalence class of f just intersected with some V_{α} for α least. Doing this, one still has that [f] = [g] for all $f \approx g$, and thus $x \mapsto [const_x]$ is a legitimate (class) function. So the result above also holds with proper classes too under this variant definition.

The existence of such an elementary embedding, however, doesn't tell you that it's nontrivial.

- 12B•2. Result — Let U be a *principal* ultrafilter over a set K. Therefore $A \cong Ult(A, U)$ by the canonical embedding.

Proof .:.

It suffices to show that the canonical embedding of Theorem 12 B • 1 is surjective. Let $u \in Ult(A, U)$ be arbitrary. We know that u = [f] for some $f : K \to A$. As U is principal, there is some $a \in A$ where $\{a\} \in U$. Hence $g \approx f$ iff g(a) = f(a). In particular, $u = [const_{f(a)}]$. Hence $x \mapsto [const_x]$ is a bijective embedding, and thus an isomorphism.

As the transitive collapse of a well-founded structure is unique, it follows that the collapsed version of the ultrapower Ult(V, U) (if well-founded) is precisely V, and it's not difficult to show inductively that in this case, $[const_x]$ is collapsed to x.

The importance of using ultrapowers, however, is when they are well-founded, because set theory in general is more concerned with transitive models. Transitive models are easier to work with, and are related to the actual universe of sets. Hence studying transitive models allows us to learn about the actual set-theoretic universe—though perhaps only conditional on large cardinal assumptions or other hypotheses. Well-founded models that satisfy extensionality are just an isomorphism away from transitive models by a mostowski collapse. In general, not all ultrapowers will be well-founded. Crucially, if we can take countably many conjunctions, then the ultrapower must be well-founded: we can collect together the countable amount of information that $f_0 \ni f_1$, and $f_1 \ni f_2$, and so on all at once. It's nice that we then have a characterization of the ultrapower being well-founded.

- 12B•3. Theorem

Let U be an ultrafilter in V. Therefore $Ult(V, U)$ is well-founded iff U is \aleph_1 -compl	ete.
--	------

Proof .:.

(←) Suppose U is \aleph_1 -complete, but Ult = Ult(V, U) is ill-founded. Let $\langle f_n : n \in \omega \rangle \in V$ be one such descending \in ^{Ult}-sequence in Ult: for every $n \in \omega$, Ult \models " $[f_{n+1}] \in [f_n]$ ". As U is \aleph_1 -complete in V,

$$\mathbf{V} \vDash \bigwedge_{n \in \omega} \forall^* \alpha \ (f_{n+1}(\alpha) \in f_n(\alpha))^n \quad \text{iff} \quad \mathbf{V} \vDash \forall^* \alpha \ \left(\bigwedge_{n \in \omega} f_{n+1}(\alpha) \in f_n(\alpha)\right)^n$$

But any such α yields an infinite, decreasing sequence $\langle f_n(\alpha) : n \in \omega \rangle$ in V, contradicting foundation.

 (\rightarrow) Now suppose Ult(V, U) is not \aleph_1 -complete. Let $\{X_n : n \in \omega\} \in \mathcal{P}(U)$ with $\bigcap_{n \in \omega} X_n \notin U$. From this,

viii To see this, let $v \in V$ be arbitrary. As U is an ultrafilter over K, let $X \in U$ be such that $X \neq K$ so that there is some $x \in K \setminus X$. Now for any $v \in V$, consider f_v to be the map sending every $y \in K$ to x except for x itself, which is sent to v, i.e. $f_v = (\text{const}_X \setminus \{\langle x, x \rangle\}) \cup \{\langle x, v \rangle\}$. Note that $\forall y \in X$ ($f_v(y) = \text{const}_X(y)$) so that $f_v \in [\text{const}_X]$. Also note that $f_v \neq f_{v'}$ for $v \neq v' \in V$ so that $[\text{const}_X]$ is a proper class.

we will construct an infinite, decreasing \in^{Ult} -sequence in Ult, contradicting foundation. Without loss of generality, assume $X_n \supseteq X_{n+1}$ just by replacing each X_n with $\bigcap_{i \le n} X_i$. Without loss of generality, U is an ultrafilter over a cardinal κ .

For each $\alpha < \kappa$, let index(α) be the least *n* for which $\alpha \notin X_n$. If there is no such *n*, then write index(α) = 0. For each $n \in \omega$, define the function $f_n : \kappa \to \omega$ by taking, for $\alpha < \kappa$,

$$f_n(\alpha) = \begin{cases} \operatorname{index}(\alpha) - n & \text{if index}(\alpha) \ge n \\ 0 & \text{otherwise.} \end{cases}$$

So in essence, $\langle f_n(\alpha) : n \in \omega \rangle$ will start at index (α) and decrease by 1 until it is eventually, constantly 0. As a result, if $\alpha \in X_n$, then index $(\alpha) > n$ and so $f_n(\alpha) > f_{n+1}(\alpha)$. As almost every α is in X_n , it follows that $\forall^* \alpha$ $(f_n(\alpha) > f_{n+1}(\alpha))$. So for each $n \in \omega$, Ult $\models [f_n] \in [f_{n+1}]^n$. Consequently, $\langle [f_n] : n \in \omega \rangle$ is a decreasing \in^{Ult} -sequence, meaning Ult is ill-founded.

Of course, this doesn't say that Ult(V, U) or V is necessarily *actually* well-founded, just that if V is well-founded—if we start from a well-founded class—then we still remain well-founded after taking the ultrapower. You might think that this result is obvious, since if Ult(V, U) is a class of V, and V thinks itself is well-founded, surely it must think this class is too. But the issue is the difference in interpretation of ' \in '. If Ult(V, U) is well-founded, then we can identify it with a transitive class of V, but otherwise, it's just some structure whose universe is a class of V.

- 12B•4. Definition Let U be an \aleph_1 -complete ultrafilter. Define cUlt(V, U) be transitive collapse of Ult(V, U) via π_U . Set $j_U : V \rightarrow$ cUlt(V, U) to be the canonical embedding: $j_U(x) = \pi_U([\text{const}_x])$.

This doesn't inherently tell us that this (collapsed) ultrapower is different from V, however, which was more obviously the case when U was principal. If V has no measurable cardinals, it will turn out that V has no \aleph_1 -complete ultrafilters, as such ultrafilters will actually be λ -complete for some (maximal) λ that turns out to be a measurable cardinal.^{ix}

The notion of completeness is also important as it determines the critical point of the canonical embedding.

- 12B•5. Theorem

Let U be a non-principal ultrafilter over some K that is κ -complete, but not κ^+ -complete for some cardinal $\kappa > \aleph_0$. Let $j : V \to \text{cUlt}(V, U)$ be the canonical embedding. Therefore, $j \neq \text{id}$ is traditional and $\text{cp}(j) = \kappa$.

Proof .:.

Let π : Ult(V, U) \rightarrow cUlt(V, U) be the collapsing isomorphism. First we show that $j \upharpoonright \kappa = id \upharpoonright \kappa$. To see this, suppose inductively that $j \upharpoonright \xi = id \upharpoonright \xi$ for some $\xi < \kappa$. We aim to show $j(\xi) = \xi$. By elementarity and the inductive hypothesis, $\alpha \in \xi$ iff $j(\alpha) = \alpha \in j(\xi)$ so that $j(\xi) \supseteq \xi$. So it suffices to show $\pi([const_{\xi}]) = j(\xi) \le \xi$.

So let $\zeta < j(\xi)$ be arbitrary. ζ can be represented in the ultrapower by some $f : K \to V$: $\zeta = \pi([f])$. Since Ult(V, U) \models "[f] < [const_{ξ}]", it follows that

$$\forall^* x \ (f(x) < \operatorname{const}_{\xi}(x)) \quad \text{iff} \quad \forall^* x \ (f(x) < \xi) \quad \text{iff} \quad \forall^* x \ \left(\bigvee_{\varepsilon < \xi} f(x) = \varepsilon\right).$$

Suppose for each $\varepsilon < \xi$ that $\neg \forall^* x \ (f(x) = \varepsilon)$ iff $\forall^* x \ (f(x) \neq \varepsilon)$. As $\xi < \kappa$, by κ -completeness,

$$\forall^* x \left(\bigwedge_{\varepsilon < \xi} f(x) \neq \varepsilon \right) \quad \text{iff} \quad \forall^* x \left(\neg \bigvee_{\varepsilon < \xi} f(x) = \varepsilon \right) \quad \text{iff} \quad \forall^* x \ (f(x) \neq \xi)$$

a contradiction. Hence there must be some $\varepsilon < \xi$ where $\forall^* x \ (f(x) = \varepsilon = \text{const}_{\varepsilon}(x))$. For this ε , we then have $\text{Ult}(\mathbf{V}, U) \models ``[f] = [\text{const}_{\varepsilon}]$ '' so after collapsing, $\zeta = \pi([f]) = j(\varepsilon) = \varepsilon < \xi$. This shows $j(\xi) \subseteq \xi$ and thus equality.

^{ix}This can be proven just by translating the ultrafilter U to a separate ultrafilter over λ according to how elements disappear from a λ -length \subseteq -decreasing sequence of elements of U.

To see that $j \neq id$, it suffices to show $j(\kappa) > \kappa$. This also shows that $cp(j) = \kappa$, since we already know $j \upharpoonright \kappa = id \upharpoonright \kappa$. To do this, we find a function $f : K \to \kappa$ sitting between every $[const_{\alpha}]$ and $[const_{\kappa}]$ in the ultrapower. To construct f, proceed as follows. As U is not κ^+ -complete, let $\langle X_{\alpha} : \alpha < \kappa \rangle$ witness this: $X_{\alpha} \in U$ for each $\alpha < \kappa$, but $\bigcap_{\alpha < \kappa} X_{\alpha} \notin U$. By subtracting this intersection we can assume without loss of generality that $\bigcap_{\alpha < \kappa} X_{\alpha} = \emptyset$. Furthermore, by using κ -completeness, each $\bigcap_{\xi < \alpha} X_{\xi} \in U$ so we can without loss of generality obtain a sequence where $X_{\alpha} \subseteq X_{\beta} \in U$ for $\beta < \alpha < \kappa$. Consider the map $f : K \to \kappa$ sending $x \in K$ to the least $\alpha < \kappa$ with $x \notin X_{\alpha}$. Now consider [f] in the ultrapower.

Note that for each α , almost every $x \in K$ is in X_{α} . In particular, for any fixed α , almost every $x \in K$ has $f(x) > \alpha$. So in the ultrapower, for each $\alpha < \kappa$, $Ult(V, U) \models "[f] > [const_{\alpha}]$ " so in taking the transitive collapse, $\pi([f]) > j(\alpha) = \alpha$. In particular, $\pi([f]) \ge \kappa$. But clearly, as f is a function from K to κ , $\forall^*x \ (f(x) < const_{\kappa}(x))$ and therefore $Ult(V, U) \models "[f] < [const_{\kappa}]$ ", meaning $\pi([f]) < j(\kappa)$. Hence $\kappa \le \pi([f]) < j(\kappa)$ so that $cp(j) = \kappa$.

We now aim to prove two main theorems about ultrapowers dealing with how we can factor embeddings through ultrapowers, and how we may represent the elements of ultrapowers. The idea is that an arbitrary traditional $j : V \to M$ can be partially coded through an ultrafilter U_j over cp(j) and thus through the ultrapower.



12B.6. Figure: Factoring through the ultrapower embedding

To derive an ultrafilter from j, note that for $\kappa = cp(j)$, most subsets of κ will be shot up beyond κ in the sense that $A \subseteq \kappa$ will likely have j(A) be unbounded in $j(\kappa) > \kappa$. In this sense, j(A) will have many more elements above those in A. The key thing for us is whether κ is in this stretched version of A, j(A). This clearly is answerable for any given subset of A, and by elementarity, will be preserved under the necessary operations.

12B•7. Definition Let $j : V \to M$ be elementary. Define the *ultrafilter derived from* j to be $U_j = \{A \subseteq cp(j) : cp(j) \in j(A)\}$.

As described above, it's not difficult to see that U_j is an ultrafilter. More importantly, U_j is actually a measure over cp(j).

12B•8. Result Let $j : V \to M$ be elementary with $cp(j) = \kappa$. Thus U_j is a measure over κ . (Moreover, the club filter $Club_{\kappa} \subseteq U_j$.)

Proof .:.

That U_j is an ultrafilter is easy enough to see as j is elementary: for $A \subseteq \kappa$ in $V, A \notin U_j$ iff $\kappa \notin j(A)$ iff $\kappa \in j(\kappa) \setminus j(A) = j(\kappa \setminus A)$ iff $\kappa \setminus A \in U_j$.

- U_j is easily seen to be non-principal. Otherwise, for some $\alpha < \kappa$, we'd have $\{\alpha\} \in U_j$ and hence $\kappa \in j(\{\alpha\})$. By elementarity of $j, j(\{\alpha\})$ has just one element: $j(\alpha) = \alpha \neq \kappa$, a contradiction.
- For κ -completeness, consider $\{A_{\alpha} : \alpha < \lambda\} \subseteq U_j$ in V for $\lambda < \kappa$. Since $\kappa \in j(A_{\alpha})$ for each $\alpha < \lambda$, $\kappa \in \bigcap_{\alpha < \lambda} j(A_{\alpha})$. Now since $\lambda < \kappa$, $j(\lambda) = \lambda$ and hence

$$\kappa \in \bigcap_{\alpha < \lambda} j(A_{\alpha}) = \bigcap_{\alpha < j(\lambda)} j(A_{\alpha}) = j\left(\bigcap_{\alpha < \lambda} A_{\alpha}\right)$$

• To show that U_i is normal, let $f: \kappa \to \kappa$ be such that $\forall^* \alpha$ ($f(\alpha) < \alpha$). This means

$$\kappa \in j(\{\alpha < \kappa : f(\alpha) < \alpha\}) = \{\alpha < j(\kappa) : j(f)(\alpha) < \alpha\}$$

and thus $j(f)(\kappa) < \kappa$. So there is some $\beta < \kappa$ with $j(f)(\kappa) = \beta = j(\beta)$ and hence $\forall^* \alpha \ (f(\alpha) = \beta)$.

 U_j extends the club filter, since being a club is a first-order property. Hence $C \in \text{Club}_{\kappa}$ has j(C) containing a club of $j(\kappa)$. Since $C \subseteq \kappa$, $C = j(C) \cap \kappa$, which contains a club of κ . As j(C) is closed, $\kappa = \sup C \in j(C)$. This means that $C \in U_j$ and thus $\text{Club}_{\kappa} \subseteq U_j$.

This actually proves the earlier claim of Figure 11 D•3: if U is a κ -complete ultrafilter over κ , then $j : V \to \text{cUlt}(V, U)$ has $\text{cp}(j) = \kappa$ and thus its derived ultrafilter U_j is a measure over κ .

— 12B•9. Theorem (Factoring)

Let $j : V \to M$ be elementary with $cp(j) = \kappa$. Let U_j be the derived ultrafilter, and let $j_{U_j} : V \to Ult(V, U_j)$ be the canonical ultrapower embedding. Therefore there is a (unique) elementary $k : Ult(V, U_j) \to M$ such that $j = k \circ j_{U_j}$ and $k([f]_{U_j}) = j(f)(\kappa)$.

Proof .:.

Write **Ult** for **Ult**(**V**, U_j). For each [f], define $k([f]) = j(f)(\kappa)$. Note that this is independent on the choice of f, since if $\forall^* \alpha$ ($f(\alpha) = g(\alpha)$), then by definition of U_j , $\kappa \in \{\alpha < j(\kappa) : j(f)(\alpha) = j(g)(\alpha)\}$ and so $k([f]) = j(f)(\kappa) = j(g)(\kappa) = k([g])$. Note also that $j = k \circ j_{U_j}$, since $k \circ j_{U_j}(x) = k([const_x]) = j(const_x)(\kappa) = const_{j(x)}(\kappa) = j(x)$.

To see that k as defined is elementary, let $\varphi(x)$ be a FOL(\in)-formula and suppose Ult \models " $\varphi([f])$ " for some $[f] \in$ Ult. By Łoś's Theorem (12 • 2), this happens iff $\forall^* \alpha \varphi(f(\alpha))$. By definition of U_j , this means $\kappa \in j(\{\alpha < \kappa : \varphi(f(\alpha))\})$, i.e. $\mathbf{M} \models$ " $\varphi(j(f)(\kappa))$ ". Rewritten, this says $\mathbf{M} \models$ " $\varphi(k([f]))$ ". Thus k is elementary. \dashv

There are a number of corollaries to this. Firstly, we have a nice theorem of how we can break down ultrapowers.

- 12B·10. Lemma Let U be a measure over κ . Therefore $\pi([\text{id} \upharpoonright \kappa]) = \kappa$ and $U_{j_U} = U$.

Proof .:.

• We know from Łoś's Theorem (12 • 2) that $Ult(V, U) \models "[id \upharpoonright \kappa] > [const_{\alpha}]$ " for each $\alpha < \kappa$. Hence in the collapse (as κ is the critical point), $cUlt(V, U) \models "\pi([id \upharpoonright \kappa]) > \alpha$ " for each $\alpha < \kappa$ and thus $\pi([id \upharpoonright \kappa]) \ge \kappa$. To show that $\pi([id \upharpoonright \kappa]) \le \kappa$, we appeal to normality.

Let $\alpha < \pi([\text{id} \upharpoonright \kappa])$ be arbitrary. Therefore, by Łoś's Theorem (12 • 2), $\forall^*\beta$ (const $_{\alpha}(\beta) < \beta$). So applying normality to const $_{\alpha}$, we get that there must be some particular $\gamma < \kappa$ where $\forall^*\beta$ (const $_{\alpha}(\beta) = \gamma = \text{const}_{\gamma}(\beta)$) so that $[\text{const}_{\alpha}] = [\text{const}_{\gamma}]$ and thus $\alpha = \gamma < \kappa$. Therefore $\pi([\text{id} \upharpoonright \kappa]) \subseteq \kappa$, and hence equal.

• It suffices to show that $A \in U$ iff $\kappa \in j_U(A)$. Rewritten, $A \in U$ says $\forall^* \alpha \ (\alpha \in \text{const}_A(\alpha))$ which is equivalent to $\mathsf{Ult}(\mathsf{V}, U) \vDash [\operatorname{id} \upharpoonright \kappa] \in [\operatorname{const}_A]^n$, meaning $\pi([\operatorname{id} \upharpoonright \kappa]) = \kappa \in j_U(A)$.

This has the consequence of showing a trivial version of the factor lemma when M is the ultrapower by a measure. But this allows us to think about the ultrapower and the "M" in the same way.

- 12B•11. Corollary

Let U be a measure over κ . Therefore $cUlt(V, U) = \{j_U(f)(\kappa) : f \in {}^{\kappa}V\}.$

Proof .:.

Let $j : V \to \text{cUlt}(V, U)$ be the canonical ultrapower embedding. By Factoring (12B•9), there is a unique, elementary $k : \text{Ult}(V, U_j) \to \text{cUlt}(V, U)$ which obeys $k([f]_{U_j}) = j(f)(\kappa)$. Since $U_j = U$ by Lemma 12B•10, $\text{cUlt}(V, U_j) = \text{cUlt}(V, U)$ so that k must just be the collapsing isomorphism. Hence every element of cUlt(V, U) can be represented in this way. \dashv

A slight generalization of this can be used for U that are merely κ -complete and not actually measures. The argument just replaces κ with [id $\restriction \kappa$]. In fact, Corollary 12 B • 11 is equivalent to a κ -complete ultrafilter U being normal.

- 12B•12. Theorem

Let $j : V \to M$ be traditional and a class with $cp(j) = \kappa$. Therefore, there is some ultrafilter U where M = Ult(V, U) with j as the canonical embedding iff there is some $s \in M$ where

$$\mathbf{M} = \{ j(f)(s) : f \in {}^{\kappa}\mathbf{V} \},\$$

in which case $U = \{A \subseteq \kappa : s \in j(A)\}.$

Proof .:.

Suppose M = cUlt(V, U) with j as the canonical embedding. Set s (the *seed*) to be $\pi_U([\text{id} \upharpoonright \kappa])$ where $\pi_U : \text{Ult}(V, U) \rightarrow \text{cUlt}(V, U)$ is the collapsing map. We know already that $M = \{\pi_U([f]_U) : f \in \kappa V\}$ so for $\pi_U([f]_U) \in M$ arbitrary, it suffices to show that $\pi_U([f]_U) = j(f)(s)$. Note that $\forall^* \alpha (f(\alpha) = f(\alpha))$. We can think of $f(\alpha)$ as coming from the map $\alpha \mapsto f(\alpha)$ or coming from the map $\alpha \mapsto (\text{const}_f(\alpha))(\text{id}(\alpha))$. Using these two interpretations, by Łoś's Theorem (12 • 2), we have that $\text{Ult}(V, U) \models "[f] = [\text{const}_f]([\text{id} \upharpoonright \kappa])$ ", meaning that in the collapse, recalling that $j(x) = \pi_U([\text{const}_x])$,

 $\pi_U([f]) = \pi_U([\operatorname{const}_f]([\operatorname{id} \upharpoonright \kappa])) = \pi_U([\operatorname{const}_f])(\pi_U([\operatorname{id} \upharpoonright \kappa])) = j(f)(s).$

Now suppose there is some $s \in M$ with $M = \{j(f)(s) : f \in {}^{\kappa}V\}$. Consider the ultrafilter $U = \{A \subseteq \kappa : s \in j(A)\}$. As in Result 12 B • 8, U can be easily shown to be an ultrafilter (by elementarity), and κ -complete (by elementarity and that $cp(j) = \kappa$). As in Factoring (12 B • 9), consider the map $k : Ult(V, U) \to M$ defined by $k([f]_U) = j(f)(s)$. To see that this is well defined, note that $[f]_U = [g]_U$ means $\forall^*\alpha (f(\alpha) = g(\alpha))$ implying $s \in j(\{\alpha < \kappa : f(\alpha) = g(\alpha)\})$ so that $s \in \{\alpha < j(\kappa) : j(f)(\alpha) = j(g)(\alpha)\}$ and thus j(f)(s) = j(g)(s).

This k is elementary by the same reasoning as in Factoring $(12 B \cdot 9)$:

$$\begin{aligned} \mathsf{Ult}(\mathsf{V},U) \vDash ``\varphi([f]_U" & \text{iff} \quad \forall^* \alpha \, \varphi(f(\alpha)) \\ & \text{iff} \quad s \in j(\{\alpha < \kappa : \varphi(f(\alpha))\} = \{\alpha < j(\kappa) : \mathsf{M} \vDash ``\varphi(j(f)(\alpha))"\} \\ & \text{iff} \quad \mathsf{M} \vDash ``\varphi(j(f)(s))". \end{aligned}$$

This implies k is injective and an embedding. So k is actually an ismorphism since it's clearly surjective: $M = \{j(f)(s) : f \in {}^{\kappa}V\}$ allow us to merely consider $k([f]_U) = j(f)(s)$ for any $f : \kappa \to V$. Hence cUlt(V, U) = M and by uniqueness, k is just the collapsing isomorphism, meaning j is the canonical ultrapower embedding. \dashv

As stated before, the $s = \pi_U([id \upharpoonright \kappa])$ being κ is equivalent to normality for κ -complete ultrafilters over $\kappa > \aleph_0$.

- 12B•13. Result

Let U be a non-principal, κ -comlete ultrafilter over $\kappa > \aleph_1$. Let $\pi_U : \text{Ult}(V, U) \to \text{cUlt}(V, U)$ be the collapsing isomorphism. Therefore U is normal iff $\pi_U([\text{id} \upharpoonright \kappa]) = \kappa$.

Proof .:.

Let $j : V \to Ult(V, U)$ be the canonical embedding which then has $cp(j) = \kappa$. Since U is κ -complete, it's clearly unbounded and thus $\forall^* \alpha \ (\alpha > \beta)$ for each β . Restated, this says $Ult(V, U) \models "[id \upharpoonright \kappa] > [const_{\beta}]$ " for each $\beta < \kappa$. So after collapsing,

 $\pi_U([\mathrm{id} \upharpoonright \kappa]) \supseteq \{\pi_U([\mathrm{const}_\beta]) : \beta < \kappa\} = \{j(\beta) : \beta < \kappa\} = \kappa.$

Note that a function f being regressive on a set in U is equivalent to $\forall^* \alpha$ (f(α) < α), meaning Ult(V, U) \models "[f] < [id \ κ]". Thus $\pi_U([id \ \kappa]) = \{\pi_U([f]) : f \text{ is regressive on a set in } U\}.$

So suppose U is normal. Note that every regressive function $f : \kappa \to \kappa$ has some $\beta < \kappa$ where $[f] = [\text{const}_{\beta}]$ and thus $\pi_U([f]) = \pi_U([\text{const}_{\beta}]) = j(\beta) = \beta$. Therefore $\pi_U([\text{id} \upharpoonright \kappa]) \subseteq \kappa$, and so we have equality.

Now suppose $\pi_U([id \upharpoonright \kappa]) = \kappa$. Thus every element of $\pi_U([id \upharpoonright \kappa])$ is an ordinal less than κ , meaning that every regressive function $f : \kappa \to \kappa$ has $\pi_U([f]) = \beta = \pi_U([const_\beta])$ for some $\beta < \kappa$. So $[f] = [const_\beta]$ and thus $\forall^* \alpha$ $(f(\alpha) = \beta)$, meaning U is normal. \dashv

§ 12 C. Properties of ultrapowers

So we have ultrapowers, and we know what they look like thanks to Theorem $12B \cdot 12$. What are some of their properties, however? The main goal of this subsection is now to look at what happens with the critical point of the ultrapower embedding: Where is it sent? How close can the ultrapower be to V? What are the combinatorial effects of taking an ultrapower? A complete answer to these questions won't be given here (if there even is such an answer). Instead, we will consider the following results.

— 12 C • 1. Result

Let U be a measure over κ . Let $j : V \to cUlt(V, U)$ be the canonical embedding. Therefore,

- 1. $\operatorname{cUlt}(V, U)$ is closed under κ -length sequences, meaning κ $\operatorname{cUlt}(V, U) \subseteq \operatorname{cUlt}(V, U)$.
- 2. $\operatorname{cUlt}(V, U)$ and V agree up to $\kappa + 1$, meaning $V_{\kappa+1} \subseteq \operatorname{cUlt}(V, U)$ but $V_{\kappa+2} \not\subseteq \operatorname{cUlt}(V, U)$;
- 3. $\mathfrak{P}(\kappa) = \mathfrak{P}(\kappa) \cap \text{cUlt}(V, U);$
- 4. $U \notin \text{cUlt}(V, U)$; and
- 5. $j(\kappa)$ is not a cardinal of V: $\kappa < 2^{\kappa} \le (2^{\kappa})^{\mathsf{cUlt}(\mathsf{V},U)} < j(\kappa) < (2^{\kappa})^+$.

To prove these from the ground up, we need some results about *measurable* cardinals which we have not introduced yet. Instead, just assume the following lemma.

- 12C·2. Lemma — Let κ have a measure U over it. Therefore, κ is strongly inaccessible.

Proof .:.

 κ is regular by κ -completeness of its measure. κ is uncountable by elementarity of j. To show that κ is a strong limit, suppose not, and let $\lambda < \kappa$ have $2^{\lambda} \ge \kappa$.

So consider family $\Lambda \subseteq \mathcal{O}(\lambda)$ be of size $|\Lambda| = \kappa$. Take a corresponding ultrafilter $W \subseteq \mathcal{O}(\Lambda)$ with U and the bijection with κ . This W, however, will not be κ -complete, contradicting that U is. To see this, for each $\alpha < \lambda$, consider

$$X_{\alpha} = \begin{cases} \{x \in \Lambda : \alpha \in x\} & \text{if this is in } U \\ \{x \in \Lambda : \alpha \notin x\} & \text{otherwise.} \end{cases}$$

By construction, $X_{\alpha} \in U$ for all $\alpha < \lambda$. The intersection of all these, by κ -completeness of U, is in W. But $\bigcap_{\alpha < \lambda} X_{\alpha}$ is a single subset of λ , contradicting nonprincipality.

- Proof of Result $12 C \cdot 1 \dots$

To save space, write M = cUlt(V, U) and Ult = Ult(V, U) with $\pi : Ult \to M$ the collapsing isomorphism.

1. Let $\vec{x} = \langle x_{\alpha} \in M : \alpha < \kappa \rangle$ be a κ -length sequence (in V). Represent $x_{\alpha} = \pi([f_{\alpha}])$ for $f_{\alpha} : \kappa \to V$. Consider the sequence (also in V) $\vec{f} = \langle f_{\alpha} : \alpha < \kappa \rangle$. Now we consider $j(\vec{f})$. By elementarity, $j(\vec{f})$ is a sequence of length $j(\kappa)$. Moreover, for every $\beta < \kappa$, $\forall^* \alpha$ ($\vec{f}(\beta)(\alpha) = f_{\beta}$) so by Łoś's Theorem (12 • 2),

 $\mathsf{Ult} \vDash ([\mathrm{const}_{\vec{f}}]([\mathrm{const}_{\beta}]) = [f_{\beta}], \quad \text{iff} \quad j(\vec{f})(\beta) = \pi([\mathrm{const}_{\vec{f}}])(\pi([\mathrm{const}_{\beta}])) = \pi([f_{\beta}]) = x_{\beta}.$

Thus $j(\vec{f}) \upharpoonright \kappa = \vec{x}$. As κ , $j(\vec{f}) \in M$, it then follows that $\vec{x} \in M$.

- 2. This follows by Result 12 A 8 and (4) below.
- 3. This follows by Result 12 A 8.
- 4. Every α < j(κ) has a representation [f] in Ult which then obeys ∀*β (f(β) < const_κ(β)), meaning we can assume without loss of generality that f : κ → κ. Since there are only 2^κ many such f, we have our surjection: F mapping f ↦ π([f]). Suppose U ∈ M so that for any f ∈ (^κκ)^M = ^κκ, we can form [f] and thus the map F within M. Hence M ⊨ "κ < j(κ) ≤ κ^κ = 2^κ", contradicting Lemma 12 C 2 since by elementarity, j(κ) is also strongly inaccessible.
- 5. By (3), 2^κ ≤ (2^κ)^M. We of course know κ < 2^κ by Cantor's theorem. We have j(κ) > (2^κ)^M because κ is a strong limit in V so that j(κ) is a strong limit in M. We have j(κ) < (2^κ)⁺ since the argument given in (4) tells us that there's a surjection from κ^κ = 2^κ to j(κ) in V.

Now all of this has been a kind of coded way of talking about measurable cardinals by way of their measures.

§ 12 D. Measurable cardinals

Although we have mentioned measurable cardinals before, they should be given a formal introduction. Measurable cardinals are important for their two equivalent characaterizations: having a measure, and being the critical point of an elementary embedding. Measurable cardinals will be quite large, and their importance is partly for the ultrapowers mentioned in the rest of this section, but also in motivating a canonical inner model L[U] to be introduced later.

— 12 D•1. Definition —

A cardinal $\kappa > \aleph_0$ is *measurable* iff there is a non-principal, κ -complete ultrafilter over κ .

Note that by the results above, there are several different characterizations of this.

— 12 D • 2. Result

Let $\kappa \geq \aleph_0$ be a cardinal. Therefore, the following are equivalent:

- 1. κ is measurable, i.e. $\kappa > \aleph_0$ has a non-principal, κ -complete ultrafilter over it.
- 2. κ has a measure over it.

3. κ is the critical point of an elementary $j : V \to M$, where M is a transitive class of V.

Proof .:.

Clearly (2) implies (1) with the only thing to check being that κ is uncountable. But normality implies this: suppose $\kappa = \aleph_0$ with U a measure over \aleph_0 . Consider $f : \omega \to \omega$ defined by f(0) = 0 and f(n) = n - 1 for n > 0. As $f(n) \ge n$ iff n = 0, it follows by uniformity that $\forall^*n \ (f(n) < n)$. So by normality, there is some $m < \omega$ where $\forall^*n \ (f(n) = m)$. But $f^{-1}(m) \subseteq \{m, m + 1\} \notin U$ by uniformity. Therefore $\kappa \neq \aleph_0$.

So suppose (1) holds: κ is measurable as witnessed by U. Therefore Ult(V, U) is well-founded since $\kappa \ge \aleph_1$: κ -completeness implies \aleph_1 -completeness. Hence the canonical embedding $j : V \to cUlt(V, U)$ has $cp(j) = \kappa$ by Theorem 12 B • 5 showing (3).

If (3) holds, the derived ultrafilter U_j is a measure over κ by Result 12 B • 8, yielding (2).

This equivalence of measurability and being a critical point is an important one in the sense that each characterization has various corollaries, and when combined they give a clearer picture of measurable cardinals. Consider the following consequences, for example, showing just how large measurables need to be. We already know that just one inaccessible goes beyond what ZFC can prove. In fact, the consistency of just any number of inaccessibles can't be proven relative to the consistency of any fewer number of them. Now consider how strong the existence of measurables is.

- 12 D•3. Corollary -

Let κ be measurable. A cardinal λ is *mahlo* iff $\{\theta < \lambda : \theta = |\theta| \text{ is inaccessible}\}$ is a stationary subset of λ . Therefore

- 1. κ is strongly inaccessible by Lemma 12 C 2;
- 2. κ is the κ th (strongly) inaccessible cardinal;
- 3. κ is the κ th mahlo cardinal;
- 4. κ has a measure by Result 12 D 2; and
- 5. κ has a measure that extends the club filter Club_{κ} by Result 12 B 8.

Proof .:.

Let U be a measure over κ , and let $j : V \to M$ be elementary with $M \subseteq V$ a transitive class.

Note that a cardinal λ being strongly inaccessible is downward absolute. So if κ is strongly inaccessible in V, then it is in M, meaning that M thinks that j(κ) has an inaccessible below it: κ. So for each α < κ, M ⊨ "∃x(x is inaccessible and α < x < j(κ))". So by elementarity, for each α < κ, V ⊨ "∃x(x is inaccessible and α < x < κ)". So the set of inaccessible cardinals below κ is unbounded in κ. As κ is regular, κ is the κth inaccessible.

3. Firstly, to see that κ is mahlo, take $j : V \to M \subseteq V$ elementary with $cp(j) = \kappa$. For any club

-

 $C \subseteq \kappa$, $j(C) \subseteq j(\kappa)$ is also club, and since $C = j(C) \cap \kappa$, it follows that $\kappa \in j(C)$ and thus $\mathbf{M} \models "j(C)$ has an inaccessible member". By elementarity and absoluteness, C has an inaccessible member so that the set of inaccessibles below κ is stationary and κ is mahlo.

 κ is still mahlo in **M**, since $\mathfrak{P}(\kappa) = \mathfrak{P}(\kappa) \cap \mathbf{M}$ meaning that **M** contains every club of κ as well as the stationary set of inaccessibles above. Hence being a stationary subset of κ is absolute between **M** and **V**. Thus the above j(C) contains a mahlo cardinal in **M**. By elementarity, *C* contains a mahlo cardinal in **V**, and thus the set of mahlos below κ is stationary, and thus κ is the κ th mahlo cardinal. \dashv

One might be tempted to apply the reasoning of Corollary $12 D \cdot 3$ to the property of being measurable, which would seem to indicate that any measurable cardinal κ would need to be the κ th measurable cardinal, or it seems at least there can't be a least measurable. To simplify the issue, let κ be the least measurable cardinal, and let $j : V \to M$ be traditional. It would seem that $j(\kappa)$ has a measurable below it, and thus κ does too, contradicting that κ is the least measurable. The issue is that κ might not be measurable in **M**, because we've thinned out the universe to $M \subseteq V$ such that it no longer contains a measure, as seen in Result $12 C \cdot 1$.

Moreover, **M**, being the collapsed ultrapower, has further properties that present limitations on the kinds of embeddings that can be realized by ultrapowers. The properties of being inaccessible, mahlo, and so forth could be used with the above reasoning, since they deal only at the level of V_{κ} and $V_{\kappa+1}$, but issues creep in if we try going beyond this, like the statement of being measurable. This is again a result of the agreement between the ultrapower and V as seen in Result $12 \text{ C} \cdot 1$.

Now despite the fact that the reasoning of Corollary 12 D • 3 breaks down when we try to apply them to the property of, for example, being measurable, the reasoning *does* apply when V = L. This is because of L being the smallest inner model: cUlt(V, U) = V = L which forces M to still recognize κ as measurable.

- 12 D \cdot 4. Theorem (L Has No Measurable Cardinals) -

Let κ be measurable. Therefore $V \neq L$.

Proof .:.

Without loss of generality, let κ be the least measurable cardinal, and assume V = L. By Result 12 D • 2, there is an elementary embedding $j : L \to M$ with a transitive class $M \subseteq L$. By elementarity,

 $M \models ZFC + "V = L" + "j(\kappa)$ is the least measurable".

Condensation implies M = L, and thus the two agree on κ : $M \models "\kappa$ is the least measurable", which contradicts that κ is the critical point of $j: \kappa \neq j(\kappa)$.

This is a relatively easy proof due to condensation, but there is a more complicated proof due to a more general result.

- 12 D • 5. Theorem (Kunen's Inconsistency Theorem) - Let $j : V \rightarrow V$ be traditional and a class. Therefore j = id.

Proof .:.

Assume $j \neq id$. By Result 12 A • 3, there is some critical point $\kappa = cp(j)$. By repeatedly applying j, we get the sequence $\langle j^n(\kappa) : n \in \omega \rangle$. Let $\theta = \sup_{n \in \omega} j^n(\kappa)$. By applying j to the sequence, by elementarity, we get that $j(\langle j^n(\kappa) : n \in \omega \rangle) = \langle j^{n+1}(\kappa) : n \in \omega \rangle$, and that $j(\theta) = \sup_{n \in \omega} j^{n+1}(\kappa) = \theta$. As a fixed point of j, this is good. Unfortunately, θ isn't regular. So instead consider the next cardinal, which by elementarity is also fixed: $j(\theta^+) = j(\theta)^+ = \theta^+$.

As θ^+ is regular, consider the stationary subset of ordinals with cofinality ω : $S = S_{\omega}^{\theta^+} = \{\alpha < \theta^+ : cof(\alpha) = \omega\}$. This can be closed under fix-points of j, since $j(\alpha) = sup(j''\alpha)$ for $cof(\alpha) < cp(j)$. The resulting set is also unbounded since $j''\theta^+ = \theta^+$. What this means is that

$$C = \{ \alpha < \theta^+ : \operatorname{cof}(\alpha) = \omega \land j(\alpha) = \alpha \}$$

is almost a club. In particular, C^+ —the closure of C under all sequences—is club in θ^+ with no new elements of cofinality ω . As a result, any stationary subset of S will intersect C.

But any stationary set of θ^+ may be partitioned into θ^+ stationary subsets. In particular, we can consider subsets $S_{\alpha} \subseteq S$ for $\alpha < \kappa$ —just take the union of S_0 with the guaranteed S_{α} for $\kappa \leq \alpha < \theta^+$ and make this the new S_0 —where all the S_{α} s are stationary and pairwise disjoint. Applying j, we get another sequence, this time of length $j(\kappa)$, of pairwise disjoint, stationary subsets of θ^+ : $\langle Z_{\alpha} : \alpha < j(\kappa) \rangle = j(\langle S_{\alpha} : \alpha < \kappa \rangle)$. By the above ideas on C^+ , $Z_{\kappa} \cap C^+ \neq \emptyset$. So there is some element $\zeta \in Z_{\kappa} \cap C^+$. As the S_{α} s partition S, there is also some $\alpha < \kappa$ with $\zeta \in S_{\alpha} \cap C^+$. But then $\zeta = j(\zeta) \in j(S_{\alpha}) = Z_{j(\alpha)}$. As $\alpha < \kappa = \operatorname{cp}(j)$, $j(\alpha) = \alpha$, yielding that $Z_{\alpha} \cap Z_{\kappa} \neq \emptyset$, a contradiction.

It's a good exercise to see where this proof breaks down for traditional $j : V \to M$. Note that this doesn't say that there can be no (non-trivial) $j : W \to W$ for $W \models ZFC$ a proper class,^x just that no j can exist as a class of V in this case. There are several other ways to state Kunen's Inconsistency Theorem (12 D • 5), one that is closer to the original form of the proof is below.

- 12 D • 6. Theorem (Kunen's Inconsistency Theorem Version 2) Let $j : V \to M$ be traditional and a class. Let $j(\lambda) = \lambda > cp(j)$. Therefore $j \upharpoonright \lambda \notin M$ and hence $V \neq M$.

Proof .:.

It suffices to show that $j"\lambda \notin M$ since as an embedding, $j \upharpoonright \lambda$ is just the increasing enumeration of $j"\lambda$. Without loss of generality, assume λ is the least fixed point above κ which, as before, takes the form $\sup_{n < \omega} j^n(\kappa)$ where $j^n(\kappa) = j(j(\cdots(j(\kappa))\cdots))$, just applying j n times. It follows that $(2^{<\lambda})^M = \lambda$ because λ is the limit of Mmeasurable cardinals. Since the cofinality of λ is ω , we have that $2^{\lambda} = \lambda^{\aleph_0}$ (any sequence in $^{\lambda}2$ is the ω -length supremum of things in $^{<\lambda}2 = \lambda$). The following useful claim is a theorem of Erdős and Hajnal.

There is a function $f : [\lambda]^{\omega} \to \lambda$ such that for every $X \subseteq \lambda$ with $|X| = \lambda$, $f''[X]^{\omega} = \lambda$.

Proof .:.

Claim 1

Consider eventual equivalence over $[\lambda]^{\omega}$, i.e. for $r, s \in [\lambda]^{\omega}$, $r \approx s$ iff $r \setminus \alpha = s \setminus \alpha \neq \emptyset$ for some $\alpha < \lambda$. It's not hard to check that \approx is an equivalence relation. So let c choose representatives for equivalence classes: $[r]_{\approx} = [c([r]_{\approx})]_{\approx}$ and let $g : [\lambda]^{\omega} \to \lambda$ be max $(r \bigtriangleup c([r]_{\approx}))$ (and 0 if they are the same), hence $c([r]_{\approx}) \setminus (g(r) + 1) = r \setminus (g(r) + 1)$. We work primarily with $g \upharpoonright [\Lambda]^{\omega}$ for some $\Lambda \subseteq \lambda$ and then may instead consider f by way of a bijection between Λ and λ .

Suppose there is no λ -sized $\Lambda \subseteq \lambda$ such that every $X \subseteq \Lambda$ with $|X| = |\Lambda|$ has $g''[X]^{\omega} \supseteq \Lambda$. For $n < \omega$, inductively define

- 1. $\Lambda_0 = \lambda \setminus 1$ and $\alpha_0 = 0$.
- 2. $\Lambda_{n+1} \subseteq \Lambda_n$ is arbitrary of size λ .
- 3. $\alpha_{n+1} \in \Lambda_n$ with $\alpha_{n+1} > \alpha_n$ witnessing the hypothesis for $\Lambda = \Lambda_n$ and $X = \Lambda_{n+1}$: $\alpha_{n+1} \notin g''[\Lambda_{n+1}]^{\omega}$.

But then $r = \{\alpha_n : n < \omega\} \in [\lambda]^{\omega}$ and we can consider $s = r \setminus c([r]_{\approx}) \in [\lambda]^{\omega}$. Since $s \approx r$, $c([s]_{\approx}) = c([r]_{\approx})$ and hence the place where they differ must be an element of $s \subseteq r$, not $c([s]_{\approx})$. In fact, g(s) is just $\max(c([r]_{\approx}) \setminus r) \in r$. So let $g(s) = \alpha_n$ and note that then $\alpha_n \in g''[\Lambda_n]^{\omega}$, a contradiction with (3) (and (1)).

So such a $\Lambda \subseteq \lambda$ of size λ exists. Through a bijection $b : \lambda \to \Lambda$ we can define f by f(x) = g(b''x). \dashv

So let f be as in Claim 1. Since j is elementary and λ, ω are both fixed points of j, j(f) also acts as in Claim 1 in M. So if we consider $j"\lambda \subseteq \lambda$ which has size λ , we get some $r \in [j"\lambda]^{\omega}$ such that j(f)(r) = cp(j). Since $r \subseteq j"\lambda$, we can enumerate $r = \{j(t_n) : n < \omega\} = j(\{t_n : n < \omega\}) = j(t)$ for $t = \{t_n : n < \omega\} \in [\lambda]^{\omega}$. It

^xIn particular, if there is a measurable cardinal, then there are such embeddings from inner models into themselves.

follows that j(f)(r) = j(f)(j(t)) = j(f(t)) = cp(j), which is a contradiction: f(t) would be an ordinal but $j(\alpha) < cp(j)$ for $\alpha < cp(j)$ while j(cp(j)) > cp(j).

§ 12 E. A first look at iterated ultrapowers

If we have a measure U, we can get an elementary embedding $j_{0,1} : V \to M_1 = \text{cUlt}(V, U) \subseteq V$ where $M_1 \models$ " $j_{0,1}(U)$ is a measure". Hence we get an elementary embedding $j_{1,2} : M_1 \to M_2 \subseteq M_1$ which then has

 $M_2 \models "j_{1,2}(j_{0,1}(U))$ is a measure".

Defining $j_{0,2}$ as $j_{1,2} \circ j_{0,1}$ means that $j_{0,2} : V \to M_2$ is elementary. We can keep doing this procedure for all $n < \omega$: getting a directed system of embeddings and transitive models $(M_n, j_{n,m} : n \le m \in \omega)$ such that for $n, m < \omega$, we define

- $M_0 = V, U = U_0$ is a measure over κ_0 in V;
- $M_{n+1} = \text{cUlt}^{M_n}(M_n, U_n)$ with $j_{n,n+1} : M_n \to M_{n+1}$ the canonical embedding;
- $j_{n,n} = \text{id}$, and $j_{n,m+1} = j_{m,m+1} \circ j_{n,m}$, thus defining $j_{n,m}$ whenever $n \le m < \omega$; and
- $\kappa_n = j_{0,n}(\kappa) = \operatorname{cp}(j_{n,n+1})$, and $U_n = j_{0,n}(U)$ which M_n thinks is a measure over κ_n .

This can be visuallized with the following figure.



12 E \cdot 1. Figure: Iterated Ultrapowers by U

From techniques of model theory, this system yields a direct limit, \mathbf{M}_{ω} , and corresponding limit embeddings, which then show that $\mathbf{M}_{\omega} \models "j_{0,\omega}(\kappa)$ is measurable". The non-trivial fact that this direct limit is well-founded is what allows us to continue up through all of the ordinals.

Well-foundedness of the limit isn't trivial, and it's important that we're using measures on different cardinals: U_{n+1} is a measure over $\kappa_{n+1} > \kappa_n$. More precisely, if we had a directed system of transitive models and their elementary embeddings $\langle \mathbf{M}_n, j_{n,m} : n < m \in \omega \rangle$ where each $j_{n,m}$ has the same critical point κ , then with direct limit embeddings $j_{n,\omega} : \mathbf{M}_n \to \mathbf{M}_\omega$, $\langle j_{n,\omega}(\kappa) : n < \omega \rangle$ is an infinite decreasing sequence of ordinals of \mathbf{M}_ω , showing \mathbf{M}_ω isn't well-founded.^{xi}

If the α th (linear) iterated ultrapower of **M** by U is well-founded, we will usually write the collapse as $\text{cUlt}_{\alpha}(\mathbf{M}, U)$ with $\text{cUlt}_{0}(\mathbf{M}, U) = \mathbf{M}$. The corresponding embedding will be $j_{0,\alpha}$. For the sake of space, we will often write just cUlt_{α} or $\text{cUlt}_{\alpha}^{\mathbf{M}}$ if the meaning is clear. Similarly, we will often write $\kappa_{\alpha} = j_{0,\alpha}(\kappa)$ if U is a measure over $\kappa \in \mathbf{M}$, and $U_{\alpha} = j_{0,\alpha}(U)$ for the corresponding measure.

This whole process defines a *linear* iteration: M_n embedds into M_m for $n \le m$. Later on, we'll consider iteration trees, where M_{n+1} might not be the ultrapower of M_n , but of some M_{n*} for an $n^* < n$.

^{xi}Such a sequence isn't in the direct limit, it just shows in V (and M_n for every $n < \omega$) that there's an infinite \in^{M_ω} -decreasing sequence.

- 12 E • 2. Definition Let κ be measurable with measure U. Let $\lambda \leq Ord$. Therefore the λ -length linear iteration of V by U is the system $\langle \mathsf{cUlt}_{\alpha}(\mathsf{V}, U) = \mathsf{cUlt}_{\alpha}, j_{\alpha,\beta} : \alpha \leq \beta < \lambda \rangle$ such that for all $\xi \leq \alpha \leq \beta < \lambda$,

- $cUlt_0 = V;$
- $\text{cUlt}_{\alpha+1} = \text{cUlt}^{\text{cUlt}_{\alpha}}(\text{cUlt}_{\alpha}, j_{0,\alpha}(U))$, and $j_{\alpha,\alpha+1}$ is the canonical ultrapower embedding;
- $cUlt_{\alpha} = dir \lim_{\xi < \alpha} cUlt_{\xi}$ for α a limit, and $j_{\xi,\alpha}$ is the direct limit embedding; and
- $j_{\alpha,\beta}$: cUlt_{α} \rightarrow cUlt_{β} is elementary with $j_{\alpha,\beta} \circ j_{\xi,\alpha} = j_{\xi,\beta}$ (and $j_{\alpha,\alpha} = id$).

Note that if we have an λ -length iteration of V by U a measure over κ , $\mathsf{cUlt}_{\alpha} \models "j_{0,\alpha}(U)$ is a measure over $j_{0,\alpha}(\kappa)$ " for all $\alpha < \lambda$. We know by Theorem 12 B • 3 that cUlt_{α} is well-founded for successor α , but cUlt_{α} is also well-founded for limit α . The proof of this is not immediate, however, and relies on the idea of taking ultrapowers within ultrapowers. So if $\mathbf{M} \subseteq \mathbf{V}$, the ultrapowers of \mathbf{M} by $U \in \mathbf{M}$ as calculated by \mathbf{M} are written $\mathsf{cUlt}_{\alpha}^{\mathbf{M}}$.

Here $\text{Ult}^{\mathsf{M}}(N, U) = \{[f]_{\approx_U} : f \in \mathsf{M} \land f : \kappa \to N^{\mathsf{M}}\}\)$, so we consider functions in M. For the most part, we will focus on ultrapowers of the form $\text{Ult}^{\mathsf{M}}(\mathsf{M}, U)$. The major result about taking ultrapowers within ultrapowers is that the resulting sequence is the tail of the sequence starting from V .

— 12 E • 3. Lemma (The Factor Lemma)

Let U be a measure over κ . Let $\alpha, \beta \in \text{Ord.}$ Therefore $\mathsf{cUlt}_{\alpha}^{\mathsf{cUlt}_{\beta}} = \mathsf{cUlt}_{\beta+\alpha}$ and $j_{0,\alpha}^{\mathsf{cUlt}_{\beta}} = j_{\beta,\beta+\alpha}$. In particular, $j_{0,\alpha} = j_{0,\beta} \circ j_{0,\alpha-\beta}^{\mathsf{cUlt}_{\beta}}$ when $\alpha > \beta$.

Proof .:.

Proceed by induction on α . Clearly the result holds for $\alpha = 0$, since by definition $\text{cUlt}_{\beta} = \text{cUlt}_{0}^{\text{cUlt}_{\beta}}$. Similarly, for α a limit, by the inductive hypothesis, the equalities hold for the direct limits.

For $\alpha + 1$ a successor, by the inductive hypothesis $\mathsf{cUlt}_{\beta+\alpha} = \mathsf{cUlt}_{\alpha}^{\mathsf{cUlt}_{\beta}}$. Let $U_{\beta+\alpha} = j_{0,\beta+\alpha}(U) = j_{0,\alpha}^{\mathsf{cUlt}_{\beta}}(U)$ where $j_{0,\beta+\alpha} : V \to \mathsf{cUlt}_{\beta+\alpha}$. Note that both ultrapowers use this measure:

$$\begin{aligned} \mathsf{cUlt}_{\alpha+1}^{\mathsf{cUlt}_{\beta}} &= \mathsf{cUlt}^{\mathsf{cUlt}_{\alpha}^{\mathsf{cUlt}_{\beta}}}(\mathsf{cUlt}_{\alpha}^{\mathsf{cUlt}_{\beta}}, j_{0,\alpha}^{\mathsf{cUlt}_{\beta}}(U)) \\ &= \mathsf{cUlt}^{\mathsf{cUlt}_{\beta+\alpha}}(\mathsf{cUlt}_{\beta+\alpha}, j_{0,\beta+\alpha}(U)) = \mathsf{cUlt}_{\beta+\alpha+1}. \end{aligned}$$

Moreover, the ultrapower embeddings are the same: $j_{\alpha,\alpha+1}^{\text{cUlt}_{\beta}} = j_{\beta+\alpha,\beta+\alpha+1}$, since we're just taking constant maps from the same universe $\text{cUlt}_{\beta+\alpha}$, and the collapsing map π : $\text{Ult}^{\text{cUlt}_{\beta+\alpha}}(\text{cUlt}_{\beta+\alpha}, j_{0,\beta+\alpha}(U)) \rightarrow \text{cUlt}_{\beta+\alpha+1}$ is unique. By the inductive hypothesis, $j_{0,\alpha}^{\text{cUlt}_{\beta}} = j_{\beta,\beta+\alpha}$ and so finally

$$j_{0,\alpha+1}^{\mathsf{cUlt}_{\beta}} = j_{\alpha,\alpha+1}^{\mathsf{cUlt}_{\beta}} \circ j_{0,\alpha}^{\mathsf{cUlt}_{\beta}} = j_{\beta+\alpha,\beta+\alpha+1} \circ j_{\beta,\beta+\alpha} = j_{\beta,\beta+\alpha+1}. \quad \dashv$$

As a result, we can get a better understanding of the direct limit ultrapowers, because we can approach them from ostensibly different ways. The Factor Lemma $(12 \text{ E} \cdot 3)$ tells us that these ways are all equivalent. So when we pull back the direct limit to a previous ultrapower, we can push forward and end up back at the same place. This idea is exemplified in the proof that the direct limits are well-founded.

12 E • 4. Theorem (The Wellfoundedness of Iterated Ultrapowers) Let U be a measure over κ , and let $\alpha \in Ord$. Therefore the α th ultrapower $Ult_{\alpha}(V, U)$ is well-founded.

Proof .:.

For $\beta \leq \gamma$, let $e_{\beta,\gamma}$: $\text{cUlt}_{\beta}(V, U) \rightarrow \text{Ult}_{\gamma}(V, U)$ be elementary. Let $j_{\beta,\gamma}$: $\text{cUlt}_{\beta}(V, U) \rightarrow \text{cUlt}_{\gamma}(V, U)$ be the usual elementary map so that if π_{γ} : $\text{Ult}_{\gamma}(V, U) \rightarrow \text{cUlt}_{\gamma}(V, U)$ is the collapsing map, $j_{\beta,\gamma} = \pi_{\gamma} \circ e_{\beta,\gamma}$. We unfortunately can't consider $j_{\beta,\alpha}$ if $\text{Ult}_{\alpha} = \text{Ult}_{\alpha}(V, U)$ is ill-founded, so we must consider $e_{\beta,\alpha}$ instead. Note we still have the same sort of factoring: $e_{\beta,\alpha} = e_{\xi,\alpha} \circ j_{\beta,\xi}$ whenever $\beta \leq \xi \leq \alpha$.

All successors are well-founded clearly by Theorem 12B•3, so assume α is the least limit where Ult_{α} is

Without loss of generality (just take a sufficiently large ordinal), let $x_0 = e_{0,\alpha}(\xi)$ where $\xi \in \text{Ord}$ is the least such that $\langle e_{0,\alpha}(\xi), \in^{\text{Ult}_{\alpha}} \rangle$ is ill-founded (from the perspective of **V**). Thus

$$\mathbf{V} \vDash ``\forall \alpha' \leq \alpha \ \forall \xi' < \xi \ \left(\langle e_{0,\alpha'}^{\mathsf{V}}(\xi'), \in^{\mathsf{Ult}_{\alpha'}} \rangle \text{ is well-founded} \right)`'.$$

Note by The Factor Lemma (12 E • 3) that $e_{0,\alpha'}^{\mathsf{cUlt}_{\beta}} = e_{\beta,\beta+\alpha'}^{\mathsf{V}}$. So by elementarity, for any $\beta < \alpha$,

$$\mathsf{cUlt}_{\beta} \vDash ``\forall \alpha' \le j_{0,\beta}^{\vee}(\alpha) \ \forall \xi' < j_{0,\beta}^{\vee}(\xi) \ \left(\left\langle e_{\beta,\beta+\alpha'}^{\vee}(\xi'), \in^{\mathsf{cUlt}_{\beta+\alpha'}} \right\rangle \text{ is well-founded} \right) ``. (*)$$

As the direct limit, let $x_1 = e_{\beta,\alpha}(\xi')$ for some $\xi' \in \text{cUlt}_{\beta}$ and $\beta < \alpha$. By the factor lemma, this means

$$\mathsf{lt}_{\alpha} \vDash "e_{\beta,\alpha}(\xi') = x_1 < x_0 = e_{0,\alpha}(\xi) = e_{\beta,\alpha} \circ j_{0,\beta}(\xi)"$$

So by elementarity, $\xi' < j_{0,\beta}(\xi)$. Write α' for the ordinal such that $\beta + \alpha' = \alpha$ so $\alpha' \le \alpha \le j_{0,\beta}(\alpha)$. Therefore $\alpha' \le \alpha$ and $\xi' < j_{0,\beta}(\xi)$ yield by (*)

$$\mathsf{cUlt}_{\beta} \models ``\langle j^{\mathsf{V}}_{\beta,\beta+\alpha'}(\xi'), \in^{\mathsf{cUlt}_{\beta+\alpha'}} \rangle = \langle j^{\mathsf{V}}_{\beta,\alpha}(\xi'), \in^{\mathsf{Ult}_{\alpha}} \rangle \text{ is well-founded''.}$$
(**)

But $\langle x_n : n \in \omega \setminus 1 \rangle$ is $\in^{\mathsf{Ult}_{\alpha}}$ -decreasing with $x_1 = j_{\beta,\alpha}^{\mathsf{V}}(\xi')$. Since well-foundedness is absolute between transitive models like $\mathsf{cUlt}_{\beta} \subseteq \mathsf{V}$, we can't have (**). \dashv



12E.5. Figure: Proof of The Wellfoundedness of Iterated Ultrapowers (12E.4)

The general idea of the very notation-heavy proof can be seen with Figure 12 E • 5.

Due to the concreteness of the definition of the iterated ultrapowers, we have a useful characterization of the resulting ultrafilters—at least at limit stages. Of course, these sets won't be ultrafilters in V, but the point stands that we have a conceptually simple way of identifying sets in them. We know already that $x \in U_0$ iff $\kappa_0 \in j_{U_0}(x)$, and there is a natural generalization of this.

12E•6. Lemma Let U be a measure over κ . Let $\lambda \in Ord$ be a limit ordinal and $x \in cUlt_{\lambda}(V, U) = cUlt_{\lambda}$. Write $\kappa_{\alpha} = j_{0,\alpha}(\kappa)$, $U_{\alpha} = j_{0,\alpha}(U)$, and $x_{\alpha} = j_{\alpha,\lambda}^{-1}(x)$ (if it exists) for $\alpha \leq \lambda$. Therefore the following are equivalent. 1. $x \in U_{\lambda}$. 2. $\kappa_{\alpha} \in x$ for some $\alpha < \lambda$ such that $j_{\alpha,\lambda}^{-1}(x)$ exists. 3. $\kappa_{\alpha} \in x$ for all sufficiently large $\alpha < \lambda$. *Proof* .:. (1) \rightarrow (2) Note that λ is a limit ordinal, which means $x_{\alpha} = j_{\alpha,\lambda}^{-1}(x)$ exists for some $\alpha < \lambda$. Applying $j_{\alpha,\lambda}^{-1}$,

 $x \in U_{\lambda}$ iff $\text{cUlt}_{\lambda} \models "\kappa_{\lambda} \in j_{\lambda,\lambda+1}(x)$ " iff $\text{cUlt}_{\alpha} \models "\kappa_{\alpha} \in j_{\alpha,\alpha+1}(x_{\alpha})$ ". Note that $\kappa_{\alpha} < \text{cp}(j_{\alpha+1,\lambda})$. So by applying $j_{\alpha+1,\lambda}$ and The Factor Lemma (12 E • 3), we get the desired result: $\kappa_{\alpha} \in j_{\alpha+1,\lambda}(j_{\alpha,\alpha+1}(x_{\alpha})) = x$.

(2) \rightarrow (3) Proceed by induction. Consider the least $\alpha < \lambda$ such that $\kappa_{\alpha} \in x$ and x can be pulled back to cUlt_{α} .

Сн II §12 Е

Let $j_{\alpha,\lambda}(x_{\alpha}) = x$. Applying $j_{\alpha+1,\lambda}^{-1}$, we get by elementarity

 $\mathsf{cUlt}_{\lambda} \vDash ``\kappa_{\alpha} \in x"$ iff $\mathsf{cUlt}_{\alpha+1} \vDash ``\kappa_{\alpha} \in x_{\alpha+1}"$

since $\kappa_{\alpha} < cp(j_{\alpha+1,\lambda})$. Therefore $x_{\alpha} \in U_{\alpha}$. So applying $j_{\alpha,\beta}$, we get that $x_{\beta} \in U_{\beta}$ for all $\alpha < \beta < \lambda$. But then $\kappa_{\beta} \in x_{\beta+1}$ for all $\alpha < \beta < \lambda$. Applying $j_{\beta+1,\lambda}$ yields by elementarity that $\kappa_{\beta} \in x$ since $\kappa_{\beta} < cp(j_{\beta+1,\lambda})$. Thus $\kappa_{\beta} \in x$ for all $\beta \ge \alpha$.

(3) \rightarrow (1) As λ is a limit, x can be pulled back to some $x_{\alpha+1} \in \text{cUlt}_{\alpha+1}$ meaning $j_{\alpha+1,\lambda}(x_{\alpha+1}) = x$. By (3), we can assume without loss of generality that $\kappa_{\alpha} \in x$ and hence $\kappa_{\alpha} \in x_{\alpha+1}$, just by applying $j_{\alpha+1,\lambda}$. Therefore $x_{\alpha} \in U_{\alpha}$ so by elementarity, $x \in U_{\lambda}$.

We also get some expected properties of the sequence $\langle \kappa_{\alpha} : \alpha \in \text{Ord} \rangle$.

- 12 E • 7. Result Let U be a measure over κ . Let $\kappa_{\alpha} = j_{0,\alpha}(\kappa)$ for $\alpha \in \text{Ord}$. Therefore the sequence $\langle \kappa_{\alpha} : \alpha \in \text{Ord} \rangle$ is increasing, and continuous.

Proof .:.

That the sequence is increasing is clear by Theorem 12 B • 5: $\kappa_{\alpha} = cp(j_{\alpha,\alpha+1})$ and thus $\kappa_{\alpha+1} = j_{\alpha,\alpha+1}(\kappa_{\alpha}) > \kappa_{\alpha}$. To see that the sequence is continuous, we already know since the sequence is increasing that $\kappa_{\alpha} \ge sup_{\beta < \alpha} \kappa_{\beta}$ for α a limit. So let $\lambda < \kappa_{\alpha}$ be arbitrary. As the direct limit, write $\lambda = j_{\beta,\alpha}(\gamma)$ for some $\beta < \alpha$ and γ . By elementarity, $\gamma < \kappa_{\beta}$ and thus $\lambda = j_{\beta,\alpha}(\gamma) = \gamma < \kappa_{\beta} < \kappa_{\alpha}$.

The final result of this section, analogous to Result $12 \text{ A} \cdot 8$, is an easy corollary to The Factor Lemma ($12 \text{ E} \cdot 3$).

- 12E•8. Corollary Let *U* be a measure over κ . Let $\kappa_{\alpha} = j_{0,\alpha}(\kappa)$ where $\text{cUlt}_{\alpha} = \text{cUlt}_{\alpha}(V, U)$. Let $x \in \mathcal{P}(\kappa_{\alpha}) \cap \text{cUlt}_{\alpha}$. Therefore $x = \kappa_{\alpha} \cap j_{\alpha,\beta}(x)$.

Section 13. Introducing Extenders

Put mildly, an extender is a system of ultrafilters that work nicely with each other. The point of extenders is to witness various large cardinal properties in the same way that a measure witnesses the existence of a measurable. Again, we will have an association with elementary embeddings, and in some sense the extender gives a natural way of extending certain smaller models to larger ones.

As many of the standard resources for extenders do not go into such detail, we will attempt to be fairly thorough here, proving in at least a hand-wavy detail each of the claims made. Further results, as usual, are left as guided exercises at the end of the section. The general approach with this section on extenders is to give two types of extenders: extenders derived from an elementary embedding, and extenders which merely satisfy certain first-order properties. Then we show these two types are really the same thing. Then we investigate some basic properties. To begin, we consider extenders derived from elementary embeddings before considering them in general.

To motivate extenders, consider a set of measures U_r on cardinals κ_r associated to each finite subset $r \in [\lambda]^{<\omega}$. We can then consider the ultrapowers $\text{cUlt}(V, U_r)$ for each $r \in [\lambda]^{<\omega}$. How do these ultrapowers interact, however? An extender has these measures as "nice" in that if $r \subseteq s$, then we have a very natural translation between the sets of U_r and the sets of U_s in a way that induces an elementary embedding $j_{r,s}$: $\text{cUlt}(V, U_r) \rightarrow \text{cUlt}(V, U_s)$. The fact that we're using finite subsets of λ tells us that we have a directed system: for any two $r, s \in [\lambda]^{<\omega}$, there's a common extension $r, s \subseteq t \in [\lambda]^{<\omega}$. This will tell us that we have a directed system of elementary embeddings between the ultrapowers and subsequently we can take the direct limit which we call $\text{cUlt}_E(V)$ where E is the extender composed of these $U_r s$. The importance of this extender ultrapower is that it can have nicer properties than the ultrapower by any measure, and derived extenders are able to talk about *more* of an elementary embedding than a derived measure is.

Depending on what properties we want the extender ultrapower to have, it can often suffice to use some fixed $\kappa = \kappa_r$ for every $r \in [\lambda]^{<\omega}$.^{xii} Such extenders are called *short* extenders because we end up with $\lambda \leq j_E(\kappa)$. Long extenders have $\kappa_r > \kappa$ and as a result can have $j_E(\kappa) > \lambda$.

§ 13 A. Extenders derived from an elementary embedding

The formal definition of an extender is both complicated, and ill-motivated at this point. To grasp some of the fundamental properties, we will begin with a simple "example" of an extender, being one derived from an elementary embedding. In the following motivation, the requirement $\lambda \leq j(\kappa)$ makes this a short extender, and is here mostly for simplicity.

Suppose $j : V \to M$ is traditional and a class. Let $\kappa = cp(j)$ and let λ be such that $\kappa < \lambda \le j(\kappa)$. The usual definition of the (κ, λ) -extender derived from j is merely

$$E_{\lambda}^{j} = \{ \langle r, X \rangle \in [\lambda]^{<\omega} \times \mathcal{P}([\kappa]^{<\omega}) : r \in j(X) \},\$$

 $E = E_{\lambda}^{j}$ generalizes the derived measure which can be thought of as $U_{j} = \{\langle \{\kappa\}, X \rangle \in [\kappa+1]^{<\omega} \times \mathcal{O}(\kappa) : \kappa \in j(X)\}$. For each $r \in [\lambda]^{<\omega}$, we get the slice E_{r} of E as an ultrafilter over $[\kappa]^{<\omega}$, as we will show. In fact, each E_{r} will be κ -complete. To get a better intuition on what these E_{r} look like, it's not difficult to show that $\forall_{E_{r}}^{*}t \ (|t| = |r| < \aleph_{0})$.

But the point is that an elementary embedding gives all sorts of κ -complete, non-principal ultrafilters defined in this way and moreover, the resulting ultrapowers are well-founded. Let's take a moment to examine these ultrapowers without actually proving any of the statements yet. Again, as with Factoring (12 B • 9), we can define an embedding k_r : cUlt(V, E_r) \rightarrow M by $k_r(\pi_r([f]_{E_r})) = j(f)(r)$, when f is a function with domain $[\kappa]^{<\omega}$ and π_r is the collapsing

^{xii}or at least for technical reasons $r \in [\lambda]^{<\omega} \setminus [\kappa]^{<\omega}$

isomorphism for $Ult(V, E_r)$. The result is again that $j = k_r \circ j_r$, where $j_r : V \to cUlt(V, E_r)$ is the canonical embedding: the transitive collapse applied to the constant function.



13 A • 1. Figure: Factoring with extenders

A benefit of using finite subsets of λ is that they give a more complex web of ultrapowers than with the linear ultrapowers of Subsection 12 E. In particular, for $r \subseteq s \in [\lambda]^{<\omega}$, we get embeddings $j_{r,s}$: cUlt(V, E_r) \rightarrow cUlt(V, E_s). And these will commute and allow us to form a directed system of the ultrapowers, giving a well-founded direct limit which will be referred to as cUlt_E(V), as seen in Figure 13 A • 1 (which writes cUlt_r for cUlt(V, E_r) to save space, and also assumes $\lambda > \kappa + \kappa$).

To establish all of these facts in a more general setting, where we involve ultrafilters on potentially more than just κ , we must give a formal definition of a derived extender. Note that κ_r might be very different from κ . Indeed, for $r = \{\alpha\}$ for $\alpha < \kappa$, it follows that $\kappa_r = \alpha + 1$, which isn't even a limit ordinal, let alone a cardinal.

— 13 A • 2. Definition

Let $j : V \to M$ be traditional and a class. Let $\kappa = cp(j)$, and let λ be an ordinal of V with $\kappa < \lambda$.

- For $r \in [\lambda]^{<\omega}$, let κ_r be the least ordinal such that $j(\kappa_r) > \max(r)$ (so $\kappa_r \le \max(r) + 1$).
- Define the (κ, λ) -extender derived from j to be

$$E_{\lambda}^{J} = \{ \langle r, X \rangle \in [\lambda]^{<\omega} \times \mathcal{O}([\kappa_{r}]^{<\omega}) : r \in j(X) \}.$$

- For $E = E_{\lambda}^{j}$ and for a finite $r \subseteq \lambda$, write $E_{r} = \{X \subseteq [\kappa_{r}]^{<\omega} : \langle r, X \rangle \in E\}.$
- Call *E* short iff $\lambda \leq j(\kappa)$ and long otherwise.

This distinction between long and short can be instead thought of in terms of these κ_r s. Note that we can also think of these κ_r s as instead indexed by ordinals below λ : it's easy to see that

 $\{\kappa_r : r \in [\lambda]^{<\omega}\} = \{\kappa_{\{\alpha\}} : \alpha < \lambda\} = \{\mu : \exists \alpha < \lambda \ (\mu \text{ is the least such that } j(\mu) > \alpha)\}.$

We also get a very simple re-characterization of short extenders that motivates why we could simply use κ and forgo these κ_r s in the motivating idea before. So any reader overwhelmed by notation can instead just assume each κ_r is κ for simplicity and realize they are working with short derived extenders.

- 13 A • 3. Result

Let $j : V \to M$ be traditional and a class. Let $E = E_{\lambda}^{j}$ be the derived (κ, λ) -extender. Therefore E is short iff $\kappa = \kappa_r$ for every $r \in [\lambda]^{<\omega} \setminus [\kappa]^{<\omega}$.

Proof .:.

Clearly $\kappa \leq \kappa_r$ for $\max(r) > \kappa$ since $\operatorname{cp}(j) = \kappa$. Note that for $r \in [\lambda]^{<\omega} \setminus [\kappa]^{<\omega}$, $\kappa_r = \kappa$ iff $j(\kappa) > \max(r)$. Hence $\kappa_r = \kappa$ for every $r \in [\lambda]^{<\omega} \setminus [\kappa]^{<\omega}$ iff $j(\kappa) > \max(r)$ for every $r \in [\lambda]^{<\omega} \setminus [\kappa]^{<\omega}$ iff $j(\kappa) > \alpha$ for every $\kappa \leq \alpha < \lambda$ iff $\lambda \leq j(\kappa)$.

Each E_r is analogous to the ultrafilter defined in Result 12 B • 8. Now although κ -completeness was stated for ultra-

filters over cardinals, the defining property easily generalizes to ultrafilters over an arbitrary set: it's κ -complete iff the intersection of $< \kappa$ -many sets in the ultrafilter is also in the ultrafilter. In particular, we can say that each E_r is a κ -complete ultrafilter.

– 13 A•4. Setup -

Let $j : V \to M$ be traditional. Let $\kappa = cp(j) < \lambda$. Let $E = E_{\lambda}^{j}$ be the (κ, λ) -extender derived from j.

- **13A**•**5**. **Result** - Assume Setup 13 A • 4. Therefore, for each $r \in [\lambda]^{<\omega}$.

- $E_r \in E$ is a κ -complete ultrafilter over $[\kappa_r]^{<\omega}$ in V.
- κ_r is the least ordinal with $[\kappa_r]^{<\omega} \in E_r$.
- E_r is non-principal iff $r \notin [\kappa]^{<\omega}$. And in either case, $Ult(V, E_r)$ is then well-founded.

Proof .:.

Firstly, $\emptyset \neq E_r \subsetneq \Theta([\kappa_r]^{<\omega})$, the second strict inequality following since $r \notin j(\emptyset)$, and the first following by $r \in [j(\kappa_r)]^{<\omega} = j([\kappa_r]^{<\omega})$ implying $[\kappa_r]^{<\omega} \in E_r$. By minimality of κ_r , we actually have that κ_r is the least such ordinal. That E_r is closed upward and under intersections follows from elementarity: $r \in j(X) \subseteq j(Y)$ for $Y \supseteq X \in E_r$ and $r \in j(X) \cap j(Y) = j(X \cap Y)$ whenever $X, Y \in E_r$. So E_r is a filter. Being an ultrafilter is similarly easy: let $X \subseteq [\kappa_r]^{<\omega}$. If $X \notin E_r$ then $r \notin j(X)$. Thus $r \in [j(\kappa_r)]^{<\omega} \setminus j(X) = j([\kappa_r]^{<\omega} \setminus X)$, meaning $[\kappa_r]^{<\omega} \setminus X \in E_r$.

 E_r is κ -complete for the same sort of reason as in Result 12 B • 8: let $\theta < \kappa$ and $\{X_\alpha : \alpha < \theta\} \in \mathcal{O}(E_r)$. Since $r \in j(X_\alpha)$ for each $\alpha < \theta$ and $j(\theta) = \theta$, it follows that $r \in \bigcap_{\alpha < \theta} j(X_\alpha) = j(\bigcap_{\alpha < \theta} j(X_\alpha))$.

 E_r is non-principal whenever $\max(r) \ge \kappa$ since if there were some $a \in [\kappa]^{<\omega}$ where $r \in j(X)$ iff $a \in X$, then $r \in j(\{a\}) = \{j(a)\}$ implies r = j(a), which is impossible, since a is a finite subset of κ : apply elementarity to $\alpha \in a$ iff $\bigvee_{i \le n} \alpha = a_i$ for parameters $a_0, \dots, a_n < \kappa$. To see that E_r is principal whenever $\max(r) < \kappa$, note that $\kappa_r < \max(r) + 1 = j(\max(r) + 1)$ so that for $X \subseteq [\kappa_r]^{<\omega}$, j(X) = X and hence $r \in j(X)$ iff $r \in X$ and so E_r is principal.

All of these resulting ultrafilters also work nicely together. But to really define what this means, we need to introduce some translations. The idea is that if $r \subseteq s$, we can identify r as, say, the first, second, and fifth entries of s in increasing enumeration. For example, consider $s = \{s_0, s_1, s_2, s_3, s_4, s_5, s_6\}$ in increasing order and $r = \{s_0, s_1, s_4\}$. Then we can generally project 7-sized sets down to 3-sized sets in the same way s projects down to r: define $\operatorname{proj}_{s,r}(\{t_0, \dots, t_6\}) = \{t_0, t_1, t_4\}$ whenever $t_0 < \dots < t_6$. The result is that if we take the pre-image of some $X \subseteq [\operatorname{Ord}]^{|r|}$ under $\operatorname{proj}_{s,r}$, then we effectively fill in the spaces to transform X as a family of |r|-sized sets into a family $X^{r,s}$ of |s|-sized sets, just by only caring about the information of r in s:

$$X^{r,s} = \operatorname{proj}_{s,r}^{-1} "X = \{t : \operatorname{proj}_{s,r}(t) \in X\}.$$

So in the case of $r = \{s_0, s_1, s_4\}$ above, if $\{t_0, \dots, t_5, t_6\} \in X^{r,s}$, then $\{t_0, \dots, t_5, \alpha\}$ is also in $X^{r,s}$ for more-or-less any α : the 6th coordinate is effectively a dummy variable since the 6th coordinate isn't in r. This translation also will work with functions to transform their domains just by first projecting the input with

 $\begin{array}{l} - \text{ 13A+6. Definition} \\ \text{Let } r \subseteq s \in [\text{Ord}]^{<\omega}. \text{ Let } \kappa_r, \kappa_s \in \text{Ord. Define } \operatorname{proj}_{s,r} : [\text{Ord}]^{|s|} \to [\text{Ord}]^{|r|} \text{ by} \\ \text{proj}_{s,r} \left(\{\xi_0, \cdots, \xi_n\} \right) = \{\xi_{i_0}, \cdots, \xi_{i_m}\} \\ \text{such that } \operatorname{proj}_{s,r}(s) = r \text{ and } \xi_0 < \cdots < \xi_n. \text{ For } X \subseteq [\kappa_r]^{|r|}, \text{ and } f \text{ a function with } \operatorname{dom}(f) \subseteq [\kappa_r]^{|r|}, \text{ define} \\ f^{r,s} = f \circ \operatorname{proj}_{s,r} \upharpoonright [\kappa_s]^{|s|} \qquad X^{r,s} = [\kappa_s]^{|s|} \cap \operatorname{proj}_{s,r}^{-1} "X = \{t \in [\kappa_s]^{|s|} : \operatorname{proj}_{s,r}(t) \in X\}. \end{array}$

The usefulness of this definition will be in defining the elementary maps $j_{r,s}$: $cUlt(V, E_r) \rightarrow cUlt(V, E_s)$ when $r \subseteq s \in [\lambda]^{<\omega}$.^{xiii}

xiii This also hints at an alternative characterization of extenders using increasing finite sequences of ordinals instead of finite sets where we preserve the order in larger sets. The two approaches are basically equivalent, but I find the notation to be easier with sets, particularly with taking a union.

- 13A•7. Lemma (Coherence) Under Setup 13 A•4, for $r \subseteq s \in [\lambda]^{<\omega}$, $X \in E_r$ iff $X^{r,s} \in E_s$.

Proof .:.

Firstly, it's easy to see that $\forall_{E_r}^* t(|t| = |r|)$ as we're dealing with finite sets, and this obviously generalizes, so we may intersect with $[\kappa_r]^{|r|}$ or $[\kappa_s]^{|s|}$ to assume all the subsets we're working with are of the appropriate size. Secondly, as proj_{s,r} is easily definable from the relative ordering of s and r, proj_{s,r} is fixed by j and thus

$$j(X^{r,s}) = \{t \in [j(\kappa_s)]^{|s|} : \operatorname{proj}_{s,r}(t) \in j(X)\} = j(X)^{r,s}.$$

Now suppose $X \in E_r$. As $r \in j(X)$, it follows that $\operatorname{proj}_{s,r}(s) \in X$ and thus $s \in j(X)^{r,s}$, meaning $X^{r,s} \in E_s$. Conversely, if $X^{r,s} \in E_s$, then $s \in j(X^{r,s}) = j(X)^{r,s}$ and so $r \in j(X)$.

It turns out that this will define an elementary $\tilde{J}_{r,s}$: Ult(V, E_r) \rightarrow Ult(V, E_s). Coherence (13 A • 7) will then give us a means of looking at the direct limit of these Ult(V, E_r), yielding the notion of an "ultrapower" of a short extender in a way that naturally extends the usual definition of an ultrapower. In particular, $\tilde{J}_{r,s}$ can be worked with as follows:

$$\begin{aligned} \mathsf{Ult}(\mathsf{V}, E_r) &\models ``\varphi([f]_{E_r})`' \quad \text{iff} \quad \{t \in [\kappa]^{|r|} : \varphi(f(t))\} \in E_r \\ &\text{iff} \quad \{t \in [\kappa]^{|s|} : \varphi(f(\operatorname{proj}_{s,r}(t)))\} \in E_s \\ &\text{iff} \quad \mathsf{Ult}(\mathsf{V}, E_s) \models ``\varphi([f^{r,s}]_{E_s})`'. \end{aligned}$$

So $\tilde{j}_{r,s}([f]_{E_r}) = [f^{r,s}]_{E_s}$ is elementary. Although this argument is simple, we can get a bit more information out of how these maps work together. In particular, since the structure $\langle [\lambda]^{<\omega}, \subseteq \rangle$ is upward directed, there is a direct limit $\mathsf{Ult}_E(\mathsf{V})$ of the system of ultrapowers $\langle \mathsf{Ult}(\mathsf{V}, E_r), \tilde{j}_{r,s} : r \subseteq s \in [\lambda]^{<\omega} \rangle$. As with the linear iterations before in The Wellfoundedness of Iterated Ultrapowers (12 E • 4), this direct limit will be well-founded. But despite this characterization of $\mathsf{Ult}_E(\mathsf{V})$ as a direct limit, this isn't too satisfying, since it doesn't directly tell us much about the structure of $\mathsf{cUlt}_E(\mathsf{V})$. Luckily, we can give a more concrete presentation of $\mathsf{Ult}_E(\mathsf{V})$ in a way similar to ultrapowers by ultrafilters in Definition 12 • 1. This definition is also intuitively the resulting direct limit.

- ── 13 A 8. Definition
 - Let $E \subseteq [\lambda]^{<\omega} \times \mathcal{P}(\text{Ord})$ for some $\lambda \in \text{Ord}$ such that each $E_r = \{X : \langle r, X \rangle \in E\}$ is an ultrafilter over some $[\kappa_r]^{<\omega}$ such that Coherence (13 A 7) holds^{xiv} (e.g. a derived extender).
 - Let σ be a FOL-signature.
 - Let **M** be a (possibly class) $FOL(\sigma)$ -model.

Define the $FOL(\sigma)$ -model $Ult_E(M)$ as follows. For $r, s \in [\lambda]^{<\omega}$ and $f : [\kappa_r]^{|r|} \to M$ and $g : [\kappa_s]^{|s|} \to M$, write

- $\langle r, f \rangle \approx_E \langle s, g \rangle$ iff $\forall^*_{E_{r \cup s}} t \ (f^{r, r \cup s}(t) = g^{s, r \cup s}(t)).$
- $\langle r, f \rangle R_E \langle s, g \rangle$ iff $\forall_{E_{r \cup s}}^* t (f^{r,r \cup s}(t) R^{\mathsf{M}} g^{s,r \cup s}(t))$, for R a σ -relation symbol.

Take $\text{Ult}_E(M)$ to be the resulting model of equivalence classes: $\{[r, f]_E : r \in [\lambda]^{<\omega} \land f : [\kappa_r]^{|r|} \to M\}$ with functions and relations interpreted in the natural way from the above: $R^{\text{Ult}_E(M)} = R_E$ and $F^{\text{Ult}_E(M)}([r, f]_E) = [r, F^{\text{M}} \circ f]_E$ for $f a \sigma$ -function symbol.

Note that the definition given here could really be a proper class, but just as before with Definition 12 • 1, we can consider an equivalent formulation via Scott's Trick (9 C • 1): $\langle s, g \rangle \in [r, f]$ iff $\langle s, g \rangle \approx_E \langle r, f \rangle$ and $g \in V_{\alpha}$ for the least α such that $\exists s', g' (\langle s', g' \rangle \approx_E \langle r, f \rangle)$.

To prove that this actually results in the direct limit of the ultrapowers, we first need to establish a very nice fact about $Ult_E(V)$: a version of Łoś's Theorem (12 • 2). Not only is this nice to have, but it also furthers the association of $Ult_E(V)$ to an actual ultrapower. The proof of this is essentially the same as with Łoś's Theorem (12 • 2), but with just a little care about the ultrafilters we're using.

With x, y, the place that contains the information of both x and y is just $x \cup y$. With tuples, there is no such notation, and would have to just write " $x \cup y$ " or some other notation as shorthand for the increasing enumeration of $ran(x) \cup ran(y)$, the union of the entries of x and the entries of y regarding them as elements of $^{<\omega}$ Ord.

^{xiv}This is necessary to show that \approx_E is indeed an equivalence relation. Specifically, it's needed for transitivity.

— 13 A • 9. Theorem (Łoś's Theorem for Extenders)

- Let $E \subseteq [\lambda]^{<\omega} \times \mathcal{P}(\text{Ord})$ for some $\lambda \in \text{Ord}$ such that each $E_r = \{X : \langle r, X \rangle \in E\}$ is an ultrafilter over some $[\kappa_r]^{<\omega}$ such that Coherence (13 A 7) holds^{xv} (e.g. a derived extender).
- Let σ be a FOL-signature.
- Let **M** be a (possibly class) $FOL(\sigma)$ -model.
- Let φ be a FOL (σ) -formula.

Thus $\mathsf{Ult}_E(\mathsf{M}) \models ``\varphi([r_0, f_0], \cdots, [r_n, f_n])$ '' iff for $E_{\bigcup_{i \le n} r_i}$ -almost every $t, \mathsf{M} \models ``\varphi(f_0(t), \cdots, f_n(t))$ ''.

Proof .:.

This can be proven by structural induction similar to Łoś's Theorem (12 • 2). In particular, the result clearly holds for the atomic formulas by definition. Conjunctions follow easily from this using Coherence (13 A • 7) to transform everything into the same context. Negations follow easily as well since the E_r s are ultrafilters. For existential quantification, suppose $\text{Ult}_E(\mathbf{M}) \models \exists x \varphi(x, [r_0, f_0], \dots, [r_n, f_n])$ ". Since there's a witness $[r, f] \in \text{Ult}_E(\mathbf{M})$, by the inductive hypothesis we have that for $s = \bigcup_{i \le nr_i \cup r}$, for E_s -almost every t, $\mathbf{M} \models "\varphi(f^{r,s}(t), f_0^{r_0,s}(t), \dots, f_n^{r_n,s}(t))$ ". Thus for E_s -almost every t, $\mathbf{M} \models "\exists x \varphi(f_0^{r_0,s}(t), \dots, f_n^{r_n,s}(t))$ ". By Coherence (13 A • 7), we can reduce s in this statement to $\bigcup_{i \le n} r_i$. This shows the (\rightarrow) direction for the existential case.

For the other direction, write $r = \bigcup_{i \le n} r_i$ and suppose for E_r -almost every t, $\mathbf{M} \models \exists x \varphi(x, f_0(t), \dots, f_n(t))$ ". Using AC, for each $t \in [\kappa]^{|r|}$, let f(t) be such an x if there is one, or else f(t) is some fixed element of M. By the inductive hypothesis, $\mathbf{Ult}_E(\mathbf{M}) \models ``\varphi([r, f], [r_0, f_0] \cdots, [r_n, f_n])$ ".

Using the embeddings $\tilde{j}_{r,s}$, we can then show that $Ult_E(V)$ is indeed a direct limit.

— 13 A•10. Corollary

- Let $E \subseteq [\lambda]^{<\omega} \times \mathcal{P}(\text{Ord})$ for some $\lambda \in \text{Ord}$ such that each $E_r = \{X : \langle r, X \rangle \in E\}$ is an ultrafilter over some $[\kappa_r]^{<\omega}$ such that Coherence (13 A 7) holds^{xvi} (e.g. a derived extender).
- Let σ be a FOL-signature.
- Let **M** be a (possibly class) $FOL(\sigma)$ -model.
- For $r, s \in [\lambda]^{<\omega}$, let $\tilde{j}_{r,s}$: Ult(M, E_r) \rightarrow Ult(M, E_s) be elementary, defined by $\tilde{j}_{r,s}([f]_{E_r}) = [f^{r,s}]_{E_s}$.

Therefore $\mathsf{Ult}_E(\mathsf{M})$ is (isomorphic to) the direct limit of the system of ultrapowers { $\mathsf{Ult}(\mathsf{M}, E_r), \tilde{j}_{r,s} : r \subseteq s \in [\lambda]^{<\omega}$ } with limit embeddings $\tilde{j}_{r,\infty} : \mathsf{Ult}(\mathsf{M}, E_r) \to \mathsf{Ult}_E(\mathsf{M})$ defined by $\tilde{j}_{r,\infty}([f]_{E_r}) = [r, f]_E$.

Proof .:.

Each $\tilde{j}_{r,\infty}$ is an elementary embedding, since by Łoś's Theorem (12 • 2) and Łoś's Theorem for Extenders (13 A • 9), for any FOL-formula φ ,

 $\mathsf{Ult}(\mathsf{M}, E_r) \vDash ``\varphi([f]_{E_r})'' \quad \text{iff} \quad \forall_{E_r}^* t \ (\mathsf{M} \vDash ``\varphi(f(t))'') \quad \text{iff} \quad \mathsf{Ult}_E(\mathsf{M}) \vDash ``\varphi([r, f]_E)''.$ Moreover, $\tilde{j}_{s,\infty} \circ \tilde{j}_{r,s} = \tilde{j}_{r,\infty}$ for $r \subseteq s$ since $\tilde{j}_{r,s}([f]_{E_r}) = [f^{r,s}]_{E_s}$ and $[s, f^{r,s}]_E = [r, f]_E.$

To prove that $Ult_E(M)$ is the direct limit, it then suffices to show that it is the "least" such: any other A such that

1. there are embeddings h_r : Ult(M, E_r) $\rightarrow A$, and

2. the embeddings obey $h_r = h_s \circ \tilde{j}_{r,s}$;

carries with it an embedding $h : \text{Ult}_E(M) \to A$ such that each $h_r = h \circ \tilde{j}_{r,\infty}$.

To see this, set $h([r, f]) = h_r([f]_r)$, which clearly satisfies $h_r = h \circ \tilde{j}_{r,\infty}$. This is well-defined, since if [r, f] = [s, g], then

$$\tilde{j}_{r,r\cup s}([f]_{E_r}) = \left[f^{r,r\cup s}\right]_{E_r\cup s} = \left[g^{s,r\cup s}\right]_{E_r\cup s} = \tilde{j}_{s,r\cup s}([g]_{E_s}),$$
 and thus (2) above tells us that

 $h([r, f]) = h_r([f]_{E_r}) = h_{r \cup s} \circ \tilde{j}_{r, r \cup s}([f]_{E_r}) = h_{r \cup s} \circ \tilde{j}_{s, r \cup s}([g]_{E_s}) = h_s([g]_{E_s}) = h([s, g]).$

Since it's clear that h as defined is an embedding, it follows that $Ult_E(M)$ is the direct limit.



13 A • 11. Figure: Elementary embeddings with derived extenders

From now on, we will adopt the following notational conventions as displayed in Figure 13 A • 11. Although these embeddings and the commutativity of the diagram haven't been proven yet, this serves as a neat way of presenting the information and the conventions that, as far as I'm aware, are specific to this work. All maps displayed in the figure are elementary. Firstly, we will, independent of the previous results, define the following maps.

13A•12. **Definition** Under Setup 13 A • 4, define the following: for $x \in V, r \subseteq s \in [\lambda]^{<\omega}$, and $f : [\kappa_r]^{<\omega} \to V$, • $\tilde{j}_r(x) = [\text{const}_x]_{E_r}$; • $\tilde{j}_{r,s}([f]_{E_r}) = [f^{r,s}]_{E_s}$; • $\tilde{j}_{r,\infty}([f]_{E_r}) = [r, f]$; • $\tilde{j}_E(x) = [\emptyset, \text{const}_x]$; • $\tilde{k}_r([f]_{E_r}) = j(f)(r)$; • $\tilde{k}_E([r, f]) = j(f)(r)$; • π_r and π_E are transitive collapse isomorphisms assuming $\mathsf{Ult}_E(\mathsf{V})$ is well-founded; • $j_r = \pi_r \circ \tilde{j}_r$ and $j_{r,s} = \pi_s \circ \tilde{j}_{r,s} \circ \pi_r^{-1}$ and $j_E = \pi_E \circ \tilde{j}_E$; • $k_r = \tilde{k}_r \circ \pi_r^{-1}$, and $k_E = \tilde{k}_E \circ \pi_E^{-1}$.

The fact that these are well-defined is not too difficult. Again, the following are the important properties of these maps, beyond their explicit definitions above, basically establishing the elementarity of the maps in Figure 13 A • 11 and the commutativity of the diagram.

13A•13. Theorem Assume Setup 13 A •4. Therefore, for $r, s, t \in [\lambda]^{<\omega}$, 1. $\tilde{j}_r : V \to Ult(V, E_r)$ is elementary. 2. $\tilde{j}_{r,s} : Ult(V, E_r) \to Ult(V, E_s)$ is elementary and $\tilde{j}_{r,t} = \tilde{j}_{s,t} \circ \tilde{j}_{r,s}$ for $r \subseteq s \subseteq t$. 3. $\tilde{j}_{r,\infty} : Ult(V, E_r) \to Ult_E(V)$ is elementary and $\tilde{j}_{r,\infty} = \tilde{j}_{s,\infty} \circ \tilde{j}_{r,s}$ for $r \subseteq s$. 4. $\tilde{j}_E : V \to Ult_E(V)$ is elementary and $\tilde{j}_E = \tilde{j}_{r,\infty} \circ \tilde{j}_r$. 5. $\tilde{k}_r : Ult(V, E_r) \to M$ is elementary and $j = \tilde{k}_r \circ \tilde{j}_r$. 6. $\tilde{k}_E : Ult_E(V) \to M$ is elementary with $j = \tilde{k}_E \circ \tilde{j}_E$. 7. $Ult(V, E_r)$ and $Ult_E(V)$ are well-founded. 8. $j_r : V \to cUlt(V, E_r)$ is elementary and traditional when E_r is non-principal. 9. $j_E : V \to cUlt_E(V)$ is traditional with $cp(j_E) = \kappa$. In fact, $cUlt_E(V) = \{j(f)(r) : r \in [\lambda]^{<\omega} \land f : [\kappa_r]^{<\omega} \to V\}$. 10. $k_r : cUlt(V, E_r) \to M$ is elementary.

11. k_E : cUlt_E(V) \rightarrow M is elementary with cp(k_E) $\geq \lambda$ if $k_E \neq$ id.

 \dashv

Proof .:.

- 1. This holds by Theorem 12 B 1.
- 2. It should be clear that $\operatorname{proj}_{t,s} \circ \operatorname{proj}_{s,r} = \operatorname{proj}_{t,r}$ so that $(f^{r,s})^{s,t} = f^{r,t}$. This tells us $\tilde{j}_{r,t} = \tilde{j}_{s,t} \circ \tilde{j}_{r,s}$ for $r \subseteq s \subseteq t$. It's also easy to see that the domain and range of $\tilde{j}_{r,s}$ are as indicated above, so all that remains is elementarity which follows easily by Coherence (13 A \cdot 7):

 $\mathsf{Ult}(\mathsf{V}, E_r) \vDash ``\varphi([f]_{E_r})" \quad \text{iff} \quad \forall_{E_r}^* t \ \varphi(f(t)) \quad \text{iff} \quad \forall_{E_s}^* t \ \varphi(f^{r,s}(t)) \quad \text{iff} \quad \mathsf{Ult}(\mathsf{V}, E_s) \vDash ``\varphi([f^{r,s}]_{E_s})".$

- 3. This follows from the proof of Corollary 13 A 10.
- 4. This follows by (1) and (3).
- 5. For any x, $\tilde{k}_r \circ \tilde{j}_r(x) = \tilde{k}_r([\text{const}_x]_{E_r})$ which by Definition 13 A 12 is equal to $j(\text{const}_x)(r) = \text{const}_{j(x)}(r) = j(x)$. So $j = \tilde{k}_r \circ \tilde{j}_r$. To see that \tilde{k}_r is elementary, proceed as in Factoring (12 B 9): by Loś's Theorem for Extenders (13 A 9),

$$\begin{aligned} \mathsf{Ult}(\mathsf{V}, E_r) &\models ``\varphi([f]_{E_r})'' \quad \text{iff} \quad \forall_{E_r}^* t \ \varphi(f(t)) \quad \text{iff} \quad \{t \in [\kappa_r]^{<\omega} : \varphi(f(t))\} \in E_r \\ &\text{iff} \quad r \in j \ \left(\{t \in [\kappa_r]^{<\omega} : \varphi(f(t))\}\right) \\ &\text{iff} \quad r \in \left\{t \in [j(\kappa_r)]^{<\omega} : \mathsf{M} \models ``\varphi(j(f)(t))''\right\} \quad \text{iff} \quad \mathsf{M} \models ``\varphi(j(f)(r))''. \end{aligned}$$

- 6. For any x, $\tilde{k}_E \circ \tilde{j}_E(x) = \tilde{k}_E([\emptyset, \text{const}_x]) = j(\text{const}_x)(\emptyset) = \text{const}_{j(x)}(\emptyset) = j(x)$. So it suffices to show \tilde{k}_E is elementary, and for this, we proceed exactly as in (5) using Coherence (13 A 7) to translate parameters to a single space.
- 7. The well-foundedness of Ult(V, E_r) follows from the κ -completeness of E_r (Result 13 A 5) by Theorem 12 B 3. The well-foundedness of Ult_E(V) follows from the elementarity of \tilde{k}_E and the well-foundedness of M: any infinite $\in^{\text{Ult}_E(V)}$ -decreasing sequence $\langle x_n : n < \omega \rangle$ yields that $\langle \tilde{k}_E(x_n) : n < \omega \rangle$ is an infinite $\in^{\text{M}} = \epsilon$ -decreasing sequence, contradicting well-foundedness in V.
- 8. The elementarity of j_r follows from the elementarity of \tilde{j}_r and that π_r is an isomorphism. That j_r is traditional follows from Theorem 12 B 5 (note, however, that we don't know $cp(j_r) = \kappa$).
- 9. The elementarity is easy as the composition of elementary embeddings. It suffices to show that $\text{cUlt}_E(V) = \{j(f)(r) : r \in [\lambda]^{<\omega} \land f : [\kappa_r]^{<\omega} \to V\}$. And so inductively we show that $\pi_E([r, f]) = j(f)(r)$ for all $[r, f] \in \text{Ult}_E(V)$. So suppose this holds for all $[s, g] \in ^{\text{Ult}_E(V)}[r, f]$. Therefore

 $\pi_E([r, f]) = \{\pi_E([s, g]) : \mathsf{Ult}_E(\mathsf{V}) \models "[s, g] \in [r, f]"\} = \{j(g)(s) : \mathsf{Ult}_E(\mathsf{V}) \models "[s, g] \in [r, f]"\}.$

As a result, $x \in \pi_E([r, f])$ iff there's some g, s where x = j(g)(s) and $\mathsf{Ult}_E(\mathsf{V}) \models "[s, g] \in [r, f]$ ". For $t = r \cup s$, by Łoś's Theorem for Extenders (13 A • 9), this is equivalent to $\forall_{E_t}^* t' (g^{s,t}(t') \in f^{r,t}(t'))$ which then says

 $t \in j(\{t' \in [\kappa_t]^{<\omega} : g^{s,t}(t') \in f^{r,t}(t')\}) = \{t' \in [j(\kappa_t)]^{<\omega} : j(g^{s,t})(t') \in j(f^{r,t})(t')\}.$

And this is equivalent to $x = j(g)(s) = j(g^{s,t})(t) \in j(f^{r,t})(t) = j(f)(r)$. Thus $\pi_E([r, f]) \subseteq j(f)(r)$. The other direction is similar and also carried out by induction: if $x \in j(f)(r)$ then inductively, x = j(g)(s) for some appropriate g, s and we follow the reverse of the reasoning above, telling us $j(f)(r) = \pi_E([r, f])$. Hence $\text{cUlt}_E(V) = \pi_E$ "Ult $_E(V)$ has the above form.

- 10. This follows as the composition of elementary functions.
- 11. Again, elementarity follows by composition. To see that $cp(k_E) \ge \lambda$, it suffices to show $\lambda \subseteq k_E$ " $cUlt_E(V) = \tilde{k}_E$ " $Ult_E(V)$. To see this, consider the max function (or really the supremum function, taking max $\emptyset = \emptyset$): for any $\alpha < \lambda$, $\tilde{k}_E([\{\alpha\}, \max]) = j(\max)(\{\alpha\}) = \max\{\alpha\} = \alpha \in k_E$ " $cUlt_E(V)$. It follows that we cannot skip any values with k_E : otherwise inductively, if $k_E(\alpha) > \alpha > \beta = k_E(\beta)$ for

INTRODUCING EXTENDERS

13A·14. Corollary -

Сн II §13 А

 \dashv

 \dashv

every $\beta < \alpha$, then $\alpha \notin k_E$ " cUlt_E (V).

As with ultrapowers by measures, we call \tilde{j}_r , j_r , \tilde{j}_E , and j_E the canonical (extender) embeddings. It's not too difficult to show that the extender derived from the extender embedding is just the original extender.

Assume Setup 13 A • 4. Therefore the (κ, λ) -extender derived from j_E is $E_{\lambda}^{j_E} = E$.

Proof .:.

By Theorem 13 A • 13 (11), $k_E(r) = r$ for $r \in [\lambda]^{<\omega}$ so that $[\lambda]^{<\omega} \subseteq \text{cUlt}_E(V)$. Thus by elementarity, $k_E(r) = r \in j(X) = k_E(j_E(X))$ iff $k_E(r) \in k_E(j_E(X))$ iff $r \in j_E(X)$. This tells us $E = E_{\lambda}^{j_E}$, as desired.

Let us list some further important properties of E_{λ}^{j} . Returning back to properties of the E_r s as ultrafilters, note that normality can't be translated directly to ultrafilters on $[\kappa]^{<\omega}$. That said, there is still a notion of normality for them, and it will allow us to call these ultrafilters "measures", or at least the short extender "normal".

- 13A • 15. Lemma Under Setup 13 A • 4, let $r \in [\lambda]^{<\omega}$ and suppose $f : [\kappa_r]^{|r|} \to V$ is such that $\forall_{E_r}^* t \ (f(t) < \max t)$. Therefore there is some $\alpha < \max r$ with $\forall_{E_r \cup \{\alpha\}}^* t \ \left(\left\{ f^{r,r \cup \{\alpha\}}(t) \right\} = \operatorname{proj}_{r \cup \{\alpha\},\{\alpha\}}(t) \right).$

Proof .:.

That $\forall_{E_r}^* t \ (f(t) < \max t)$ is just to say that $r \in j(\{t \in [\kappa_r]^{|r|} : f(t) < \max t\})$, which means $j(f)(r) < \max r$. In particular, let $\alpha = j(f)(r)$. Note that $\operatorname{proj}_{r \cup \{\alpha\}}(r \cup \{\alpha\}) = \{\alpha\}$ and $j(f^{r,r \cup \{\alpha\}})(r \cup \{\alpha\}) = j(f)(r) = \alpha$. Hence

$$\{J(f^{r,r\cup\{\alpha\}})\} = \operatorname{proj}_{r\cup\{\alpha\},\{\alpha\}}(r\cup\{\alpha\})$$

and thus $r\cup\{\alpha\} \in \{t \in [\kappa_r]^{r\cup\{\alpha\}} : \{f^{r,r\cup\{\alpha\}}(t)\} = \operatorname{proj}_{r\cup\{\alpha\},\{\alpha\}}(t)\}$, meaning that we have the result. \dashv

Let us collect some of the major results about derived extenders. These will be used to define extenders absent any discussion about embeddings.

— 13 A • 16. Result

Let $j: V \to M$ be traditional. Let $cp(j) = \kappa < \lambda$ and let $E = E_{\lambda}^{j}$ be the derived (κ, λ) -extender. Therefore,

- 1. each E_r is a κ -complete ultrafilter over $[\kappa_r]^{<\omega}$;
- 2. there is an E_r which is not κ^+ -complete;
- 3. for each $\alpha < \kappa$, there is an E_r with $\forall_{E_r}^* t \ (\alpha \in t)$;
- 4. if $f : [\kappa_r]^{|r|} \to V$ is such that $\forall_{E_r}^* t \ (f(t) < \max t)$, then for some $s \supseteq r, \forall_{E_s}^* t \ (f^{r,s}(t) \in t);$
- 5. for $r \subseteq s, X \in E_r$ iff $X^{r,s} \in E_s$, yielding elementary maps $\tilde{j}_{r,s}([f]_{E_r}) = ([f]_{E_r} \mapsto [f^{r,s}]_{E_s})$; and
- 6. The direct limit $Ult_E(V)$ of the ultrapowers by these E_r s is well-founded.

Proof .:.

All of these have been proven already with the exception of (2). In particular,

- (1) follows from Result 13 A 5;
- (3) follows from the proof of Result 13 A 5: take $E_{\{\alpha\}}$ as a principal ultrafilter;
- (4) follows from Lemma 13 A 15;
- (5) follows from Coherence (13 A 7) along with Theorem 13 A 13 (2); and
- (6) follows from Theorem 13 A 13 (7).

Onto (2), if each E_r is κ^+ -complete, $\tilde{j}_r(\kappa)$ is collapsed to κ in the ultrapower: for every r, f, there is an α where $\forall_{E_r}^* u(f(u) < \kappa \rightarrow f(u) = \alpha)$, which is equivalent to, by Łoś's Theorem for Extenders (13 A • 9), the statement that for every $[r, f] \in \text{Ult}_E(V)$, there is an $\alpha < \kappa$ where

 $\mathsf{Ult}_E(\mathsf{V}) \vDash [r, f] < [\emptyset, \mathrm{const}_{\kappa}] \rightarrow [r, f] = [\emptyset, \mathrm{const}_{\alpha}]^n$.

So in the transitive $\text{cUlt}_E(V)$, $[\emptyset, \text{const}_{\kappa}]$ is collapsed to κ itself. But this contradicts Theorem 13 A • 13 (11), since we would have $j(\kappa) = \tilde{k}_E \circ \tilde{j}_E(\kappa) = \tilde{k}_E([\emptyset, \text{const}_{\kappa}]) = \kappa < j(\kappa)$.

So far, all of this has just been setting up a more general definition of extenders, which is basically that Result 13 A • 16 applies. It's important to recognize the limits of the consequences of this. In principle, a measure U ensures the existence of such an E_{λ}^{j} , though ostensibly the existence of such a short extender is stronger, giving the existence of many κ -complete ultrafilters and encoding information up to their *length* λ . One of the major benefits of extenders is the ability to phrase certain large cardinal properties into the existence of certain extenders, or sequences of extenders. So a slightly more general theory should be introduced without making reference to proper classes like elementary embeddings from V into some inner model M.

§13 B. Characterizing extenders

First, we more properly state the definition of a (κ , λ)-extender purely in the language of ZFC, absent any knowledge of classes like elementary embeddings or inner models.

- 13B•1. Definition

- Let $\kappa < \lambda$ with κ an infinite cardinal.
- For $r \in [\lambda]^{<\omega}$, let κ_r^E be an ordinal, usually written as just " κ_r " if E is implied by context.
- Let $E \subseteq \bigcup_{r \in [\lambda]^{<\omega}} \{r\} \times \mathcal{O}([\kappa_r^E]^{<\omega}).$
- Write $E_r = \{X \subseteq [\kappa_r^E]^{<\omega} : \langle r, X \rangle \in E\}.$

We call *E* a (κ , λ)*-extender* iff

1. every E_r is a κ -complete ultrafilter over $[\kappa_r^E]^{<\omega}$ with κ_r^E the least ordinal such that $(\kappa_r^E)^{|r|} \in E_r$.

- 2. there is an E_r that is not κ^+ -complete;
- 3. for each $\alpha < \kappa$, there is an E_r with $\forall_{E_r}^* t \ (\alpha \in t)$;
- 4. (coherency) for $r \subseteq s \in [\lambda]^{<\omega}$, $X \in E_r$ iff $X^{r,s} \in E_s$;
- 5. (normality) for $r \in [\lambda]^{<\omega}$, if $f : [\kappa_r^E]^{|r|} \to V$ is such that $\forall_{E_r}^* t \ (f(t) < \max t)$, then for some $s \supseteq r$, $\forall_{E_s}^* t \ (f^{r,s}(t) \in t)$;
- 6. (well-foundedness) for every sequence $\langle r_n, X_n : n \in \omega \rangle$ with $X_n \in E_{r_n}$, there is an order preserving function $\pi : \bigcup_{n \in \omega} r_n \to \bigcup_{r \in [\lambda]^{\leq \omega}} \kappa_r^E$ such that for each $n \in \omega, \pi^{"}r_n \in X_n$.
- κ is then called the *critical point* of *E*, denoted cp(*E*), and λ is called the *length* of *E*.

The requirement of (6) is a bit odd, but it is equivalent to $Ult_E(V)$ being well-founded so that we can consider the collapsed ultrapower as an inner model. The π in (6) plays the role of the intersection as with \aleph_1 -completeness, and in fact (6) is sometimes called \aleph_1 -completeness. To show that (6) and well-foundedness are equivalent, we must use ultrapowers again.

Firstly, we may use Definition 13 A • 8 to form the ultrapower. Coherency implies a form of Łoś's theorem and hence tells us $Ult_E(V)$ is the direct limit of the ultrapowers $Ult(V, E_r)$ for $r \in [\lambda]^{<\omega}$. In other words, we still have Łoś's Theorem for Extenders (13 A • 9), the embeddings $\tilde{j}_{r,s}$: Ult(V, E_r) \rightarrow Ult(V, E_s) for $r \subseteq s \in [\lambda]^{<\omega}$, and Corollary 13 A • 10—that $Ult_E(V)$ is the direct limit of ultrapowers. From here, it's a straight-forward argument that (6) ensures well-foundedness. That this *characterizes* well-foundedness is a bit trickier.

- 13 B • 2. Result

Let *E* be as in Definition 13 B • 1 (1)–(5), i.e. satisfies all requirements of being an extender except possibly (6). Therefore *E* is a (κ, λ) -extender—i.e. satisfies (6)—iff Ult_{*E*}(V) is well-founded.
Proof .:.

For the sake of notation, write s_n for $\bigcup_{i < n} r_i$.

 (\rightarrow) Assume *E* is a (κ, λ) -extender. To show that $Ult_E(V)$ is well-founded, suppose $\langle [r_n, f_n] : n \in \omega \rangle$ is a \in^{Ult_E} -decreasing sequence. Set for $t \in [\kappa_{s_n}]^{<\omega}$,

$$t \in X_n \quad \text{iff} \quad \bigwedge_{i < n} f_{i+1}^{r_{i+1}, s_n}(t) \in f_i^{r_i, s_n}(t).$$

Thus $X_n \in E_{s_n}$ for each *n* by Łoś's Theorem for Extenders (13 A • 9). Now the π as guaranteed in (6) has π " $s_n \in X_n$ for each $n < \omega$. But note that $\operatorname{proj}_{s_n,r_i}(\pi$ " $s_n) = \pi$ " r_i for each i < n, because π is order-preserving. Hence

$$f_{n+1}^{s_{n+1},r_{n+1}}(\pi"s_{n+1}) = f_{n+1}(\pi"r_{n+1}) \in f_n(\pi"r_n) = f_n^{s_{n+1},r_n}(\pi"s_{n+1}).$$

So the sequence $\langle f_n(\pi^*r_n) : n \in \omega \rangle$ is \in -decreasing, contradicting the well-foundedness of V.

(←) Suppose $\langle r_n, X_n : n \in \omega \rangle$ with $X_n \in E_{r_n}$ is a counter-example to (6) of Definition 13 B•1. We will show that Ult_E(V) is ill-founded. Consider a tree of consisting of approximations to π in (6): write $\tau \in T$ iff there is some $n < \omega$ with $\tau : s_n \to \bigcup_{r \in [\lambda] < \omega} \kappa_r$ order preserving and $\tau"r_i \in X_i$ for each $i \leq n$.

One can see that $T \neq \emptyset$ and in fact the height of $\langle T, \subseteq \rangle$ is ω which can be seen as follows: for $n < \omega$, let $t \in X_n^{r_n, s_n} \cap \bigcap_{i < n} X_i^{r_i, s_n} \in E_{s_n}$ be arbitrary with size $|s_n|$. Get a increasing bijection $\tau : s_n \to t$ and note that $\tau \in T$ and in fact $\langle \tau \upharpoonright s_i : i < n \rangle$ is a chain of length n + 1 in T.

Under our assumption that (6) fails, $\langle T, \supseteq \rangle$ is well-founded: any infinite, \subseteq -increasing branch $\langle \tau_n : n < \omega \rangle$ has $\pi = \bigcup_{n < \omega} \tau_n$ as in (6), a contradiction. So there is some rank function, rank_T, on $\langle T, \supseteq \rangle$. To give a $\in^{\text{Ult}_E(V)}$ -decreasing sequence, define $f_n : [\kappa_{s_n}]^{<\omega} \to \bigcup_{r \in [\lambda]^{<\omega}} \kappa_r$ by

$$f_n(t) = \begin{cases} \operatorname{rank}_T(\tau) & \text{if } t = \operatorname{im}(\tau) \text{ for some } \tau : s_n \to \bigcup_{r \in [\lambda]^{<\omega}} \kappa_r \text{ with } \tau \in T \\ 0 & \text{otherwise.} \end{cases}$$

Note that if $t = \operatorname{im}(\tau)$ for some $\tau \in T$, then τ is unique just as an order-preserving bijection between s_n and t. So there's no worry about which τ we take to compute $f_n(t)$. Now let $n < \omega$ and consider an arbitrary $t \in X_{n+1}^{r_{n+1},s_{n+1}} \cap \bigcap_{i \leq n} X_i^{r_i,s_{n+1}} \in E_{s_{n+1}}$ which is the image of s_{n+1} by some $\tau \in T$ by the previous argument showing the height of T is ω . We have $f_{n+1}(t) = \operatorname{rank}_T(\tau)$ for some $\tau \in T$ with $\tau^{"}s_{n+1} = t$. Since τ is order preserving, $\operatorname{im}(\tau \upharpoonright s_n) = \operatorname{proj}_{s_{n+1},s_n}(t)$. $\tau \upharpoonright s_n$ is below τ in T so that (recall rank here uses reverse inclusion)

$$f_n^{s_n,s_{n+1}}(t) = \operatorname{rank}_T(\tau \upharpoonright s_n) > \operatorname{rank}_T(\tau) = f_{n+1}(t)$$

Hence this holds for $E_{s_{n+1}}$ -almost every t, and so $Ult_E(V) \vDash "[s_n, f_n] > [s_{n+1}, f_{n+1}]$ ". Since n was arbitrary, the sequence $\langle [s_n, f_n] : n < \omega \rangle$ witnesses that $Ult_E(V)$ is ill-founded.

So far, this basically shows that derived extenders are actually extenders. It's not difficult to verify that we have all of the embeddings of Figure $13 \text{ A} \cdot 11$ as a result, reproduced below. So we adopt here the same notational conventions and get similar results as with Theorem $13 \text{ A} \cdot 13$.



13 B • 3. Figure: Elementary embeddings with extenders

— 13 B • 4. Theorem ——

Let *E* be a (κ, λ) -extender. Therefore, for $r, s, t \in [\lambda]^{<\omega}$, there are collapsing maps $\pi_r : \text{Ult}(V, E_r) \to \text{cUlt}(V, E_r)$ and $\pi_E : \text{Ult}_E(V) \to \text{cUlt}_E(V)$. Moreover,

- 1. There is an elementary $\tilde{j}_r : V \to Ult(V, E_r)$ defined by $\tilde{j}_r(x) = [const_x]_{E_r}$.
- 2. There is an elementary $j_r : V \to \text{cUlt}(V, E_r)$ defined by $j_r = \pi_r \circ \tilde{j}_r$, traditional with $\text{cp}(j_r) \ge \kappa$ when E_r is non-principal.
- 3. There are elementary $\tilde{j}_{r,s}$: Ult(V, E_r) \rightarrow Ult(V, E_s) defined by $\tilde{j}_{r,s}([f]_{E_r}) = [f^{r,s}]_{E_s}$ whenever $r \subseteq s \subseteq t$ so that $\tilde{j}_{r,t} = \tilde{j}_{s,t} \circ \tilde{j}_{r,s}$.
- 4. There are elementary $j_{r,s}$: cUlt(V, E_r) \rightarrow cUlt(V, E_s) defined by $j_{r,s} = \pi_s \circ \tilde{j}_{r,s} \circ \pi_r^{-1}$.
- 5. There is an elementary $\tilde{j}_{r,\infty}$: Ult(V, E_r) \rightarrow Ult_E(V) defined by $\tilde{j}_{r,\infty}([f]_{E_r}) = [r, f]_E$.
- 6. There is an elementary $j_{r,\infty}$: cUlt(V, E_r) \rightarrow cUlt_E(V) defined by $j_{r,\infty} = \pi_E \circ \tilde{j}_{r,\infty} \circ \pi_r^{-1}$.
- 7. There is an elementary $\tilde{j}_E : V \to \text{Ult}_E(V)$ defined by $\tilde{j}_E(x) = [\emptyset, \text{const}_x]_E$ such that $\tilde{j}_E = \tilde{j}_{r,\infty} \circ \tilde{j}_r$ for each $r \in [\lambda]^{<\omega}$.
- 8. There is an elementary $j_E : V \to \text{cUlt}_E(V)$ defined by $j_E = \pi_E \circ \tilde{j}_E$ and j_E is traditional with $\text{cp}(j_E) = \kappa$ and $j_E = j_{r,\infty} \circ j_r$ for each $r \in [\lambda]^{<\omega}$.

Proof .:.

- 1. This follows by Theorem 12 B 1.
- 2. This is obvious from (1) because π_r is an isomorphism. That j_r is traditional follows from Theorem 12 B 5.
- 3. This follows as in Theorem 13 A 13 (2).
- 4. This is obvious from (3) because π_r^{-1} and π_s are both isomorphisms.
- 5. This follows from the proof of Corollary 13 A 10.
- 6. This is obvious from (5) because π_r^{-1} and π_E are both isomorphisms.
- 7. This follows from (1) and (5), or (2) and (6). That \tilde{j}_E factors as $\tilde{j}_{r,\infty} \circ \tilde{j}_r$ follows from the fact that it's the direct limit embedding in Corollary 13 A 10.
- 8. This follows from (7) because π_E is an isomorphism. That j_E is traditional follows from the fact that everything is done internal to V: $\text{cUlt}_E(V)$ is an inner model and $j_E \neq \text{id}$ since $\text{cp}(j_E) = \kappa$. The reason why $\text{cp}(j_E) = \kappa$ follows from the fact that j_E is the direct limit embedding: it factors as $j_E = j_{r,\infty} \circ j_r$ for any $r \in [\lambda]^{<\omega}$. Hence $\text{cp}(j_E) \leq \text{cp}(j_r) = \kappa$ whenever $r \in [\lambda]^{<\omega}$ is such that E_r is κ -complete but not κ^+ -complete. The fact that $\text{cp}(j_E) \geq \kappa$ follows from κ -completeness of the ultrafilters. Inductively, let $\alpha < \kappa$ such that $j_E \upharpoonright \alpha = \text{id}$. To show $j_E(\alpha) \leq \alpha$, suppose $\pi_E([r, f]) < \pi_E([\emptyset, \text{const}_{\alpha}])$. By Los's Theorem for Extenders (13 A • 9), $\forall_{E_r}^* t$ ($f(t) < \alpha$). By κ -completeness, there is some $\beta < \alpha$ where $\forall_{E_r}^* t$ ($f(t) = \beta$) and hence $\pi_E([r, f]) = \pi_E([r, \text{const}_{\beta}]) = \pi_E([\emptyset, \text{const}_{\beta}])$ which is inductively β . Hence $j_E(\alpha) \leq \alpha \leq j_E(\alpha)$ showing $\text{cp}(j_E) \geq \kappa$.

More importantly, this overly technical definition isn't extremely necessary, as we can show that any extender E, with sufficient simplification, is just the derived extender E_{λ}^{jE} , where $j_E : V \rightarrow \text{cUlt}_E(V)$ is the direct limit embedding. In effect, the extenders derived from elementary embeddings are the only kind of extenders anyway. To show this, let's investigate some of the properties of the direct limit Ult_E(V) and the embedding j_E .

– 13 B•5. Lemma

Let *E* be a (κ, λ) -extender. Let π_E : Ult_{*E*}(V) \rightarrow cUlt_{*E*}(V) be the collapsing map. Therefore, (taking max $(\emptyset) = \emptyset$ so that really, max = sup is just the union).

- 1. for each $\alpha < \lambda$, if $\pi_E([\{\alpha\}, \max]) = \alpha$.
- 2. for each $r \in [\lambda]^{<\omega}$, $\pi_E([r, id]) = r$.

Proof .:.

Proceed by induction on α < λ. As max is a function into ordinals, [{α}, max] is collapsed into an ordinal α'. First we show α ≤ α': let ξ < α be arbitrary, aiming to show ξ = π_E([{ξ}, max]) < π_E([{α}, max]). For any pair t ∈ [Ord]²,

$$\max^{\{\xi\},\{\xi,\alpha\}}(t) = \max(\operatorname{proj}_{\{\xi,\alpha\},\{\xi\}}(t)) = \min(t) < \max(t) = \max(\operatorname{proj}_{\{\xi,\alpha\},\{\alpha\}}(t)) = \max^{\{\alpha\},\{\xi,\alpha\}}(t).$$

It follows that $\mathsf{Ult}_E(\mathsf{V}) \models ``[\{\xi\}, \max] < [\{\alpha\}, \max]''$ and hence after collapsing, $\xi = \pi_E([\{\xi\}, \max])$ is less than $\alpha' = \pi_E([\{\alpha\}, \max])$. So $\alpha \leq \alpha'$.

To show that $\alpha \geq \alpha'$, let $\xi < \alpha'$ be arbitrary, say $\xi = \pi_E([r, f])$ and without loss of generality, $\alpha \in r$ (consider $[r \cup \{\alpha\}, f^{r,r \cup \{\alpha\}}]$). This means that $\forall_{E_r}^* t$ $(f(t) < \max^{r,\{\alpha\}}(t) \le \max t)$. If we don't already have $\forall_{E_r}^* t$ $(f(t) = \max(t))$ then by normality, for some $r \subseteq s \in [\lambda]^{<\omega}$, we get $\forall_{E_s}^* t$ $(f^{r,s}(t) \in t)$, and so we get this in either case. In particular by κ -completeness, for some $n < |s| < \omega$, $\forall_{F_n}^* t$ $(f^{r,s} = t_n)$ where $\{t_0, \dots, t_{|s|-1}\}$ is an increasing enumeration of *t*. In particular,

$$\forall_{E_s}^* t \ (\{f^{r,s}(t)\} = \operatorname{proj}_{s,\{s_n\}}(t)).$$

But since we already know $\forall_{E_r}^* t \ (f(t) < \max^{r, \{\alpha\}}(t))$, coherency yields that

$$\forall_{E_s}^* t \left(f^{r,s}(t) = \max(\operatorname{proj}_{s,\{s_n\}}(t)) < \max(\operatorname{proj}_{r,\{a\}}(\operatorname{proj}_{s,r}(t))) = \max(\operatorname{proj}_{s,\{a\}}(t)) \right)$$

and thus $s_n < \alpha$. Furthermore, the above tells us that $\xi = \pi_E([s, f^{r,s}]) = \pi_E([\{s_n\}, \max])$. As $s_n < \alpha$, the inductive hypothesis gives $\pi_E([\{s_n\}, \max]) = s_n$ and thus $\xi = s_n < \alpha$. Hence $\alpha' \le \alpha$, and $\alpha' = \alpha$.

2. This is easily shown by induction on |r|. So for $r = \{r_0\}$, note that $[\{r_0\}, id]$ is a singleton in $Ult_E(V)$, since by coherency, for each $s \supseteq r$, $\mathrm{id}^{\{r_0\},s}(t) = \mathrm{proj}_{s,\{r_0\}}(t)$ (which is clearly a singleton) for E_s -almost every t. But then note that $\max(t) \in t$ for every t so that $Ult_E(V) \models "[\{r_0\}, \max] \in [\{r_0\}, \operatorname{id}]"$, meaning $r_0 \in \pi_E([\{r_0\}, id])$, as desired.

And for $|r| > 1, \xi \in \pi_E([r, id])$ means $[\{\xi\}, \max] \in ^{Ult_E(V}[r, id]]$. Writing $r' = r \cup \{\xi\}$, this is equivalent to $\forall_{E_{r'}}^* t \, (\max^{\{\xi\}, r'}(t) \in \operatorname{proj}_{r', r}(t))$. In particular by κ -completeness, this is equivalent to the existence of some $\zeta \in r$ where $\forall_{E_{r'}}^* t (\max^{\{\xi\}, r'}(t) \in \operatorname{proj}_{r', \{\zeta\}}(t) = \operatorname{id}^{\{\zeta\}, r'}(t))$. This is equivalent to the existence of some $\zeta \in r$ with $\xi = \pi_E([\xi, \max]) \in \pi_E([\{\zeta\}, \operatorname{id}]) = \{\zeta\} \subseteq r$. Hence $\pi_E([r, \operatorname{id}]) = r$. \dashv

This allows us to conclude the following, which yields a very useful characterization of inner models of the form $cult_E(V)$ for some short extender E. Such a characterization is arguably more useful than the presentations given above, but it's not so easily representable in ZFC alone.

- 13B•6. Lemma Let E be a (κ, λ) -extender. Therefore, j_E is elementary with $cp(j_E) = \kappa, \pi_E([r, f]) = j_E(f)(r)$ for all appropriate r, f, and

$$\operatorname{Ult}_{E}(\mathbf{V}) = \{ j_{E}(f)(r) : r \in [\lambda]^{<\omega} \land f : [\lambda]^{|r|} \to \mathbf{V} \}.$$

Proof .:.

In essence, we just need to show that $\pi_E([r, f]) = j_E(f)(r)$. By (2) of Lemma 13 B • 5, $\pi_E([r, id]) = r$ and $j_E(f) = \pi_E(\tilde{j}_E(f))$ so we need to show $\pi_E([r, f]) = \pi_E(\tilde{j}_E(f)([r, id]))$, removing the π_E s, equivalently $\mathsf{Ult}_E(\mathsf{V}) \vDash "[r, f] = \tilde{j}_E(f)([r, \mathrm{id}]) = [\emptyset, \mathrm{const}_f]([r, \mathrm{id}])".$

But using Łoś's Theorem for Extenders (13 A • 9), this is immediate: $\forall_{F_u}^* t$ ($f(t) = (\text{const}_f(t))(\text{id}(t))$). \dashv

Already this hints at the relation between short extenders derived from elementary embeddings: if we start with an extender E, take the ultrapower embedding $j_E : V \to \text{cUlt}_E(V)$ and then derive an extender $E_{\text{lh}(E)}^{j_E}$, we get the same ultrapower. But more than this, we actually get *equality* between E and $E_{lh(E)}^{j_E}$. Note that the result doesn't rely on Lemma 13 B • 6, but just Lemma 13 B • 5.

– 13 B•7. Theorem –

Let *E* be a (κ, λ) -extender. Therefore, the derived (κ, λ) -extender $E_{\lambda}^{j_E} = E$.

Proof .:.

Write F for $E_{\lambda}^{j_E}$. First we must show that each $\kappa_r^E = \kappa_r^F$. From here the proof is an easy identification of E_r with F_r by simple calculations.

Claim 1 For each $r \in [\lambda]^{<\omega}$, $\kappa_r^E = \kappa_r^F$. Proof .:. This is given by the minimality of the κ_r s: κ_r^E is the minimal ordinal such that $\kappa_r^{|r|} \in E_r$ and κ_r^F is the minimal ordinal such that $j_E(\kappa_r) > \max(r)$. For the sake of readability, let $r_0 = \max(r)$. For every α , $j_E(\alpha) > r_0$ is equivalent to, by Theorem 13 B • 4 and Lemma 13 B • 5, that $\pi_E(\tilde{j}_E(\alpha)) > r_0$ $\pi_E([\{r_0\}, \max])$. So in the ultrapower, using Łoś's Theorem for Extenders (13 A • 9), this is equivalent to $\mathsf{Ult}_{E}(\mathsf{V}) \models "[\emptyset, \mathrm{const}_{\alpha}] > [\{r_{0}\}, \mathrm{max}] = [r, \mathrm{max}^{\{r_{0}\}, r}] = [r, \mathrm{max}]"$ iff $\forall_{E_n}^* t \ (\alpha > \max(t))$ iff $\{t \in [\kappa_r^E]^{|r|} : \max(t) < \alpha\} = [\min(\kappa_r^E, \alpha)]^{|r|} \in E_r$ iff $\alpha \geq \kappa_r^E$, by the minimality of κ_r^E . The minimal such α being κ_r^F gives the desired equality: $\kappa_r^F = \kappa_r^F$. \neg So we can refer to both $\kappa_r^E = \kappa_r^F$ as just κ_r . Let $r \in [\lambda]^{<\omega}$ and let $X \subseteq [\kappa_r]^{<\omega}$. Therefore by Lemma 13 B • 5, $X \in E_r$ iff $\forall_{E_r}^* t \ (t \in X)$ iff $\forall_{E_r}^* t \ (id(t) \in const_X(t))$ iff $\mathsf{Ult}_E(\mathsf{V}) \vDash "[r, \mathrm{id}] \in [r, \mathrm{const}_X] = [\emptyset, \mathrm{const}_X]$ " iff $\mathsf{cUlt}_E(\mathsf{V}) \vDash "r = \pi_E([r, \mathrm{id}]) \in \pi_E([\emptyset, \mathrm{const}_X]) = j_E(X)"$ iff $X \in F_r$. Hence $E_r = F_r$ and so the derived extender F = E. \neg

So far, we have concluded an equivalence between two types of short extenders: those derived from an elementary embedding, and those more generally following Definition 13 B • 1. As we are often more interested in the ultrapowers $cUlt_E(V)$ rather than the extenders themselves, it can also be useful to understand when an elementary $j : V \to M$ is the result of an extender.

- 13B•8. Corollary Let $j : V \to M$ be traditional with $cp(j) = \kappa$. Therefore $M = \{j(f)(r) : r \in [\lambda]^{<\omega} \land f : [\lambda]^{<\omega} \to V\}$ for some $\lambda \in Ord$ iff $j = j_E$ and $M = cUlt_E(V)$ for some (κ, λ) -extender E.

Proof .:.

Lemma 13 B • 6 gives the (\leftarrow) direction. So suppose M has the above form. By Theorem 13 A • 13 (9), M = $\operatorname{cUlt}_E(V)$ for $E = E_{\lambda}^{j}$. To see $j = j_E$, first factor $j = \tilde{k}_E \circ \tilde{j}_E$ by Theorem 13 A • 13 (6). Note that \tilde{k}_E : $\operatorname{Ult}_E(V) \to M = \operatorname{cUlt}_E(V)$ is surjective and elementary by Definition 13 A • 12 and the hypothesis on M. Hence \tilde{k}_E must be the unique transitive collapse map π_E : $j = \pi_E \circ \tilde{j}_E$ which is just j_E by Theorem 13 B • 4 (8). \dashv

On the topic of looking at the underlying universe of ultrapowers, we get a nice corollary that shows precisely what the ultrapowers $cUlt(V, E_r)$ look like. Of course, we know they take the form

$$\operatorname{cUlt}(\mathsf{V}, E_r) = \{ j_r(f)(s) : f : [\kappa_r]^{|r|} \to \mathsf{V} \}.$$

for some $s \in \text{cUlt}(V, E_r)$ by Theorem 12 B • 12. But in $\text{cUlt}_E(V)$, their *copy* (which isn't necessarily transitive) takes the following unsurprising form.

— 13 B•9. Corollary –

Let *E* be a (κ, λ) extender, and let $r \in [\lambda]^{<\omega}$. Therefore, $j_{r,\infty}$ " $\operatorname{cUlt}(V, E_r) = \{j_E(f)(r) : f : [\kappa_r]^{|r|} \to V\}$.

INTRODUCING EXTENDERS

Proof .:.

An arbitrary element $\pi_r([f]_{E_r}) \in \text{cUlt}(V, E_r)$ for some $f : [\kappa_r]^{<\omega} \to V$ has by Theorem 13 B • 4 (7) and (5), and Lemma 13 B • 6 that

$$j_{r,\infty}(\pi_r([f]_{E_r})) = \pi_E(\tilde{j}_{r,\infty}([f]_{E_r})) = \pi_E([r, f]) = j_E(f)(r). \quad \dashv$$

In the vein of showing results similar to those for measures and ultrafilters, we have the following, analogous to Result $12 \text{ C} \cdot 1$ (4) with a proof due to Farmer Schlutzenberg [10].

− 13B•10. Lemma Let *E* be an extender. Therefore $E \notin \text{cUlt}_E(V)$.

Proof .:.

Let $cp(E) = \kappa$ and $lh(E) = \lambda$. Let $\sigma = \sup\{\kappa_r + 1 : r \in [lh(E)]^{<\omega}\}$. Proceed by induction on σ . Assume $E \in cUlt_E(V)$. Note that this implies $\mathcal{O}(\kappa_r)^{cUlt_E(V)} = \mathcal{O}(\kappa_r)$ since $\mathcal{O}([\kappa_r]^{<\omega}) = \{X, [\kappa_r]^{<\omega} \setminus X : X \in E_r\}$.

- Write $M_0 = V$, $M_1 = cUlt_E(V)$ and $M_2 = cUlt_E^{M_1}(M_1)$.
- Let $j_E^{\mathbf{M}_0} : \mathbf{M}_0 \to \mathbf{M}_1$ and $j_E^{\mathbf{M}_1} : \mathbf{M}_1 \to \mathbf{M}_2$ be the ultrapower maps with similar notation for π_E and so forth.

We use Kunen's Inconsistency Theorem Version 2 (12 D • 6) while showing that $j \upharpoonright \eta \in M$ for some fixed point $\eta > \kappa$ to get a contradiction.

Suppose $\sigma = \varsigma + 1$ is a successor. By definition of σ , ς is then κ_r for some $r \in [\lambda]^{<\omega}$. It follows that $\mathcal{O}(\varsigma) \in M_0$ and hence $|\varsigma|^+ = (|\varsigma|^+)^{M_0}$. Since $|\varsigma|^+$ is regular, any function from $[\varsigma]^{<\omega} \to |\varsigma|^+$ is bounded and is thus in $|[\varsigma]^{<\omega}|_{\varsigma}$ which can be coded by $|\varsigma|^{|\varsigma|} = 2^{|\varsigma|}$. In other words, M_1 contains every $f : [\varsigma]^{<\omega} \to |\varsigma|^+$. Since ordinals $\alpha \leq |\varsigma|^+$ in the ultrapowers M_0 and M_1 are the collapsed versions of these functions, we get the following.

Claim 1

$$j_E^{\mathbf{M}_0} \upharpoonright (|\varsigma|^+ + 1) = j_E^{\mathbf{M}_1} \upharpoonright (|\varsigma|^+ + 1) \text{ and moreover, } j_E^{\mathbf{M}_0}(|\varsigma|^+) = \sup j_E^{\mathbf{M}_0} ||\varsigma|^+.$$

Proof ::.

 $j_E^{\mathsf{M}_0}(|\varsigma|^+)$ is just the ordertype of $A = \{[r, f]_E^{\mathsf{M}_0} : r \in [\lambda]^{<\omega} \land f : [\kappa_r]^{<\omega} \to |\varsigma|^+\}$ under $\in_E^{\mathsf{M}_0}$. But because M_1 contains all such r and f, it follows that $[r, f]_E^{\mathsf{M}_0} \mapsto [r, f]_E^{\mathsf{M}_1} = [r, f]_E^{\mathsf{M}_0} \cap \mathsf{M}_1$ is an isomorphism between the $\in_E^{\mathsf{M}_0}$ -predecessors of $[\emptyset, \operatorname{const}_{\alpha}]_E^{\mathsf{M}_0}$ and the $\in_E^{\mathsf{M}_1}$ -precedessors of $[\emptyset, \operatorname{const}_{\alpha}]_E^{\mathsf{M}_1}$, meaning the two are collapsed to the same place.

Note that $j_E^{\mathsf{M}_0}|_{\mathcal{S}}|^+$ is unbounded in $j_E^{\mathsf{M}_0}(|_{\mathcal{S}}|^+)$. To see this, any f and r such that $\forall_{E_r}^* t$ $(f(t) < \operatorname{const}_{|_{\mathcal{S}}|^+}(t))$ has f bounded by some $\alpha < |_{\mathcal{S}}|^+$ and thus $\pi_E^{\mathsf{M}_0}([r, f]) < j_E^{\mathsf{M}_0}(\alpha)$ for some $\alpha < |_{\mathcal{S}}|^+$. \dashv

Now because $j_E^{M_0"}|\varsigma|^+ = j_E^{M_1"}|\varsigma|^+ \in M_1$ is cofinal in $j_E^{M_0}(|\varsigma|^+)$, we have that the cofinality of $j_E^{M_0}(|\varsigma|^+)$ is $|\varsigma|^+$. But by elementarity, $j_E^{M_0}(|\varsigma|^+)$ is regular in M_1 , meaning the two must be equal: $j_E^{M_0}(|\varsigma|^+) = |\varsigma|^+$ and thus $|\varsigma|^+$ is a fixed point of $j_E^{M_0}$. This contradicts Kunen's Inconsistency Theorem Version 2 (12 D • 6) given that Claim 1 tells us that (as $j_E^{M_1}$ is a class of M_1) $j_E^{M_0} \upharpoonright |\varsigma|^+ = j_E^{M_1} \upharpoonright |\varsigma|^+$ is in M_1 .

Now suppose σ is a limit. It follows that $\mathcal{P}(\varsigma) \in M_1$ for each $\varsigma < \sigma$. In just the same way as with Claim 1, $j_E^{\mathsf{M}_0} \upharpoonright |\varsigma|^+ = j_E^{\mathsf{M}_1} \upharpoonright |\varsigma|^+$ for every $\varsigma < \sigma$ and hence $j_E^{\mathsf{M}_0} \upharpoonright \sigma = j_E^{\mathsf{M}_1} \upharpoonright \sigma \in \mathsf{M}_1$. This will imply $\mathcal{P}(\sigma) \in \mathsf{M}_1$ since for any $X \subseteq \sigma$, we have $\alpha \in X$ iff $j_E^{\mathsf{M}_0}(\alpha) = j_E^{\mathsf{M}_1}(\alpha) \in j_E^{\mathsf{M}_0}(X) \in \mathsf{M}_1$ so that $X = \{\alpha < \sigma : j_E^{\mathsf{M}_1}(\alpha) < j_E^{\mathsf{M}_1}(X)\} \in \mathsf{M}_1$. But then just as before, $|\sigma|^+ = (|\sigma|^+)^{\mathsf{M}_1}$ and so we also get Claim 1 for σ instead of just for each $\varsigma < \sigma$. By the same argument as before, $|\sigma|^+$ is a fixed point of $j_E^{\mathsf{M}_0}$ but $j_E^{\mathsf{M}_0} \upharpoonright |\sigma|^+ = j_E^{\mathsf{M}_1} \upharpoonright |\sigma|^+ \in \mathsf{M}_1$ contradicting Kunen's Inconsistency Theorem Version 2 (12 D \cdot 6).

 \dashv

§13C. The benefit of extenders, and their properties

What is the point of extenders? Can we just get by just with measures? Note that the inner models given by measures are certainly encompassed by those given by extenders in the following sense.

- 13C•1. Result

Let U be a measure over κ . Therefore there is a $(\kappa, \kappa + 1)$ -short extender E with $\mathsf{cUlt}(\mathsf{V}, U) = \mathsf{cUlt}_E(\mathsf{V})$.

Proof .:.

As each $r \in [\kappa]^{<\omega}$ has E_r as principle, $\mathsf{cUlt}(\mathsf{V}, E_r) = \mathsf{V}$ in these cases. We will see that it suffices to show $\mathsf{Ult}(\mathsf{V}, E_r) \cong \mathsf{Ult}(\mathsf{V}, E_{\{\kappa\}})$ for all $\kappa \in r \in [\kappa + 1]^{<\omega}$. In this case, $\mathsf{Ult}_E(\mathsf{V}) \cong \mathsf{Ult}(\mathsf{V}, E_{\{\kappa\}})$.

 $- \text{ Claim 1} \qquad - \text{ Ult}(\mathbf{V}, E_r) \cong \text{ Ult}(\mathbf{V}, E_{\{\kappa\}}) \text{ for each } r \in [\kappa + 1]^{<\omega} \setminus [\kappa]^{<\omega}.$

Proof .:.

Consider the elementary $\tilde{j}_{\{\kappa\},r}$: Ult(V, $E_{\{\kappa\}}$) \rightarrow Ult(V, E_r). It suffices to show surjectivity. Write r' for $r \setminus \{\kappa\}$. As $r' \in [\kappa]^{<\omega}$, $E_{r'}$ is principle, generated by r', and thus by coherency,

 $\forall_{E_{r'}}^* t \ (\mathrm{id}(t) = \mathrm{const}_{r'}(t)) \ \text{ implies } \ \forall_{E_r}^* t \ (\mathrm{proj}_{r,r'}(t) = r').$

So let $f : [\kappa]^{<\omega} \to V$ be arbitrary. Define $f' : [\kappa]^1 \to V$ by $f'(\{\alpha\}) = f(r' \cup \{\alpha\})$. Hence $f'^{\{\kappa\},r}(t) = f(r' \cup \{\max t\})$. So in the transitive collapse, $j(f'^{\{\kappa\},r})(r) = j(f)(j(r') \cup \{\kappa\}) = j(f)(r' \cup \{\kappa\}) = j(f)(r' \cup \{\kappa\}) = j(f)(r)$. Therefore,

$$\pi_E \circ \tilde{j}_{r,\infty} \left(\tilde{j}_{\{\kappa\},r} \left(\left[f' \right]_{E_{\{\kappa\}}} \right) \right) = \left[r, f'^{\{\kappa\},r} \right] = \left[r, f \right] = \pi_E \circ \tilde{j}_{r,\infty} \left(\left[f \right]_{E_r} \right),$$

is surjective

meaning $\tilde{J}_{\{\kappa\},r}$ is surjective.

Now we give a $(\kappa, \kappa+1)$ -short extender E with $Ult(V, E_{\{\kappa\}}) \cong Ult(V, U)$. Firstly, for $X \subseteq [\kappa]^1$ and $f : [\kappa]^1 \to V$, define (s for *singleton*) $X^s \subseteq \kappa$ and $f^s : \kappa \to V$ in the natural way: $X^s = \bigcup X$ and $f^s(\alpha) = f(\{\alpha\})$. Now define the following:

- let $j: V \to cUlt(V, U)$ be the canonical embedding with $E_{\kappa+1}^{J}$ the derived short extender;
- let s : Ult(V, $E_{\{\kappa\}}$) \rightarrow Ult(V, U) be defined by s($[f]_{E_{\{\kappa\}}}$) = $[f^s]_U$.

Thus using Lemma 12 B • 10 with Definition 12 B • 7, $X \in E_{\{\kappa\}}$ iff $\{\kappa\} \in j(X)$ iff $\kappa \in j(X^s)$ iff $X^s \in U$. And so s is well-defined, and an embedding. Surjectivity follows since $f = \{\langle \{\alpha\}, y \rangle : f(\alpha) = y \}^s$ for any $f : \kappa \to V$. \dashv

There are similarities between measures and short extenders that go beyond this. For example, for a measure U, $U \notin \text{cUlt}(V, U)$; and similarly for an extender $E, E \notin \text{cUlt}_E(V)$, which tells us that—because we can code E as an element of $V_{\text{lh}(E)+1}$ — $V_{\text{lh}(E)+1} \not\subseteq \text{cUlt}_E(V)$. This motivates some important concepts.

— 13C•2. Definition —

Let *E* be a (κ, λ) -extender.

- κ is the *critical point* of E, cp(E);
- λ is the *length* of *E*, lh(*E*); and
- the strength of E, str(E), is the largest $\alpha \in \text{Ord such that } E$ is α -strong, meaning $V_{\alpha} \subseteq \text{cUlt}_{E}(V)$.
- the completeness of E, cpl(E) is the largest cardinal $\alpha \in Ord$ such that E is $< \alpha$ -complete, meaning $<^{\alpha} cUlt_E(V) \subseteq cUlt_E(V)$.

So we can in principle do better than just measures with extenders, since a measure overly has, in essence, a strength of $\kappa + 1$ by Result 12 C • 1 (4), as expected by Result 13 C • 1 where $cp(E) = \kappa < str(E) \le \kappa + 1 = lh(E)$ for the extender defined there. In particular, we can assert the existence of an extender with arbitrary strength. This leads to the idea of the *strength* of a cardinal according to how strong its extenders are.

— 13C·3. Corollary -

Suppose *E* is a κ + 2-strong (κ , λ)-extender. Therefore κ is the limit of measurable cardinals, and in fact, there's a measure *U* such that *U*-almost every cardinal below κ is measurable.

Proof .:.

Consider the derived measure from the extender embedding: $U = U_{j_E}$ where $j_E : V \to \text{cUlt}_E(V)$. Note that $U \subseteq \mathcal{O}(\kappa)$ and hence $U \in V_{\kappa+2}$. Since E is $\kappa + 2$ -strong, $U \in \text{cUlt}_E(V)$ and hence

 $cUlt_E(V) \models$ "there's a measure over a cardinal $< j_E(\kappa)$ ".

This means $\kappa \in j(\{\delta < \kappa : \delta \text{ is measurable}\})$ and so $\{\delta < \kappa : \delta \text{ is measurable}\} \in U$.

Similar results happen for κ + 3-strong cardinals: κ is the limit of κ + 2-strong cardinals and so on. There is, however, a limit on how strong an extender can be, as evidenced by Kunen's Inconsistency Theorem Version 2 (12 D • 6).

- **13C**•4. **Result** Let *E* be an extender. Therefore 1. $\operatorname{cp}(E) < \operatorname{str}(E) \le \operatorname{lh}(E)$ and $\operatorname{cpl}(E) \le |\operatorname{sup}_{r \in [\lambda]} < \omega \kappa_r^E|^+$. 2. if *E* is short, $\operatorname{lh}(E) \le j_E(\operatorname{cp}(E))$ and $\operatorname{cpl}(E) \le \operatorname{cp}(E)^+$.

Proof .:.

Write $\operatorname{cp}(E) = \kappa$, $\operatorname{lh}(E) = \lambda$, $\operatorname{str}(E) = \rho$, and $\operatorname{cpl}(E) = \delta$. Write σ for $|\operatorname{sup}_{r \in [\lambda] \le \omega} \kappa_r^E|$.

1. Clearly $\rho \ge \kappa + 1$ by Result 12 A • 8. To see that $\rho \le \lambda$, note that *E* can be coded as a subset of V_{λ} and so by an element of $V_{\lambda+1}$. Hence if $\rho \ge \lambda + 1$, then $E \in \text{cUlt}_E(V)$, which contradicts Lemma 13 B • 10.

To see that $\delta \leq \sigma^+$, suppose otherwise: that σ^+ cUlt_E(V) \subseteq cUlt_E(V). It follows that $j_E \upharpoonright \sigma^+ : \sigma^+ \to Ord$ is in cUlt_E(V) and hence $j_E "\sigma^+ \in cUlt_E(V)$ and so has the form $\pi_E([r, f]) = j_E(f)(r)$ for some $r \in [\lambda]^{<\omega}$ and $f : [\kappa_r]^{<\omega} \to Ord$. As E_r is an ultrafilter, either $\forall_{E_r}^* t \ (|f(t)| \leq \sigma)$ or $\forall_{E_r}^* t \ (|f(t)| > \sigma)$.

- a. In the first case, $\sigma^+ \setminus \bigcup_{t \in [\kappa_r] \le \omega} f(t)$ is a σ^+ -sized set minus a $|\kappa_r| \le \sigma$ -sized set and is hence non-empty. Such an $\alpha < \sigma^+$ has $\forall t \ (\alpha \notin f(t))$ and hence $j_E(\alpha) \notin j_E(f)(r)$, a contradiction.
- b. In the second case, where $\forall_{E_r}^* t \ (|f(t)| > \sigma)$, enumerate $[\kappa_r]^{<\omega}$ by σ and then inductively choose elements of f(t) not chosen before (whenever $|f(t)| > \sigma$). The result is an injective choice function c such that $c(t) \in f(t)$ whenever $|f(t)| > \lambda$. It follows that $\forall_{E_r}^* t \ (\text{const}_x(t) \neq c(t) \in f(t))$ for every $x \in V$ and hence $j_E(x) \neq j_E(c)(r) \in j_E(f)(r)$ so that $j_E(c)(r) \in j_E(f)(r) \setminus j_E$ "V and so $j_E(f)(r) \neq j_E$ " λ^+ , a contradiction.
- 2. If *E* is short, Result 13 A 3 implies $\lambda \leq j_E(\kappa)$: $\kappa = \kappa_{\{\alpha\}}$ for every $\kappa \leq \alpha < \lambda$ and hence $j_E(\kappa) > \alpha$ for every such α . This also tells us that $\sup_{r \in [\lambda]^{<\omega}} \kappa_r = \kappa$ so $\delta \leq \kappa^+$ by (1). \dashv

This partially highlights the importants of long extenders: if we want more closure of the ultrapower, we need larger and larger κ_r s. For the most part though, we will be focused on short extenders, because the current inner model theory that focuses on them is richer.

Thinking about κ 's strength leads to the idea of a *strong* cardinal, one which is arbitrarily strong. Strong cardinals can be defined with them in a way that they can't with mere measures, or at least in a way using ultrapowers from measures.

13C · 5. Definition Let κ be a cardinal. κ is *strong* iff for every set $x \in V$, there is an elementary $j : V \to M$ with $M \subseteq V$ an inner model such that $cp(j) = \kappa$ and $x \in M$.

We shouldn't expect to be able to express this through a some special measure U, because taking the ultrapower loses $U: U \in \{W : \mathbf{V} \models ``W \text{ is a measure over } \kappa''\} \notin \text{cUlt}(\mathbf{V}, U)$ as in Result 12 C • 1 (4)^{xvii}. Moreover, in the ultrapower,

xviiThis, of course, is not a proof that a measure U on κ can't be in M: any strong cardinal will be measurable, and will have an elementary

 $j(\kappa) \le 2^{\kappa}$ by Result 12 C • 1 (4). But the following characterization tells us that we will need embeddings with $j(\kappa) > 2^{\kappa}$, and so measures aren't sufficient on their own. Extenders, however, allow us to characterize the property of being strong in much the same way that measures characterize measurables.

13C•6. Theorem — Let κ be a cardinal. Therefore, κ is strong iff for each $\alpha > \kappa$ there is an α -strong short extender E with $cp(E) = \kappa$.

This equivalence is not difficult to show. But a major thing to investigate in general is why and for which αV_{α} is contained in cUlt_{*E*}(V). Indeed, it's not entirely obvious that we can define what it means for an extender to be α -strong. We will prove Theorem 13 C • 6 after we have introduced a lemma and investigated the first-order definability of strength.

- 13C·7. Result — The strength of an extender E is FOL-definable from E.

Proof .:.

Having $V_{\alpha} \subseteq \text{cUlt}_{E}(V)$ requires any $x \in V_{\alpha}$ to be represented in the collapse of the ultrapower. But it's not obvious in general how to find a representative of x. So instead, note that $x \in \text{cUlt}_{E}(V)$ iff $\text{trcl}(\{x\}) \in$ $\text{cUlt}_{E}(V)$. What's useful about this is that $\text{trcl}(\{x\})$ is already transitive and so if $\langle \text{trcl}(\{x\}), \in \rangle$ is isomorphic to the $\in^{\text{Ult}_{E}(V)}$ -predecessors of [r, f] ordered by $\in^{\text{Ult}_{E}(V)}$, then the collapse [r, f] in $\text{cUlt}_{E}(V)$ is precisely $\text{trcl}(\{x\})$ by the uniqueness of The Mostowski Collapse (4 • 1). Hence E is α -strong iff for every $x \in V_{\alpha}$, there is an $r \in [\lambda]^{<\omega}$, an $f : [\kappa_{r}]^{<\omega} \to V$, and an isomorphism F between $\langle \text{trcl}(\{x\}), \in \rangle$ and $\langle \{[s, g] \in \text{Ult}_{E}(V) : \langle s, g \rangle \in_{E} \langle r, f \rangle \}, \in_{E} \rangle$. Since by Scott's Trick the equivalence class [s, g] refers to an actual set, this is all first-order definable. \dashv

Returning to Theorem 13 C • 6, in essence, a strong cardinal has short extenders of arbitrarily large strength. In dealing more generally with $j : V \to M$, for $E = E_{\lambda}^{j}$, we can replace $\text{cUlt}_{E}(V)$ in Definition 13 C • 2 with M in the sense that E is at least α -strong whenever $V_{\alpha} \subseteq M$ and $|V_{\alpha}|^{+} < j(\kappa)$.

— 13C•8. Lemma

Let $j : V \to M$ be traditional with $cp(j) = \kappa$. Suppose $V_{\alpha} \subseteq M$ with $|V_{\alpha}|^+ < \lambda \leq j(\kappa)$. Therefore $E = E_{\lambda}^{j}$ is an α -strong (κ, λ) -short extender.

Proof .:.

For $\alpha \leq \kappa + 1$, we already know that any such *E* is α -strong. So assume $\alpha > \kappa + 1$. In essence, we will work with a coded version of V_{α} to show that $k_E(V_{\alpha}) = V_{\alpha} \in \text{cUlt}_E(V)$ where $k_E : \text{cUlt}_E(V) \rightarrow M$ is given by Theorem 13 A • 13 (11): $j = k_E \circ j_E$. We do this as follows. Let $\mu = |V_{\alpha}| < \lambda$ and consider the corresponding subset of $\mu \times \mu$ isomorphic to $\langle V_{\alpha}, \epsilon \rangle$. Because $k_E \upharpoonright \lambda = \text{id}$, this means

$$\mathbf{M} \vDash ``\exists R \subseteq k(\mu) \times k(\mu) (\langle k(\mu), R \rangle \cong \langle \mathbf{V}_{k(\alpha)}, \epsilon \rangle)'$$

so by elementarity,

$$\mathsf{cUlt}_F(\mathsf{V}) \models ``\exists R \subseteq \mu \times \mu (\langle \mu, R \rangle \cong \langle \mathsf{V}_{\alpha}, \in \rangle)$$
''

So $\langle \mu, k(R) \rangle \cong V_{k(\alpha)}^{\mathsf{M}} = V_{\alpha}$. Again, since $k_E \upharpoonright \lambda = \operatorname{id}, k(R) = R \in \operatorname{cUlt}_E(V)$ (since $\xi \in k(R)$ iff $k(\xi) \in k(R)$ iff $\xi \in R$) so that $\langle \mu, R \rangle \cong \langle V_{\alpha}^{\operatorname{cUlt}_E(V)}, \epsilon \rangle \cong \langle V_{\alpha}, \epsilon \rangle$, showing that $V_{\alpha} \subseteq \operatorname{cUlt}_E(V)$. \dashv

A combinatorial result is the following.

– 13 C • 9. Corollary

Let $j : V \to M$ be traditional with $cp(j) = \kappa < \lambda \le j(\kappa)$ with λ a cardinal that is \square -closed (i.e. $\eta < \lambda$ implies $|V_{\eta}| < \lambda$) and $V_{\lambda} \subseteq M$. Therefore there is a λ -strong (κ, λ) -short extender *E*.

Proof .:.

Since λ is a strong limit, for any $\alpha < \lambda$, $|V_{\alpha}|^+ < \lambda$ and hence applying Lemma 13 C • 8, $V_{\alpha} \subseteq \text{cUlt}_E(V)$ where

 $j: V \to M$ with $U \in M$, for example. It's just that in these cases, $cUlt(V, U) \neq M$.

Н.

 $E = E_{\lambda}^{j}$. Taking the union of such V_{α} s yields that $V_{\lambda} \subseteq cUlt_{E}(V)$

As a result, we can easily characterize strong cardinals in terms of sufficiently strong short extenders. $Proof of Theorem 13 C \cdot 6 ...$

Suppose κ is strong. To generate an α -strong short extender, let $j : V \to M$ be traditional with $cp(j) = \kappa$ and $V_{\alpha+\omega} \in M$. Note that then $|V_{\alpha}|^+ < j(\kappa)$ since $j(\kappa)$ is inaccessible in **M** and **M** correctly computes the cardinality of V_{α} due to $V_{\alpha+\omega} \subseteq M$. As a result, using Lemma 13 C • 8 gives that $E_{j(\kappa)}^{j}$ is an α -strong $(\kappa, j(\kappa))$ short extender.

For the other direction, let $x \in V$ be arbitrary. Let $\alpha = \operatorname{rank}(x)$ so by hypothesis, there is an $\alpha + 1$ -strong short extender E with $\operatorname{cp}(E) = \kappa$. Thus $j_E : V \to \operatorname{cUlt}_E(V)$ has $\operatorname{cp}(j_E) = \kappa$ and $x \in V_{\alpha+1} \subseteq \operatorname{cUlt}_E(V)$, as desired. \dashv

Strength also allows us to properly calculate cardinality.

- 13C·10. Theorem Let *E* be a (κ, λ) -extender with $\kappa < \operatorname{str}(E) \leq \lambda$. Therefore $\aleph_{\alpha}^{\mathsf{V}} = \aleph_{\alpha}^{\operatorname{cUlt}_{E}(\mathsf{V})}$ for $\alpha < \operatorname{str}(E)$.

Proof .:.

Proceed by induction on α . Trivially, $\aleph_{\alpha}^{\mathsf{V}} \geq \aleph_{\alpha}^{\mathsf{cUlt}_{E}(\mathsf{V})}$ since $\mathsf{cUlt}_{E}(\mathsf{V}) \subseteq \mathsf{V}$ is a transitive class, so it suffices to show $\aleph_{\alpha}^{\mathsf{V}} \leq \aleph_{\alpha}^{\mathsf{cUlt}_{E}(\mathsf{V})}$. For $\alpha < \kappa$ this is easy. For $\alpha = \kappa$, this is also easy since κ is still inaccessible in both models. For limit α , if $\xi < \aleph_{\alpha}^{\mathsf{V}}$ then $\xi \leq \aleph_{\beta}^{\mathsf{V}}$ so inductively, $\xi \leq \aleph_{\beta}^{\mathsf{cUlt}_{E}(\mathsf{V})}$ for some $\beta < \alpha$. Hence $\aleph_{\alpha}^{\mathsf{cUlt}_{E}(\mathsf{V})} = \aleph_{\alpha}^{\mathsf{V}}$.

For $\alpha + 1 > \kappa$, write \aleph_{α} for both interpretations since they're inductively equal. Note that every $\xi < (\aleph_{\alpha}^{+})^{\vee}$ is in bijection with \aleph_{α} and thus induces a well-order $\langle \aleph_{\alpha} \times \aleph_{\alpha}, R \rangle$ of order-type ξ . $R \subseteq \aleph_{\alpha} \times \aleph_{\alpha}$ can be coded by a subset of \aleph_{α} . Since $\alpha < \operatorname{str}(E)$, $\alpha + 1 \leq \operatorname{str}(E)$ and hence $\operatorname{V}_{\alpha+1} \subseteq \operatorname{cUlt}_{E}(\operatorname{V})$, i.e. $\mathcal{P}(\aleph_{\alpha}) \subseteq \operatorname{cUlt}_{E}(\operatorname{V})$. And hence this subset coding R is in $\operatorname{cUlt}_{E}(\operatorname{V})$, meaning $R \in \operatorname{cUlt}_{E}(\operatorname{V})$ and hence $\operatorname{cUlt}_{E}(\operatorname{V})$ knows $|\xi| = \aleph_{\alpha}$, as desired.

In particular, a κ + 2-strong extender has the ultrapower correctly compute κ^{++} , and similarly a κ + σ -strong extender correctly computes $\kappa^{+\sigma}$ (the σ th cardinal larger than κ).

We also have some closure properties of the ultrapower, assuming the extender is "nice".

- 13C • 11. Theorem Let *E* be a (κ, λ) -short extender with str $(E) = \ln(E) = \lambda$ where $\lambda = \kappa + \delta$ and $\operatorname{cof}(\kappa^{+\delta}) > \kappa$ for some δ . Therefore $\operatorname{cpl}(E) = \kappa^+$ in that κ cUlt_{*E*}(V) \subseteq cUlt_{*E*}(V).

Proof .:.

For any κ -length sequence $\langle X_{\alpha} : \alpha < \kappa \rangle$, $X_{\alpha} \in \text{cUlt}_{E}(V)$, we have a representation $X_{\alpha} = j_{E}(f_{\alpha})(r_{\alpha})$ for various $f_{\alpha} : [\kappa]^{<\omega} \to V$ and $r_{\alpha} \in [\lambda]^{<\omega}$ for $\alpha < \kappa$. Note that $j_{E}(\langle f_{\alpha} : \alpha < \kappa \rangle) = \langle F_{\alpha} : \alpha < j(\kappa) \rangle$ where $F_{\alpha} = j_{E}(f_{\alpha})$ for $\alpha < \kappa$. So restricting this sequence down to κ yields $j_{E}(\langle f_{\alpha} : \alpha < \kappa \rangle) \models \kappa = \langle j_{E}(f_{\alpha}) : \alpha < \kappa \rangle$. Thus it suffices to show that $\langle r_{\alpha} : \alpha < \kappa \rangle \in \text{cUlt}_{E}(V)$. To do this, becase we can code finite subsets of λ just as elements of λ , we can regard $\langle r_{\alpha} : \alpha < \kappa \rangle$ as a κ -length sequence of ordinals in $\lambda = \kappa + \delta < \kappa^{+\delta}$. Therefore $\sup_{\alpha < \kappa} r_{\alpha} < \kappa^{+\delta}$ by the cofinality restriction. It follows by the λ -strength of E and Result 12 A • 9 that $\langle r_{\alpha} : \alpha < \kappa \rangle \in H_{\kappa+\delta}^{\vee} = H_{\kappa+\delta}^{\text{cult}_{E}(V)}$. Hence the composition $\langle j_{E}(f_{\alpha})(r_{\alpha}) : \alpha < \kappa \rangle \in \text{cUlt}_{E}(V)$, as desired. \dashv

In particular, measures—which can be identified with $(\kappa, \kappa + 1)$ -extenders—have ultrapowers that are $\leq \kappa$ -closed. So they are closed under ω -length sequences, and thus are well-founded (which, of course, we already knew).

As a final note for this subsection, let's investigate ultrapowers of set-sized models, and in particular V_{α} for $E \in V_{\alpha}$.

Firstly, note that since $\mathcal{O}([\kappa]^{<\omega})$ and $\mathcal{O}([\lambda]^{<\omega})$ are absolute between V and V_{α} for $\alpha \geq \max(\kappa, \lambda) + \omega$, being an extender is absolute between V and V_{α} for $\alpha \geq \lambda + \omega$.

- 13C·12. Result Let $E \in V_{\alpha}$ be a (κ, λ) -short extender. Let $j_E : V \to \text{cUlt}_E(V)$ be the canonical embedding. Therefore $\text{cUlt}_E^{V_{\alpha}}(V_{\alpha}) = \bigcup_{\beta < \alpha} V_{j_E(\beta)}^{\text{cult}_E(V)}$.

Proof .:.

Looking back to Definition 13 A • 8, note that $\langle r, f \rangle \approx_E \langle s, g \rangle$ and $\langle r, f \rangle \in_E \langle s, g \rangle$ is absolute between V_{α} and V whenever $f, g \in V_{\alpha}$ since it only references E, r, and s which are all in V_{α} whenever $E \in V_{\alpha}$. Moreover, if $\langle r, f \rangle \in_E \langle s, g \rangle$ and $g \in V_{\alpha}$, then $\langle r, f \rangle \approx_E \langle r, f' \rangle$ for some $f' \in V_{\alpha}$: just set

$$f'(t) = \begin{cases} 0 & \text{if } f(t) \notin \mathcal{V}_{\alpha} \lor f(t) \notin g(t) \\ f(t) & \text{if } f(t) \in g(t) \in \mathcal{V}_{\alpha}. \end{cases}$$

Since $g \in V_{\alpha}$, $f(t) \in V_{\alpha}$ for E_r -almost all $t \in [\kappa]^{|r|}$. Since $\langle r, f \rangle \in_E \langle s, g \rangle$ already, it follows that $\langle r, f \rangle \approx_E \langle r, f' \rangle$. It should also be clear that $f' \in V_{\alpha}$ since $f \subseteq [\kappa]^{|r|} \times (\operatorname{im}(g) \cup \{0\}) \in V_{\alpha}$. All of this tells us that $[r, f]^{\mathsf{V}} \cap V_{\alpha} = [r, f]^{\mathsf{V}_{\alpha}}$ whenever $f \in V_{\alpha}$, and modulo this translation, their predecessors are the same too: $\mathsf{Ult}_E(\mathsf{V}) \models "s, g] \in [r, f]$ " iff $\mathsf{Ult}_E^{\mathsf{V}_{\alpha}}(\mathsf{V}_{\alpha}) \models "[s, g] \cap V_{\alpha} \in [r, f] \cap V_{\alpha}$ ".

Let $j = j_E : \mathbb{V} \to \operatorname{cUlt}_E(\mathbb{V})$. Suppose $f \in \mathbb{V}_{\alpha}$ so that im $f \subseteq \mathbb{V}_{\beta}$ for some $\beta < \alpha$. It follows that for every $r \in [\lambda]^{<\omega}$, $j(f)(r) \subseteq \mathbb{V}_{j(\beta)}^{\operatorname{cUlt}_E(\mathbb{V})}$. Hence $\operatorname{cUlt}_E^{V_{\alpha}}(\mathbb{V}_{\alpha}) \subseteq \bigcup_{\beta < \alpha} \mathbb{V}_{j(\beta)}^{\operatorname{cUlt}_E(\mathbb{V})}$. Similarly, if $x \in \bigcup_{\beta < \alpha} \mathbb{V}_{j(\beta)}^{\operatorname{cUlt}_E(\mathbb{V})}$ is represented by $j(f)(r) \in \mathbb{V}_{j(\beta)}^{\operatorname{cUlt}_E(\mathbb{V})}$ for some $\beta < \alpha$ then E_r -almost every t has $\operatorname{rank}(f(t)) < \beta$. Without loss of generality, this happens for all t, meaning $f \in \mathbb{V}_{\alpha}$, meaning [r, f] also represents x. Therefore $\bigcup_{\beta < \alpha} \mathbb{V}_{j(\beta)}^{\operatorname{cUlt}_E(\mathbb{V})} \subseteq \operatorname{cUlt}_E^{V_{\alpha}}(\mathbb{V}_{\alpha})$.

In particular, for regular $\lambda > \kappa$, $\mathsf{cUlt}_E^{\mathsf{V}_{\lambda}}(\mathsf{V}_{\lambda}) = \mathsf{V}_{j_E(\lambda)}^{\mathsf{cUlt}_E(\mathsf{V})}$. More generally, we have the following.

- 13C·13. Corollary Let *E* be a-short extender with $cp(E) = \kappa$. Let $\alpha \in Ord$. Therefore $cof(\alpha) \neq \kappa$ implies $j_E(\alpha) = \sup\{j_E(\beta) : \beta < \alpha\}$. In particular, if $cof(\alpha) \neq \kappa$, and $\alpha > \kappa$, $cUlt_E^{V\alpha}(V_{\alpha}) = V_{j_E(\alpha)}^{cUlt_E(V)}$.

Proof .:.

Let $j = j_E : V \to \text{cUlt}_E(V)$. To show that $\{j(\beta) : \beta < \alpha\}$ is unbounded in α , let $\xi < j(\alpha)$ be arbitrary. It suffices to find a $\beta < \alpha$ with $j(\beta) \ge \xi$. Represent $\xi = j(f)(r)$ by [r, f] for $r \in [\text{lh}(E)]^{<\omega}$ and $f : [\kappa]^{|r|} \to V$. Since $\text{Ult}_E(V) \models ``[r, f] < [\emptyset, \text{const}_{\alpha}]$ '', without loss of generality, im $f \subseteq \alpha$.

- First suppose $\operatorname{cof}(\alpha) < \kappa$. Let $\alpha = \sup\{\beta_{\gamma} : \gamma < \operatorname{cof}(\alpha)\}$ and now for $t \in [\kappa]^{|r|}$, let $\gamma_t < \operatorname{cof}(\alpha) < \kappa$ be the least such that $f(t) < \beta_{\gamma_t}$. Since there are at most $\operatorname{cof}(\alpha) < \kappa$ many options, κ -completeness gives some γ such that $\forall_{E_r}^* t \ (\gamma_t = \gamma)$. Hence $\operatorname{Ult}_E(\mathbf{V}) \models "[r, f] < [\emptyset, \operatorname{const}_{\beta_{\gamma}}]$ " and thus $\xi = j(f)(r) < j(\beta_{\gamma})$.
- Now suppose $cof(\alpha) > \kappa$ so that f is bounded in α by some β . It follows that $Ult_E(V) \models$ " $[r, f] < [\emptyset, const_{\beta}]$ " and hence $\xi = j(f)(r) < j(\beta)$.

Note that if $\lambda > \kappa$ is regular, $\text{cUlt}_E(V) \models "j_E(\lambda)$ is regular". This might seem to conflict with Corollary 13 C • 13, which says $\operatorname{cof}(j_E(\lambda)) = \operatorname{cof}(\lambda) = \lambda$, but note that $\operatorname{cUlt}_E(V)$ doesn't have access to j_E and—being closed only under κ -length sequences, not λ -length ones—doesn't know about the λ -length sequence $\langle j_E(\alpha) : \alpha < \lambda \rangle$ whereas V does.

§ 13 D. Generators and length

Beyond these $\kappa_r s$, there's the notion of a *generator* of an extender and the *natural length* of an extender, being the limit of the generators. Note that we can trivially get extenders with arbitrarily large length that fundamentally are the same as an extender with a smaller length. The natural length is the minimal such length.

— 13 D•1. Example -

Let *E* be a $(\kappa, \kappa + 1)$ -short extender and let $\lambda > \kappa$ be arbitrarily large. Let $F = E_{\lambda}^{j_E}$, the (κ, λ) -short extender derived from $j_E : V \to \text{cUlt}_E(V)$. Therefore $\text{cUlt}_E(V) = \text{cUlt}_F(V)$.

To prove the statement of Example $13 D \cdot 1$, we need to examine what happens when we "cut off" extenders to yield other extenders.

— 13 D•2. Lemma

Let *E* be an extender and let ξ be such that $cp(E) < \xi \le lh(E)$. Therefore $E \upharpoonright \xi = \{\langle r, X \rangle \in E : r \in [\xi]^{<\omega}\}$

is a $(cp(E), \xi)$ -extender.

Proof .:.

Write $cp(E) = \kappa$, $lh(E) = \lambda$. That $E \upharpoonright \xi$ meets Definition 13 B • 1 is trivial by previous results:

- 1. Trivially each E_r is still a κ -complete ultrafilter over $[\kappa]^{<\omega}$ with $[\kappa]^{|r|} \in E_r$ the least such.
- 2. We have $\{\kappa\} \in [\xi]^{<\omega}$ which gives that $E_{\{\kappa\}}$ is a κ -complete ultrafilter that is not κ^+ -complete by Result 13 A 5, noting that $\kappa_{\{\kappa\}} = \kappa$ and $E_{\{\kappa\}}$ cannot be κ^+ -complete over κ without being principal. That $E_{\{\kappa\}}$ is non-principal follows from the fact that the derived ultrafilter $U = \{X \subseteq \kappa : \kappa \in j(X)\}$ is non-principal—Result 12 B 8—and that U is essentially the same as $E_{\{\kappa\}} = \{X \subseteq [\kappa]^1 : \{\kappa\} \in j(X)\}$ by the map $X \mapsto \{\{x\} : x \in X\} \subseteq [\kappa]^1$.
- 3. That every $\alpha < \kappa$ has an r with $\forall_{E_r}^* t \ (\alpha \in t)$ follows from the fact that such an r is just any $r \in [\kappa]^{<\omega}$ with $\alpha \in r$ by the proof of Result 13 A 5.
- 4. Coherency is immediate from the coherency of E.
- 5. Normality is immediate from the normality of *E* in conjunction with Lemma 13 A 15 to tell us that the required $s \supseteq r$ is still contained in ξ if $r \subseteq \xi$.
- 6. Well-foundedness is immediate from the well-foundedness of E.

 \dashv

With this in mind, we can also define the limit ultrapowers for $E \upharpoonright \xi$ just as with E itself.

─ 13 D • 3. Theorem

Let *E* be an extender and $cp(E) < \xi \le lh(E)$. Therefore

- 1. Ult_{*E* \\ \xi}(V) is the direct limit of the system of ultrapowers {Ult(V, *E_r*), $\tilde{j}_{r,s}$: $r \subseteq s \in [\xi]^{<\omega}$ }.}
- 3. Hence there's an elementary k_{ξ} : $\operatorname{cUlt}_{E \upharpoonright \xi}(V) \rightarrow \operatorname{cUlt}_{E}(V)$ such that $j_{E} = k_{\xi} \circ j_{E \upharpoonright \xi}$ and defined by $k_{\xi}(j_{E \upharpoonright \xi}(f)(r)) = j_{E}(f)(r)$.
- 4. $cp(k_{\xi})$, if it exists, is at least ξ .
- 5. Thus k_{ξ} " cUlt_{*E* \\color ξ}(V) = { $j_E(f)(r) : r \in [\xi]^{<\omega} \land f : [\kappa_r]^{|r|} \to V$ }.
- 6. Moreover, for $r \in [\xi]^{<\omega}$, $j_{r,\infty} = k_{\xi} \circ j_{r,\xi}$ where $j_{r,\infty}$: $\operatorname{cUlt}(V, E_r) \to \operatorname{cUlt}_E(V)$ and $j_{r,\xi}$: $\operatorname{cUlt}(V, E_r) \to \operatorname{cUlt}_{E \upharpoonright \xi}(V)$ are the direct limit embeddings.

- 1. The first statement holds by Corollary 13 A 10.
- 2. The second holds by Łoś's Theorem for Extenders (13 A 9) when we consider the embedding defined by $\tilde{k}_{\xi}([r, f]_{E \upharpoonright \xi}) = [r, f]_{E}$.
- 3. The third holds by considering the factor map $k_{\xi} = \pi_E \circ \tilde{k}_{\xi} \circ \pi_{E \uparrow \xi}^{-1}$ where $\pi_E : \text{Ult}_E(V) \to \text{cUlt}_E(V)$ is the collapsing isomorphism and similarly for $\pi_{E \uparrow \xi}$. That $j_E = k_{\xi} \circ j_{E \uparrow \xi}$ is immediate:

 $k_{\xi}(j_{E \upharpoonright \xi}(x)) = \pi_E(\tilde{k}_{\xi}([\emptyset, \operatorname{const}_x]_{E \upharpoonright \xi})) = \pi_E([\emptyset, \operatorname{const}_x]_E) = j_E(x).$

- 4. If $\alpha < \xi$, then by Lemma 13 B 5, $\pi_{E \upharpoonright \xi}([\{\alpha\}, \max]_{E \upharpoonright \xi}) = \alpha = \pi_E([\{\alpha\}, \max]_E)$ and hence $\alpha = \pi_E \circ \tilde{k}_{\xi} \circ \pi_{E \upharpoonright \xi}^{-1}(\alpha) = k_{\xi}(\alpha)$.
- 5. This follows by Lemma 13 B 6 and applying k_{ξ} , noting that $k_{\xi}(r) = r$ for $r \in [\xi]^{<\omega}$.

Proof .:.

6. Note that $j_{r,\infty}(x) = j_E(f)(r)$ whenever $\pi_r([f]_{E_r}) = x$ and similarly for $j_{r,\xi}$: $j_{r,\xi}(x) = j_{E \mid \xi}(f)(r)$. Moreover, since $k_{\xi}(r) = r$, $k_{\xi}(j_{E \mid \xi}(f)(r) = k_{\xi} \circ j_{E \mid \xi}(f)(k_{\xi}(r)) = j_E(f)(r)$. Thus for any $x \in \text{cUlt}(V, E_r)$, there's some f where $j_{r,\infty}(x) = j_E(f)(r) = k_{\xi} \circ j_{r,\xi}(x)$.

Of course, there's no reason to think that k_{ξ} needs to have a critical point ξ . For example, $k_{\kappa+1}$ for any extender of length $\geq \kappa + 2$, does not have critical point $\kappa + 1$, since this isn't even a limit ordinal in $\text{CUlt}_{E \upharpoonright \kappa + 1}(V)$. Moreover, even if $\xi = \text{cp}(k_{\xi})$, there's no reason to think ξ is a cardinal in V, just in $\text{CUlt}_{E \upharpoonright \xi}(V)$.

Why do we care about the critical points of these k_{ξ} s? They allow us to give a more natural characterization of length in the following sense.

– 13 D•4. Theorem

Let E be an extender. Let $\lambda = \sup\{\kappa + 1, \xi + 1 \le \ln(E) : \operatorname{cp}(k_{\xi}) = \xi\}$. Therefore $\operatorname{cUlt}_{E \upharpoonright \lambda}(V) = \operatorname{cUlt}_{E}(V)$.

Proof .:.

Consider \hat{k} : Ult_{E \black \lambda}(V) \rightarrow Ult_E(V) defined by $\hat{k}([r, f]_{E \black \b$

$$(\forall r \in [\rho]^{<\omega}) \ (\forall f : [\kappa_r]^{|r|} \to \mathsf{V}) \ (\exists s \in [\lambda]^{<\omega}) \ (\exists g : [\kappa_s]^{|s|} \to \mathsf{V}) \ (j_E(f)(r) = j_E(g)(s))$$
(*)

Clearly if $\rho \leq \lambda$ already we're done. So assume (*) inductively holds for all $\rho' < \rho$. If ρ is a limit, the result for ρ is immediate by induction.

So assume $\rho = \rho' + 1$ and without loss of generality, $\max(r) = \rho'$. Because $\rho' \ge \lambda$, $\rho' > \xi$ for any ξ such that $\xi = \operatorname{cp}(k_{\xi})$. So if $k_{\rho'}$ is non-trivial, then $\operatorname{cp}(k_{\rho'}) \ge \rho' + 1 = \rho$. Hence $k_{\rho'}(r) = r$ and thus $k_{\rho'}(j_{E \upharpoonright \rho'}(f)(r)) = j_E(f)(r)$ and $j_{E \upharpoonright \rho'}(f)(r) \in \operatorname{cUlt}_{E \upharpoonright \rho'}(V)$. Hence inductively, $k_{\lambda}^{-1}(j_E(f)(r)) \in \operatorname{cUlt}_{E \upharpoonright \lambda}(V)$ exists, i.e. $j_E(f)(r)$ can be represented by $j_E(g)(s)$ for $s \in [\lambda]^{<\omega}$ and $g : [\kappa_s]^{|s|} \to V$, as desired.

So these critical points are interesting to investigate. We have the following definition.

- 13 D • 5. Definition

Let *E* be an extender. An ordinal ξ with $cp(E) \le \xi \le lh(E)$ is a generator of *E* iff $\xi = cp(E)$ or $\xi = cp(k_{\xi})$. The *natural length* of *E* is $sup\{\xi + 1 : \xi \text{ is a generator of } E\} \le lh(E)$.

Note that although the generators contain the information of the extender's embedding, they are not necessarily the κ_r s associated with extender. An alternative, more first-order definition is the following, motivating why we should include the special case of cp(*E*).

– 13 D•6. Result

Let *E* be an extender and $\xi \in [cp(E), lh(E)]$. Therefore ξ is a generator iff there is no $r \in [\xi]^{<\omega}$ and $f : [\kappa_r]^{|r|} \to V$ such that $j_E(f)(r) = \xi$.

Proof .:.

Suppose ξ is a generator. Let $r \in [\xi]^{<\omega}$ and $f : [\kappa_r]^{|r|} \to V$. Since $k_{\xi}(r) = r$, it follows that $j_E(f)(r) = k_{\xi}(j_{E \upharpoonright \xi}(f)(r))$. Since $\operatorname{cp}(k_{\xi}) = \xi$, $\xi \notin \operatorname{im} k_{\xi}$ and hence $j_E(f)(r) \neq \xi$. Conversely, let $\xi = j_{E \upharpoonright \xi}(f)(r)$ in $\operatorname{cult}_{E \upharpoonright \xi}(V)$ for some $r \in [\xi]^{<\omega}$ and $f : [\kappa_r]^{|r|} \to V$. It follows that $k_{\xi}(j_{E \upharpoonright \xi}(f)(r)) = j_E(f)(r) \neq \xi$ meaning $k_{\xi}(\xi) \neq \xi$ and thus as the least possible ordinal where this happens, $\operatorname{cp}(k_{\xi}) = \xi$.

So this tells us why $cp(E) = \kappa$ should be considered a generator of E: j(f)(r) for some finite subset $r \subseteq \kappa$ and $f: [\kappa]^{|r|} \to V$ would yield $j_E(f)(r) = f(r)$. So if this were κ , then we'd have $\forall_{E_r}^* t$ $(f(t) \ge const_{\kappa}(t))$ meaning $j_E(f)(r) \ge j_E(\kappa) > \kappa$, a contradiction.

— 13 D•7. Corollary

Let E be an extender. Therefore the natural length of E is the least ρ such that $\text{cUlt}_{E \uparrow \rho}(V) = \text{cUlt}_{E}(V)$.

Proof .:.

We have $\text{cUlt}_{E \upharpoonright \rho}(V) = \text{cUlt}_E(V)$ by Theorem 13 D•4. That ρ is the least such follows from the fact that any $\alpha < \rho$ has some generator $\alpha < \xi < \rho$ with then k_{ξ} non-trivial, meaning $k_{\alpha} : \text{cUlt}_{E \upharpoonright \alpha}(V) \rightarrow \text{cUlt}_E(V)$ is non-trivial.

Another corollary tells us that we can therefore artificially *increase* the length of an extender without harm, showing Example $13 D \cdot 1$. The basic idea is that if we derive an extender from an ultrapower embedding, taking the length of the new extender to be long enough doesn't add anything to the natural length.

– 13 D•8. Corollary

Let E be an extender. Therefore there are extenders F with arbitrarily large lengths such that $cUlt_E(V) = cUlt_F(V)$.

Proof .:.

So that

Without loss of generality, assume $\ln(E)$ is the natural length of E. Let $j_E : V \to \text{cUlt}_E(V)$ be the canonical embedding, and let $F = E_{\lambda}^{j_E}$ for an arbitrarily large λ . Hence $E = F \upharpoonright \ln(E)$ and so $\kappa_r^E = \kappa_r^F$ for $r \in [\ln(E)]^{<\omega}$. Note that by Theorem 13 A • 13 (where the extender embedding is j_F and $j = j_E$ is the given traditional embedding) that

$$\operatorname{cUlt}_{E}(\mathsf{V}) = \{ j_{E}(f)(r) : r \in [\operatorname{lh}(E)]^{<\omega} \wedge f : [\kappa_{r}^{F}]^{<\omega} \to \mathsf{V} \}$$
$$\operatorname{cUlt}_{F}(\mathsf{V}) = \{ j_{E}(f)(r) : r \in [\lambda]^{<\omega} \wedge f : [\kappa_{r}^{F}]^{<\omega} \to \mathsf{V} \} \subseteq \operatorname{cUlt}_{E}(\mathsf{V}).$$
$$\operatorname{cUlt}_{E}(\mathsf{V}) \subseteq \operatorname{cUlt}_{F}(\mathsf{V}) \subseteq \operatorname{cUlt}_{E}(\mathsf{V}).$$

 \dashv

Section 14. External Ultrapowers

So far, there's nothing stopping us from considering ultrapowers for models that don't have the model in them. Indeed, in model theory, ultrapowers are taken as in Definition 12 • 1 with no issue. But for us, Definition 12 • 1 can be more of an issue because it can allow too many functions, and so we'd have little control over what the ultrapower looked like in different background universes that might have more or fewer functions. As with Corollary 13 C • 13, we can consider ultrapowers of small models, but this also includes inner models which may not have the measure inside them.

One major motivation for external ultrapowers is from external elementary embeddings. It may be that there is an elementary embedding $j : L \rightarrow L$, but this embedding cannot exist as a class of L itself. External elementary embeddings will have derived measures and extenders that allow us to make similar conclusions as previously, but require slightly more finesse to work with. Nevertheless, external elementary embeddings of models are generally more common than the internal ones (which require measurable cardinals), and so understanding and working with them will be quite useful in understanding iterations and comparisons between iterates.

§14A. External measures

For the most part, we will only consider ultrapowers by extenders and measures that are "close" to being inside the relevant models, enough so that the usual arguments about extenders apply. In particular, we require a condition called *amenability*.

— 14A•1. Definition

Let **M** be a model of (some fragment of) set theory. Let $U \subseteq \mathcal{P}(\kappa) \cap \mathbf{M}$ for some κ .

- U is weakly amenable to M iff $x \cap U \in M$ for any $x \in M$ such that $M \models "x \subseteq \mathcal{O}(\kappa) \land |x| = \kappa$ ".
- U is an M-measure iff $(M, \in, U) \models "U$ is a measure over κ ". We similarly define M-ultrafilters, M- κ completeness, M-normality, and so on.
- U is a *weakly amenable* M-*measure* iff U is an M-measure weakly amenable to (M, \in, U) .

External measures refer to U that are M-measures for some M, and where stereotypically $U \notin M$.

There are several things to consider here. Firstly, the usual non-weak version of amenability would say $x \cap U \in M$ for any $x \in M$, not just those of size κ in **M**. That of course would be too strong for us, because then $U \cap \mathcal{O}(\mathcal{O}(\kappa))^{\mathsf{M}} = U \cap \mathsf{M} = U$ would be in M. Secondly, the model $\langle \mathsf{M}, \in, U \rangle$ does not necessarily work nicely with U (excluding weak amenability). The sentence that "U is a measure over κ " is stated in the language of FOL (\in, U) , meaning we are able to ask questions about what is or is not in U including things like κ -completeness, which is stated as "for all $\beta < \kappa$, for all $f : \beta \to \mathsf{M}$ (in M) if $\forall \alpha < \beta$ ($f(\alpha) \in U$) then $\bigcap_{\alpha < \beta} f(\alpha) \in U$ " for example. This doesn't mean we can form $U = \{x \in \mathcal{O}(\kappa) : x \in U\}$ in $\langle \mathsf{M}, \in, U \rangle$ however, as there is no reason to believe that $\langle \mathsf{M}, \in, U \rangle$ satisfies the comprehension scheme lifted to the language of FOL (\in, U) . All of this is just to say that U need not be in M, which is the entire point of this discussion. This will be useful for us, especially when we consider small "toy" models of set theory like from Subsection 7 C.

The weak amenability condition defined above is a less useful formulation compared to the following.

— 14A•2. Lemma

Let **M** be a transitive model of some (suitable) fragment of set theory. Let U be an M-measure. Therefore U is weakly amenable to **M** iff for every $f : \kappa \to M$ in M, $\{x < \kappa : f(x) \in U\} = f^{-1} U \in M$.

EXTERNAL ULTRAPOWERS

Proof .:.

For one direction, suppose U is weakly amenable and $f \in {}^{\kappa}M \cap M$. Thus $|\operatorname{im}(f) \cap \mathcal{O}(\kappa)| \leq \kappa$ (just by Σ_0 -replacement to remove repeated values and reindex to get a bijection between it and κ) so by weak amenability $\operatorname{im}(f) \cap U \in M$. By Σ_0 -comprehension, $\{x < \kappa : f(x) \in \operatorname{im}(f) \cap U\} = f^{-1} U \in M$. For the other direction, suppose $f^{-1} U \in M$ for every $f \in {}^{\kappa}M \cap M$. If $x \subseteq \mathcal{O}(\kappa)$ has size $|x| \leq \kappa$, then there is some surjection $f : \kappa \to x$. It follows that $f^{-1} U = \{\alpha < \kappa : f(\alpha) \in x \cap U\}$ is in M by hypothesis. Thus $f''(f^{-1}U) = x \cap U \in M$ by Σ_0 -replacement, as desired, meaning U is weakly amenable.

The word "suitable" is used above in a vague way for the amount of ZFC that **M** satisfies. This is really just to say that we don't need to consider only inner models, but also certain set models of the form L_{α} , H_{λ} , or V_{α} . Usually this just requires enough replacement and comprehension to carry out Transfinite Recursion (3 C • 2) and define certain sets, and the reader can take it to be ZFC – P, which is more than enough. Later on, we'll use this with fine structural models of the form $J_{\alpha}[\mathcal{E}]$, which often don't satisfy a full version of comprehension or replacement, but which nevertheless satisfy enough set theory to be relevant in this discussion, although perhaps in a restricted way.

External elementary embeddings, regardless of whether they are classes of M, give rise to these M-measures.

— 14 A • 3. Result

- Let M, N be transitive models of (some fragment of) set theory such that $\mathcal{P}(\kappa)^{\mathsf{M}} = \mathcal{P}(\kappa)^{\mathsf{N}}$.
- Let $j : M \to N$ be elementary with $cp(j) = \kappa$.

Therefore, $U_j = \{X \in \mathcal{O}(\kappa)^{\mathsf{M}} : \kappa \in j(X)\}$ is a weakly amenable M-measure over κ .

Proof .:.

The fact that U is an M-measure is exactly Result 12 B • 8. So it suffices to show weak amenability. Using Lemma 14 A • 2, let $f : \kappa \to \mathcal{O}(\kappa)^{\mathsf{M}}$ in M be arbitrary. We want to show $\{\alpha < \kappa : f(\alpha) \in U\} \in \mathsf{M}$. Since this is a subset of κ and $\mathcal{O}(\kappa)$ is the same in both M and N, it suffices to show this set is in N. Note by elementarity, $j(f(\alpha)) = j(f)(j(\alpha))$ for any α . Since $\operatorname{cp}(j) = \kappa$, $j(f(\alpha)) = j(f)(\alpha)$ whenever $\alpha < \kappa$. Clearly $j(f) \in \mathsf{N}$ so that

$$\{\alpha < \kappa : f(\alpha) \in U\} = \{\alpha < \kappa : \kappa \in j(f(\alpha))\} = \{\alpha < \kappa : \kappa \in j(f)(\alpha)\} \in \mathbb{N}.$$

Hence this set $f^{-1} U \in \mathcal{O}(\kappa) \subseteq \mathbb{M}$.

Now, given such an M-measure, we can form the external ultrapower of M, but not in the sense of Definition 12 • 1, which considers $Ult(M, U) = \{[f]_U : f \in {}^{\kappa}M\}$. Again, the reason we don't consider this is that frequently, ${}^{\kappa}M \not\subseteq M$, and this can result in some things being harder to control and argue about. So instead, we consider only functions in M, and call the result Ult[M, U]. The interpretation $\in^{Ult[M,U]}$ remains as before, but now M may not be able to form the resulting model because, again, U might not be in M.

— 14A•4. Definition

Let **M** be a model of (some fragment of) set theory. Let U be an M-measure over κ . In this case, we define everything exactly as in Definition 12 • 1, but additionally

$$\mathsf{Jlt}[\mathsf{M}, U] = \{ [f]_U : f \in \mathsf{M} \cap {}^{\kappa}\mathsf{M} \}.$$

If $U \in M$ then we get back the original notion of an ultrapower interpreted in the model itself: $Ult[M, U] = Ult^{M}(M, U)$. This also gives another kind of measure over the ultrapower according to the language of FOL(\in , U): $U^{Ult[M,U]}$ should be a Ult[M, U]-measure by elementarity assuming we have weak amenability. Speaking of elementarity, we should establish Łoś's Theorem (12 • 2) for these external ultrapowers, which also features weak amenability.

– 14A•5. Theorem (Łoś's Theorem) -

- 1. Let σ be a signature.
- 2. Let $M \models AC$ be a transitive $FOL(\in)$ -model of some (suitable) fragment of set theory.
- 3. Let U be an M-measure over κ .
- 4. Let $\varphi(x)$ be a FOL(\in)-formula.

Therefore, for any $[f] \in \text{Ult}[M, U]$,

 $\mathsf{Ult}[\mathsf{M}, U] \vDash ``\varphi([f])`` iff \forall_U^* x (\mathsf{M} \vDash ``\varphi(f(x))``).$

If in addition, U is weakly amenable to M, then this also holds for $FOL(\in, U)$ -formulas.

Proof .:.

The proof of this is almost exactly the same as with $\text{Los's Theorem (12 \cdot 2)}$ with the only two hiccups:

1. The atomic formulas (specifically those asserting membership in $U^{Ult[M,U]}$) require weak amenability.

2. We need AC to hold in **M** to find witnesses for existential statements.

For (1), we'll use \dot{U} as a symbol to distinguish it from U. In this case, we have by Definition 12 • 1 that

$$\mathsf{Ult}[\mathsf{M}, U] \vDash ``[f] \in U`` \quad \text{iff} \quad \{x < \kappa : f(x) \in U\} \in U.$$

In this way, we need to have $\{x < \kappa : f(x) \in U\} \in M$ to even talk about this. Weak amenability proves this by Lemma 14 A • 2. Hence the atomic case is makes sense, and is therefore trivially true by Definition 12 • 1.

For (2), only one direction changes: suppose $\forall^* x \ \mathbf{M} \models ``\exists y \ \varphi(f(x), y)''$. We want to show $\mathsf{Ult}[\mathbf{M}, U] \models$ $``\exists y \ \varphi([f], y)''$ and thus find a $g \in {}^{\kappa} \mathbf{M} \cap \mathbf{M}$ such that $\mathsf{Ult}[\mathbf{M}, U] \models ``\varphi([f], [g])''$. To do this, we have some $X \subseteq \kappa$ in U such that $\mathbf{M} \models ``\forall x \in X \exists y \ \varphi(f(x), y)''$. So by AC let $g : \kappa \to \mathbf{M}$ be defined in \mathbf{M} by choosing some such y for each $x \in X$, and define g(x) = 0 for $x \in \kappa \setminus X$. We have $g \in \mathbf{M}$ and $\mathbf{M} \models ``\forall x \in X \ \varphi(f(x), g(x))''$ so that by structural induction on formulas, $\mathsf{Ult}[\mathbf{M}, U] \models ``\varphi([f], [g])''$ and so $\mathsf{Ult}[\mathbf{M}, U] \models ``\exists y \ \varphi([f], y)''$, as desired.

The rest of the induction and proof is exactly the same as $\text{Los's Theorem (12 \cdot 2)}$.

 \neg

 \neg

We still get Łoś's Theorem $(14 \text{ A} \cdot 5)$ for FOL(\in)-formulas if we don't have weak amenability, but if we do, then the interpretation of the M-measure is again an (external) measure. This is at the heart of why we want weak amenability: to be able to use the same measure and consider iterated ultrapowers.

– 14 A • 6. Corollary

- Let $M \models AC$ be a transitive model of some (large) fragment of set theory.
- Let U be a weakly amenable M-measure over κ .
- Suppose Ult[M, U] is well-founded with transitive collapse map $\pi : Ult[M, U] \rightarrow cUlt[M, U]$.

Write \dot{U} for a symbol with $\dot{U}^{\mathsf{M}} = U$. Therefore $\dot{U}^{\mathsf{cUlt}[\mathsf{M},U]}$ is a weakly amenable $\mathsf{cUlt}[\mathsf{M},U]$ -measure over $\pi([\mathsf{const}_{\kappa}])$.

Proof .:.

Write W for $\dot{U}^{\text{cult}[M,U]}$. We know that $\mathbf{M} \models "\dot{U}$ is a measure over $\text{const}_{\kappa}(\alpha)$ " for every $\alpha < \kappa$. So by Łoś's Theorem (14 A • 5),

 $\mathsf{Ult}[\mathsf{M}, U] \models "\dot{U}$ is a measure over $[\mathrm{const}_{\kappa}]$ ".

Hence W is a cUlt[M, U]-measure over $\pi([const_{\kappa}])$, and it suffices to show weak amenability. Suppose that $X \in \mathcal{O}(\pi([const_{\kappa}]))^{cUlt[M,U]}$ has size $|X|^{cUlt[M,U]} = \pi([const_{\kappa}])$. We'd like to show $X \cap U \in cUlt[M, U]$. Firstly, let $F \in cUlt[M, U]$ be a bijection $F : \pi([const_{\kappa}]) \to X$. We have that $X = \pi([\chi])$ and $F = \pi([f])$ for some $\chi, f \in {}^{\kappa}M \cap M$. Note that therefore

Ult[M, U] \models "[f] is a bijection with domain $[\text{const}_{\kappa}]$ and range $[\chi] \land [\chi]$ is a subset of $\mathcal{O}([\text{const}_{\kappa}])$ ". So by Łoś's Theorem (14 A • 5), $f(\alpha) : \kappa \to \chi(\alpha)$ is a bijection and $\chi(\alpha)$ is a subset of $\mathcal{O}(\kappa)^{\mathsf{M}}$ for almost every α (so without loss of generality, for every α). So consider $\bigcup_{\alpha < \kappa} \chi(\alpha) \subseteq \mathcal{O}(\kappa)^{\mathsf{M}}$ which still has size κ in M . Write $A = U \cap \bigcup_{\alpha < \kappa} \chi(\alpha)$ so that $A \in \mathsf{M}$ by weak amenability. Therefore the map $\chi' : \kappa \to \mathsf{M}$ defined by $\alpha \mapsto \chi(\alpha) \cap A = \chi(\alpha) \cap U$ is in M . We can show that $\pi([\chi']) = X \cap W$: for any $g \in {}^{\kappa}\mathsf{M} \cap \mathsf{M}$,

$$\pi([g]) \in X \cap W \quad \text{iff} \quad \mathsf{Ult}[\mathsf{M}, U] \vDash "[g] \in [\chi] \cap \dot{U}"$$

$$\text{iff} \quad \forall^* \alpha \ (g(\alpha) \in \chi(\alpha) \cap U = \chi'(\alpha))$$

$$\text{iff} \quad \mathsf{Ult}[\mathsf{M}, U] \vDash "[g] \in [\chi']" \quad \text{iff} \quad \pi([g]) \in \pi([\chi']).$$

It follows that $X \cap W = \pi([\chi']) \in \text{cUlt}[M, U]$. Hence W is weakly amenable to cUlt[M, U].

Another corollary is that we get an elementary embedding from M into Ult[M, U]. One crucial point, however, is that this embedding is not necessarily traditional: Ult[M, U] need not be an inner model of M. The proof here is again the exact same as before with the analogous result for an internal ultrapower from Theorem 12 B • 1.

— 14 A • 7. Corollary

Let $\mathbf{M} \models AC$ be a transitive model of some (suitable) fragment of set theory. Let U be an M-measure over κ . Therefore 1. $\tilde{j} : \mathbf{M} \rightarrow \text{Ult}[\mathbf{M}, U]$ is elementary, defined by $\tilde{j}(x) = [\text{const}_x]$.

2. Moreover, if $i : M \to N$ is elementary with N transitive and such that U is the derived measure, then there is an elementary $k : \text{Ult}[M, U] \to N$ defined by $k([f]) = i(f)(\kappa)$. In this case, Ult[M, U] is well-founded, and the transitive collapse cUlt $[M, U] = \{i(f)(\kappa) : f \in {}^{\kappa}M \cap M\}$.

Proof .:.

- 1. This follows just by Łoś's Theorem (14 A 5): $\mathbf{M} \models "\varphi(x)"$ iff $\forall \alpha \ (\mathbf{M} \models "\varphi(const_x(\alpha)")$ iff $\mathsf{Ult}[\mathbf{M}, U] \models "\varphi(\tilde{j}(x))"$.
- 2. The proof that k exists and takes this form is just Factoring (12 B 9). This implies Ult[M, U] is well-founded, because any ill-founded sequence $\langle x_n : n < \omega \rangle$ (in V) yields an illfounded sequence $\langle k(x_n) : n < \omega \rangle$

We also get a number of similar results as with internal ultrapowers, similar to Result 12 C • 1 and Lemma 12 B • 10.

— 14A•8. Definition –

Let $M \models AC$ be a transitive model of some (suitable) fragment of set theory. Let U be an M-measure over $\kappa \in Ord$ such that Ult[M, U] is well-founded.

- Write cUlt[M, U] for the transitive class, and let π_U : Ult $[M, U] \rightarrow cUlt[M, U]$ be the collapsing isomorphism.
- Write $\tilde{j}_U : M \to \text{Ult}[M, U]$ for the canonical embedding defined by $\tilde{j}(x) = [\text{const}_x]_U$.
- Write $j_U : \mathbf{M} \to \mathrm{cUlt}[\mathbf{M}, U]$ for $\pi_U \circ j_U$.
- If $j : M \to N$ is elementary and $U = U_j$, then write $\tilde{k}_U : \text{Ult}[M, U] \to N$ for the map defined by $\tilde{k}_U([f]_U) = j(f)(\kappa)$. Write $k_U : \text{cUlt}[M, U] \to N$ for $\tilde{k}_U \circ \pi_U^{-1}$.

— 14A•9. Lemma

Let $M \models AC$ be a transitive model of some (suitable) fragment of set theory. Let U be a weakly amenable M-measure over $\kappa \in Ord$ such that Ult[M, U]. Therefore,

- 1. $\operatorname{cp}(j) = \kappa$.
- 2. $V_{\kappa}^{\mathsf{M}} = V_{\kappa}^{\mathsf{cUlt}[\mathsf{M},U]}$ and $\mathcal{P}(\kappa)^{\mathsf{M}} = \mathcal{P}(\kappa)^{\mathsf{cUlt}[\mathsf{M},U]}$.
- 3. κ is strongly inaccessible in **M**.
- 4. $U \notin \text{cUlt}[M, U]$.
- 5. If M is a set, $|\mathbf{M}| = |\operatorname{cUlt}[\mathbf{M}, U]|$.

Proof .:.

This follows just as with Theorem 12 B • 5. In particular, suppose inductively that j \ ξ = id \ ξ, aiming to show j(ξ) = ξ for ξ < κ. Thus α ∈ ξ implies α = j(α) ∈ j(ξ), meaning ξ ≤ j(ξ). For the other direction, let ζ < j(ξ) be arbitrary. ζ = π([f]) for some f : κ → M with f ∈ M, and j(ξ) = π([const_ξ]), meaning by Łoś's Theorem (14 A • 5) that

 $\mathsf{cUlt}[\mathsf{M}, U] \models "\pi([f]) < \pi([\mathsf{const}_{\xi}])"$ iff $\mathsf{Ult}[\mathsf{M}, U] \models "[f] < [\mathsf{const}_{\xi}]"$ iff $\forall^* \alpha (f(\alpha) < \xi)$. It follows that f is constant on a large set by κ -completeness since $\xi < \kappa$. Hence $\forall^* \alpha (f(\alpha) = \varepsilon)$ for some $\varepsilon < \xi$, and therefore $[f] = [\mathsf{const}_{\varepsilon}]$. Thus $\zeta = \pi([\mathsf{const}_{\varepsilon}]) = j(\varepsilon) = \varepsilon < \xi$ by the inductive hypothesis. So $j(\xi) < \xi$.

To see that $\kappa \neq j(\kappa)$ —and thus that $cp(j) = \kappa$ —consider the identity function. Note that $\forall \xi < \kappa \ \forall^* \alpha \ (\alpha > \xi)$, i.e. for every $\xi < \kappa, \ \forall^* \alpha \ (id(\alpha) > const_{\xi}(\alpha))$. Hence $\pi([id]) > \pi([const_{\xi}]) = j(\xi) = \xi$ for every $\xi < \kappa$. This means $\kappa \leq \pi([id])$. On the other hand, $\forall^* \alpha \ (id(\alpha) < const_{\kappa}(\alpha))$, meaning $\pi([id]) < \pi([const_{\kappa}]) = j(\kappa)$. Thus $\kappa \leq \pi([id]) < j(\kappa)$ so that $cp(j) = \kappa$.

This is less obvious than one might think at first glace. Generally speaking, cp(j) = κ doesn't imply M and cUlt[M, U] agree up to κ unless the ultrapower is contained in the original model. Instead, we only get V^M_κ ⊆ V^{cUlt[M,U]} because j "V^M_κ = V^M_κ ⊆ cUlt[M, U]. For the other direction, assume inductively that M and cUlt[M, U] agree on V_α for α < κ. To see that the two agree on V_{α+1}, let x ∈ V^{cUlt[M,U]}_{α+1} be arbitrary. We can write x = π([f]) for some f : κ → M with f ∈ M. Since V^M_α = V^{CUlt[M,U]}_α, x ⊆ M. Because cp(j) = κ > α, j"x = x, and so by Łoś's Theorem (14 A • 5),

 $\pi([\text{const}_y]) = j(y) = y \in x = \pi([f])$ iff $\mathsf{Ult}[\mathsf{M}, U] \models "[\text{const}_y] \in [f]"$ iff $\forall^* \xi \ (y \in f(\xi))$. Hence we can define x by

$$\pi([f]) = \{ y \in V_{\alpha}^{\mathsf{M}} : \{ \xi < \kappa : y \in f(\xi) \} \in U \}.$$

This seems like it's in M but to check this, we need to use weak amenability, which means calculating cardinality.

– Claim 1

We have $x \in M$ and therefore, as x was arbitrary, $V_{\alpha+1}^{\mathsf{cult}[\mathsf{M},U]} \subseteq V_{\alpha+1}^{\mathsf{M}}$.

Proof .:.

Ostensibly, we'd set $F : V_{\alpha}^{\mathsf{M}} \to \mathcal{P}(\kappa)$ by $F(y) = \{\xi < \kappa : y \in f(\xi)\}$ and then note $x = F^{-1} "U$, motivated Lemma 14 A • 2 to show $F^{-1} "U \in \mathsf{M}$ if U is weakly amenable to M . But showing this requires V_{α}^{M} to have size $\leq \kappa$ in M . We can restrict ourselves to only considering $y \in \bigcup \operatorname{im}(f)$ since $y \in x$ iff $\forall^* \xi \ (y \in f(\xi))$. So now it suffices to show $\operatorname{im}(f)$ has size $\leq \kappa$ in M , which is less of an issue. To see that $|\bigcup \operatorname{im}(f)|^{\mathsf{M}} < \kappa$, suppose otherwise and by AC get an injection $g : \kappa \to \bigcup \operatorname{im}(f)$. It follows that $g(\xi)$ has rank $\leq \alpha < \kappa$ for every $\xi < \kappa$, meaning $\pi([g]) \in V_{\alpha}^{\mathsf{cult}[\mathsf{M},U]} = V_{\alpha}^{\mathsf{M}}$. As $\operatorname{cp}(j) = \kappa > \alpha, \pi([g]) = z = j(z) = \pi([\operatorname{const}_z])$ for some $z \in V_{\alpha}^{\mathsf{M}}$. This means $\forall^* \xi \ (g(\xi) = z)$, which contradicts that g is injective. Hence $|\bigcup \operatorname{im}(f)| < \kappa$ and so taking $F(y) = \{\xi < \kappa : y \in$ $f(\xi)\}$ yields $F : \bigcup \operatorname{im}(f) \to \mathfrak{P}(\kappa)$ with $x = \pi([f]) = F^{-1} "U \in \mathsf{M}$ by weak amenability. \dashv

Hence $V_{\alpha+1}^{\text{cUlt}[M,U]} = V_{\alpha+1}^{M}$. This completes the successor stage of the induction, and as limit stages are just unions, this implies by induction that $V_{\kappa}^{M} = V_{\kappa}^{\text{cUlt}[M,U]}$.

We also get that **M** and **cUlt**[**M**, U] agree on $\mathcal{P}(\kappa)$. One direction is as before: $j(x) \cap \kappa = x$ for $x \in \mathcal{P}(\kappa)^{\mathsf{M}}$ since every $\alpha < \kappa$ has $\alpha \in x$ iff $j(\alpha) = \alpha \in j(x)$. This implies $\mathcal{P}(\kappa)^{\mathsf{M}} \subseteq \mathcal{P}(\kappa)^{\mathsf{cUlt}[\mathsf{M},U]}$. So suppose $x \in \mathcal{P}(\kappa)^{\mathsf{cUlt}[\mathsf{M},U]}$ with $x = \pi([f])$ for some $f \in {}^{\kappa}\mathsf{M} \cap \mathsf{M}$. As previously,

 $x = \{ \alpha \in \kappa : \{ \xi < \kappa : \alpha < f(\xi) \} \in U \}.$

By weak amenability, the version from Lemma 14 A • 2, this implies $x \in M$: let $F : \kappa \to \mathcal{O}(\kappa)$ be defined by $F(\alpha) = \{\xi < \kappa : \alpha < f(\xi)\}$ so that $x = F^{-1}$ "U.

3. The fact that κ is regular follows by κ-completeness of U: if ⟨γ_α : α < λ⟩ ∈ M is cofinal in κ with λ < κ, consider the statement ∀α < λ ∀*ξ (γ_α < ξ) which implies by κ-completeness ∀*ξ ∀α < λ (γ_α < ξ) and in particular, ∃ξ < κ ∀α < λ (γ_α < ξ), which contradicts that the γ_αs are cofinal. The fact that κ is uncountable in M is proven as follows. Note j(ω) = ω—so that ω < cp(j) = κ—and if x ⊆ ω codes a well-order of length κ, then j(x) codes a well-order of length j(κ). But j(x) = x because ω < κ implies ω + 1 < κ and so j ↾ V^M_{ω+1} = id ↾ V^M_{ω+1}. This is a contradiction because x cannot code a well-order of length both κ and j(κ) ≠ κ.

The fact that κ is strongly inaccessible follows in just the same way as Lemma $12 \,\mathbb{C} \cdot 2$, but with some careful checking to ensure we're still working in M. Suppose $\lambda < \kappa$ has $2^{\lambda} \ge \kappa$. Let $\Lambda \subseteq \mathcal{O}(\lambda)^{\mathsf{M}}$ have size κ in **M** as witnessed by a bijection $g : \Lambda \to \kappa$. Consider $U_g = \{x \in \mathcal{O}(\Lambda)^{\mathsf{M}} : g^{"}x \in U\}$ which need not be in M. Now we can specify any element of Λ with just λ -many pieces of information, and we can use the κ -completeness of U to bring all of these pieces together into a single $x \in \Lambda$.

Consider the function $f : \lambda \to \mathcal{P}(\kappa)$ that decides whether $\alpha \in x$: define $f(\alpha) = g''\{y \in \Lambda : \alpha \in y\}$. By weak amenability, $f^{-1}''U \in M$, and this will be our x. Now we can consider $X_{\alpha} \subseteq \Lambda$ incorporating the

EXTERNAL ULTRAPOWERS

information about whether $\alpha \in x$ or $\alpha \notin x$: define for $\alpha < \lambda$,

$$X_{\alpha} = \begin{cases} \{y \in \Lambda : \alpha \in y\} & \text{if } \alpha \in f^{-1}"U \text{ i.e. if } \{y \in \Lambda : \alpha \in y\} \in U_g \\ \{y \in \Lambda : \alpha \notin y\} & \text{if } \alpha \notin f^{-1}"U. \end{cases}$$

As a result, $X_{\alpha} \in U_g$ for every $\alpha < \lambda$ because if $\alpha \in f^{-1}"U$, $f(\alpha) = g"X_{\alpha} \in U$ so $X_{\alpha} \in U_g$, and otherwise if $\alpha \notin f^{-1}"U$ then $f(\alpha) = g"\{y \in \Lambda : \alpha \in y\} \notin U$ so that as an M-ultrafilter, $\kappa \setminus g"\{y \in \Lambda : \alpha \in y\} \notin U$. As a bijection, this is equal to $g"\{y \in \Lambda : \alpha \notin y\} = g"X_{\alpha}$ so that $X_{\alpha} \in U_g$. Note that $\langle X_{\alpha} : \alpha < \lambda \rangle \in M$. So by κ -completeness, $\bigcap_{\alpha < \lambda} g"X_{\alpha} \in U$. But as a bijection, $\bigcap_{\alpha < \lambda} g"X_{\alpha} = g"\bigcap_{\alpha < \lambda} X_{\alpha}$ and $\bigcap_{\alpha < \lambda} X_{\alpha}$ is at most a singleton: it is either empty or $\{f^{-1}"U\}$, which means either $g"\emptyset = \emptyset$ or the singleton $\{g(f^{-1}"U)\}$ is in U, contradicting that U is non-principal.

- 4. If U ∈ cUlt[M, U], then consider in cUlt[M, U] the relation ≺ on (κ^κ)^{cUlt[M,U]} = (κ^κ)^M—which are equal by 2—defined by f ≺ g iff {α < κ : f(α) < g(α)} ∈ U. We know this relation is a well-order, because it is precisely the ordering ∈^{Ult[M,U]} on the predecessors of [const_κ] in Ult[M, U] whose transitive collapse is the membership relation on the ordinal π([const_κ]) = j(κ). Hence in cUlt[M, U], there is a surjection from 2^κ onto j(κ), which contradicts by elementarity that j(κ) is strongly inaccessible in cUlt[M, U].
- 5. Clearly every element of cUlt[M, U] can be identified with an element of ${}^{\kappa}M \cap M \subseteq M$. So $|cUlt[M, U]| \le |M|$. On the other hand, $j : M \to cUlt[M, U]$ is an injection, meaning $|M| \le |cUlt[M, U]|$.

(2) and (4) together tell us that **M** and **cUlt**[**M**, U] agree on $V_{\kappa+1}$ as well since $|V_{\kappa}|^{\mathsf{M}} = \kappa$ and so $V_{\kappa+1}^{\mathsf{M}}$ can be identified with $\mathcal{P}(\kappa)^{\mathsf{M}}$. This also tells us κ is the κ th inaccessible, mahlo, the κ th mahlo, and so forth for the same reason as in Corollary 12 D • 3.

Now we come to a point where the theory of internal and external ultrapowers diverge. Here is a list of several things that can go differently:

- 1. cUlt[M, U] need not be contained in M.
- 2. Ord \cap M can be strictly smaller than Ord \cap cUlt[M, U], meaning the height of M is strictly less than the ultrapower cUlt[M, U].
- 3. κ need not be measurable in M, just a property called *weak compactness*.
- 4. Ult[M, U] need not be well-founded.

The last point is the most important for us. The issue is that even if U is κ -complete for sequences in M, V may still contain sequences that witness that the ultrapower is ill-founded. This is made formal with the following, which has the same proof as with ultrapowers by $U \in M$, Theorem 12 B • 3. Note that we do not get an equivalence as with Theorem 12 B • 3.

- 14A•10. Result

Let **M** be a transitive model of some (suitable) fragment of set theory. Let U be an M-measure. Therefore Ult[M, U] is well-founded if $\bigcap_{n < \omega} X_n \neq \emptyset$ for any family $\{X_n : n < \omega\} \in \mathcal{P}(U)^{\mathsf{V}}$ (this weakens \aleph_1 -completeness which requires $\bigcap_{n < \omega} X_n \in U$)

Proof .:.

Suppose U satisfies the hypothesis but Ult[M, U] is ill-founded. Let $\langle f_n : n \in \omega \rangle \in M$ be a descending $\in^{\text{Ult}[M,U]}$ -sequence in Ult[M, U]: for every $n \in \omega$, Ult[M, U] $\models "[f_{n+1}] \in [f_n]$ ". It follows by Łoś's Theorem (14 A • 5),

$$\bigwedge_{n \in \omega} \forall^* \alpha \ (\mathbf{M} \vDash "f_{n+1}(\alpha) \in f_n(\alpha)")$$

By the hypothesis, the intersection of these large sets is non-empty in V:

$$\exists \alpha \, \bigwedge_{n \in \omega} \mathbf{M} \vDash f_{n+1}(\alpha) \in f_n(\alpha)^n.$$

But any such α yields an infinite, decreasing sequence $\langle f_n(\alpha) : n \in \omega \rangle$ in M and V, a contradiction.

This leads to the idea of *iterability*. Even if the external ultrapower is well-founded, the iterated ultrapowers as in

 \neg

 \neg

Subsection 12 E may not be. So we'd like to know when a model is "iterable" in the sense that its external ultrapowers are well-founded. This is actually closed tied with weak amenability as ultrapowers of weakly amenable measures over κ preserve $\mathcal{P}(\kappa)$ by Lemma 14 A • 9 (2) which is a kind of partial converse of Result 14 A • 3. The weak amenability portion is necessary to get another measure over the ultrapower and continue the iteration as per Corollary 14 A • 6.

– 14A•11. Lemma

Let $\mathbf{M} \models AC$ be a transitive model of some (suitable) fragment of set theory. Let U be an M-measure such that $Ult[\mathbf{M}, U]$ is well-founded. Therefore the derived measure $U_j = U$ where $j : \mathbf{M} \rightarrow cUlt[\mathbf{M}, U]$ is the canonical embedding.

Proof .:.

By Corollary 14 A • 7, there is a unique factor embedding \tilde{k} : Ult[M, U] \rightarrow cUlt[M, U] with $\tilde{k}([f]) = j(f)(\kappa)$. Since the transitive collapse is unique, $\tilde{k} = \pi_U$ so that $\pi_U([f]) = j(f)(\kappa)$. As a result, for $X \subseteq \kappa$,

 $\begin{aligned} X \in U_j & \text{iff} \quad \kappa \in j(X) \\ & \text{iff} \quad j(\text{id})(\kappa) = k([\text{id}]) = \pi_U([\text{id}]) \in \pi_U([\text{const}_X]) \\ & \text{iff} \quad \mathsf{Ult}[\mathsf{M}, U] \vDash ``[\text{id}] \in [\text{const}_X]'' \\ & \text{iff} \quad \forall^* \alpha \ (\text{id}(\alpha) = \alpha \in X = \text{const}_X(\alpha)), \quad \text{by Loś's Theorem (14 A • 5),} \\ & \text{iff} \quad X \in U. \end{aligned}$

– 14A•12. Corollary -

Let $\mathbf{M} \models \mathbf{AC}$ be a transitive model of some (suitable) fragment of set theory. Let U be an M-measure such that $Ult[\mathbf{M}, U]$ is well-founded. Therefore U is weakly amenable to \mathbf{M} iff $\mathcal{O}(\kappa)^{\mathbf{M}} = \mathcal{O}(\kappa)^{cUlt[\mathbf{M}, U]}$.

Proof .:.

Note that Lemma 14 A • 9 (2) gives one direction: if U is weakly amenable to **M**, then the two models agree on $\mathcal{O}(\kappa)$. For the other direction, assume $\mathcal{O}(\kappa)^{\mathsf{M}} = \mathcal{O}(\kappa)^{\mathsf{cUlt}[\mathsf{M},U]}$. By Result 14 A • 3, the derived measure U_{j_U} is weakly amenable to **M**. By Lemma 14 A • 11, $U_{j_U} = U$.

This is sometimes incorrectly interpreted as saying that weak amenability is equivalent to preserving the powerset (in the sense that if the measure derived from $j : \mathbf{M} \to \mathbf{N}$ is weakly amenable then **M** and **N** agree on $\mathcal{O}(\kappa)$), but this isn't true generally.[?gitmanwelchiterablecards] Instead, we only get the equivalence for **N** as the ultrapower, and more broadly speaking, preserving the powerset of κ is stronger than weak amenability as shown with Result 14 A • 3.

§14 B. Linear iterations of external measures

We proceed in a similar way to Subsection 12 E. The downside to using external measures, however, is that the direct limit need not be well-founded anymore unless we have something like the \aleph_1 -completeness from Result 14 A • 10. In such a case, we can present a similar proof of well-foundedness for limit ultrapowers as in The Wellfoundedness of Iterated Ultrapowers (12 E • 4). But we will give an alternative route that we could have also used there.

14B•1. Definition Let $\lambda \in \text{Ord.}$ Let $\mathcal{M} = \langle \mathsf{M}_{\alpha} : \alpha < \lambda \rangle$ and $\mathcal{U} = \langle U_{\alpha} : \alpha < \lambda \rangle$. We call $\langle \mathcal{M}, \mathcal{U} \rangle$ a *linear iteration* iff for all $\alpha < \lambda$,

- 1. $\mathbf{M}_{\alpha}^{\mathcal{U}}$ is a transitive model of some (suitable) fragment of set theory.
- 2. U_{α} is an $M^{\mathcal{U}}_{\alpha}$ -measure such that $\mathsf{Ult}[\mathsf{M}^{\mathcal{U}}_{\alpha}, U_{\alpha}]$ is well-founded.
- 3. $\mathbf{M}_{\alpha+1}^{\mathcal{U}} = \mathbf{cUlt}[\mathbf{M}_{\alpha}^{\mathcal{U}}, U_{\alpha}].$
- 4. For limit α , $\mathbf{M}^{\mathcal{U}}_{\alpha}$ is the direct limit of $\mathbf{M}^{\mathcal{U}}_{\beta}$ for $\beta < \alpha$, where the embeddings $j_{\beta,\gamma} : \mathbf{M}^{\mathcal{U}}_{\beta} \to \mathbf{M}^{\mathcal{U}}_{\gamma}$ are given by composition: $j_{\beta,\beta+1}$ is the canonical embedding, $j_{\beta,\gamma} = j_{\delta,\gamma} \circ j_{\beta,\delta}$ for any δ with $\beta \le \delta \le \gamma < \alpha$.

In this case, we can define $\mathsf{M}^{\mathcal{U}}_{\infty}$ as the direct limit of \mathcal{M} using the embeddings above. We say \mathcal{U} is *weakly amenable* to \mathcal{M} iff each U_{α} is weakly amenable to $\mathsf{M}^{\mathcal{U}}_{\alpha}$. So this generalizes Definition 12 E • 2 in two major respects: the measures U_{α} need not be in $M_{\alpha}^{\mathcal{U}}$, and U_{α} need not be $j_{0,\alpha}(U_{\alpha})$. Later on, we'll generalize this further into iteration trees where $M_{\alpha+1}^{\mathcal{U}}$ need not be $\text{cUlt}(M_{\alpha}^{\mathcal{U}}, U_{\alpha})$ but instead $\text{cUlt}(M_{\alpha}^{\mathcal{U}}, U_{\alpha}^*)$ for some $\alpha^* \leq \alpha$.^{xviii} For now, though, we indeed only care about when U_{α} does equal $j_{0,\alpha}(U_{\alpha})$, because it helps us understand what properties of U we need to be able to iterate it α steps.

As before, it's not clear when $\mathbf{M}_{\infty}^{\mathcal{U}}$ is well-founded. For example, U could be as nice as we'd like, but if take $\mathcal{U} = \langle U : \alpha < \omega \rangle$, then $\mathbf{M}_{\infty}^{\mathcal{U}}$ is always illfounded. Indeed, just having the same critical point implies illfoundedness: if $\mathcal{U} = \langle U_n : n < \omega \rangle$ with $\operatorname{cp}(j_{n,n+1}) = \kappa$ for all n, then $\mathbf{M}_{\infty}^{\mathcal{U}}$ is illfounded. This can be seen with Figure 14 B • 2.



14B.2. Figure: An Illfounded Iteration

What matters for us is how long iterations can go on. It's consistent that $M_n^{\mathcal{U}}$ exists for $n < \omega$, but the direct limit $M_{\infty}^{\mathcal{U}}$ is ill-founded. Indeed, one sufficient condition to being fully iterable—in that the α th linear iterate exists for all $\alpha \in \text{Ord}$ —is a form of countable completeness from Result 14 A • 10.

A set U is
$$\sigma$$
-complete iff for any $\{X_n : n < \omega\} \subseteq U, \bigcap_{n < \omega} X_n \neq \emptyset$

Again, Result $14 \text{ A} \cdot 10$ tells us that being σ -complete implies the ultrapower is well-founded. Note that if U is an ultrafilter in V, this idea doesn't make any difference from \aleph_1 -completeness before, but it can if U is an M-measure and M does not contain all such countable sequences in ωU .

- 14B•4. Result Let U be an ultrafilter (in V). Then U is σ-complete iff U is \aleph_1 -complete.

Proof .:.

Suppose U is σ -complete. If $\bigcap_{n < \omega} X_n = Y \notin U$ then as an ultrafilter, $Z_n = X_n \setminus Y \in U$ and so $\bigcap_{n < \omega} Z_n = \emptyset$ which contradicts σ -completeness. This shows U is also \aleph_1 -complete. The other direction is trivial. \dashv

Let us give an overview of why the iterates of a model M by a σ -complete, weakly amenable M-measure U are well-founded.



14B•5. Figure: Realizing Iterates Back into M

^{xviii}So the production of ultrapowers is still "linear" in these iteration trees in that we proceed one-by-one, but the resulting graph of embeddings is not a line but instead a tree as there is ostensibly no embedding from M_{α} to $M_{\alpha+1}$, just from $M_{\alpha*}$ to $M_{\alpha+1}$.

- 1. If **M** has an iteration of length α , then so do its countable hulls.
- 2. Moreover, if **M** doesn't have an iteration of length α , then a countable hull has an ill-founded iteration.
- 3. Ultrapowers of hulls can be realized back into M.
- 4. So if M isn't fully iterable, then some countable hull N isn't fully iterable, but nevertheless, every iterate can be realized back into M, meaning the direct limit can be realized back into M. This shows the direct limit should be well-founded because M is transitive, and we get a contradiction. This is shown in Figure 14 B 5

The ideas here require weak amenability mostly just for us to be able to define the iterated ultrapower by the same measure (translated by the embedding of course). The property of σ -completeness comes into play in realizing hulls back into M. All of this is quite vague for now, but let us get started in proving these.

We have another characterization of σ -completeness, which is useful for us in the context of taking hulls. First let's show that hulls still have measures over them when the original models do too. This marks a shift in how we should think about the ultrapowers of a model being well-founded according to whether hulls are well-founded, and eventually this changes our perspective from how sets are measured to instead what the resulting embeddings look like.

– 14B•6. Lemma

- Let M be a transitive model of a (sufficient) fragment of set theory.
- Let U be an M-measure on some κ .
- Let $\text{Hull} \preccurlyeq_{\in} M$ (not necessarily a skolem hull, though this is what you should have in mind, also recall the notation of \preccurlyeq_{σ} from Definition 6 A 1) with $\kappa \in \text{Hull}$ (e.g. if $\text{Hull} \preccurlyeq_{\in, \dot{U}} M$).
- Let π : Hull \rightarrow N be the collapsing isomorphism.

Therefore π "*U* is an N-measure over $\pi(\kappa)$.

Proof .:.

Suppose $\pi^{"}U$ is not an N-measure over $\pi(\kappa)$. Clearly $\pi^{"}U$ is still a filter for subsets in $\mathcal{O}(\kappa) \cap \mathbb{N}$: if $x, y \in \pi^{"}U$ then $\pi^{-1}(x) \cap \pi^{-1}(y) \in U$. As an isomorphism, $\pi^{-1}(x) \cap \pi^{-1}(y) = \pi^{-1}((x \cap y)^{\mathsf{Hull}}) \in \mathsf{Hull}$. Since $\mathsf{Hull} \preccurlyeq \mathsf{M}$, $(x \cap y)^{\mathsf{Hull}} = x \cap y$ and so $x \cap y \in \pi^{"}U$. We similarly get upward closure under \subseteq .

 π "*U* is an N-ultrafilter because if $x \in \mathcal{O}(\pi(\kappa)) \setminus \pi$ "*U*, then $x = \pi(y)$ where by elementarity, Hull \models " $y \subseteq \kappa$ ". Since Hull \preccurlyeq M and M is transitive, y is truly a subset of κ and $y \notin U$. Thus $\kappa \setminus y \in U$. By the same sort of elementarity argument from before, $\pi(\kappa \setminus y) = \pi(\kappa) \setminus x \in \pi$ "*U*. Thus π "*U* is an N-ultrafilter.

N- $\pi(\kappa)$ -completeness and N-normality are straightforward too. To see that $\pi^{"}U$ is N- $\pi(\kappa)$ -complete, let f: $\gamma \to \pi^{"}U$ be in N with $\gamma < \pi(\kappa)$. Note by elementarity, $\pi^{-1}(f) : \pi^{-1}(\gamma) \to U$ with $\pi^{-1}(\gamma) < \kappa$. It follows that $\bigcap_{\alpha < \gamma} \pi^{-1}(f)(\alpha) \in U$. By elementarity, this is also $\pi^{-1}(\bigcap_{\alpha < \gamma} f(\alpha))$ and it follows that $\bigcap_{\alpha < \gamma} f(\alpha) \in \pi^{"}U$. To see that $\pi^{"}U$ is N-normal, let $f : \pi(\kappa) \to \pi(\kappa)$ be regressive on a set in $\pi^{"}U$. It follows that $\pi^{-1}(f)$ is regressive and hence constant on a set in U. It follows that f is therefore constant on a set in $\pi^{"}U$.

Weak amenability is a more difficult condition as weak amenability to N requires weak amenability to Hull, and ostensibly we might only have weak amenability to M. Nevertheless, countable hulls are also important more generally for understanding when M is iterable, and it tells us that countable iterations suffice to show full iterability.

— 14B•7. Lemma

Let **M** be a transitive model of a (sufficient) fragment of set theory. Let U be an M-measure on κ . Therefore, the following are equivalent:

1. U is σ -complete

2. For any countable Hull $\preccurlyeq_{\in} M$ with $\kappa \in$ Hull (or for any Hull $\preccurlyeq_{\in, \dot{U}} M$) and with collapse map π : Hull $\rightarrow N$, there is an elementary embedding k: Ult $[N, \pi^{"}U] \rightarrow M$ such that $\pi^{-1} = k \circ j_{\pi^{"}U} : N \rightarrow M$ as below:



Proof .:.

• $(1 \rightarrow 2)$ Suppose that U is σ -complete. The factor map is defined outside of N by noting that since N is countable, we can intersect all elements in $\pi^{"}U$: consider Hull $\cap U$ which is countable and therefore $\bigcap_{x \in \text{Hull } \cap U} x \neq \emptyset$ by σ -completeness. If α is in this set, we can define k : Ult $[N, \pi^{"}U] \rightarrow M$ by $k([f]) = \pi^{-1}(f)(\alpha)$. This is well-defined by the choice of α :

$$\begin{aligned} [f] &= [g] \quad \leftrightarrow \quad \forall_{\pi^{*}U}^{*} \xi \ (f(\xi) = g(\xi)) \\ &\leftrightarrow \quad \forall_{\text{Hull} \cap U}^{*} \xi \ (\pi^{-1}(f)(\xi) = \pi^{-1}(g)(\xi)) \\ &\to \quad k([f]) = \pi^{-1}(f)(\alpha) = \pi^{-1}(g)(\alpha) = k([g]) \end{aligned}$$

The fact that k is elementary follows in the same way as with other ultrapower factor maps, but now using Hull. Note by Lemma 14B•6 that $\pi^{"}U$ is indeed an N-measure. So if $\text{Ult}[N, \pi^{"}U] \models "\varphi([f])$ " then by Loś's Theorem (14A•5), $N \models "\varphi(f(\xi))$ " for $\pi^{"}U$ -almost every ξ and hence Hull $\models "\varphi(\pi^{-1}(f)(\xi))$ " for $U \cap$ Hull-almost every ξ . In particular, this means it holds for $\xi = \alpha$ and since Hull $\preccurlyeq_{\epsilon} M$, we get $M \models "\varphi(\pi^{-1}(f)(\alpha))$ ", i.e. $M \models "\varphi(k([f]))$ ". So

Ult[N, π "U] \models " $\varphi([f])$ " implies M \models " $\varphi(k([f]))$ ".

Note that this suffices for elementarity since if $\mathbf{M} \models "\varphi(k([f]))"$ but $\mathsf{Ult}[\mathbf{N}, \pi"U] \not\models "\varphi([f])"$ then it models the negation " $\neg \varphi([f])$ " and so the above implication gives $\mathbf{M} \models "\neg \varphi(k([f]))"$, a contradiction.

• $(2 \to 1)$ Let $\{X_n : n < \omega\} \subseteq U$ be arbitrary. Consider Hull = Hull^M($\{\kappa, X_n : n < \omega\}$) $\preccurlyeq_{\{\epsilon,\kappa,X_n:n<\omega\}}$ M so that Hull is countable, and $\{\kappa, X_n : n < \omega\} \subseteq$ Hull (as per Taking a Skolem Hull (6A • 2)). Let π : Hull \rightarrow cHull be the collapsing map so that by Lemma 14B • 6, π "U is a cHull-measure over $\pi(\kappa)$. We can then define Ult[cHull, π "U]. By (2), there is an elementary k : Ult[cHull, π "U] \rightarrow M such that $\pi^{-1} = k \circ j_{\pi"U}$: cHull \rightarrow M.

Now with all of that set up, we get onto the meat of the proof. Let $f : \pi(\kappa) \to \text{cHull}$ be arbitrary. Thus $\forall_{\pi^*U}^* \alpha \ (f(\alpha) = f(\alpha))$. By Łoś's Theorem (14 A • 5), we can view this as equivalent to

 $\mathsf{Ult}[\mathsf{cHull}, \pi"U] \vDash "[f] = [\mathsf{const}_f]([\mathsf{id}])".$

After applying k we get by elementarity that k([f]) is the function $k([\text{const}_f])$ evaluated at the point k([id]). Note that $k([\text{const}_f]) = k(j_{\pi^*U}(f)) = \pi^{-1}(f)$ and since [id] is an ordinal of **Ult[cHull**, π^*U], $k([\text{id}]) = \alpha$ is some particular ordinal. Hence we can generally say $k([f]) = \pi^{-1}(f)(\alpha)$ for some fixed α .

Now consider in **M** the characteristic function f_n for X_n : $f_n(\xi) = 1$ if $\xi \in X_n$ and $f_n(\xi) = 0$ otherwise. Thus $f_n \in$ cHull for each $n < \omega$ and $g_n = \pi(f_n)$ is the characteristic function for $\pi(X_n)$. Note that $[g_n] = [\text{const}_1]$ for every n because $g_n(\xi) = 1$ for all $\xi \in \pi(X_n) \in \pi^{"}U$. Thus $k([g_n]) = k([\text{const}_1]) = \pi^{-1}(\text{const}_1)(\alpha) = 1$. But since $\bigcap_{n < \omega} X_n = \emptyset$, there is some n such that $\alpha \notin X_n$ in which case $k([g_n]) = \pi^{-1}(g_n)(\alpha) = f_n(\alpha) = 0$, a contradiction.

This idea is at the heart of showing that the limit iterations of **M** by a σ -complete M-measure are well-founded. If instead of the collapsing map π : Hull \rightarrow N as fundamental, we regard the inverse *un*collapsing map τ : N \rightarrow Hull as fundamental, the factor map k : Ult[N, π "U] \rightarrow Hull is sometimes called the τ -realization of Ult[N, π "U], because properties of the ultrapower can be realized back into properties of Hull (and of M) through k. And crucially, this idea can be iterated as long as the critical points of the measures are in the hulls.

- 14B•8. Lemma

- Let ${\bf M}$ be a transitive model of a (sufficient) fragment of set theory.
- Let $\langle \mathcal{M}, \mathcal{U} \rangle$ be a linear iteration with first model $\mathsf{M}_0^{\mathcal{U}} = \mathsf{M}$ and length λ , where $\mathcal{U} = \langle U_\alpha \subseteq \mathcal{P}(\kappa_\alpha) : \alpha < \lambda \rangle$.

Therefore,

- 1. There is a linear iteration $\langle \mathcal{N}, \mathcal{W} \rangle$ with first model $N_0^{\mathcal{W}} = N$ and length λ' where $1 \le \lambda' \le \lambda$.
- 2. Moreover, there is an FOL(\in)-elementary (or FOL(\in , \dot{U})-elementary if τ is) τ_{α} : $N_{\alpha}^{W} \to M_{\alpha}^{U}$ for all $\alpha \leq \lambda'$ such that $W = \langle W_{\alpha} = \tau_{\alpha}^{-1} | | U_{\alpha} : \alpha < \lambda' \rangle$, and these commute with the iterations of $\langle \mathcal{M}, \mathcal{U} \rangle$ and $\langle \mathcal{N}, W \rangle$: if $\alpha \leq \beta < \lambda'$, then $j_{\alpha,\beta}^{\mathcal{U}} \circ \tau_{\alpha} = \tau_{\beta} \circ j_{\alpha,\beta}^{\mathcal{W}}$ as pictured:

$$\mathbf{M} = \mathbf{M}_{0}^{\mathcal{U}} \xrightarrow{j_{0,1}^{\mathcal{U}}} \mathbf{M}_{1}^{\mathcal{U}} \xrightarrow{j_{1,2}^{\mathcal{U}}} \mathbf{M}_{2}^{\mathcal{U}} \xrightarrow{j_{2,3}^{\mathcal{U}}} \cdots \longrightarrow \mathbf{M}_{\lambda'}^{\mathcal{U}} \xrightarrow{j_{\lambda',\lambda'+1}^{\mathcal{U}}} \cdots \longrightarrow M^{\mathcal{U}}$$

$$\xrightarrow{\tau_{0}} \overbrace{\bigcirc} \tau_{1} \uparrow \qquad \tau_{2} \uparrow \qquad \uparrow \qquad \tau_{\lambda'} \uparrow$$

$$\mathbf{N} = \mathbf{N}_{0}^{\mathcal{W}} \xrightarrow{j_{0,1}^{\mathcal{W}}} \mathbf{N}_{1}^{\mathcal{W}} \xrightarrow{j_{1,2}^{\mathcal{W}}} \mathbf{N}_{2}^{\mathcal{W}} \xrightarrow{j_{2,3}^{\mathcal{W}}} \cdots \longrightarrow \mathbf{N}_{\infty}^{\mathcal{W}}$$

3. And if $\kappa_{\alpha} \in \operatorname{im}(\tau_{\alpha})$ at each stage (e.g. if $\kappa_{\alpha} = j_{0,\alpha}^{\mathcal{U}}(\kappa_0)$ for each $\alpha < \lambda$), then we can ensure $\lambda' = \lambda$.

Proof .:.

Regard $\tau_0 = \tau$ and $N_0 = N$. Now proceed by induction on $\alpha < \lambda$, stopping and declaring $\alpha = \lambda'$ if $\kappa_\alpha \notin \operatorname{im}(\tau_\alpha)$. For $\alpha = 0$, the result is just the hypothesis. For limit α , the claim follows using the direct limit embeddings. For the successor stage, suppose the elementary $\tau_\alpha : N_\alpha^{\mathcal{W} \upharpoonright \alpha} \to M_\alpha^{\mathcal{U}}$ exists where $\mathcal{W} \upharpoonright \alpha = \langle W_{\xi} : \xi < \alpha \rangle$ as been defined so far. nductively set $W_\alpha = \tau_\alpha^{-1} U_\alpha$. Regard Hull $= \tau_\alpha'' N_\alpha^{\mathcal{W} \upharpoonright \alpha}$ so that Hull $\leq M_\alpha^{\mathcal{U}}$. Collapsing Hull yields, by uniqueness of the transitive collapse, the original model $N_\alpha^{\mathcal{W} \upharpoonright \alpha}$ so that by Lemma 14 B • 6, $\tau_\alpha^{-1} U_\alpha$ is an $N_\alpha^{\mathcal{W} \upharpoonright \alpha}$ -measure over $\tau^{-1}(\kappa_\alpha)$.

So it suffices to show there is an elementary

 $\tilde{\tau}_{\alpha+1} : \mathrm{Ult}[\mathrm{N}^{\mathcal{W}\restriction\alpha}_{\alpha}, W_{\alpha}] \to \mathrm{Ult}[\mathrm{M}^{\mathcal{U}}_{\alpha}, U_{\alpha}].$

We define $\tilde{\tau}_{\alpha+1}$ in the only reasonable way, and check that this works: for $f \in \tau_{\alpha}^{-1}(\kappa_{\alpha}) N_{\alpha}^{W \upharpoonright \alpha} \cap N_{\alpha}^{W \upharpoonright \alpha}$, set

$$\mathcal{I}_{\alpha+1}([f]_{W_{\alpha}}) = [\tau_{\alpha}(f)]_{U_{\alpha}}$$

This is well-defined because if $f, g \in \tau_{\alpha}^{-1}(\kappa_{\alpha}) \mathbb{N}_{\alpha}^{W \upharpoonright \alpha}$ and f and g agree on a set $X \in W_{\alpha}$ then $\tau_{\alpha}(f)$ and $\tau_{\alpha}(g)$ agree on the set $\tau(X) \in U_{\alpha}$. This is elementary due to Loś's Theorem (14 A • 5):

$$\begin{aligned} \mathsf{Ult}[\mathsf{N}_{\alpha}^{\mathfrak{W}\restriction\alpha}, W_{\alpha}] &\models ``\varphi([f]_{W_{\alpha}})'' \quad \text{iff} \quad \forall_{W_{\alpha}}^{\ast} \xi \left(\mathsf{N}_{\alpha}^{\mathfrak{W}\restriction\alpha} \models ``\varphi(f(\xi))''\right) \\ &\quad \text{iff} \quad \forall_{U_{\alpha}}^{\ast} \xi \left(\mathsf{M}_{\alpha}^{\mathfrak{U}} \models ``\varphi(\tau_{\alpha}(f)(\xi))''\right) \\ &\quad \text{iff} \quad \mathsf{Ult}[\mathsf{M}_{\alpha}^{\mathfrak{U}}, U_{\alpha}] \models ``\varphi([\tau_{\alpha}(f)]_{U_{\alpha}})'' \\ &\quad \text{iff} \quad \mathsf{Ult}[\mathsf{M}_{\alpha}^{\mathfrak{U}}, U_{\alpha}] \models ``\varphi(\tilde{\tau}_{\alpha+1}([f]_{W_{\alpha}}))''.\end{aligned}$$

Note that this means $\mathsf{Ult}[\mathsf{N}_{\alpha}^{\mathfrak{W}\restriction\alpha}, W_{\alpha}]$ is well-founded because $\mathsf{Ult}[\mathsf{M}_{\alpha}^{\mathfrak{U}}, U_{\alpha}] \cong \mathsf{M}_{\alpha+1}^{\mathfrak{U}}$ is. Hence $\mathsf{N}_{\alpha+1}^{\mathfrak{W}\restriction\alpha+1}$ makes sense, and we define τ_{α} as induced by $\tilde{\tau}_{\alpha}$ and the collapsing maps on $\mathsf{Ult}[\mathsf{M}_{\alpha}^{\mathfrak{U}}, U_{\alpha}]$ and $\mathsf{Ult}[\mathsf{N}_{\alpha}^{\mathfrak{W}\restriction\alpha}, W_{\alpha}]$. So we have defined $\langle \mathcal{N} \upharpoonright \alpha + 1, \mathfrak{W} \upharpoonright \alpha + 1 \rangle$ as an $\alpha + 1$ -length linear iteration, and $\tau_{\alpha+1}$ as desired.

Finally, we can show the map $\tau_{\alpha+1}$ commutes with the iteration maps for $\langle \mathcal{M}, \mathcal{U} \rangle$ and $\langle \mathcal{N} \upharpoonright \alpha + 1, \mathcal{W} \upharpoonright \alpha + 1 \rangle$. To see this, we first show just the case that $j^{\mathcal{U}}_{\alpha,\alpha+1} \circ \tau_{\alpha} = \tau_{\alpha+1} \circ j^{\mathcal{W}}_{\alpha,\alpha+1}$. Let $\pi^{\mathcal{U}}_{\alpha}$ be the collapsing map with image $M^{\mathcal{U}}_{\alpha}$ and similarly for $N^{\mathcal{W}}_{\alpha}$ so that $\tau_{\alpha} = \pi^{\mathcal{U}}_{\alpha} \circ \tilde{\tau}_{\alpha} \circ (\pi^{\mathcal{W}}_{\alpha})^{-1}$. Thus

$$\tau_{\alpha+1} \circ j_{\alpha,\alpha+1}^{\mathcal{W}}(x) = \pi_{\alpha+1}^{\mathcal{U}} \circ \tilde{\tau}_{\alpha+1} \circ (\pi_{\alpha+1}^{\mathcal{W}})^{-1} \circ j_{\alpha,\alpha+1}^{\mathcal{W}}(x)$$
$$= \pi_{\alpha+1}^{\mathcal{U}} \circ \tilde{\tau}_{\alpha+1}([\text{const}_x]_{W_{\alpha}})$$
$$= \pi_{\alpha+1}^{\mathcal{U}}([\tau_{\alpha}(\text{const}_x)]_{U_{\alpha}})$$
$$= \pi_{\alpha+1}^{\mathcal{U}}([\text{const}_{\tau_{\alpha}(x)}]_{U_{\alpha}}) = j_{\alpha,\alpha+1}^{\mathcal{U}} \circ \tau_{\alpha}(x).$$

Now for any $\xi < \alpha$, and any $x \in \mathbf{N}^{\mathcal{W}}_{\xi}$, we get inductively

$$j_{\xi,\alpha+1}^{\mathcal{U}} \circ \tau_{\xi}(x) = j_{\alpha,\alpha+1}^{\mathcal{U}} \circ j_{\xi,\alpha}^{\mathcal{U}} \circ \tau_{\xi}(x) = j_{\alpha,\alpha+1}^{\mathcal{U}} \circ \tau_{\alpha} \circ j_{\xi,\alpha}^{\mathcal{W}}(x) = \tau_{\alpha+1} \circ j_{\alpha,\alpha+1}^{\mathcal{W}} \circ j_{\xi,\alpha}^{\mathcal{W}}(x) = \tau_{\alpha+1} \circ j_{\xi,\alpha+1}^{\mathcal{W}}.$$

And this completes the induction and proof.

Let's now shift to using the same measure, translated by the embeddings, every time we iterate. In other words, we have \mathbf{M}, U and form $\mathbf{M}_1 = \mathbf{cUlt}[\mathbf{M}, U]$ if this is well-founded now with measure $U_1 = j_{0,1}$ "U where $j_{0,1}$ is the canonical embedding. Then we can take another ultrapower to get \mathbf{M}_2 now with measure $U_2 = j_{0,2}$ " $U = j_{1,2}$ " U_1 , and so on. Then we can This simplifies the situation considerably, and gives a good understanding of when the direct limit ultrapowers are well-founded, similar to The Wellfoundedness of Iterated Ultrapowers (12 E • 4). What Lemma 14 B • 8 then shows is that weak amenability to \mathbf{M} doesn't need to translate to weak amenability to \mathbf{N} for \mathbf{N} to continue its iteration. This is very nice for us. For example, if we use the same measure, then ω_1 -iterability implies full iterability in the following sense.

- 14B•9. Definition

Let **M** be a transitive model of some fragment of set theory. Let U be a weakly amenable M-measure over $\kappa \in M$.

- A λ -length linear iteration of \mathbf{M} by U (if it exists) is the iteration $\langle \mathcal{M}, \mathcal{U} \rangle$ where $\mathbf{M}_0^{\mathcal{U}} = \mathbf{M}$ and $\mathcal{U} = \langle U_{\alpha} : \alpha < \lambda \rangle$ is defined by $U_{\alpha} = \dot{U}^{\mathbf{M}_{\alpha}^{\mathcal{U}}}$ where each $\mathbf{M}_{\alpha}^{\mathcal{U}}$ is a FOL(\in, \dot{U})-model and $U = \dot{U}^{\mathbf{M}}$. We frequently write $\mathbf{M}_{\alpha}^{\mathcal{U}}$ for $\mathbf{M}_{\alpha}^{\mathcal{U}}$ in this case.
- We say \mathbf{M} , U is λ -linearly iterable iff \mathbf{M}_{α}^{U} exists for every $\alpha < \lambda$.
- We say M, U is *fully linearly iterable* iff it is λ -iterable for all $\lambda <$ Ord.

Again, we need weak amenability at the start to ensure $j_{0,1}^{\mathcal{U}}$ "U is an $\mathbf{M}_1^{\mathcal{U}}$ -measure (that is also weakly amenable), but Lemma 14 B • 8 tells us we don't need weak amenability for the hulls.

– 14B•10. Lemma

Let **M** be a transitive set model of some (suitable) fragment of set theory, and let U be a weakly amenable M-measure over κ . Therefore the following are equivalent:

- 1. \mathbf{M}, U is fully linearly iterable.
- 2. **M**, U is ω_1 -linearly iterable.

3. N, \dot{U}^{N} is ω_1 -linearly iterable for any countable, transitive N with FOL(\in , \dot{U})-elementary $\tau : N \to M$.

Proof .:.

Clearly (1) implies (2). For (2) implies (3), suppose (2) holds. For any λ -length linear iteration of **M** by U, $\lambda < \omega_1$, the critical point of U_{β} is $j_{0,\beta}^U(\kappa)$ for any $\beta < \alpha$. Let **N** be countable and transitive with $FOL(\in, \dot{U})$ -elementary $\tau : N \to M$ so that $\kappa \in im(\tau)$. By Lemma 14 B • 8, because the hypothesis of (3) is satisfied, **N** has an α -length linear iteration $\langle \mathcal{N}, W \rangle$. Hence **N**, \dot{U}^{N} is ω_1 -linearly iterable.

For (3) implies (1), suppose (1) is false: \mathbf{M}, U is not fully linearly iterable. Let λ be the length of a linear iteration $\langle \mathcal{M}, \mathcal{U} \rangle$ where $\mathbf{M}_{\infty}^{\mathcal{U}}$ is ill-founded. Let η be large enough such that $\langle \mathcal{M}, \mathcal{U} \rangle \in \mathbf{H}_{\eta}$, and consider the skolem hull $\mathbf{Hull} = \mathbf{Hull}^{\mathbf{H}_{\eta}}(\{\mathbf{M}, U, \lambda, \mathcal{M}, \mathcal{U}\})$ which is countable and an $\mathrm{FOL}(\in, \dot{U})$ -elementary submodel of \mathbf{H}_{η} . The transitive collapse map π yields a transitive model **cHull**. By elementarity, it follows that **cHull** $\models "\pi(\mathbf{M}), \pi(U)$ is not $\pi(\lambda)$ -iterable" because it knows $\langle \mathcal{M}, \mathcal{U} \rangle$ is a $\pi(\lambda)$ -length linear iteration with ill-founded last model. As ill-foundedness is absolute between transitive models of $\mathsf{ZF} - \mathsf{P}$, it follows that in \mathbf{H}_{η} (and hence in V) $\pi(\mathbf{M}), \pi(U)$ is not $\pi(\lambda)$ -iterable. Note that as cHull is countable, $\pi(\lambda) < \omega_1$. Note also that $\pi^{-1} \upharpoonright N : N \to M$ is $\mathsf{FOL}(\in, \dot{U})$ -elementary by the $\mathsf{FOL}(\in, \dot{U})$ -elementarity of **Hull** in \mathbf{H}_{η} :

$$x \in N$$
 and $\mathbf{N} \models "\varphi(x)$ " iff $\mathsf{cHull} \models "x \in N \land \mathbf{N} \models "\varphi(x)$ ""

iff
$$\mathsf{cHull} \models ``\pi^{-1}(x) \in \pi^{-1}(N) = M \land \pi^{-1}(\mathsf{N}) = \mathsf{M} \models ``\varphi(\pi^{-1}(x))````.$$

Elementarity also gives $\dot{U}^{N} = \pi(U)$. But then the hypotheses of (3) hold, but N, $\pi(U)$ is not $\pi(\lambda)$ -iterable, and hence not ω_1 -iterable. This shows (3) is false and hence we've shown the contrapositive. \dashv

This also proves the result for class models due to ill-foundedness being a result of only a bounded amount of infor-

mation from M; if M is a class, we can bound all of the problems in some set, and then do the same proof. Of course, this requires quite a bit of ugly technical problem checking, and so it is left to a diligent reader who is for some reason interested.

As a result, if we continually use only a single measure, the well-foundedness of the external ultrapowers becomes clear.

- 14B•11. Theorem (Wellfoundedness of Iterations) -

Let **M** be a transitive model of a (sufficient) fragment of set theory. Let U be a σ -complete, weakly amenable **M**-measure over κ . Therefore **M**, U is fully lineary iterable.

Proof .:.

Suppose \mathbf{M}, U is not linearly iterable. By Lemma 14 B • 10, there is some countable $\mathbf{N}, \dot{U}^{\mathbf{N}}$ and $\tau : N \to \mathbf{M}$ that is FOL(\in , \dot{U})-elementary, but \mathbf{N} is not ω_1 -iterable. This results in some $\alpha < \omega_1$ -length iteration $\langle \mathcal{N}, \mathcal{W} \rangle$ with $\mathbf{N}_0^{\mathcal{W}} = \mathbf{N}$, and $\mathcal{W} = \langle W_{\xi} : \xi < \alpha \rangle$ with $\dot{U}^{\mathbf{N}} = W_0$ such that $\mathbf{N}_{\infty}^{\mathcal{W}}$ is ill-founded. But note that by Lemma 14 B • 7, each $\mathbf{N}_{\xi}^{\mathcal{W}}$ can be realized back into \mathbf{M} : there is some FOL(\in)-elementary $\tau_{\xi} : N_{\xi}^{\mathcal{W}} \to \mathbf{M}$. But then by definition of the direct limit, there is an elementary $\tau_{\infty} : N_{\infty}^{\mathcal{W}} \to \mathbf{M}$ which means that $N_{\infty}^{\mathcal{W}}$ is well-founded, a contradiction. \dashv

§14C. External Extenders

Section 15. Exercises

Exercises

Chapter III. Relative Definability and Constructibility*

Inner model theory deals with inner models of set theory, so it makes sense to start with the first non-trivial example: the constructible hierarchy L. The usual definition Definition 8 A • 1 of L will be used here, although there is an alternative characterization with fine-structure in mind. In particular, $L_0 = \emptyset$, $L_{\gamma} = \bigcup_{\alpha < \gamma} L_{\alpha}$ for γ a limit ordinal, and

 $\mathsf{L}_{\alpha+1} := \mathrm{FOLp}(\mathsf{L}_{\alpha}) = \{ s \subseteq \mathsf{L}_{\alpha} : s \text{ is } \mathrm{FOLp}(\in) \text{-definable over } \langle \mathsf{L}_{\alpha}, \in \rangle \}.$

The resulting structure $L := \bigcup_{\alpha \in Ord} L_{\alpha}$ will satisfy ZFC by Theorem 8 A • 7 and Theorem 8 A • 8 and will have strong canonicity properties, including Condensation (8 B • 3), which yields nice, useful combinatorial properties like GCH by Theorem 8 C • 5.

Generally speaking, the way we form L is very useful and the results of Section 8 generalize if we are trying to get a "canonical" inner model that includes some set X. For instance, X might be a measure or an extender, something incompatible with L by L Has No Measurable Cardinals ($12D \cdot 4$). In the resulting model, L[X], we have a model which acts very much like L, having lots of the same desirable properties, but which is now compatible with certain large cardinals where L was not.

We will be working with a slightly different language here than $FOL(\in)$. For the sake of being explicit, for any language extending $FOL(\in)$, we make the following definition for how to modify the axioms of ZFC.

– 15•1. Definition

Let σ be consist of non-logical symbols, including the symbol ' \in '. Define the FOL(σ)-theory ZFC(σ) and its variants as follows.

• (FOL(σ)-Comprehension, Comp(σ)) for each x, and for each FOL(σ)-formula $\varphi(v, \vec{w})$,

 $\forall w_0 \cdots \forall w_n \ \forall x \ \exists z \ \forall v \ (v \in z \leftrightarrow v \in x \land \varphi(v, \vec{w})).$

• $(FOL(\sigma)$ -Replacement, $Rep(\sigma)$) for each $FOL(\sigma)$ -formula φ ,

 $\forall w_0 \cdots \forall w_n \forall D \left(\forall x (x \in D \to \exists ! y \varphi(x, y, \vec{w})) \to \exists R(y \in R \leftrightarrow \exists x (x \in D \land \varphi(x, y, \vec{w}))) \right).$

 $ZFC(\sigma)$ consists of $ZFC + Comp(\sigma) + Rep(\sigma)$. We similarly define $ZF(\sigma)$, and so forth. Typically for ease of notation, if $\sigma = \{ \in, A, B, \dots, Z \}$ then we just write $ZFC(A, B, \dots, Z)$ rather than $ZFC(\{ \in, A, B, \dots, Z \})$.

Section 16. Relative Constructibility

The general idea behind relative constructibility is just that want to build inner models, and generally we want some "least" inner model that contains some desired set X, and the way this is done is by starting with X and then forming L as usual from X. This, however, is a vague idea and there are a few ways to interpret this.

- 1. We allow X as a predicate, essentially taking $FOLp(\in, X)$ -definable subsets of the previous stage to form the next stage where 'X' as a symbol is interpreted as $J \cap X$ for each stage J. In this way, we form things constructible using X as a predicate. Call the result L[X].
- 2. We start with trcl(X) as the first stage and then take $FOLp(\in)$ -definable subsets of the previous stage to form the next stage. In this way, we form things constructible after being given access to all of X. Call the result L(X).

The formal definitions of these things are unfortunately technical, but aren't that much more than slight variants of Definition $8 \text{ A} \cdot 1$.

As before there is a little bit of technical worry here about what it means to be truly F0Lp-definable, but just note that we are interpreting all of this relative to the background model V. So if we consider a non-standard universe M that is not transitive or not well-founded or some other such oddity, then we interpret definability according to this model: we

RELATIVE CONSTRUCTIBILITY

allow potentially non-standard "formulas" in "definitions" over previous stages. More precisely, we can use Theorem $6 B \cdot 6$ and disregard the restriction that we need to code a real-world formula, taking whatever the model thinks of as a formula as a legitimate.

Note that defining L(E) requires that *E* is a set rather than a proper class, but we will want to consider proper class \mathcal{E} when defining $L[\mathcal{E}]^{i}$.

− 16 • 1. Definition Let \mathcal{E} be an arbitrary class. Regarding ' \mathcal{E} ' as a symbol, for any transitive model M, the interpretation of this symbol is just $\mathcal{E} \cap M$. We frequently write $\langle M, \in, \mathcal{E} \rangle$ for $\langle M, \in, \mathcal{E} \cap M \rangle$. Now define

$$\begin{split} & L_0[\mathcal{E}] = \emptyset & & L_0(\mathcal{E}) = \operatorname{trcl}(\mathcal{E}) \\ & L_{\gamma}[\mathcal{E}] = \bigcup_{\xi < \gamma} L_{\xi}[\mathcal{E}] & & L_{\gamma}(\mathcal{E}) = \bigcup_{\xi < \gamma} L_{\xi}(\mathcal{E}), \quad \text{ for } \gamma \text{ a limit,} \end{split}$$

and crucially, for successor stages,

 $L_{\alpha+1}[\mathcal{E}] = \{ s \subseteq L_{\alpha}[\mathcal{E}] : s \text{ is } FOLp(\in, \mathcal{E}) \text{-definable over } \langle L_{\alpha}[\mathcal{E}], \in, \mathcal{E} \rangle \}$

 $L_{\alpha+1}(\mathfrak{E}) = \{ s \subseteq L_{\alpha}(\mathfrak{E}) : s \text{ is FOLp}(\in) \text{-definable over } \langle L_{\alpha}(\mathfrak{E}), \in \rangle \}.$

This gives $L[\mathcal{E}]$ as $L_{Ord}[\mathcal{E}]$ and similarly for $L(\mathcal{E})$. As a model, we usually regard these in the language $FOL(\in, \mathcal{E})$.

With those technical consideration aside, note that the two interpretations can be different.

- 16.2. Example If $\mathcal{P}(\omega)^{\mathsf{V}} \neq \mathcal{P}(\omega)^{\mathsf{L}}$, then $\mathsf{L} = \mathsf{L}[\mathcal{P}(\omega)] \neq \mathsf{L}(\mathcal{P}(\omega))$.

Proof .:.

In this case, the additional predicate of being an element of $\mathcal{P}(\omega)$ is already $\text{FOLp}(\in)$ -definable. So we should have $L[\mathcal{P}(\omega)]$ is just L: we never gain anything new from having access to an oracle that tells us whether or not something is a subset of ω . More precisely, clearly $L_{\alpha} \subseteq L_{\alpha}[\mathcal{P}(\omega)]$ for all α . For equality, $L_0 = L_0[\mathcal{P}(\omega)]$. For $\alpha > 0$, $L_{\alpha}[\mathcal{P}(\omega)] \cap \mathcal{P}(\omega) = \mathcal{P}(\omega) \cap L_{\alpha}$ inductively. But $\mathcal{P}(\omega) \cap L_{\alpha}$ is already $\text{FOLp}(\in)$ -definable over $\langle L_{\alpha}, \in \rangle$ (regardless of whether $\omega \in L_{\alpha}$ or not):

 $x \in \mathcal{P}(\omega) \cap L_{\alpha}$ iff $L_{\alpha} \models "x \subseteq \text{Ord} \land$ there is no limit ordinal $\langle \sup(x)"$.

Hence we can always just replace the new predicate for being in $\mathcal{P}(\omega) \cap L_{\alpha}$ with this above definition to get $L_{\alpha+1}[\mathcal{P}(\omega)] \subseteq L_{\alpha+1}$ and hence equality. Thus $L[\mathcal{P}(\omega)] = L$.

On the other hand, $\mathfrak{P}(\omega) = \operatorname{trcl}(\mathfrak{P}(\omega)) \subseteq L(\mathfrak{P}(\omega))$ and thus $\mathfrak{P}(\omega)^{\mathsf{V}} = \mathfrak{P}(\omega)^{\mathsf{L}(\mathfrak{P}(\omega))} \neq \mathfrak{P}(\omega)^{\mathsf{L}}$ shows that $L[\mathfrak{P}(\omega)] = L \neq L(\mathfrak{P}(\omega))$.

We also get standard results similar to L. Note that often people require E to be transitive to consider L(E), so in such cases, trcl(E) = E.

— 16•3. Result –

For any class \mathcal{E} and any $\alpha \in \text{Ord}$, $L_{\alpha}[\mathcal{E}]$ is transitive and $\text{FOL}(\in, \mathcal{E})$ -definable in a way that is absolute between transitive models of $\text{ZF}(\mathcal{E}) - P$. Assuming E is a set, we have the same for $L_{\alpha}(E)$ but with E replaced by trcl(E) above.

Proof .:.

Definability and absoluteness follows in just the same way as Corollary 8 A • 4 and Absoluteness of L (8 B • 1).

ⁱPart of the reason why is the following. We will eventually consider adding extenders to L to form L[E] that witnesses various large cardinal properties inside an L-like model. Adding more extenders allows us to witness multiple large cardinals. This all works out fine and dandy if we only work with $L_{\kappa}[E]$ with κ sufficiently larger than the E, and we can use similar techniques to L like condensation in conjunction with skolem hulls like Corollary 8 C • 4 for stages $\geq \kappa$. But this might not be true for smaller κ s because we can't easily define L[E] without access to all of E: if $E \notin L_{\kappa}[E]$, what is L[E]^{L_{\kappa}[E]}? We can't argue about the absoluteness of construction to say that the theory determines what it means to be an initial segment of L[E]. Why is this really a problem? Ideally, we'd also like the ability to work with a proper class of measurable cardinals, and for a proper class, there is no such κ that will be sufficiently large to work with. In such cases, we need to be able to talk about mere portions of the class.

For transitivity, $L_0 = L_0[\mathcal{E}] = \emptyset$ and $L_0(\mathcal{E}) = \operatorname{trcl}(\mathcal{E})$ are by definition transitive. Limit stages are obvious: x in the limit stage α must be in a previous stage ξ so inductively is a subset of stage ξ , and hence a subset of the union: stage α . For transitivity for successor stages, we only prove the result for $L_{\alpha}[\mathcal{E}]$ since the result similarly holds for $L_{\alpha}(\mathcal{E})$. First we show $L_{\alpha}[\mathcal{E}] \subseteq L_{\alpha+1}[\mathcal{E}]$. Let $x \in L_{\alpha}[\mathcal{E}]$ be arbitrary so that inductively $x \subseteq L_{\alpha}[\mathcal{E}]$. Then we can define

 $y \in x$ iff $L_{\alpha}[\mathcal{E}] \models "y \in x"$.

So x is FOLp(\in)-definable over $L_{\alpha}[\mathcal{E}]$ and hence $x \in L_{\alpha+1}[\mathcal{E}]$. So $L_{\alpha}[\mathcal{E}] \subseteq L_{\alpha+1}[\mathcal{E}]$. Since each element $s \in L_{\alpha+1}[\mathcal{E}]$ has $s \subseteq L_{\alpha}[\mathcal{E}] \subseteq L_{\alpha+1}[\mathcal{E}]$, transitivity follows. \dashv

This implies that $L[\mathcal{E}]$ and L(E) are stratified and hence satisfy a pretty significant portion of ZF. We state the following without proof as the proofs are identical to the case for L in Subsection 8 A.

Let \mathcal{E} be a class. Assume when working with $L(\mathcal{E})$ that $\mathcal{E} = E$ is a set.

- 1. $L[\mathcal{E}] \models \mathsf{ZFC} \text{ and } L(E) \models \mathsf{ZF}.$
- 2. $L \subseteq L[\mathcal{E}]$ and $L \subseteq L(E)$.
- 3. If \mathcal{E} is a set, $L[\mathcal{E}]$ is the least inner model M of ZF P such that $\mathcal{E} \cap M \in M$.
- 4. More generally, $L[\mathcal{E}]$ is the least inner model of the theory $ZF(\mathcal{E}) P$.
- 5. L(E) is the least inner model M of ZF P such that $trcl(E) \in M$.

Proof .:.

- That both model ZF(E) follows from the same proofs as Lemma 8 A 6, Theorem 8 A 7, and using Result 7 D 3. That L[E] ⊨ AC follows by the same proof as Theorem 8 A 8. This doesn't work for L(E) because of the potential lack of an ability to definably order the starting point, trcl(E). Indeed from the existence of sufficiently large cardinals, it's possible that L(ℝ) ⊨ ¬AC.
- 2. This follows from Absoluteness of L (8 B 1).
- 3. Clearly $L[\mathscr{E}]$ is such an inner model (if $\mathscr{E} \subseteq V_{\alpha}$ then $\mathscr{E} \cap L[\mathscr{E}] = \mathscr{E} \cap V_{\alpha} \cap L[\mathscr{E}] = \mathscr{E} \cap V_{\alpha}^{L[\mathscr{E}]}$ and so $\mathscr{E} \cap L[\mathscr{E}]$ is a FOLp(\in, \mathscr{E})-definable subset of $V_{\alpha}^{L[\mathscr{E}]}$ and hence in $L[\mathscr{E}]$). So we need to show that $L[\mathscr{E}]$ is the least such inner model. Let $M \models ZF P$ be an inner model such that $\mathscr{E} \cap M \in M$. By the absoluteness of $L[\mathscr{E} \cap M]$, we get that $L[\mathscr{E} \cap M]^M = L[\mathscr{E} \cap M] \subseteq M$. We now show that $L[\mathscr{E} \cap M] = L[\mathscr{E}]$. We show a level by level equivalence by induction: $L_0[\mathscr{E} \cap M] = L_0[\mathscr{E}]$. Limit stages are immediate. For successor stages, assume inductively that $L_{\alpha}[\mathscr{E} \cap M] = L_{\alpha}[\mathscr{E}]$ so that both interpret '\mathscr{E}' the same way, and hence have the same FOLp(\in, \mathscr{E})-definable subsets, meaning $L_{\alpha+1}[\mathscr{E} \cap M] = L_{\alpha+1}[\mathscr{E}]$. It follows that $L[\mathscr{E} \cap M] = L[\mathscr{E}] \subseteq M$, as desired.
- 4. $L[\mathcal{E}]$ is such an inner model of the modified theory by the same reasoning as with Theorem 8A•7 and Result 7D•3. Let $\mathbf{M} \models \mathsf{ZF}(\mathcal{E}) - \mathsf{P}$ be an inner model. Since $\mathcal{E} \cap \mathsf{M}$ is a class of M , we can define $L[\mathcal{E} \cap \mathsf{M}]$ and the exact same reasoning as in (3) goes through to tell us $L[\mathcal{E}] = L[\mathcal{E} \cap \mathsf{M}] = L[\mathcal{E} \cap \mathsf{M}]^{\mathsf{M}} \subseteq \mathsf{M}$.
- 5. Clearly L(E) has this property. So suppose $M \vDash ZF P$ is an inner model with $trcl(E) \in M$. We can thus form L(E) in an absolute way as with Absoluteness of $L(8B \cdot 1)$ and hence $L(E)^{M} = L(E) \subseteq M$. \dashv

In particular, if $E \subseteq L$, then L[E] = L(E) which need not be L. For example, suppose $\mathcal{P}(\omega) \neq \mathcal{P}(\omega)^{L}$ with $x \in \mathcal{P}(\omega) \setminus L$. It follows that $L[x] \neq L$ but nevertheless, since both are the least inner model containing $x \cap \omega = x$, L[x] = L(x). This is despite Example 16 • 2 which showed that $L[\mathcal{P}(\omega)] = L$ always: knowing whether something is a subset of ω isn't helpful, but knowing whether something is in some undefinable, complicated subset of ω does give us something new: x. So even if $X \in Y$ or $X \subseteq Y$, we don't necessarily have $L[X] \subseteq L[Y]$.ⁱⁱ Nevertheless, we always have $X \in Y$ implies $L(X) \subseteq L(Y)$.

- 16 • 5. Corollary

Let $X \in Y$ be arbitrary sets. Therefore $L(X) \subseteq L(Y)$.

ⁱⁱFor example, above, $x \subseteq \omega$ but $L[x] \neq L[\omega] = L$.

Relative Constructibility

Proof .:.

Since
$$X \in trcl(Y) \subseteq L(Y)$$
 we get $trcl(X) \in L(Y)$. As the least such model, $L(X) \subseteq L(Y)$. \dashv

We unfortunately cannot get any such relation for L[X] and L[Y] in general for the examples stated above. More generally, for each X, there is a Y such that $X \subseteq Y$ and $X \in Y$ but L[Y] = L: consider sufficiently large α such that $X \in V_{\alpha}$ so that $V_{\alpha} \cap L = V_{\alpha}^{L} \in L$ implies $L[V_{\alpha}] = L$. If X is a proper class, take Y = V for the same result. The issue really is that often small sets carry more specific information than larger sets, and L[X] uses this information rather than any surrounding context.

- 16.6. Theorem (Relative Condensation) – Suppose $M \models ZF(\mathcal{E}) - P$ is transitive.

- 1. Suppose $\mathbf{M} \models "\mathbf{V} = \mathbf{L}[\mathcal{E}]$ ". Therefore $\mathbf{M} = \mathbf{L}_{\text{Ord} \cap \mathbf{M}}[\mathcal{E}]$.
- 2. Suppose $\mathbf{M} \models "\mathbf{V} = \mathbf{L}(E)$ " for $\mathcal{E} = E \in \mathbf{M}$. Therefore $\mathbf{M} = \mathbf{L}_{\operatorname{Ord} \cap \mathbf{M}}(E)$.

Proof .:.

By the definability and absoluteness of the construction of $L[\mathcal{E}]$, it follows that $V^{\mathsf{M}} = L[\mathcal{E}]^{\mathsf{M}} = L_{\text{Ord}\cap\mathsf{M}}[\mathcal{E}]$, and similarly for $L(\mathcal{E})$.

And we can use this with skolem hulls as with L to get similar properties. Recall that for L, if κ is regular in L, then $L_{\kappa} \models ZFC - P + "V = L"$ by Result 8 C • 2. Hence taking the skolem hull of cHull^{L_{\karkappi}</sup>(A) for some $A \subseteq L_{\kappa}$ yields by condensation another level L_{α} for some $\alpha < \kappa$. We'd like the same sort of situation to happen for L[\mathcal{E}] and L(E).</sup>}

- **16•7.** Lemma Let \mathscr{E} be a class. Let $\kappa > \aleph_0$ be regular. Therefore 1. $|L_{\alpha}[\mathscr{E}]| = |\alpha|$ for every infinite $\alpha \in \text{Ord}$; and

2. $L_{\kappa}[\mathcal{E}] \models ZFC - P + "V = L[\mathcal{E}]".$

Proof .:.

- 1. Note that $L_{\alpha+1}[\mathcal{E}]$ is given by the closure of $L_{\alpha}[\mathcal{E}]$ under countably many operations, just as with getting $L_{\alpha+1}$ from L_{α} but with with the ability to check membership in $\mathcal{E} \cap L_{\alpha}[\mathcal{E}]$. As a result, $|L_{\alpha+1}[\mathcal{E}]| \leq |L_{\alpha}[\mathcal{E}]| \cdot \aleph_0$ which for $\alpha \geq \omega$ clearly has $\omega \subseteq L_{\alpha}[\mathcal{E}]$ meaning $|L_{\alpha}[\mathcal{E}]| \cdot \aleph_0 = |L_{\alpha}[\mathcal{E}]|$ and so $|L_{\alpha+1}[\mathcal{E}]| = |L_{\alpha}[\mathcal{E}]|$. Inductively this is $|\alpha| = |\alpha+1|$. For limit γ , $|L_{\gamma}[\mathcal{E}]| \leq \gamma \cdot \sup_{\alpha < \gamma} |L_{\alpha}[\mathcal{E}]| \leq \gamma \cdot \gamma = \gamma$. Clearly $\gamma \subseteq L_{\gamma}[\mathcal{E}]$ implies $|L_{\gamma}[\mathcal{E}]| \geq |\gamma|$ and so we have equality.
- 2. $L_{\kappa}[\mathcal{E}] \models ZF(\mathcal{E}) P$ by the same proof as in Result 8 C 2, using (1) for wRep(\mathcal{E}). So by Result 16 4 (4) and absoluteness from Result 16 3,

$$\mathsf{L}_{\kappa}[\mathscr{E}] \subseteq \mathsf{L}_{\kappa}[\mathscr{E} \cap \mathsf{L}_{\kappa}[\mathscr{E}]]^{\mathsf{L}_{\kappa}[\mathscr{E}]} = \mathsf{L}_{\kappa}[\mathscr{E} \cap \mathsf{L}_{\kappa}[\mathscr{E}]] = \mathsf{L}_{\kappa}[\mathscr{E}].$$

When working with mere $FOL(\in)$ -models like $\langle L_{\kappa}[\mathcal{E}], \in \rangle$ (rather than $\langle L_{\kappa}[\mathcal{E}], \in, \mathcal{E} \rangle$) we might not be able to define \mathcal{E} as a class in $L_{\kappa}[\mathcal{E}]$, and so the ability to define $L[\mathcal{E}]$ inside $L_{\kappa}[\mathcal{E}]$ isn't true, and in such cases, we require κ to be sufficiently large such that $L[\mathcal{E}] \cap \mathcal{E} \in L_{\kappa}[\mathcal{E}]$ to argue as above.

This allows us to prove things like GCH in a similar way to Theorem 8 C • 5. But there is a little bit of a hiccup. For example, if R codes \aleph_2 -many real numbers, then we should not expect $L[R] \models CH$. So instead, we only get GCH for cardinals left unaffected by R, i.e. cardinals above |trcl(R)|.

16.8. Corollary Let *E* be a set, and write $E' = E \cap L[E] \in L[E]$. Therefore, $L[E] \models "\forall \kappa \ge \aleph_0 \cdot |\operatorname{trcl}(E')| \ (2^{|\kappa|} = |\kappa|^+)"$. *Proof* ...

Argue in L[E]. Let $\kappa \geq |\operatorname{trcl}(E')|, \aleph_0$. It follows that $E' \in L_{\mu}[E]$ for some sufficiently large, regular μ . Let

 $A \subseteq \kappa$ be in L[E] and without loss of generality, μ is large enough that $A \in L_{\mu}[E]$. Consider the FOL(\in)-hull

$$\mathbf{H} = \mathbf{Hull}^{\mathbf{L}_{\mu} \lfloor E \rfloor}(\{A\} \cup \kappa \cup \operatorname{trcl}(E') \cup \{E'\}).$$

Since $L_{\mu}[E] \models ZF(E) - P + "V = L[E]"$ by Lemma 16 • 7, H does the same. Thus when we collapse by some $\pi : H \to M$, where *M* is transitive, we get that $M \models ZF(\pi(E')) - P + "V = L[\pi(E')]"$. Recalling properties of the transitive collapse from Definition 4 • 6, note that $trcl(E') \cup \{trcl(E'), E'\} \subseteq H$ and so $\pi \upharpoonright trcl(E') = id \upharpoonright trcl(E')$ by Corollary 6 C • 2. Thus

$$\pi(E') = \{\pi(x) : x \in E' \subseteq \operatorname{trcl}(E')\} = \{x : x \in E'\} = E'.$$

Thus $\mathbf{M} \models \mathsf{ZF}(E) - \mathsf{P} + "\mathsf{V} = \mathsf{L}[E]$ ". So by Relative Condensation (16 • 6), $\mathbf{M} = \mathsf{L}_{\alpha}[E]$ for some $\alpha < \mu$. In fact, since $|H| = |M| \le \kappa + |\operatorname{trcl}(E')| = \kappa$, by size restrictions—(1) of Lemma 16 • 7—we get $|\alpha| = \kappa$ and so $\alpha < \kappa^+$. Moreover, $\pi(A) = A \in M = \mathsf{L}_{\alpha}[E]$ by the same reasoning as with $\pi(E') = E'$, meaning that $A \in \mathsf{L}_{\kappa^+}[E]$. As $A \subseteq \kappa$ was arbitrary, it follows that $\mathcal{O}(\kappa) \subseteq \mathsf{L}_{\kappa^+}[E]$, which has only κ^+ -many elements by (1) of Lemma 16 • 7.

The restriction on $\kappa > \text{trcl}(E')$ is necessary, since by techniques of forcing, we can force the failure of $2^{\kappa} = \kappa^+$ for smaller cardinals. Similarly, the requirement on *E* being a set is necessary since it's consistent for a proper class *E* to yield that GCH fails at arbitrarily large cardinals in L[*E*]. This is discussed more in the context of *forcing* in a later chapter.

Unfortunately, we cannot have Corollary 16 • 8 for L(E) so easily. It's consistent relative to sufficiently large cardinals that AC is false in $L(\mathcal{P}(\omega))$, meaning that we can't really consider the cardinality of $\mathcal{P}(\kappa)$ as a cardinal, i.e. as an element of Ord.

Lastly, it will be very useful to simply assume $E \in L[E]$ by way of replacing E with $E' = E \cap L[E]$ and deducing $E' \in L[E'] = L[E]$.

- **16**•**9**. **Result** For any class \mathcal{E} , $L[\mathcal{E}] = L[\mathcal{E} \cap L[\mathcal{E}]]$.

Proof .:.

Write $\mathcal{E}' = \mathcal{E} \cap L[\mathcal{E}]$. Since \mathcal{E}' is a class of $L[\mathcal{E}]$, we can define $L[\mathcal{E}']$ in $L[\mathcal{E}]$ in an absolute way: $L[\mathcal{E}']^{L[\mathcal{E}]} = L[\mathcal{E}'] \subseteq L[\mathcal{E}]$. As a result,

 $\mathcal{E} \cap L[\mathcal{E}'] = \mathcal{E} \cap L[\mathcal{E}] \cap L[\mathcal{E}'] = \mathcal{E}' \cap L[\mathcal{E}']$

is a class of $L[\mathcal{E}']$ and thus $L[\mathcal{E}'] \vDash ZFC(\mathcal{E})$ by Result 16 • 4 (1), meaning as the minimal such inner model, $L[\mathcal{E}] \subseteq L[\mathcal{E}']$ by Result 16 • 4 (4). Hence we have equality. \dashv

§16 A. Constructibility and measures

The first important fact about relative constructibility, which is partially our aim in introducing it, is its compatibility with measurable cardinals. In particular, if we have a measure U in V, U will (more or less) still be a measure in L[U].

```
- 16A • 1. Theorem - Let U be a measure on \kappa in V. Therefore L[U] \models "U \cap L[U] is a measure". In particular, \kappa is measurable in L[U].
```

Proof .:.

We must show that $U' = U \cap L[U]$ is

1. an ultrafilter;

- 2. non-principal;
- 3. κ -complete; and
- 4. normal
- in L[U]. These are mostly immediate from the fact that U is a measure in V.

Exercises

- 1. That U' is an ultrafilter over $\mathfrak{P}(\kappa) \cap \mathbb{L}[U]$ is almost immediate from U being an ultrafilter: if $x, y \in U' \subseteq U$ then $x \cap y \in U \cap \mathbb{L}[U] = U'$. $x \notin U'$ for $x \in \mathbb{L}[U]$ means $x \notin U$ so $\kappa \setminus x \in U'$, and vice versa. Upward closure and that $\emptyset \notin U$, $\kappa \in U$ are straightforward.
- 2. This follows since U is non-principal: every bounded subset of κ has its complement in $U \cap L[U] = U'$.
- 3. If $\langle x_{\alpha} \in U : \alpha < \gamma \rangle \in L[U]$ for $\gamma < \kappa$, then $\bigcap_{\alpha < \gamma} x_{\alpha} \in U \cap L[U] = U'$.
- 4. If $f : X \to \kappa$ is decreasing for $X \in U'$ and $f \in L[U]$, then there is some α such that $f^{-1}\{\alpha\} \in U$. But clearly $f^{-1}\{\alpha\} \in L[U]$ and so it's in U'.

So this already shows the compatibility of measurable cardinals with L-like inner models. Note that there are some interesting consequences of the above. In particular, we can consider countable models with (what they believe are) measurables in them. We use Result 16 • 9 to get rid of the annoying notation of $U \cap L[U]$ as in Theorem 16 A • 1, replacing this just with U without issue.

— 16 A • 2. Corollary –

```
Let U be a measure in L[U]. Therefore there are W, \alpha < \omega_1 such that L_{\alpha}[W] \models ZF - P + "W is a measure".
```

Proof .:.

Let μ be sufficiently large and regular such that $U \in L_{\mu}[U]$. Consider the FOL(\in)-hull **cHull**^{L_{\mu}[U]}({U}) which is therefore a countable, transitive model of ZFC – P+ "there is a measurable cardinal" +"V = L[W]" for some W a measure (as understood in the hull). W itself is just given by $\pi(U)$ for the transitive collapse map π : Hull^{L_{\mu}[U]}({U}) \rightarrow cHull^{L_{\mu}[U]}({U}). By Relative Condensation (16 • 6), the hull is L_{α}[W] for some countable α .

This guarantees the existence of a "mouse" relative to the existence of a measurable cardinal. A "mouse" is a vague notion that can kind of be thought of as a building block for an inner model. There are lots of standard techniques involving mice, but one of the most fundamental is that of *comparison*, which basically means that if we have two "mice" and continually iterate them using their measures, we'll eventually arrive at the same place: L[E] for some E. In this place, we can then *compare* them as initial segments of each other.

Section 17. Sharp Objects

Section 18. Multiple Measures and Comparison

Section 19. Extenders and Sequences of Extenders

Section 20. Exercises

Chapter IV. Descriptive Set Theory*

Modern descriptive set theory is a diverse and complicated field that interacts with a wide variety of other fields of mathematics, especially group theory, measure theory, computability theory, and of course set theory. We will not be so interested in these connections for now. Instead, we will focus on the basics needed for inner model theory.

Section 21. Relevant Topology

Formally, \mathbb{R} is usually taken as a set of dedekind cuts of \mathbb{Q} , itself a set of equivalence classes formed from either ω or \mathbb{Z} . This formal notion is not easy to work with, and is basically abandoned outside of demonstrating the power of ZFC as a mathematical foundation. So instead, we will work with \mathbb{R} as the branches of the binary tree ${}^{<\omega}2$, or the branches of ${}^{<\omega}\omega$, or as the space $\omega \times {}^{\omega}\omega \times {}^{\omega}\omega$, and so on. The part of the structure that we care about will not change with these technically different spaces, since there is a unique standard borel space that is uncountable. The proof of this highly nontrivial theorem can be found in classic books on descriptive set theory like [18].

§ 21 A. Metric and topological spaces

We begin with some very basic topological and metric space results. Many of these will not be proven here as the reader is assumed to be somewhat familiar with them.ⁱ Nevertheless, we will still preset the results if only to be thorough. A reader unfamiliar with these concepts is encouraged to read a standard analysis book like [28] and [29].

— 21 A • 1. Definition

Let X be set. A *metric* on X is a function $d : X \times X \to \mathbb{R}$ such that

• d(x, x) = 0;

• d(x, y) > 0 for all $x \neq y$;

- d(x, y) = d(y, x); and
- (the triangle inequality) $d(x, y) + d(y, z) \ge d(x, z)$.
- For $d : X^2 \to \mathbb{R}$ a metric, the open ball around x of radius r is just the set $B_r(x) = \{y \in X : d(x, y) < r\}$.

A metric space is just a set with a "metric" on it, which really just gives the notion of a *distance* between two elements of the set. In the case of \mathbb{R} , the standard metric is just the absolute value of one number minus the other: d(x, y) = |x-y|. But there are many others for \mathbb{R} : $\langle x, y \rangle \mapsto \frac{|x-y|}{|x-y|+1}$ and $\langle x, y \rangle \mapsto 3|x-y|^2$, for example. On any given set, we can take the discrete metric, defined by

$$d(x, y) = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{if } x = y. \end{cases}$$

This is called discrete because all points are "separated" from each other. Whereas \mathbb{R} always has infinitely many points in the open interval (a, b) $(a, b \in \mathbb{R})$, the open ball $B_1(x) = \{x\}$ in the discrete topology has just one element, as d(x, y) < 1 implies d(x, y) = 0 and thus x = y.

In some sense, the triangle inequality just ensures that the "distance" understanding makes sense. More precisely, it ensures that the actual distance from x to z is the least of all paths: if we go from x to y and then from y to z, the distance from x to z should be no more than the total distance of this detour through y. In other words, "distance"

ⁱIndeed, it's quite rare for someone to have come into contact with advanced mathematics without having encountered these ideas in a basic real analysis course.

should be measuring the shortest path. Of course, this has all been abstracted away, and there need not be any sensible notion of "path" anymore, but that is the motivation behind the triangle inequality. For the most part, we will have very little direct need of the triangle inequality beyond establishing the basic results about metric spaces. These results will have far reaching consequences which then rely on the inequality, but it is rarely if ever directly referenced later.

The notion of distance also gives the notion of convergence: a sequence that gets arbitrarily close to a point in the space (and it doesn't eventually go further way). For the discrete topology, this only happens with eventually constant sequences, as again, the only way to get closer than a distance of 1 to a point is to actually be that point. The notion of convergence can also be defined more generally for any topological space, but the use of a metric makes the idea more concrete.

— 21 A • 2. Definition -

Let $\langle X, d \rangle$ be a metric space. Let $\langle x_n \in X : n < \omega \rangle$ be a sequence.

• $\langle x_n : n < \omega \rangle$ converges iff there is a unique x (also called $\lim_{n\to\infty} x_n$) where for every $\varepsilon > 0$ in \mathbb{R} , $d(x_n, x) < \varepsilon$ for sufficiently large $n < \omega$. Equivalently,

 $\exists ! x \in X \ \forall \varepsilon \in \mathbb{R} \ (\varepsilon > 0 \to \exists N \in \omega \ \forall n \in \omega \ (n > N \to d(x, x_n) < \varepsilon)).$

• $\langle x_n : n < \omega \rangle$ is *cauchy* iff for every $\varepsilon > 0$ in \mathbb{R} , $d(x_n, x_m) < \varepsilon$ for sufficiently large $n, m < \omega$. Equivalently,

 $\forall \varepsilon \in \mathbb{R} \ (\varepsilon > 0 \to \exists N \in \omega \ \forall n, m \in \omega \ (n, m > N \to d(x_n, x_m) < \varepsilon)).$

• $\langle X, d \rangle$ is *complete* iff every cauchy sequence converges.

– 21 A · 3. Result

Let $\langle X, d \rangle$ be a metric space and let $\vec{x} \in {}^{\omega}X$ be a convergent sequence in $\langle X, d \rangle$. Therefore \vec{x} is cauchy.

Proof Sketch .:.

Let $\vec{x} \in {}^{\omega}X$ converge to x. Let $\varepsilon > 0$ be arbitrary and consider $\varepsilon/2 > 0$ with the definition of convergence: if for all n > N we get $d(x, x_n) < \varepsilon/2$, then in particular, for all n, m > N we get $d(x_n, x_m) \le d(x_n, x) + d(x, x_m) < \varepsilon/2 + \varepsilon/2 = \varepsilon$, showing \vec{x} is cauchy. \dashv

Moreover, for any metric space, we can consider the *completion* of it. In some sense, this means all cauchy sequences *should* converge, but might not because the metric space is "missing" some points. For example, consider the open interval $(0, 1) = \{x \in \mathbb{R} : 0 < x < 1\}$ with the usual metric d(x, y) = |x - y|. It's not hard to see that the sequence $\langle 1/n : n < \omega \rangle$ is cauchy, but doesn't converge in $\langle (0, 1), d \rangle$, precisely because the point it *should* converge to, 0, isn't an element of (0, 1).

21 A • 4. Theorem

- \mathbb{R} with the usual metric is complete.
- Every set *X* with the discrete metric is complete.
- Every metric space $\langle X, d_X \rangle$ has a completion $\langle Y, d_Y \rangle$, i.e. a complete metric space $\langle Y, d_Y \rangle$ such that $Y \supseteq X$ and $d_Y \upharpoonright X \times X = d_X$.

Proof Sketch .:.

That \mathbb{R} is complete follows from the fact that it is *defined* as the completion of \mathbb{Q} with its usual metric: d(x, y) = |x-y|. Every discrete space is complete as any cauchy sequence is easily seen to be eventually constant (consider $\varepsilon = 1$ in Definition 21 A • 2). That every metric space has a completion just follows by considering first

$$Cauchy(X) = \{\vec{x} \in {}^{\omega}X : \vec{x} \text{ is cauchy}\}\$$

We then mod out by the equivalence relation $\vec{x} \approx \vec{y}$ iff $\lim_{n\to\infty} d(x_n, y_n) = 0$ to get the resulting set of equivalence classes. We then define Y by replacing equivalence classes that converge in $\langle X, d \rangle$ with the point they converge to:

 $Y = X \cup \{ [\vec{x}]_{\approx} : \vec{x} \in \text{Cauchy}(X) \text{ doesn't converge in } \langle X, d \rangle \}.$
Relevant Topology

We can then define the metric

$$d_Y(x, y) = \begin{cases} d_X(x, y) & \text{if } x, y \in X \\ \lim_{n \to \infty} d_X(x, y_n) & \text{if } x \in X \text{ and } y \in Y \setminus X \\ \lim_{n \to \infty} d_X(x_n, y_n) & \text{if } x, y \in Y \setminus X. \end{cases}$$

A metric, more than just a function, lays the groundwork for understanding the shape of the space. To describe such things, we use topology.

— 21 A • 5. Definition -

A *topological space* is a pair (X, \mathcal{O}) where $\mathcal{O} \subseteq \mathcal{O}(X)$ is such that

1. $X, \emptyset \in \mathcal{O};$

2. $F \subseteq \mathcal{O}$ implies $\bigcup F \in \mathcal{O}$; and

3. $A, B \in \mathcal{O}$ implies $A \cap B \in \mathcal{O}$.

We call an $A \in \mathcal{O}$ open and call \mathcal{O} the collection of open sets of X. Moreover,

- A set A is *closed* iff $X \setminus A$ is open.
- For $x \in X$, a *neighborhood* of x is just an open set $A \in \mathcal{O}$ with $x \in A$.
- We also refer to the topological space as the *topology* on *X*.

For $F \subseteq \mathcal{O}(X)$, the topology generated by F is $\langle X, \mathcal{O} \rangle$ where \mathcal{O} is \subseteq -least where $F \subseteq \mathcal{O}$ such that $\langle X, \mathcal{O} \rangle$ is a topological space. If F is closed under finite intersections, we call such an F a basis and $A \in F$ a basic open set.

So the properties of open sets are just that the whole space and the empty set are open, the intersection of two (equivalently finitely many) open sets is open, and that the union of any collection of open sets is also open. Note that since $X \setminus \emptyset = X, X \setminus X = \emptyset \in \mathcal{O}$, this implies X and \emptyset both closed and open, sometimes called *clopen*. We have two main trivial examples of topological spaces, and a way of generating topological spaces more generally.

— 21 A • 6. Example

For every set X;

- the *discrete topology* $\langle X, \mathcal{P}(X) \rangle$ is a topological space;
- $\langle X, \{\emptyset, X\} \rangle$ is a topological space; and
- For $F \subseteq \mathcal{O}(X)$, taking \mathcal{O} to be the closure of $F \cup \{\emptyset\}$ under finite intersections and arbitrary unions yields $\langle X, \mathcal{O} \rangle$ as a topological space and is the topological space generated by F.

The concept of a basis is very useful, because we can represent open sets just as unions of basic open sets.

21A•7. **Result** Let F be a basis for a topology $\langle X, \mathcal{O} \rangle$. Therefore, $Y \subseteq X$ is open iff $Y = \bigcup A$ for some $A \subseteq F$.

Proof .:.

If $Y = \bigcup A$ for some $A \subseteq F$ then clearly Y is open: $A \subseteq F \subseteq O$ implies the union of open sets $\bigcup A = Y$ is open. So suppose $Y \subseteq X$ is open. Since the topology is generated by F, Y appears at some stage of closing F under unions and finite intersections. So proceed by transfinite induction on this stage. For stage $0, Y \in F$ and we're done. For stage $\alpha \neq 0$,

- Suppose $Y = \bigcup A$ where A is a family of sets appearing at previous stages. Inductively we can write each as the union of basic open sets, and therefore Y as this union.
- Suppose Y = A ∩ B where A, B are sets appearing at previous stages. Therefore, inductively we can write A = ∪ A' and B = ∪ B' where A', B' ⊆ F. In particular, Y = ∪_{S∈A',Z∈B'} S ∩ Z. Since F is closed under finite intersections, this is the union of basic open sets. ⊢

This general notion of a topological space isn't useful itself. Mostly it just provides a concept that can be used in other settings. For example, with \mathbb{R} , we can define open sets as being generated by open intervals, themselves the result of the fact that \mathbb{R} can be regarded as a metric space. In particular, we can consider the topology generated by open balls and by *rectangles* when taking products.

21 A • 8. Definition

- Let $\langle X, d \rangle$ be a metric space. The *topology induced by* d on X is the topological space $\langle X, \mathcal{O} \rangle$ where \mathcal{O} is generated by the family of open balls.
- For an ordinal κ and topological spaces $\langle X_{\alpha}, \mathcal{O}_{\alpha} \rangle$ for $\alpha < \kappa$, the *product space* is the topology on $\prod_{\alpha < \kappa} X_{\alpha}$ generated by open sets of the form $\prod_{\alpha < \kappa} W_{\alpha}$ where each $W_{\alpha} \in \mathcal{O}_{\alpha}$ all but finitely many W_{α} are the whole corresponding space, X_{α} .

We have some trivial examples and non-examples of metric spaces just as with the trivial examples of topological spaces. This hints that a metric space really does give information about the induced topology as not every topological space comes from a metric.

— 21 A • 9. Result

For every set *X*;

- 1. The discrete metric on X induces the discrete topology on X.
- 2. There is no metric where the induced topology is $\langle X, \{\emptyset, X\} \rangle$ unless $|X| \leq 1$.
- 3. Let (X, d) be a metric space. Therefore for $U \subseteq X$, U is open iff for every $x \in U$, there is some open ball B centered on x with $B \subseteq U$.
- 4. Let (X, d) be a metric space. Therefore $Y \subseteq X$ is closed iff every $\vec{x} \in {}^{\omega}Y$ that converges in (X, d) converges to a point in Y.

Proof .:.

The key point here is that we just need to consider open balls of sufficiently small radius.

- 1. d(x, y) is either 1 or 0, meaning that d(x, y) < 1 iff x = y. In particular, for each $x \in X$, the open ball $B_1(x) = \{x\}$. As a result, for \mathcal{O} the induced topology and for any $Y \subseteq X$, $\{\{y\} : y \in Y\} \subseteq \mathcal{O}$ implies the union of this, just Y itself, is also open: $\mathcal{O}(X) \subseteq \mathcal{O} \subseteq \mathcal{O}(X)$ yields equality.
- 2. If $|X| \ge 2$, then let $x, y \in X$ be two distinct elements. Let d(x, y) = r. Therefore, the open ball $B_r(x) \in \mathcal{O}$ yet $y \in X \setminus B_r(x)$ and $x \in B_r(x)$. Thus $\mathcal{O} \neq \{X, \emptyset\}$.
- 3. If $U = \emptyset$, this is trivial. So suppose $U \neq \emptyset$. For the \leftarrow direction, if every $x \in U$ has an open ball $B_{\varepsilon_x}(x) \subseteq U$ centered on x of radius ε_x , then $U = \bigcup_{x \in U} B_{\varepsilon_x}(x)$ is open as the union of basic open sets.

So suppose U is open. Proceed by structural induction to show U as the desired property. If U is basic open, then the result holds easily: $U = B_r(x)$ is the ball of radius r centered on some x. Hence $y \in U$ with $d(x, y) = \rho$ has $B_{r-\rho}(y) \subseteq B_r(x) = U$ by the triangle inequality:

 $z \in B_{r-\rho}(y)$ implies $d(x,z) \le d(x,y) + d(y,z) < \rho + (r-\rho) = r$.

If U if the union of a family of open sets for which the result holds, then clearly the result holds for U as well. If U is the intersection of two open sets, then we just take a sufficiently small radius: $x \in U \cap W$ has $B_{\varepsilon_0}(x) \subseteq U$ and $B_{\varepsilon_1}(x) \subseteq W$ for some $\varepsilon_0, \varepsilon_1 > 0$. So then $B_{\min(\varepsilon_0,\varepsilon_1)}(x) \subseteq U \cap W$. As all (non-empty) open sets are generated in this way, it follows that all open sets of the induced topology have this property.

4. Suppose $Y \subseteq X$ is closed, but \vec{x} converges to $x \in X \setminus Y$. Since $X \setminus Y$ is open, there is some $\varepsilon > 0$ where $B_{\varepsilon}(x) \subseteq X \setminus Y$. But then there can be no entry of $\vec{x} \in {}^{\omega}Y$ within ε of x, contradicting that \vec{x} converges to x.

Suppose every convergent sequence in Y converges to a point in Y. We need to show $X \setminus Y$ is open, so let $x \in X \setminus Y$. If there is no $0 < \varepsilon \in \mathbb{R}$ where $B_{\varepsilon}(x) \subseteq X \setminus Y$, then in particular, for every $n < \omega$, there is some $x_n \in B_{1/n}(x) \cap Y$. Hence $\langle x_n : n < \omega \rangle$ converges to x and is a subsequence of y. But this contradicts $x \notin Y$. Hence some ε_x has $B_{\varepsilon_x}(x) \subseteq X \setminus Y$ and therefore $X \setminus Y$ is the union of the open balls $\bigcup_{x \in X \setminus Y} B_{\varepsilon_x}(x)$ and is then open. \dashv

─ 21 A • 10. Definition

Let (X, \mathcal{O}) be a topological space. A subset $Y \subseteq X$ is *dense* in X iff $Y \cap U \neq \emptyset$ for every $U \in \mathcal{O}$. A topological space (X, \mathcal{O}) is *separable* iff it has a countable, dense subset.

– 21 A•11. Example

 \mathbb{R} with the standard topology induced by the standard metric is separable, as witnessed by $\mathbb{Q} \subseteq \mathbb{R}$.

Now we have the central definition for descriptive set theory of a polish space, which \mathbb{R} with its usual structure is an example of by Example 21 A • 11 and Theorem 21 A • 4.

- 21 A • 12. Definition

A metric space is *polish* iff it is separable and complete.

A topological space (X, \mathcal{O}) is *metrizable* iff there is a metric $d : X \times X \to \mathbb{R}$ that induces the same topology via open balls. In this case, d is called a *compatible* metric.

A topological space is *polish* iff it is metrizable, and polish as a metric space.

We also have several other examples, as seen below.

– 21 A•13. Example

- 1. The unit interval $[0, 1] \subseteq \mathbb{R}$ with the usual topology is polish.
- 2. The discrete topology on any $n \le \omega$ is a polish space.
- 3. If X and Y are polish spaces, then the product topology $X \times Y$ is polish. More generally, if X_n is polish for each $n < \omega$, then $\prod_{n \in \omega} X_n$ is polish.

As a result, the *baire space*ⁱⁱ, $\mathcal{N} = \prod_{n \in \omega} \langle \omega, \mathcal{P}(\omega) \rangle$ is polish; and so is the *cantor space*, $\mathcal{C} = \prod_{n \in \omega} \langle 2, \mathcal{P}(2) \rangle$.

Proof ...

(1) is clear. For (2), for $|X| \leq \aleph_0$ implies separability and the discrete metric induces the discrete topology.

To show (3), the result for products of polish spaces, note that for $d : X \times X \to \mathbb{R}$ a metric, we can consider $d' = \frac{d}{1+d}$. This gives the same topology, but d'(x, y) < 1 for all $x, y \in X$. This is useful, because we can now have some assurance that certain infinite sums will converge. Write \mathfrak{X} for $\prod_{n < \omega} X_n$, and so \mathfrak{X} for $\prod_{n < \omega} X_n$.

In particular, for $n \in \omega$, let d_n be a compatible metric on X_n such that $d_n(x, y) < 1$ for all $x, y \in X_n$. Define $d : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ as follows. It's easy to check that this will define a metric on \mathcal{X} : for $\vec{x} = \langle x_n \in X_n : n < \omega \rangle$ and $\vec{y} = \langle y_n \in X_n : n < \omega \rangle$,

$$d(\vec{x}, \vec{y}) = \sum_{n \in \omega} \frac{d_n(x_n, y_n)}{2^n}.$$

The metric space is complete, since if $\langle x_n : n \in \omega \rangle$ is cauchy, then each $\langle x_n(m) : m \in \omega \rangle$ is cauchy for $n \in \omega$ and so converges to some x(n). Thus $\langle x_n : n \in \omega \rangle$ converges to $x = \langle x(n) : n \in \omega \rangle$. So it suffices to show that the induced topology on \mathcal{X} from d, call this \mathcal{X}_d , is precisely the product topology \mathcal{X} . To do this, we need to show any subset is open in the metric sense iff it is open in the other sense.

We must show two things: for each $\vec{x} \in \mathcal{X}$,

a. For $N_{\vec{x}}$ a neighborhood of \vec{x} in \mathfrak{X} , there is a neighborhood $N'_{\vec{x}} \subseteq N_{\vec{x}}$ of \vec{x} in \mathfrak{X}_d .

b. For $N_{\vec{x}}$ a neighborhood of \vec{x} in \mathfrak{X}_d , there is a neighborhood $N'_{\vec{x}} \subseteq N_{\vec{x}}$ of \vec{x} in \mathfrak{X} .

In fact, it suffices to show the above for $N_{\vec{x}}$ as a basic open set. The idea is that if $Y \subseteq \mathcal{X}$ is open in \mathcal{X} , $Y = \bigcup_{\vec{x} \in Y} N_{\vec{x}}$ and (a) then implies $Y = \bigcup_{\vec{x} \in Y} N'_{\vec{x}}$ is open in \mathcal{X}_d . And the other direction is similar. So let $\vec{x} = \langle x_n \in X_n : n < \omega \rangle$ be arbitrary.

To show (a), basic open sets in \mathfrak{X} are rectangles $R = \prod_{n < N} U_n \times \prod_{N \le n < \omega} X_n$, where $N < \omega$ and U_n is open in X_n . Without loss of generality, since R is a neighborhood of \vec{x} and we can safely take subsets, assume each U_n is a ball centered on x_n with radius r_n . Hence for each $\vec{y} \in \mathfrak{X}$, $\vec{y} \in R$ iff $\forall n < N \ (d_n(x_n, y_n) < r_n)$. So let

ⁱⁱAs far as I know, the notation \mathcal{N} comes from thinking of \mathcal{N} as the second order analogue of $\mathbf{N} = \langle \omega, 0, 1, +, \cdot \rangle$, meaning we can think of \mathcal{N} as essentially a two-sorted model of arithmetic where we have access to $\omega \omega$ in addition to ω . Clearly the notation of \mathbf{N} is motivated from the usual mathematical notation of $\mathbb{N} = \omega$.

Relevant Topology

CH IV §21 A

 $r = \min_{n < N} \frac{r_n}{2^n}$. Consider the open ball

$$B = \{ \vec{y} \in \mathcal{X} : d(\vec{x}, \vec{y}) < r \}.$$

If $\vec{y} \in B$, then $d(\vec{x}, \vec{y}) < r$. Note that if $d_n(x_n, y_n) \ge r_n$ for any n < N, then $d(\vec{x}, \vec{y}) \ge \frac{r_n}{2^n}$. Therefore $d_n(x_n, y_n) < r_n$ for each n < N, and thus $\vec{y} \in R$.

To show (b), let $B = B(\vec{x}, r) = {\vec{y} \in \mathcal{X} : d(\vec{x}, \vec{y}) < r}$ be the ball of radius *r* centered at \vec{x} . Let *N* be large enough such that $\sum_{n \ge N} \frac{1}{2^n} = \frac{1}{2^{N-1}} < r$, which means $\frac{1}{2^N} < \frac{r}{2}$. Consider now the open set

$$R = \prod_{n < N} B_n(x_n, r/4) \times \prod_{N \le n < \omega} X_n.$$

We will show that $R \subseteq B$ and (b) holds Let $\vec{y} \in R$. Note that

$$d(\vec{x}, \vec{y}) = \sum_{n < N} \frac{d_n(x_n, y_n)}{2^n} + \sum_{n \ge N} \frac{d_n(x_n, y_n)}{2^n} \le \sum_{n < N} \frac{r/4}{2^n} + \sum_{n \ge N} \frac{1}{2^n} \le \frac{r}{2} \left(1 - \frac{1}{2^{N-1}} \right) + \frac{1}{2^N} < \frac{r}{2} + \frac{r}{2} = r \quad \Rightarrow \quad \vec{y} \in R.$$

As metric spaces can also play nicely with topology, it follows that being polish does too.

- 21 A • 14. Result Let $\mathcal{M} = \langle \mathcal{M}, \mathcal{O}_{\mathcal{M}} \rangle$ be polish and let $X \subseteq \mathcal{M}$ be closed. Therefore $\mathbf{X} = \langle X, \{X \cap U : U \in \mathcal{O}_{\mathcal{M}}\} \rangle$, i.e. the topology on X inhereted from \mathcal{M} , is also polish. Moreover, this topology is induced by the same metric (restricted to X).

Proof .:.

It's not difficult to see that the topology on X inhereted from \mathcal{M} is given by $\{X \cap U : U \in \mathcal{O}_{\mathcal{M}}\}$ by the following observation. If d is a compatible metric for \mathcal{M} , every open set is the union of open balls of the form $B_{\varepsilon}(x) = \{y \in \mathcal{M} : d(x, y) < \varepsilon\}$ for some $0 < \varepsilon \in \mathbb{R}$ and $x \in \mathcal{M}$. Hence the restricted metric $d \upharpoonright X^2$ gives the open ball of radius ε centered on x as just $\{y \in X : d(x, y) < \varepsilon\} = B_{\varepsilon}(x) \cap X$. As intersections distribute over unions, this gives the result.

Let $Y \subseteq \mathcal{M}$ be countable and dense so that $Y \cap X$ is a countable, dense subset of X. To see that X is complete, let d be a compatible metric for \mathcal{M} . It's not difficult to see that $d \upharpoonright X \times X$ is a compatible metric for X. So suppose $\langle x_n \in X : n < \omega \rangle$ is cauchy. Therefore this converges in \mathcal{M} to some $x \in \mathcal{M}$. It then suffices to show $x \in X$. To see this, we use that X is closed: suppose $x \in \mathcal{M} \setminus X$, an open set. Therefore, there is some open ball $B = \{y \in \mathcal{M} : d(x, y) < r\} \subseteq \mathcal{M} \setminus X$. But this prevents $\langle x_n \in X : n < \omega \rangle$ from converging to x as $\varepsilon = r$ has that eventually an $n < \omega$ has $d(x_n, x) < r$ and therefore $x_n \in X \cap B \subseteq \mathcal{M} \setminus X$, a contradiction. Thus $x \in X$. \dashv

We cannot do exactly the same for open subsets, since using the same metric may not yield a complete metric space as with closed sets: (0, 1) is an open subset of [0, 1] with the standard metric, but this metric isn't complete on (0, 1). Nevertheless, we can find a *different* metric that yields the subspace as polish. This also tells us that being a compatible metric doesn't imply that the metric space is complete.

- 21 A • 15. Result
Let
$$\mathcal{M} = \langle \mathcal{M}, \mathcal{O}_{\mathcal{M}} \rangle$$
 be polish and let $X \subseteq \mathcal{M}$ be open. Therefore $\mathbf{X} = \langle X, \{X \cap U : U \in \mathcal{O}_{\mathcal{M}}\} \rangle$ is also polish.

Proof .:.

Let *d* be a complete, compatible metric on \mathcal{M} . For $x \in \mathcal{M}$, write D(x) for $\inf\{d(x, z) : z \notin X\}$. Now define the metric $d' : X \times X \to \mathbb{R}$ by

$$d'(x, y) = d(x, y) + \frac{|D(x) - D(y)|}{D(x) \cdot D(y)} = d(x, y) + \left|\frac{1}{D(x)} - \frac{1}{D(y)}\right|$$

The idea here is that for a fixed x, as y gets closer to $\mathcal{M} \setminus X$, d'(x, y) goes to infinity in the same way we might identify the entire real line with instead just (0, 1).

 $\langle X, d' \rangle$ is a complete metric space.

Proof .:.

It's clear that $d'(x, y) \ge d(x, y) \ge 0$ for all $x, y \in X$. Moreover, x = y implies d'(x, y) = 0. The other direction is also easy as the sum of two non-negative real numbers is 0 iff both are 0. For the triangle inequality,

$$d'(x, y) + d'(y, z) = d(x, y) + d(y, z) + \left| \frac{1}{D(x)} - \frac{1}{D(y)} \right| + \left| \frac{1}{D(y)} - \frac{1}{D(z)} \right|$$

$$\geq d(x, z) + \left| \frac{1}{D(x)} - \frac{1}{D(y)} + \frac{1}{D(y)} - \frac{1}{D(z)} \right|$$

$$\geq d(x, z) + \left| \frac{1}{D(x)} - \frac{1}{D(z)} \right| = d'(x, z).$$

Hence d' is a metric. To see that $\langle X, d' \rangle$ is complete, suppose $\vec{x} = \langle x_n \in X : n < \omega \rangle$ is cauchy. Again, since $d'(x, y) \ge d(x, y)$ this means \vec{x} is cauchy with respect to d and hence converges in \mathcal{M} to some $x \in \mathcal{M}$. Fix $\varepsilon > 0$. As a cauchy sequence for d', we have not only that $d(x_n, x_m) < \varepsilon/2$ for sufficiently large $n, m < \omega$, but also $\left| \frac{1}{D(x_n)} - \frac{1}{D(x_m)} \right| < \varepsilon/2$ for sufficiently large $n, m < \omega$ and so $\langle \frac{1}{D(x_n)} : n < \omega \rangle$ is cauchy in \mathbb{R} with its standard metric. Therefore this sequence converges in \mathbb{R} and in particular, $D(x) = \lim_{n \to \infty} D(x_n) \neq 0$, i.e. $x \in X$.

So now it suffices to show d' gives the same topology on X as $\mathcal{O} = \{X \cap U : U \in \mathcal{O}_{\mathcal{M}}\}$. Let \mathcal{O}' be the topology on X induced by d'. Write $B_{\varepsilon}(x)$ for the open ball of radius ε around x with respect to d, and similarly $B'_{\varepsilon}(x)$ is the same for d'. We aim to show $\mathcal{O}' = \mathcal{O}$.

(⊆) Suppose $A \in \mathcal{O}'$ so that $A \subseteq X$ is the union of open balls (with respect to d'): $A = \bigcup_{x \in A} B'_{\varepsilon_x}(x)$ where $0 < \varepsilon_x \in \mathbb{R}$ is such that $B'_{\varepsilon_x}(x) \subseteq A$ for each $x \in A$. It then suffices to show each $B'_{\varepsilon}(x) \in \mathcal{O}$. To do this, we want every $y \in B'_{\varepsilon}(x)$ to have an ε_y where $B_{\varepsilon_y}(y) \subseteq B'_{\varepsilon}(x) \subseteq X$. This would imply $B'_{\varepsilon}(x) = \bigcup_{y \in B'_{\varepsilon}(x)} B_{\varepsilon_y}(y) \in \mathcal{O}$.

Let $x \in X$ and $\varepsilon > 0$ be arbitrary. For any $\eta > 0$, if $y \in B_{\eta}(x)$ then $D(x) - \eta \le D(y) \le D(x) + \eta$. In particular, $d(x, y) < \eta$ implies

$$d'(x, y) = d(x, y) + \frac{|D(x) - D(y)|}{D(x) \cdot D(y)} \le \eta + \frac{\eta}{D(x) \cdot (D(x) - \eta)}.$$
(*)

Since $\lim_{\eta\to 0} (\eta + \frac{\eta}{D(x) \cdot (D(x) - \eta)}) = 0$, let ε_y be sufficiently small such that $\varepsilon_y + \frac{\varepsilon_y}{D(x) \cdot (D(x) - \varepsilon_y)} < \varepsilon$. Therefore $y \in B_{\varepsilon_y}(x)$ implies $y \in B'_{\varepsilon}(x)$. Thus $B'_{\varepsilon}(x) = \bigcup_{y \in B'_{\varepsilon}(x)} B_{\varepsilon_y}(y) \in \mathcal{O}$.

(⊇) Suppose $A = X \cap U \in \mathcal{O}$ for $U \in \mathcal{O}_{\mathcal{M}}$. We have $U = \bigcup_{x \in U} B_{\varepsilon_x}(x)$ for $0 < \varepsilon_x \in \mathbb{R}$ such that $B_{\varepsilon_x}(x) \subseteq U$. Hence $A = \bigcup_{x \in U} X \cap B_{\varepsilon_x}(x)$ and it suffices to show $X \cap B_{\varepsilon}(x) \in \mathcal{O}'$ for each $x \in \mathcal{M}$ and $0 < \varepsilon \in \mathbb{R}$. So let $y \in X \cap B_{\varepsilon}(x)$. As an open set, there is some ε_y with $B_{\varepsilon_y}(y) \subseteq X \cap B_{\varepsilon}(x)$. Since $d'(a, b) \ge d(a, b)$ for all $a, b \in X$, it follows that $B'_{\varepsilon_y}(y) \subseteq B_{\varepsilon_y}(y)$ and therefore $X \cap B_{\varepsilon}(x) = \bigcup_{y \in X \cap B_{\varepsilon}(x)} B_{\varepsilon_y}(y) = \bigcup_{y \in X \cap B_{\varepsilon}(x)} B'_{\varepsilon_y}(x) \in \mathcal{O}'$. ⊢

§ 21 B. The main takeaways

We will mainly work with spaces like \mathcal{N} and products of \mathcal{N} . But we also might identify ω with, say, $\omega \times \omega$. Note the following property which is basically stating the definition of a product space.

- 21 B • 1. Result Let \mathcal{M} be a product of copies of \mathcal{N} . Therefore if $X_0, X_1, \dots, X_n \subseteq \mathcal{N}, n < \omega$, are open in \mathcal{N} then the product of $X_0 \times X_1 \times \dots \times X_n$ (with perhaps a bunch of copies of \mathcal{N}) is open in \mathcal{M} . Similarly, the product topology on ${}^{<\omega}\omega$ is the discrete topology, and thus we can consider basic open sets as just singletons.

Now we prove some easy results about the baire space. For the most part, we will just consider products of the baire space. The following also easily generalizes to cantor space, \mathcal{C} . First we introduce some basic notation. We say τ is an *initial segment* of π , written $\tau \leq \pi$, iff τ and π are functions with dom (τ) , dom $(\pi) \in$ Ord—i.e. they are sequences—and $\tau \subseteq \pi$. As sequences, we write lh (τ) for the *length* of τ (which is just the domain).

- 21 B • 2. Result - For $\tau \in {}^{<\omega}\omega$, let $\mathcal{N}_{\tau} = \{x \in \mathcal{N} : \tau \leq x\}$. Therefore each \mathcal{N}_{τ} is open, and the collection of cones is a basis for \mathcal{N}_{τ} .

Proof .:.

Recall that \mathcal{N} is just the product topology on $\prod_{n \in \omega} \omega$, where the topology on ω is discrete. The basic open sets in \mathcal{N} are then rectangles that can be broken up into finitely many non-trivial open rectangles and co-finitely many copies of the whole space: $R = \prod_{n < N} U_n \times \prod_{N \leq n < \omega} \omega$, where $N < \omega$ and each $U_n \subseteq \omega$ (every U_n works since all subsets are open). Clearly each cone is open in this topology as we just consider $\mathcal{N}_{\tau} = \prod_{n < h(\tau)} \{\tau_n\} \times \prod_{h(\tau) \leq n < \omega} \omega$.

So it suffices to show that any rectangle $R = \prod_{n < N} U_n \times \prod_{N \le n < \omega} \omega$ can be written as a union of cones. If we consider $T = \{\tau \in {}^N \omega : \exists x \in R \ (\tau \le x)\}$, then $R = \bigcup_{\tau \in T} \mathcal{N}_{\tau}$ yields the result. \dashv

The general picture one should have in their head is the following figure.



21 B • 3. Figure: Cones in baire space

The good thing about all of these products being polish is that it's easy to characterize and check whether a given function is continuous. We will often need continuous functions between product spaces in our arguments, and want an easy way to check whether these functions are actually continuous without resorting back to arguments regarding rectangles of open sets.

```
- 21 B • 4. Corollary 
Let \mathcal{M} be polish. Therefore a function f : \mathcal{N} \to \mathcal{M} is continuous iff for any x \in \mathcal{N}, \bigcap_{n < \omega} f "\mathcal{N}_{x \upharpoonright n} = \{f(x)\}.
```

This is really just characterizing continuity in terms of the complete metric instead of open sets. Note that \mathcal{N} is particularly important for us because of the following two results.

- 21 B • 5. Theorem

Let \mathcal{M} be a polish space. Therefore, there is some continuous surjection $f: \mathcal{N} \to \mathcal{M}$.

Proof .:.

As \mathcal{M} is separable, there is some countable dense set $Q \subseteq \mathcal{M}$ which can be enumerated as $\{q_n : n < \omega\}$. Also let d be a compatible metric. For each $r \in \mathcal{N}$, we define the converging sequence $r^* \in {}^{\omega}Q$ by recursion. Firstly, set $r_0^* = q_{r(0)}$ and then take r_{n+1}^* to be $q_{r(n+1)}$ if this is close enough: if $d(r_n^*, q_{r(n+1)}) < 1/2^n$. Otherwise, we stay put: $r_{n+1}^* = r_n^*$. It's easy to see that this sequence is then cauchy and so as \mathcal{M} is complete, it has a limit $\lim r^* \in \mathcal{M}$. So take $f : \mathcal{N} \to \mathcal{M}$ to be the map $f(r) = \lim r^*$. It's easy to see that this map is continuous since if r' agrees with r up to n, i.e. $r \upharpoonright n = r' \upharpoonright n$, then $(r')^*$ agrees with r^* up to stage n and hence the limits f(r) and f(r') have a distance less than $\frac{1}{2^n} + \frac{1}{2^n} = 1/2^{n-1}$.

To see that f is surjective, because Q is dense, for any $p \in \mathcal{M}$, we can consider the function $p' : \omega \to \omega$ where p'(n) is the least such that $d(p, q_{p'(n)}) < 2^{n+1}$. This $p' \in \mathcal{N}$ then clearly has f(p') = p. \dashv

The only obstacle in containing a copy of \mathcal{N} is if the space looks like ω . More precisely, the basic open sets of \mathcal{N} all have size $|\mathcal{N}| = 2^{\aleph_0}$, meaning there are always lots of points surrounding any given real in \mathcal{N} . So in a map $f : \mathcal{N} \to \mathcal{M}$, we shouldn't map $x \in \mathcal{N}$ to a point in \mathcal{M} that isn't surrounded by lots of points too.

🖵 21 B•6. Definition —

Let \mathcal{M} be a topological space. A point $m \in \mathcal{M}$ is *isolated* iff $\{m\}$ is open in \mathcal{M} .

For a metric space, this just means that as we zoom in on *m*, eventually we notice that there are no elements around: for some $\varepsilon > 0$, the only element within ε of *m* is *m* itself. It's not difficult to see that \mathcal{N} has no isolated points whereas $\langle \omega, \mathcal{P}(\omega) \rangle$ consists *only* of isolated points.

21B•7. Theorem Let \mathcal{M} be a polish space with no isolated points. Therefore, there is some continuous injection $f : \mathcal{N} \to \mathcal{M}$.

Proof .:.

First we will show that there is a continuous injection $nc : \mathcal{N} \to \mathcal{C}$ from baire space to cantor space. Then we will show there is a continuous injection from cantorspace into \mathcal{M} , and so we reach the conclusion just by composition.

For $n < \omega$, let 0_n be the constant sequence of 0s of length n. So for $x = \langle x_n : n < \omega \rangle \in \mathcal{N}$, let nc(x) be the sequence

$$\langle 1 \rangle^{\frown} 0_{x_0}^{\frown} \langle 1 \rangle^{\frown} 0_{x_1}^{\frown} \langle 1 \rangle^{\frown} \cdots^{\frown} \langle 1 \rangle^{\frown} 0_{x_n}^{\frown} \langle 1 \rangle^{\frown} \cdots$$

This map is continuous since if we've determined the first N entries of x, then we've determined the first $\sum_{n < N} x_n \ge N$ entries of nc(x). It should be clear that this map is injective. So there is a continuous injection from baire space to cantor space. It then suffices to find a continuous injection from cantor space to \mathcal{M} .

Let *d* be a compatible metric for \mathcal{M} . Without loss of generality, we can assume $d(x, y) \leq 1$ for all $x, y \in \mathcal{M}$. We identify $\tau \in {}^{<\omega}2$ with neighborhoods of \mathcal{M} . In particular, define open sets $\mathcal{M}_{\tau} \subseteq \mathcal{M}$ by recursion. Firstly, define $\mathcal{M}_{\emptyset} = \mathcal{M}$, and for \mathcal{M}_{τ} already defined, as no points are isolated, let $x, y \in \mathcal{M}_{\tau}$ be two distinct points. Let B_0 be the ball around x of radius $r_0 < d(x, y)/2^{\ln(\tau)+1}$ while B_1 is the ball around y of radius $r_1 < d(x, y)/2^{\ln(\tau)}$. Then define $\mathcal{M}_{\tau} \frown _{(0)} = B_0 \cap \mathcal{M}_{\tau}$ and $\mathcal{M}_{\tau} \frown _{(1)} = B_1 \cap \mathcal{M}_{\tau}$.

We thus have $\mathcal{M}_{\tau} \cap \mathcal{M}_{\sigma} \emptyset$ iff $\tau \leq \sigma$ or $\sigma \leq \tau$. Also, by the restriction on the sizes of the balls, for any $x \in \mathcal{N}$, any sequence of points $\langle p_n \in \mathcal{M}_{x \upharpoonright n} : n < \omega \rangle$ is cauchy and thus converges to some (unique) point $p_x \in \mathcal{M}$. If we define $f : \mathcal{N} \to \mathcal{M}$ by $f(x) = p_x \in \bigcap_{n < \omega} \mathcal{M}_{x \upharpoonright n}$, then we have that f is injective: $x \neq y$ implies $x \upharpoonright n \neq y \upharpoonright n$ for some $n < \omega$ where then $p_x \notin \mathcal{M}_{y \upharpoonright n}$ while $p_y \in \mathcal{M}_{y \upharpoonright n}$.

To see that f is continuous, just note that if y agrees with x up to n, $y \upharpoonright n = x \upharpoonright n$, then $f(y) \in \mathcal{M}_{y \upharpoonright n} = \mathcal{M}_{x \upharpoonright n}$ meaning $d(f(y), f(x)) < 1/2^n$.

As a result, we have restrictions around the size of a polish space and so the resulting topology.

– 21 B•8. Corollary

Let \mathcal{M} be a polish space. Therefore, either $|\mathcal{M}| \leq \aleph_0$, or else $|\mathcal{M}| = 2^{\aleph_0}$. Moreover, if \mathcal{M} is countable, then \mathcal{M} is the discrete topology on \mathcal{M} .

Proof .:.

We know $|\mathcal{M}| \leq 2^{\aleph_0}$ by Theorem 21 B • 5. There can be at most countably many isolated points of \mathcal{M} , as otherwise \mathcal{M} won't be separable. Let $I \subseteq \mathcal{M}$ be the set of isolated points, which is then open and hence removing these isolated points, $\mathcal{M} \setminus I$, we have a closed subset of \mathcal{M} . By continuing this removal of isolated points (at most countably many times), we can assume $\mathcal{M} \setminus I$ has no isolated points.

There is a countable set
$$J \subseteq \mathcal{M}$$
 such that $\mathcal{M} \setminus J$ has no isolated points.

Proof .:.

To formalize "continually removing isolated points", for $X \subseteq M$, let X' be $X \setminus \{x \in X : x \text{ is isolated}\}$. Then we define by recursion

$$\mathcal{M}_0 = \mathcal{M}, \qquad \mathcal{M}_{\alpha+1} = \mathcal{M}'_{\alpha}, \qquad \text{and} \quad \mathcal{M}_{\alpha} = \bigcap_{\xi < \alpha} \mathcal{M}_{\xi} \quad \text{for limit } \alpha.$$

Note that inductively, if X is a polish space, then X' is closed and hence another polish space with the restricted topology. Since the intersection of closed sets is closed, it follows that each \mathcal{M}_{α} is closed.

We now show that at some point this process stabilizes and we get $\mathcal{M}_{\alpha} = \mathcal{M}_{\beta}$ for sufficiently large α, β . Specifically, we will get that this holds after some countable ordinal. To show this, let $U = \{U_n : n \in \omega\}$ be a countable basis for \mathcal{M} since it's separable. We can identify any closed set $X \subseteq \mathcal{M}$ with its basic neighborhoods $N_X = \{U_n : U_n \cap X \neq \emptyset\}$ in that $x \in X$ iff $x \notin \bigcup (U \setminus N_X)$. (To see this, $\mathcal{M} \setminus X$ is open so $x \in \mathcal{M} \setminus X$ iff x is in some U_n disjoint from X.) As a result, any two closed sets $X \neq Y$ have $N_X \neq N_Y$ and it's not difficult to see $X \subseteq Y$ implies $N_X \supseteq N_Y$.

Since the sequence of \mathcal{M}_{α} s is \subseteq -decreasing, the sequence of $N_{\mathcal{M}_{\alpha}}$ s is \subseteq -increasing. But since each $N_{\mathcal{M}_{\alpha}}$ is contained in the countable set U, we cannot have an uncountable strictly \subsetneq -increasing sequence of $N_{\mathcal{M}_{\alpha}}$. Thus $N_{\mathcal{M}_{\alpha}} = N_{\mathcal{M}_{\alpha+1}}$ for some $\alpha < \omega_1$ which then tells us $\mathcal{M}_{\alpha} = \mathcal{M}_{\alpha+1}$ and therefore $\mathcal{M}_{\alpha} = \mathcal{M}_{\beta}$ for all $\beta > \alpha$.

Define recursively $J_{\beta} = \mathcal{M} \setminus \mathcal{M}_{\beta}$. Clearly $J_0 = \mathcal{M} \setminus \mathcal{M}_0 = \emptyset$ is countable. If J_{β} is countable, then $J_{\beta+1} = J_{\beta} \cup \mathcal{M}_{\beta} \setminus \mathcal{M}_{\beta+1}$ is also countable since $\mathcal{M}_{\beta} \setminus \mathcal{M}_{\beta+1}$ is countable by separability. At limit stages $\beta < \omega_1, J_{\beta} = \bigcup_{\xi < \beta} J_{\xi}$ which is the countable union of countable sets and is thus countable. It follows that $J = J_{\alpha} = \mathcal{M} \setminus \mathcal{M}_{\alpha}$ is countable. Since $\mathcal{M}_{\alpha} = \mathcal{M}_{\alpha+1}, \mathcal{M} \setminus J = \mathcal{M}_{\alpha}$ has no isolated points. \dashv

If $\mathcal{M} \setminus J \neq \emptyset$, then restricting the topology of \mathcal{M} to $\mathcal{M} \setminus J$ yields another polish space with no isolated points (it's clearly separable still, and the metric is still complete as $\mathcal{M} \setminus J$ closed). Hence by Theorem 21 B • 7, $2^{\aleph_0} \leq |\mathcal{M} \setminus J|$ and therefore $2^{\aleph_0} \geq |\mathcal{M}| \geq 2^{\aleph_0} + \aleph_0 = 2^{\aleph_0}$.

The above idea supports the idea that polish spaces can basically be thought of as just copies of \mathcal{N} with countably many isolated points. The topologies *can* be different, however, as we will see. Indeed, even cantor space and baire space differ fundamentally in that one is *compact* while the other isn't.

– 21 B•9. Definition

```
A topological space (X, \mathcal{O}) is compact iff for any A \subseteq \mathcal{O}, X = \bigcup A implies X = \bigcup \Delta for some finite \Delta \subseteq A.
```

 $- 21B \cdot 10. \text{ Result}$ $\mathcal{L} \text{ is compact while } \mathcal{N} \text{ isn't.}$

Proof .:.

It's easy to see that \mathcal{N} isn't compact, since $\bigcup_{n < \omega} \mathcal{N}_{\langle n \rangle} = \mathcal{N}$, but there is no $N < \omega$ where $\bigcup_{n < N} \mathcal{N}_{\langle n \rangle} = \mathcal{N}$.

To see that \underline{C} is compact, write C_{τ} for the cone $\{x \in C : \tau \triangleleft x\}$ for $\tau \in {}^{<\omega}2$. For $A \subseteq {}^{<\omega}2$, write C_A for $\bigcup_{\tau \in A} C_{\tau}$. Suppose C isn't compact. Because open sets for \underline{C} are unions of cones, we can then assume $C = C_A$ for some $A \subseteq {}^{<\omega}2$, but for any finite $\Delta \subseteq A$, $C \neq C_{\Delta}$.

Note that since each cone is open *and* closed, C_{Δ} (as the finite union of closed sets) is also closed whenever $\Delta \subseteq A$ is finite. Hence if $x \notin C_{\Delta}$, there is a $\sigma \in A$ with $\sigma \lhd x$ with $C_{\sigma} \cap C_{\Delta} = \emptyset$. This motiviates the following tree where branches attempt to build an element of $C \setminus C_A = \emptyset$ according to the fact that $C \setminus C_{\Delta}$ is never \emptyset for

finite $\Delta \subseteq A$. Consider the tree

 $T = \{ \langle \tau, \Delta \rangle : \mathcal{C}_{\tau} \cap \mathcal{C}_{\Delta} = \emptyset \land \tau \in A \land \Delta \subseteq A \land \Delta \subseteq \{ \sigma \lhd \tau : \sigma \in A \} \},$

ordered by \triangleleft : $\langle \tau, \Delta \rangle$ preceeds $\langle \sigma, \Sigma \rangle$ iff $\tau \triangleleft \sigma$ and $\Delta \cup \{\tau\} \subseteq \Sigma$. By the argument above, any $\langle \tau, \Delta \rangle$ has an extension $\langle \sigma, \Delta \cup \{\tau\} \rangle$ in *T* (with then $\tau \triangleleft \sigma$) and so *T* is infinite. Since the levels of *T* are finite (any level of ${}^{<\omega}2$ is finite, and there can be only finitely many subsets of the finite set { $\sigma \in A : \sigma \triangleleft \tau$ } for $\tau \in {}^{<\omega}2$), it follows from König's Lemma on Trees (9 B • 5) that *T* has an infinite branch $B \subseteq T$.

Set $x = \bigcup_{\langle \tau, \Delta \rangle \in B} \tau \in C$. Such an x cannot exist: for $\langle x \upharpoonright n, \Delta \rangle \in B$, we have some extension $\langle x \upharpoonright n, \Delta \rangle \leq \langle x \upharpoonright m, \Sigma \rangle \in B$ for some $m \ge n$ and $\Sigma \supseteq \Delta \cup \{\tau\}$. As an extension, $C_{x \upharpoonright m} \cap C_{\Delta \cup \{x \upharpoonright n\}} = \emptyset$. But by definition of $x, x \in C_{x \upharpoonright m} \cap C_{\Delta \cup \{x \upharpoonright n\}}$, a contradiction. Hence no such x can exist, and therefore T can't be infinite, and therefore some Δ must have $C_{\Delta} = C$. As A was arbitrary, \mathcal{C} is compact.

Section 22. The Boldface Hierarchies

Rather than work towards the formal presentation of \mathbb{R} , we will work with a space that is homeomorphicⁱⁱⁱ to the set of irrational numbers. We restate the topological characterization gleamed from Result 21 B • 2.

- 22 · 1. Definition

The baire space \mathcal{N} is the product topology on $\prod_{n \in \omega} \langle \omega, \mathcal{P}(\omega) \rangle$. More precisely, $\mathcal{N} = {}^{\omega}\omega$, and a set $A \subseteq \mathcal{N}$ is open iff A is the union of cones: for $\tau \in {}^{<\omega}\omega$, the cone for τ is the set $\mathcal{N}_{\tau} = \{x \in {}^{\omega}\omega : \tau \subseteq x\}$.

As a result, we can see that $A \subseteq \mathcal{N}$ is open iff every $x \in A$ has some initial segment $\tau \triangleleft x$ with $\mathcal{N}_{\tau} \subseteq A$. So A being closed—meaning $\mathcal{N} \setminus A$ is open—just says that for each $x \notin A$, there is some initial segment $\tau \triangleleft x$ with $\mathcal{N}_{\tau} \subseteq \mathcal{N} \setminus A$, i.e. $\mathcal{N}_{\tau} \cap A = \emptyset$. Note that $\mathcal{N} = \mathcal{N}_{\emptyset}$ is then both closed and open.

Disregarding the formality about what exactly \mathbb{R} *is* as a set^{iv}, we get the following result. The connection is that two reals are close if they agree on a large initial segment. So this is why our topology looks the way it does: cones \mathcal{N}_{τ} for $\tau \leq x$ correspond to open balls around x in the sense that we're looking at points that are sufficiently close to x by sharing the initial segment τ . Smaller balls correspond to larger initial segments, and so on.

- 22.2. Theorem \mathcal{N} is homeomorphic to the space of real, irrational numbers, i.e. $\mathbb{R} \setminus \mathbb{Q}$, with the standard topology induced by the usual metric d(x, y) = |x - y|.

The proof of this theorem isn't relevant for us, but it does show the connection between the usual interpretation of \mathbb{R} and the sequences of natural numbers that descriptive set theory considers. With this correspondence, we will often call $x \in \mathcal{N}$ a "real". And so using the increasing enumeration (repeating the last element forever if finite), we identify subsets of ω with reals.

There are two fundamental concepts we will look at: complexity and trees. We have already introduced the open sets which will eventually give rise to the borel hierarchy and eventually the projective and analytical hierarchies. Trees will be a major way that we examine the projective and analytical hierarchies. To hint at the connection, we have the following definition and result.

— 22•3. Definition

Let X be a set. $T \subseteq {}^{<\omega}X$ is a *tree* over X iff T is closed under initial segments: $\sigma \triangleleft \tau \in T$ implies $\sigma \in T$. So $\langle T, \triangleleft \rangle$ is a tree in the usual mathematical sense.

For T a tree over X, the set of infinite branches of T is denoted $[T] = \{x \in {}^{\omega}X : \forall n < \omega \ (x \upharpoonright n \in T)\}.$

- 22 • 4. Result

A set $A \subseteq \mathcal{N}$ is closed iff there is some tree T over ω where A = [T].

Proof .:.

So suppose T is an arbitrary tree over ω . To show that A = [T] is closed, suppose $x \in \mathcal{N} \setminus [T]$. Because x is not an infinite branch of T, some initial segment $\tau \triangleleft x$ has $\tau \notin T$. But then no extension of τ is in T meaning $\mathcal{N}_{\tau} \cap [T] = \emptyset$, i.e. $\mathcal{N}_{\tau} \subseteq \mathcal{N} \setminus [T]$. Thus $\mathcal{N} \setminus [T]$ is open so A = [T] is closed.

Let A be a closed set. Consider T as the set of initial segments of elements in A:

 $T = \{ \tau \in {}^{<\omega}\omega : \exists x \in A \ (\tau \triangleleft x) \} = \{ \tau \in {}^{<\omega}\omega : \mathcal{N}_{\tau} \cap A \neq \emptyset \}.$

ⁱⁱⁱTwo topologies (X_1, N_1) and (X_2, N_2) are homeomorphic if there is a function $f : X_1 \to X_2$ such that for any $A \subseteq X_1, A \in N_1$ iff $f^{"}A \in N_2$, meaning A is open iff $f^{"}A$ is open.

^{iv}We think of \mathbb{R} as the canonical completion of \mathbb{Q} under Cauchy sequences with the usual metric. There are various ways to define this in set theory—usually using lots and lots of equivalence classes—but it's unimportant for our purposes.

It should be obvious that *T* is a tree, so it suffices to show that A = [T]. Clearly $A \subseteq [T]$, as $x \in A$ implies every initial segment is in *T* by definition, which then implies $x \in [T]$. To show $[T] \subseteq A$, we use that *A* is closed. Suppose $x \in [T]$ but $x \notin A$. As $\mathcal{N} \setminus A$ is open, there is some neighborhood $x \in \mathcal{N}_{\tau} \subseteq \mathcal{N} \setminus A$ so that $\mathcal{N}_{\tau} \cap A = \emptyset$. But $\tau \in T$ as an initial segment of $x \in [T]$, so $\mathcal{N}_{\tau} \cap A \neq \emptyset$, a contradiction.

§ 22 A. The borel hierarchy

Our first hierarchy to consider is the hierarchy of borel sets, named after Émile Borel. In general, for any collection of subsets $S \subseteq \mathcal{O}(X)$, we can close S under countable intersections, countable unions, and complements. This forms a σ -algebra, and we get a hierarchy describing how any particular set is built up from S. Note that we only need countable unions, since complements give us closure under countable intersections: $X \setminus \bigcup_{n < \omega} (X \setminus A_n) = \bigcap_{n < \omega} A_n$.

- 22A • 1. Definition For X a set, $S \subseteq \mathcal{P}(X)$ is a σ -algebra over X iff for all A, • $A \in S$ implies $X \setminus A \in S$; and • $\{A_n : n < \omega\} \subseteq S$ implies $\bigcup_{n < \omega} A_n \in S$.

We will be interested mostly in the hierarchy built up from the open sets. This generally forms the borel hierarchy on any topological space, but we will be interested in \mathcal{N} . Moreover, we want results about the whole topology of \mathcal{N} . Easily enough, we could instead consider the collection of clopen sets, which fairly trivially is seen to be a σ -algebra for \mathbb{R} but not for \mathcal{N} . The clopen sets, however, ignore much of the topology.

– 22 A • 2. Definition

For $S \subseteq \mathcal{O}(X)$, the σ -algebra generated by S is the \subseteq -least σ -algebra containing S. The set $\mathcal{B}^{\mathcal{M}}$ of *borel subsets* of a topological space \mathcal{M} is the σ -algebra generated by open sets. This induces a hierarchy: for $X \subseteq \mathcal{M}$ and $\alpha \in \text{Ord}$,

- X is $\sum_{n=1}^{\infty} \sum_{i=1}^{\infty} X^{n}$ iff X is open;
- X is $\sum_{\alpha}^{0,\mathcal{M}}$ for $\alpha > 1$ iff X is the countable union of sets in $\bigcup_{\beta < \alpha} \prod_{\beta}^{0,\mathcal{M}}$;
- X is $\Pi_{\alpha}^{0,\mathcal{M}}$ iff $\mathcal{N} \setminus X$ is $\Sigma_{\alpha}^{0,\mathcal{M}}$;
- X is $\Delta_{\alpha}^{0,\mathcal{M}}$ iff X is $\Pi_{\alpha}^{0,\mathcal{M}}$ and $\Sigma_{\alpha}^{0,\mathcal{M}}$.

These $\sum_{\alpha}^{0,\mathcal{M}}$, $\prod_{\alpha}^{0,\mathcal{M}}$, and $\Delta_{\alpha}^{0,\mathcal{M}}$ are called the *borel pointclasses*. We write just \sum_{α}^{0} for $\sum_{\alpha}^{0,\mathcal{N}}$ and similarly for the other borel pointclasses.

This yields a hierarchy similar to the arithmetical hierarchy in Appendix Section A3, but significantly longer.

$$\underline{A}_{1}^{0} \underbrace{\overset{\boldsymbol{\Sigma}_{1}^{0}}{\underset{\boldsymbol{\Omega}_{1}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}^{0}}{\overset{\boldsymbol{\Omega}_{2}^{0}}{\underset{\boldsymbol{\Omega}_{2}}{\underset$$



We can easily show the following, which yields some nice corollaries like the various containments above. Showing that the containments are strict is more difficult, of course. It should be obvious that $\Pi^0_{\alpha} \subseteq \Sigma^0_{\beta}$ for $\alpha < \beta$ since $x \in \Pi^0_{\alpha}$ has $x = \bigcup_{n < \omega} x \in \Sigma^0_{\beta}$ as a countable union of sets in $\bigcup_{\xi < \beta} \Pi^0_{\xi}$.

Let
$$1 \le \alpha < \beta < \omega_1$$
. Therefore,
 $1. \sum_{\alpha=0}^{1} \subseteq \sum_{\alpha=0}^{0} \subseteq \Sigma_{\beta}^{0}$, and $\underline{\Pi}_{\alpha}^{0} \subseteq \underline{\Pi}_{\beta}^{0}$; and

3.
$$\underline{A}^{0}_{\alpha} \subseteq \underline{\Sigma}^{0}_{\alpha} \subseteq \underline{A}^{0}_{\alpha+1}$$
, and similarly for $\underline{\Pi}^{0}_{\alpha}$;

Proof .:.

1. We actually need this in order to show (2) and (3). Suppose $X \in \sum_{i=1}^{0} i$ is open. We need to show X is the countable union of closed sets. We know X is the union of basic open sets (cones in the case of \mathcal{N}) of which there are only countably many.

So it suffices to show that each cone \mathcal{N}_{τ} is actually closed (and hence clopen). To see this, just note that $T_{\tau} = \{\sigma \in {}^{<\omega}\omega : \sigma \leq \tau \lor \tau \leq \sigma\}$ is a tree with $[T_{\tau}] = \mathcal{N}_{\tau}$. By Result 22 • 4, it follows that \mathcal{N}_{τ} is closed. So as the countable union of cones, $X \in \sum_{\tau=2}^{0} \mathbb{Z}^{0}$.

We only prove the first containment in general, as the second follows from the first (to see this, X ∈ Π⁰_α implies N \ X ∈ Σ⁰_α ⊆ Σ⁰_β so that X ∈ Π⁰_β). But this is easy: fix β and proceed by induction on α < β. For α = 1, this is just (1). For α > 1, suppose X ∈ Σ⁰_α. This means X = ⋃_{n<ω} X_n where, since α < β,

$$\{X_n : n \in \omega\} \subseteq \bigcup_{\xi < \alpha} \overline{\mathfrak{n}}_{\xi}^0 \subseteq \bigcup_{\xi < \beta} \overline{\mathfrak{n}}_{\xi}^0 \quad \text{implying} \quad \bigcup_{n \in \omega} X_n \in \Sigma_{\beta}^0.$$

3. That $\underline{\lambda}^{0}_{\alpha} \subseteq \underline{\Sigma}^{0}_{\alpha}$ is obvious by definition. For $\underline{\Sigma}^{0}_{\alpha} \subseteq \underline{\Lambda}^{0}_{\alpha+1}$, we already have $\underline{\Sigma}^{0}_{\alpha} \subseteq \underline{\Sigma}^{0}_{\alpha+1}$. To show $\underline{\Sigma}^{0}_{\alpha} \subseteq \underline{\Pi}^{0}_{\alpha+1}$, if $\mathcal{N} \setminus X \in \underline{\Pi}^{0}_{\alpha}$ then $\mathcal{N} \setminus X = \bigcup \{\mathcal{N} \setminus X\} \in \underline{\Sigma}^{0}_{\alpha+1}$ so that $X \in \underline{\Pi}^{0}_{\alpha+1}$.

For $1 < \alpha < \omega_1, X \in \sum_{\alpha=1}^{0} \inf X = \bigcup_{n < \omega} X_n$ for some $\{X_n : n < \omega\} \subseteq \prod_{\alpha=1}^{0} \bigcup_{n < \omega} X_n$

We can also easily confirm various closure properties for the borel pointclasses.

22 A • 6. Result Let $1 \le \alpha < \omega_1$. Therefore

1. $\sum_{\alpha=0}^{0}$ is closed under countable unions, and finite intersections;

- 2. Π_{α}^{0} is closed under finite unions, and countable intersections;
- 3. Δ_{α}^{0} is closed under finite unions, finite intersections, and complements.

Proof .:.

1. For countable unions, if $\{X_n : n < \omega\} \subseteq \sum_{\alpha}^0$, then for each $n < \omega$ let $\{X_{n,m} : m < \omega\} \subseteq \bigcup_{\xi < \alpha} \prod_{\xi}^0$ witness $X_n \in \sum_{\alpha}^0$, i.e. $\bigcup_{m < \omega} X_{n,m} = X_n$. Therefore, as a countable union, $\bigcup_{n,m < \omega} X_{n,m} = \bigcup_{n < \omega} X_n \in \sum_{\alpha}^0$.

For finite intersections, suppose $X = \bigcup_{n < \omega} X_n \in \sum_{\alpha}^0$ and $Y = \bigcup_{n < \omega} Y_n \in \sum_{\alpha}^0$ where $X_n, Y_n \in \bigcup_{\xi < \alpha} \prod_{\xi}^0$ for each $n < \omega$. Note that $X \cap Y = \bigcup_{n,m < \omega} X_m \cap Y_n$. Inductively, $X_m \cap Y_n \in \prod_{\xi}^0$ for some $\xi < \alpha$ (in some sense, we're proving (1)–(3) simultaneously by induction on α). It follows that $\bigcup_{n,m < \omega} X_m \cap Y_n = X \cap Y \in \sum_{\alpha}^0$.

- 2. If $\{X_n : n < \omega\} \subseteq \prod_{\alpha=0}^{0}$, then $\mathcal{N} \setminus X_n \in \sum_{\alpha=0}^{0}$ so by (1) the countable union $\bigcup_{n < \omega} \mathcal{N} \setminus X_n \in \sum_{\alpha=0}^{0}$ and therefore the complement $\mathcal{N} \setminus \bigcup_{n < \omega} \mathcal{N} \setminus X_n = \bigcap_{n < \omega} X_n \in \prod_{\alpha=0}^{0}$. Same reasoning of applying complements everywhere shows closure under finite unions.
- Closure under finite unions and intersections follows just from the fact that both ∑⁰_α and Π⁰_α are closed under these. For complements, X ∈ Δ⁰_α implies X ∈ Σ⁰_α so that N \ X ∈ Π⁰_α. But we also have X ∈ Π⁰_α, implying N \ X ∈ Σ⁰_α. Hence X, N \ X ∈ Σ⁰_α ∩ Π⁰_α = Δ⁰_α.

As we will also be concerned with product spaces, it's important to note that taking products doesn't increase complexity. In fact, we have the more general result below.

— 22 A • 7. Result

Let \mathcal{M} and \mathcal{W} be topologies. Let $f : \mathcal{M} \to \mathcal{W}$ be continuous and $\alpha > 0$. Therefore $X \in \sum_{\alpha}^{0, \mathcal{W}}$ implies $f^{-1} X \in \sum_{\alpha}^{0, \mathcal{M}}$, and similarly for $\prod_{\alpha}^{0, \mathcal{W}}$ and $\Delta_{\alpha}^{0, \mathcal{W}}$.

Proof .:.

As f is continuous, it is trivially true for open sets, i.e. $\alpha = 1$. For $\alpha > 1$, preimages works well with unions and complements, showing the result by an easy induction: $f^{-1"}\bigcup_{n<\omega} X_n = \bigcup_{n<\omega} f^{-1"}X_n$ and $f^{-1"}(W \setminus X) = \mathcal{M} \setminus f^{-1"}X$.

– 22 A • 8. Corollary –

If \mathcal{M} and \mathcal{W} are homeomorphic by some f, then $X \in \sum_{\alpha}^{0,\mathcal{M}}$ iff $f''X \in \sum_{\alpha}^{0,\mathcal{W}}$.

The above is useful, because products of baire space are all homeomorphic to baire space.

```
– 22 A • 9. Result
```

 ${}^{\omega}\mathcal{N}$ (or similarly ${}^{<\omega}\mathcal{N}$) is homeomorphic to \mathcal{N} . Hence their borel hierarchies are (in essence) the same through this coding by f.

Proof .:.

Consider the map $f : {}^{\omega} \mathcal{N} \to \mathcal{N}$ defined by $f(\langle x_n : n < \omega \rangle) = \langle x(n) : n < \omega \rangle$ where $x(\operatorname{code}(n, m)) = x_n(m)$. More precisely, the map code : $\omega \times \omega \to \omega$ is a bijection where $\operatorname{code}(x, y) \ge \max(x, y)$ (e.g. $\operatorname{code}(a, b)$ is the length of the path that spirals around the origin, skipping repeated points, and ends at the point $\langle a, b \rangle$). So really we say $x(n) = x_a(b)$ where $\operatorname{code}^{-1}(n) = \langle a, b \rangle$.

In checking that things converge, note that in the product space ${}^{\omega}\mathcal{N}$, a sequence $\langle \vec{x}_n : n < \omega \rangle$ converges to $\vec{x} = \langle x_m : m < \omega \rangle$ iff—regarding \vec{x}_n as the real sequence of reals $\langle x_n(m) < \omega \rangle$ —for each $m < \omega$, $\langle x_n(m) : n < \omega \rangle$ converges to x_m . In other words, a sequence converges iff it converges pointwise.

This f is a homeomorphism. To see this, f is clearly a bijection. f is continuous: suppose $\vec{x} = \langle \vec{x}_n \in {}^{\omega}\mathcal{N} : n < \omega \rangle$ converges to $x = \langle x_n : n < \omega \rangle \in {}^{\omega}\mathcal{N}$. We want to show $\langle f(\vec{x}_n) \in \mathcal{N} : n < \omega \rangle$ converges to f(x). Without loss of generality (just by ignoring issues of indices and passing to a subsequence), for each $N < \omega$ and each n < N, say $(\vec{x}_N)_n \upharpoonright N \leq x_n$, meaning the first N entries of \vec{x}_N each approximate the first N entries of x. If this is the case, then $f(\vec{x}_N)$ approximates x up to at least N since $\operatorname{code}(x, y) \geq \max(x, y)$. More explicitly, for $\operatorname{code}(a, b) < N$, we have a, b < N and since $(\vec{x}_N)_a \upharpoonright N \leq x_a$,

$$f(\vec{x}_N)(\text{code}(a,b)) = (\vec{x}_N)_a(b) = x_a(b) = f(x)(\text{code}(a,b)).$$

 f^{-1} is continuous: if $\vec{x} = \langle x_n \in \mathcal{N} : n < \omega \rangle$ converges to $x \in \mathcal{N}$, we may assume without loss of generality that $x_n \upharpoonright n \leq x$ for $n < \omega$. We want to show $\langle f^{-1}(x_m) \in {}^{\omega}\mathcal{N} : m < \omega \rangle$ converges to $f^{-1}(x) \in {}^{\omega}\mathcal{N}$. Note that $f^{-1}(y)_a(b) = y(\operatorname{code}(a, b))$ for all $a, b < \omega$ and $y \in \mathcal{N}$. Let $m < \omega$ and $n < \omega$ be fixed. We want to show that for sufficiently large N, $f^{-1}(x_N)_m \upharpoonright n \leq f^{-1}(x)_m$. Let $N > \operatorname{code}(m, n) > \max(m, n)$. For k < n, since $\operatorname{code}(m, k) < \operatorname{code}(m, n) < N$ has $x_N \upharpoonright N \leq x$,

$$f^{-1}(x_N)_m(k) = x_N(\operatorname{code}(m,k)) = x(\operatorname{code}(m,k)) = f^{-1}(x)_m(k).$$

Thus for any $m, n < \omega$, there is some sufficiently large N where $f^{-1}(x_N)_m \upharpoonright n \leq f^{-1}(x)$. This tells us that $\langle f^{-1}(x_n) : n < \omega \rangle$ converges to $f^{-1}(x)$. \dashv

At some point (how about now) we should prove this hierarchy *does* classify all borel sets, and in fact does this by stage ω_1 .

- 22 A · 10. Theorem $\mathcal{B} = \bigcup_{\alpha < \omega_1} \sum_{\alpha}^{0} = \bigcup_{\alpha < \omega_1} \prod_{\alpha}^{0} \alpha$.

Proof .:.

As the open sets are contained in $\bigcup_{\alpha < \omega_1} \sum_{\alpha}^{0}$, it suffices to show two things: that $\bigcup_{\alpha < \omega_1} \sum_{\alpha}^{0}$ is indeed a σ -algebra, and that $\bigcup_{\alpha < \omega_1} \sum_{\alpha}^{0}$ is the smallest that contains the open sets.

To see that $\bigcup_{\alpha < \omega_1} \sum_{\alpha}^0$ is a σ -algebra, we need to show that it's closed under complements and countable unions. But this is clear: $X \in \sum_{\alpha}^0$ has $\mathcal{N} \setminus X \in \prod_{\alpha}^0 \subseteq \sum_{\alpha+1}^0$. Moreover, any countable collection $\{X_n : n < \omega\} \subseteq \bigcup_{\alpha < \omega_1} \sum_{\alpha}^0$ has each $X_n \in \sum_{\alpha}^0 \subseteq \prod_{\alpha}^0 \subseteq \prod_{\alpha+1}^0$ for some $\alpha_n < \omega_1$. Since $\operatorname{cof}(\omega_1) > \omega$, $\{X_n : n < \omega\} \subseteq \sum_{\sup_{\alpha < \omega} \alpha_n + 2}^0$. Thus closure under countable unions gives $X \in \sum_{\sup_{\alpha < \omega} \alpha_n + 2}^0 \subseteq \bigcup_{\alpha < \omega_1} \sum_{\alpha}^0$.

To see $\bigcup_{\alpha < \omega_1} \sum_{\alpha}^0$ is the least such σ -algebra, let *S* be a σ -algebra containing \sum_{1}^0 . Inductively for $\alpha < \omega_1$, if $\bigcup_{\xi < \alpha} \sum_{\xi}^0 \subseteq S$, then closure under complements yields $\bigcup_{\xi < \alpha} \prod_{\xi}^0 \subseteq S$. Closure under countable unions yields $\sum_{\alpha}^0 \subseteq S$. Hence by induction, every $\sum_{\alpha}^0 \subseteq S$ and so $\bigcup_{\alpha < \omega_1} \sum_{\alpha}^0 \subseteq S$.

- 22 A • 11. Corollary $\sum_{\omega_1}^{0} = \bigcup_{\alpha < \omega_1} \sum_{\alpha}^{0} = \bigcup_{\alpha < \omega_1} \prod_{\alpha}^{0}$. In fact, $\sum_{\alpha}^{0} = \mathcal{B}$ for any $\alpha \ge \omega_1$.

Proof .:.

Argue by induction on $\alpha \ge \omega_1$. For $\alpha = \omega_1$, this is clear by Theorem 22 A • 10: if $X_n \in \prod_{\alpha_n}^0$ for $n < \omega$ and $\alpha_n < \omega_1$, then $\bigcup_{n < \omega} X_n \in \sum_{\beta+1}^0$ for $\beta = \sup_{n < \omega} \alpha_n < \omega_1$. Hence $\sum_{\omega_1}^0 \subseteq \bigcup_{\alpha < \omega_1} \sum_{\alpha}^0$. The other containment follows by Result 22 A • 4. This shows $\sum_{\omega_1}^0 = \mathcal{B}$ and as a σ -algebra, $\prod_{\omega_1}^0 = \mathcal{B}$. For $\alpha > \omega_1$, the result clearly holds since inductively $\bigcup_{\xi < \alpha} \prod_{\alpha}^0 = \bigcup_{\xi < \omega_1} \prod_{\alpha}^0$.

This tells us that the length of the borel hierarchy is at most ω_1 . We still need to show that all of the levels are different, however. To do this, we need the concept of a universal set, analogous to the concept from computability. We then apply a diagonalization argument. The work above tells us that there's really no difference between $\mathcal{N} \times \mathcal{N}$ and \mathcal{N} in their borel hierarchies. Hence from now on, we drop the \mathcal{M} in $\Sigma_{\alpha}^{0,\mathcal{M}}$ when \mathcal{M} is a (countable) product of copies of baire space.

- 22 A • 12. Definition Let $\underline{\Gamma}$ be a borel pointclass (or any subset of $\mathcal{P}(\mathcal{M} \times \mathcal{M})$ for \mathcal{M} polish). A set $U \subseteq \mathcal{N}$ is said to be $\underline{\Gamma}$ -universal iff $U \in \underline{\Gamma}$ and for every A in $\underline{\Gamma}$,

$$\exists r \in \mathcal{N} \, \forall x \in A(\langle r, x \rangle \in U \Leftrightarrow x \in A).$$

In other words, the set of projections of U is $\underline{\Gamma}$: for $U_r = \{x : \langle r, x \rangle \in U\}$, $\{U_r : r \in {}^{\omega}2\} = \underline{\Gamma}$. Most of the time, we will be looking at sets \sum_{α}^{0} -universal sets (or similarly for $\underline{\Pi}_{\alpha}^{0}$). With this new concept, we should show the existence of universal sets.

- 22 A • 13. Theorem For each $\alpha < \omega_1$, there is a $\sum_{\alpha \alpha}^{0}$ -universal set, and a $\prod_{\alpha \alpha}^{0}$ -universal set.

Proof .:.

It suffices to give just \sum_{α}^{0} -universal sets, since if U is such a set, writing $U_r = \{x : \langle r, x \rangle \in U\}$ yields $U' = \{\langle r, x \rangle : r \in U \land x \in \mathcal{N} \setminus U_r\}$ as a \prod_{α}^{0} -universal.

For $\alpha = 1$, \sum_{α}^{0} is just the open sets of \mathcal{N} , which are given by unions of cones. Identify ${}^{<\omega}\omega = \{\tau_n : n < \omega\}$ so that $X \in \sum_{1}^{0}$ is $\bigcup_{n \in r} \mathcal{N}_{\tau}$ for some $r \subseteq \omega$, i.e. some $r \in \mathcal{N}$. So taking $U = \bigcup_{n \in \omega} \{r \in \mathcal{N} : r(n) = 1\} \times \mathcal{N}_{\tau_n}$ yields a \sum_{1}^{0} -universal set: if $X = \bigcup_{n \in r} \mathcal{N}_{\tau_n}$, then $X = \{x : \langle r, x \rangle \in U\}$.

For $1 < \alpha < \omega_1$ write $\alpha = \sup_{n < \omega} (\alpha_n + 1)$ for some sequence of ordinals $\langle \alpha_n : n < \omega \rangle$ (possibly constant if α is a successor). For $n < \omega$, let U_n be $\prod_{\alpha_n}^0$ -universal. Each $X \in \sum_{\alpha}^0$ is the union of $\{X_n : n < \omega\} \subseteq \bigcup_{n < \omega} \prod_{\alpha_n}^0$.

By the containments of Result 22 A • 4, we can assume each $X_n \in \prod_{\alpha m}^0$ for some $m \ge n$. Thus for $n < \omega$ there is an $r_n \in \mathcal{N}$ where

$$X = \bigcup_{n < \omega} X_n = \bigcup_{n < \omega} \{ x : \langle r_n, x \rangle \in U_n \}.$$

Through coding, X identified by this sequence $\langle r_n \in \mathcal{N} : n < \omega \rangle$ can instead be identified through a single r as in Result 22 A • 9 so translating the U_n s in this way yields a \sum_{α}^{0} -univeral set. More explicitly, $f : {}^{\omega}\mathcal{N} \to \mathcal{N}$ a homeomorphism yields $U'_n = \{\langle r, x \rangle \in {}^{\omega}\mathcal{N} \times \mathcal{N} : \langle f(r_n), x \rangle \in U_n\}$ yields $U = \bigcup_{n < \omega} U'_n$ as \sum_{α}^{0} -universal. f is continuous as is the map proj_n where $\operatorname{proj}_n(r) = r_n$. So U'_n is $\prod_{\alpha}^{0} v_n$ by Result 22 A • 7, implying that U is \sum_{α}^{0} .

– 22 A • 14. Corollary —

For all $\alpha < \omega_1$, $\sum_{\alpha}^{0} \neq \prod_{\alpha}^{0}$ since any \sum_{α}^{0} -universal set is not in \prod_{α}^{0} and vice versa. In particular, all containments of Figure 22 A • 3 are strict: meaning for $\alpha < \omega_1$,

- $\Delta_{\alpha}^{0} \subsetneq \Sigma_{\alpha}^{0} \subsetneq \Delta_{\alpha+1}^{0}$, and similarly for Π_{α}^{0} ; and
- we always add sets, even at limit stages: $\bigcup_{\xi < \alpha} \sum_{\xi} \sum_{\alpha} \sum_{\alpha$

Proof .:.

Let U be \sum_{α}^{0} -universal. Let $D = \{x : \langle x, x \rangle \in U\}$ which is in \sum_{α}^{0} as the continuous preimage of U by $x \mapsto \langle x, x \rangle$. Suppose $\mathcal{N} \setminus D \in \sum_{\alpha}^{0}$ so by universality, there is some r with $U_r = \{x \in \mathcal{N} : \langle r, x \rangle \in U\} = \mathcal{N} \setminus D$. But then $r \notin D$ iff $r \in U_r$ iff $\langle r, r \rangle \in U$ iff $r \in D$, a contradiction. This shows $D \in \sum_{\alpha}^{0} \setminus \prod_{\alpha}^{0}$, and an analogous result shows $\prod_{\alpha}^{0} \setminus \sum_{\alpha}^{0} \neq \emptyset$.

- If U is \sum_{α}^{0} -universal, then $U \in \sum_{\alpha}^{0} \setminus \prod_{\alpha}^{0} = \sum_{\alpha}^{0} \setminus \Delta_{\alpha}^{0}$, showing $\Delta_{\alpha}^{0} \subsetneq \sum_{\alpha}^{0}$. Similarly, If W is \prod_{α}^{0} -universal, then $W \in \prod_{\alpha}^{0} \setminus \sum_{\alpha}^{0} \subseteq \Delta_{\alpha+1}^{0} \setminus \sum_{\alpha}^{0}$, showing $\sum_{\alpha}^{0} \subsetneq \Delta_{\alpha+1}^{0}$. The same argument above works for \prod_{α}^{0} .
- A \sum_{α}^{0} or \prod_{α}^{0} -universal set shows the strict containment for \sum_{α}^{0} and \prod_{α}^{0} . For \sum_{α}^{0} , let $\langle \alpha_{n} : n < \omega \rangle$ be such that $\sup_{n < \omega} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. In particular, take U'_{n} to be $\sum_{\alpha}^{0} (\alpha_{n} 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) = \alpha$. Take increasingly complex sets $U_{n} \in \sum_{\alpha}^{0} (\alpha_{n} + 1) \setminus \sum_{\alpha}^{0} (\alpha_{n} + 1) \in \sum_{\alpha}^{$
 - 1. $U_n \subseteq \mathcal{N}_{\langle n \rangle} \in \sum_{i=1}^{n} 0$ (by construction);
 - 2. $U_n \cap U_m = \emptyset$ for $n < m < \omega$ (by construction);
 - 3. $\mathcal{N} \setminus \bigcup_{n < \omega} U_n = \bigcup_{n < \omega} \mathcal{N}_{(n)} \setminus U_n$ (from (1) and (2));

4. $U_n, \mathcal{N}_{\langle n \rangle} \setminus U_n \in \underline{A}^0_{\alpha_n+1} \setminus \underline{A}^0_{\alpha_n} \subseteq \prod_{\alpha_n+1}^0$ (by construction and properties of borel pointclasses);

(4) implies $\bigcup_{n<\omega} U_n \in \sum_{\alpha}^0$. Clearly $\bigcup_{n<\omega} U_n \notin \bigcup_{\xi<\alpha} \Delta_{\xi}^0$ as otherwise the union is in some $\Delta_{\alpha_n}^0$ and thus $U_n = \mathcal{N}_{\langle n \rangle} \cap \bigcup_{m<\omega} U_m \in \Delta_{\alpha_n}^0$, contradicting (4). To see that $\bigcup_{n<\omega} U_n \in \prod_{\alpha}^0$, (3) tells us that $\bigcup_{n<\omega} \mathcal{N}_{\langle n \rangle} \setminus U_n \in \Sigma_{\alpha}^0$ and therefore its complement $\bigcup_{n<\omega} U_n \in \prod_{\alpha}^0$. So $\bigcup_{n<\omega} U_n \in \Delta_{\alpha}^0$.

It should be noted that not all sets of reals are borel. In particular, as the union of countably many cones, there are only 2^{\aleph_0} open sets. Similarly there are continuum many elements of Π_1^0 . One can easily show (noting that $\omega_1 \leq 2^{\aleph_0}$) that there are only 2^{\aleph_0} many borel sets of reals. But there are $2^{2^{\aleph_0}}$ sets of reals overall. Hence there must be many non borel sets. In fact, we can come up with an example of one later through showing that the *projective* hierarchy, an extension of the borel hierarchy, has distinct pointclasses.

§ 22 B. Changing the hierarchy

The primary question for this subsection is this: to what extent does the collection of borel sets on a polish space determine the topology on the polish space? That is to say, to what extent does $\mathcal{B}^{\mathcal{M}} = \bigcup_{\alpha < \omega_1} \sum_{\alpha}^{0,\mathcal{M}}$ determine $\sum_{\alpha}^{0,\mathcal{M}}$ (and subsequently the rest of the hierarchy)? The answer is "not at all". Although certainly the topology determines the borel sets, we can "refine" our topology such that any given borel subset is open and closed in the new topology.

This, however, also suggests (which we do not prove here) that there is a *unique* borel space among the uncountable polish spaces.

First we show that we can make closed sets closed and open without changing the borel sets.

- 22B•1. Lemma Let $\mathcal{M} = \langle \mathcal{M}, d_{\mathcal{M}} \rangle$ and $\mathcal{W} = \langle \mathcal{W}, d_{\mathcal{W}} \rangle$ be polish metric spaces. Therefore the *disjoint union* space $\mathcal{M} \sqcup \mathcal{W}$ is the polish space $\langle \mathcal{M} \sqcup \mathcal{W}, \mathcal{O} \rangle$ where $U \subseteq \mathcal{M} \sqcup \mathcal{W}$ is in \mathcal{O} iff $U \cap \mathcal{M} \in \mathcal{O}_{\mathcal{M}}$ and $U \cap \mathcal{W} \in \mathcal{O}_{\mathcal{W}}$; i.e. $U = A \sqcup B$ for $A \in \mathcal{O}_{\mathcal{M}}$ and $B \in \mathcal{O}_{\mathcal{W}}$.

Proof ...

We need to check that $\langle \mathcal{M} \sqcup \mathcal{W}, \mathcal{O} \rangle$ is indeed a polish space. Write $\mathcal{U} = \mathcal{M} \sqcup \mathcal{W}$. Let \mathcal{O}' be the topology induced by the metric $d : \mathcal{U}^2 \to \mathbb{R}$ where

$$d(x, y) = \begin{cases} d_{\mathcal{M}}(x, y) & \text{if } x, y \mathcal{M} \\ d_{\mathcal{W}}(x, y) & \text{if } x, y \in \mathcal{W} \\ 1 & \text{otherwise.} \end{cases}$$

We need to show that d is indeed a metric, $\langle \mathcal{U}, \mathcal{O}' \rangle$ is complete, and $\mathcal{O} = \mathcal{O}'$. That \mathcal{U} is separable then follows from the fact that both \mathcal{W} and \mathcal{M} are.

 $\begin{array}{c} - \text{ Claim 1} \\ \langle \mathcal{U}, d \rangle \text{ is a complete metric space.} \end{array}$

Proof .:.

Firstly, we show that d is a metric, which is easy: we clearly have d(x, y) = 0 iff $d_{\mathcal{M}}(x, y) = 0$ or $d_{\mathcal{W}}(x, y) = 0$ iff x = y, and d(x, y) = d(y, x). The triangle inequality is fairly clear from it holding for $d_{\mathcal{M}}$ and $d_{\mathcal{W}}$:

$$x, y, z \in \mathcal{M}$$
 implies $d(x, y) + d(y, z) = d_{\mathcal{M}}(x, y) + d_{\mathcal{M}}(y, z) \ge d_{\mathcal{M}}(x, z) = d(x, z)$
 $x, y \in \mathcal{M}$ and $z \in \mathcal{W}$ implies $d(x, y) + d(y, z) = d_{\mathcal{M}}(x, y) + 1 \ge 1 = d(x, z)$

 $x \in \mathcal{M}$ and $y, z \in \mathcal{W}$ implies $d(x, y) + d(y, z) = 1 + 1 \ge 1 = d(x, z)$,

and similarly for the other cases. This establishes that (\mathcal{U}, d) is indeed a metric space, and we merely need to show that it's complete.

So let $\vec{x} = \langle x_n : n < \omega \rangle \in {}^{\omega}\mathcal{U}$ be cauchy with respect to d. We must show that \vec{x} converges in $\langle \mathcal{U}, d \rangle$. But eventually the sequence must be in either \mathcal{M} or \mathcal{W} where it must then stay and converge due to the completeness of \mathcal{M} and \mathcal{W} . More explicitly, let $\varepsilon > 0$ be arbitrary, and without loss of generality $\varepsilon < 1$. Therefore, for sufficiently large $n, m < \omega$, $1 > \varepsilon > d(x_n, x_m)$ which is then equal to $d_{\mathcal{M}}(x_n, x_m)$ or $d_{\mathcal{W}}(x_n, x_m)$ and so the tail of \vec{x} (which is then in \mathcal{M} or \mathcal{W}) is cauchy and so converges to some x in \mathcal{M} or \mathcal{W} , and it's not hard to see that \vec{x} converges to this x in $\langle \mathcal{U}, d \rangle$ as well.

Now we need to show $\mathcal{O} = \mathcal{O}'$, where \mathcal{O} is defined as in the statement, and \mathcal{O}' is the topology induced by d. Note that elements of \mathcal{O}' are just unions of open balls. For $x \in \mathcal{U}$, write $B_{\varepsilon}(x)$ for $\{y \in \mathcal{U} : d(x, y) < \varepsilon\}$. Write $B_{\varepsilon}^{\mathcal{M}}(x) = \{y \in \mathcal{M} : d_{\mathcal{M}}(x, y) < \varepsilon\}$, and similarly for $B_{\varepsilon}^{\mathcal{W}}(x)$. Hence $X \in \mathcal{O}'$ iff there is some $F \subseteq \mathcal{U} \times \{\varepsilon \in \mathbb{R} : \varepsilon > 0\}$ such that $X = \bigcup_{(x,\varepsilon) \in F} B_{\varepsilon}(x)$.

Let $0 < \varepsilon \in \mathbb{R}$ and $x \in \mathcal{U}$. Note that $\mathcal{O}_{\mathcal{M}} \subseteq \mathcal{O}$ and $\mathcal{O}_{\mathcal{W}} \subseteq \mathcal{O}$.

• If $\varepsilon < 1$ and $x \in \mathcal{M}$, we have $B_{\varepsilon}(x) = B_{\varepsilon}^{\mathcal{M}}(x) \sqcup \emptyset$.

- If $\varepsilon < 1$ and $x \in \mathcal{W}$, we have $B_{\varepsilon}(x) = \emptyset \sqcup B_{\varepsilon}^{\mathcal{W}}(x)$.
- If $\varepsilon \geq 1$ and $x \in \mathcal{M}$, we have $B_{\varepsilon}(x) = B_{\varepsilon}^{\mathcal{M}}(x) \sqcup \mathcal{W}$.
- If $\varepsilon \ge 1$ and $x \in W$, we have $B'_{\varepsilon}(x) = \mathcal{M} \sqcup B_{\varepsilon}(x)$.

Hence open balls have this form: $B_{\varepsilon}(x) = X_{x,\varepsilon} \sqcup Y_{x,\varepsilon}$. And so if $A \in \mathcal{O}'$, then $A = \bigcup_{x \in A} B_{\varepsilon_x}(x)$ for ε_x such

that $B_{\varepsilon_X}(x) \subseteq A$, and therefore

$$A = \bigcup_{x \in A} X_{x,\varepsilon_x} \sqcup Y_{x,\varepsilon_x} = \bigcup_{x \in A \cap \mathcal{M}} X_{x,\varepsilon_x} \sqcup \bigcup_{x \in A \cap \mathcal{W}} Y_{x,\varepsilon_x}.$$

Hence $\mathcal{O}' \subseteq \mathcal{O}$. Similarly, if $A \in \mathcal{O}$, then $A = (A \cap \mathcal{M}) \sqcup (A \cap \mathcal{W})$ where $A \cap \mathcal{M} \in \mathcal{O}_{\mathcal{M}}$ is the union of open balls (with respect to d) and $A \cap \mathcal{W} \in \mathcal{O}_{\mathcal{W}}$ is too and so $A \in \mathcal{O}'$.

— 22 B•2. Lemma —

Let $\mathcal{M} = \langle \mathcal{M}, \mathcal{O} \rangle$ be a polish topological space with $X \in \mathfrak{M}_1^{0, \mathcal{M}}$. Therefore, there is an $\mathcal{O}' \supseteq \mathcal{O}$ such that

- 1. $\mathcal{M}' = \langle \mathcal{M}, \mathcal{O}' \rangle$ is a polish topological space;
- 2. $X \in \mathbf{\Delta}_{1}^{0,\mathcal{M}'}$; and
- 3. $\mathcal{B}^{\mathcal{M}} = \mathcal{B}^{\mathcal{M}'}$.

Proof .:.

As X is closed, the inhereted topology X is polish by Result 21 A • 14. Similarly, $\mathcal{M} \setminus X$ is open, giving another inhereted polish topology $\mathcal{M} \setminus X$ by Result 21 A • 15. This means we can consider the disjoint union $\mathcal{M}' = \langle \mathcal{M}, \mathcal{O}' \rangle = (\mathcal{M} \setminus X) \sqcup X$ as polish as in Lemma 22 B • 1. As the topologies are inhereted, every open subset U of \mathcal{M} has $U \cap X$ and $U \setminus X$ as open in X and $\mathcal{M} \setminus X$ respectively and therefore open in $\mathcal{M}': \mathcal{O} \subseteq \mathcal{O}'$. This gives (1) and the rest follow easiliy.

- 2. Since $\mathcal{M} \setminus X \in \mathcal{O} \subseteq \mathcal{O}' = \sum_{1}^{0,\mathcal{M}'}$, we know $X \in \prod_{1}^{0,\mathcal{M}'}$. X is also open in \mathcal{M}' , since it's the union of open balls $\bigcup_{y \in \mathcal{M} \setminus X} B_1(y)$. Hence $X \in \underline{A}_1^{0,\mathcal{M}'}$.
- 3. It's clear that $\mathcal{B}^{\mathcal{M}} \subseteq \mathcal{B}^{\mathcal{M}'}$ since (1) just says $\sum_{1}^{0,\mathcal{M}} \subseteq \sum_{1}^{0,\mathcal{M}'}$. For the other containment, we use Lemma 22 B 1: since $X \in \mathcal{B}^{\mathcal{M}}$, any open set of \mathcal{M}' is of the form $U \cup (V \cap X)$ for $U, V \in \mathcal{O}$, and hence is borel in \mathcal{M} and so the rest of the borel hierarchy on \mathcal{M}' is also borel in \mathcal{M} by an easy induction. \dashv

This, in some sense, allows us to just continually collapse certain sets down to the complexity of $\underline{A}_{1}^{0,\mathcal{M}}$, like a $\underline{\Pi}_{1}^{0,\mathcal{M}}$ -universal set for example. As a result, we can make *any* borel set $\underline{A}_{1}^{0,\mathcal{M}}$ by sufficiently refining the topology according to all the topologies that make things clopen.

22B·3. Theorem
Let M = ⟨M, Ø⟩ be a polish topological space with X ∈ B^M. Therefore there is an Ø' ⊇ Ø such that
1. M' = ⟨M, Ø'⟩ is a polish topological space;
2. X ∈ Δ^{0,M'}₁; and
3. B^M = B^{M'}.

Proof .:.

Consider the set of all *X* where the result holds:

 $\mathfrak{G} = \{X \subseteq \mathcal{M} : \text{there is a polish topology} \mathcal{M}' \text{ on } \mathcal{M} \text{ where } X \text{ is closed and open and } \mathcal{B}^{\mathcal{M}'} = \mathcal{B}^{\mathcal{M}} \}.$

Clearly $\Pi_1^{0,\mathcal{M}} \subseteq \mathfrak{G}$. Moreover, if $X \in \mathfrak{G}$, then the same topology witnessing this also witnesses that $\mathcal{M} \setminus X \in \mathfrak{G}$ so that \mathfrak{G} is closed under complementation and thus contains all the open and closed sets. It then suffices to show \mathfrak{G} is closed under countable intersections as this generates the borel hierarchy on $\mathfrak{M} \colon \mathcal{B}^{\mathcal{M}} \subseteq \mathfrak{G}$.

So suppose $\{X_n : n < \omega\} \subseteq \mathbb{B}$ with each $X_n \in \mathbb{B}$ witnessed by $\mathcal{M}_n = \langle \mathcal{M}, \mathcal{O}_n \rangle$. Consider the product $\prod_{n < \omega} \mathcal{M}_n$, which is polish by Example 21 A • 13. If we consider $j : \mathcal{M} \to \prod_{n < \omega} \mathcal{M}$ defined by $j(x) = \operatorname{const}_x \upharpoonright \omega$, this induces another polish topology on \mathcal{M} by preimages: $\mathcal{M}_\omega = \langle \mathcal{M}, \mathcal{O}_\omega \rangle$ where $U \in \mathcal{O}_\omega$ iff $U = j^{-1} W$ for some W open in $\prod_{n < \omega} \mathcal{M}_n$. It suffices to show $\bigcap_{n < \omega} X_n \in \mathbb{B}$ as witnessed by \mathcal{M}_ω . This may not be true, but we can *make it true* by refining again.

1. This is easily checked from the definition of \mathcal{M}_{ω} because $\prod_{n < \omega} \mathcal{M}_n$ is polish and $j " \mathcal{M}_{\omega}$ is closed.

- 2. $\bigcap_{n < \omega} X_n$ is closed as the intersection of the closed sets j^{-1} " $(\prod_{i < n} \mathcal{M} \times X_n \times \prod_{i > n} \mathcal{M})$ for $i < \omega$. $\bigcap_{n < \omega} X_n$ may not be open, but using Lemma 22 B • 2, it becomes clopen in a new polish topology \mathcal{M}' .
- 3. Since $\mathcal{B}^{\mathcal{M}} = \mathcal{B}^{\mathcal{M}_n}$ for each n, each open set $U \in \sum_{1}^{0,\mathcal{M}}$ has $W = \prod_{n < N} \mathcal{M} \times U \times \prod_{N \le n < \omega} \mathcal{M}$ as open in $\prod_{n < \omega} \mathcal{M}_n$ with preimage $j^{-1} W = U$ as open in \mathcal{M}_{ω} . Hence $\sum_{1}^{0,\mathcal{M}} \subseteq \sum_{1}^{0,\mathcal{M}_{\omega}}$ implies $\mathcal{B}^{\mathcal{M}} \subseteq \mathcal{B}^{\mathcal{M}_{\omega}} = \mathcal{B}^{\mathcal{M}'}$.

For the other containment, any open set of $\prod_{n<\omega} \mathcal{M}_n$ is given by unions of rectangles, so the open sets of \mathcal{M}_{ω} are given by unions of preimages of rectangles: $j^{-1"} \bigcup_{U \in F} U = \bigcup_{U \in F} j^{-1"}U$ for any F. So it suffices to show $j^{-1"}R \in \sum_{1}^{0,\mathcal{M}}$ for any open rectangle $R = \prod_{n<\omega} U_n$ where each U_n is open in \mathcal{M} and only finitely $U_n \neq \mathcal{M}$. The preimage $j^{-1"}(\prod_{n<\omega} U_n) = \bigcap_{n<\omega} j^{-1"}U_n$ is borel in \mathcal{M}_{ω} , and therefore $\mathcal{B}^{\mathcal{M}'} = \mathcal{B}^{\mathcal{M}_{\omega}} \subseteq \mathcal{B}^{\mathcal{M}}$.

As a result, if we can prove generally that a property holds for closed sets of any polish space, then we have in fact proven the result for all borel sets too. We will see examples of this sort of reasoning later.

§ 22 C. The projective hierarchy

One may have wondered what the '0' in " \sum_{α}^{0} " represented. The projective hierarchy will change this '0' to a '1'. Essentially, the 0 represents that we've really only quantified over \mathbb{N} whereas 1 represents that we're quantifying also over $\mathcal{O}(\mathbb{N})$, i.e. \mathbb{R} . This can be made precise in terms of higher order logic, but really we just note that countable unions act like existential quantification over \mathbb{N} . The analogue of existential quantification over \mathbb{R} is projection. What exactly are we projecting?

- 22C • 1. Definition Let $A \subseteq X \times Y$, usually polish spaces. The *projection* pA is just dom $A = \{x \in X : \exists y \in Y \ (\langle x, y \rangle \in A)\} \subseteq X$. Where there is ambiguity (like if X itself is a product), we also write $p_X A$ to denote this set, which can also be used to define $p_Y A = \operatorname{ran} A$, for example.

We can either talk about product spaces or coding. We code pairs of reals by $x * y = \langle x(n), y(n) : n \in \omega \rangle$, meaning x * y(2n) = x(n) and x * y(2n+1) = y(n). The end result is the same by Result 22 A • 9. So although we only state the definitions for \mathcal{N} , they in principle hold for countable products of copies of \mathcal{N} as well.

In most literature, projections of closed sets are called *analytic* sets, which will cause confusion with the *analytical* hierarchy defined later. So we abandon this terminology, instead referring to them only as $\sum_{i=1}^{1}$ -sets. The basic idea is that we never actually have to mention spaces other than \mathcal{N} , since any witness to a set being analytic that we find over some weird space yields a witness to being analytic over \mathcal{N} . We also note that these kinds of sets play nicely with continuous functions, which will show that although the borel sets are closed under continuous preimages, they are not closed under continuous images.

22C•2. Result Let $A \subseteq \mathcal{N}$. The following are equivalent. 1. $A = \{x : \exists y \ (x * y \in B)\}$ for some closed $B \subseteq \mathcal{N}$. 2. $A = \mathfrak{p}B$ for some closed $B \subseteq \mathcal{N} \times \mathcal{N}$. 3. $A = \mathfrak{p}B$ for some borel $B \subseteq \mathcal{N} \times \mathcal{N}$. 4. $A = \mathfrak{p}B$ for some polish space \mathcal{M} and closed $B \subseteq \mathcal{N} \times \mathcal{M}$. 5. $A = \operatorname{im} f$ for some continuous $f : \mathcal{N} \to \mathcal{N}$.

Proof ...

(1) \leftrightarrow (2) Suppose (1) holds. Consider $B' = \{\langle x, y \rangle : x * y \in B\}$, the continuous preimage of $\langle x, y \rangle \mapsto x * y$ so that B' is also closed, and clearly A = pB'. Similarly, if (2) holds, $B' = \{x * y : \langle x, y \rangle \in B\}$ is

the continuous preimage of $x * y \mapsto \langle x, y \rangle$ so that $A = \{x : \exists y \ (x * y \in B')\}.$

(2) ↔ (3) One direction is trivial since closed sets are borel. So suppose A = pB for some borel B. First we show that B = p_{N×N}C for some closed C ⊆ N³. This would imply p_NC = p_Np_{N×N}C = pB = A. To show this, it suffices to show the set of projections of closed sets, call this ∑¹₁, contains all open sets, all closed sets, and is closed under countable unions and intersections (one can generate the borel sets in this way).

Clearly every closed and open set is in $\sum_{i=1}^{1} j$ just by coding: for U a closed set, the set $U \times \mathcal{N}$ is also closed as a continuous preimage of U under $\langle x, y \rangle \mapsto x$. This set clearly has $\mathfrak{p}(U \times \mathcal{N}) = U$ so that $U \in \sum_{i=1}^{1}$. Since $\sum_{i=1}^{0} \subseteq \sum_{i=2}^{0} j$ (i.e. every open set is the countable union of closed sets) it suffices to show that $\sum_{i=1}^{1} j$ is closed under countable unions and intersections. So suppose $X_n \in \sum_{i=1}^{1} j$ for $n < \omega$. Let $X_n = \mathfrak{p}Y_n$ where each $Y_n \subseteq \mathcal{N} \times \mathcal{N}$ is closed.

For the countable union $\bigcup_{n < \omega} X_n$, consider $Y'_n = Y_n \times \mathcal{N}_{\langle n \rangle}$, which is also closed as the preimage by the continuous map $\langle x, y, \langle n \rangle^{\frown} z \rangle \mapsto \langle x, y, z \rangle$. Note that $\mathfrak{p}_{\mathcal{N} \times \mathcal{N}} Y'_n = Y_n$ so that $\mathfrak{p}_{\mathcal{N}} Y'_n = \mathfrak{p} Y_n = X_n$. Moreover, $\bigcup_{n < \omega} Y'_n$ is closed in \mathcal{N}^3 since the complement,

$$\mathcal{N}^3 \setminus \bigcup_{n < \omega} Y'_n = \bigcup_{n < \omega} \mathcal{N}^2 \times \mathcal{N}_{\langle n \rangle} \setminus Y'_n,$$

is the union of a bunch of open sets, and is thus open. And it's not difficult to then see that $\mathfrak{p}_{\mathcal{N}} \bigcup_{n < \omega} Y'_n = \bigcup_{n < \omega} X_n \in \Sigma_1^1$.

For the countable intersection, let $f : \mathcal{N} \to {}^{\omega}\mathcal{N}$ be a homeomorphism. Decompose f as a sequence of functions, $f(x) = \langle f_n(x) : n < \omega \rangle$, so that each $f_n : \mathcal{N} \to \mathcal{N}$ is continuous. For each $n < \omega$, consider $Y_n^* = \{\langle x, y \rangle \in \mathcal{N} \times \mathcal{N} : \langle x, f_n(y) \rangle \in Y_n\}$, which is also closed as the preimage of Y_n under the continuous map $\langle x, y \rangle \mapsto \langle x, f_n(y) \rangle$. Note that then $\bigcap_{n < \omega} Y_n^*$ is closed and it's easy to see $\mathfrak{p} \bigcap_{n < \omega} Y_n^* = \bigcap_{n < \omega} \mathfrak{p} Y_n = \bigcap_{n < \omega} X_n$.

- $(2) \rightarrow (4)$ This is trivial since $\mathcal{M} = \mathcal{N}$ is polish.
- (4) \rightarrow (5) Let $A = \mathfrak{p}B$ for $B \subseteq \mathcal{N} \times \mathcal{M}$. Knowing a bit about metric spaces, closed subsets of complete metric spaces are complete, and hence B is also a polish space. So let $f : \mathcal{N} \to B$ be a continuous surjection as per Theorem 21 B 5. The projection map $\pi : \mathcal{N} \times \mathcal{N} \to \mathcal{N}$ defined by $\pi(\langle x, y \rangle) = x$ is clearly continuous so that $\pi \circ f : \mathcal{N} \to \mathcal{N}$ is a continuous map with $\pi \circ f"\mathcal{N} = \pi"B = \mathfrak{p}B = A$.
- (5) \rightarrow (2) If $f : \mathcal{N} \rightarrow \mathcal{N}$ is continuous with A = im f, then $f \subseteq \mathcal{N} \times \mathcal{N}$ is closed (and thus so is $f^{-1} = \{\langle y, x \rangle : \langle x, y \rangle \in f\}$) with $A = \mathfrak{p} f^{-1}$, giving the result. \dashv

By continually taking projections and complements, we get another hierarchy. In principle, one could relativize this to other polish spaces, but we will focus on \mathcal{N} and countable products of copies of \mathcal{N} .

22C·3. Definition
We form the *projective hierarchy* as follows: for X ⊆ N and n < ω,
X is ∑₀¹ iff X is ∑₁⁰, i.e. X is open;
X is ∑_{n+1}¹ iff X = pA for some A ∈ Π_n¹;
X is Π_n¹ iff N \ X is ∑_n¹;
X is Δ_n¹ iff X is both ∑_n¹ and Π_n¹.
These ∑_n¹, Π_n¹, Δ_n¹ are the *projective pointclasses*, and sets in them are called *projective*.

One gets the expected properties from the notation and definitions.



22 C • 4. Figure: The projective hierarchy

$-22C \cdot 5$. Result

For $0 < n < \omega$,

- 1. $\sum_{n=1}^{1}$ is closed under countable unions, countable intersections, and projections.
- 2. $\prod_{n=1}^{n}$ is closed under countable unions, countable intersections, and co-projections (i.e. $X \mapsto \mathcal{N} \setminus \mathfrak{p}(\mathcal{N} \setminus X)$)
- 3. Δ_n^1 is closed under countable unions, and complements, and is thus a σ -algebra.

Proof .:.

Proceed by induction on *n*. For the base case n = 1, we proved the closure under countable unions and intersections in proving (2) \leftrightarrow (3) for Result 22 C • 2. Closure under projections (or co-projections) is the same as in the inductive case, which we will show now.

1. For countable unions, let $\{A_i : i < \omega\} \subseteq \sum_{n+1}^{1}$, where each $A_i = \mathfrak{p}B_i$ for $B_i \in \prod_n^{1}$. The inductive hypothesis on \prod_n^{1} yields that $B = \bigcup_{i \in \omega} B_i \in \prod_n^{1}$ so that $\mathfrak{p}B = \bigcup_{i \in \omega} \mathfrak{p}B_i = \bigcup_{i < \omega} A_i \in \sum_{n+1}^{1}$. Similarly, for countable intersections, $\bigcap_{i \in \omega} B_i$ has $\mathfrak{p} \bigcap_{i \in \omega} B_i = \bigcap_{i \in \omega} \mathfrak{p}B_i \in \sum_{n+1}^{1}$.

Projections are trivial, because the two witnesses for $x \in ppB$ can be coded as a single witness: for $B \subseteq \mathcal{N}^3$, $p_{\mathcal{N}}B = A$ has $p_{\mathcal{N}}A = p_{\mathcal{N}\times\mathcal{N}}B \in \sum_{n=1}^{n} D$.

2. This follows from the results on $\sum_{n=1}^{1} \frac{1}{n}$.

3. This follows from (1) and (2): since both $\sum_{n=1}^{1}$ and $\prod_{n=1}^{1}$ are closed under countable unions, so is $\sum_{n=1}^{1} \cap \prod_{n=1}^{1} = \Delta_{n}^{1}$. It's also easy to see that Δ_{n}^{1} is closed under complements, since $X \in \Delta_{n}^{1}$ implies $X \in \sum_{n=1}^{1}$ and so $\mathcal{N} \setminus X \in \prod_{n=1}^{1}$. But $X \in \Delta_{n}^{1}$ also implies $X \in \prod_{n=1}^{1}$ and so $\mathcal{N} \setminus X \in \sum_{n=1}^{1}$, meaning $\mathcal{N} \setminus X \in \Delta_{n}^{1}$.

We again get an analogous picture to that of the borel and arithmetical hierarchies. Although the borel and projective hierarchies both start from the open sets, the projective hierarchy reaches the borel sets almost immediately: $\mathcal{B} = \Delta_1^1$. Now unlike the borel hierarchy, which stops at stage ω_1 —meaning $\sum_{\omega_1}^0 = \bigcup_{\alpha < \omega_1} \sum_{\alpha}^0$ is closed under complements and countable unions—the projective hierarchy is much shorter, stopping by stage ω : $\bigcup_{n < \omega} \sum_n^1$ is closed under complements and projection.

- 22 C • 6. Corollary — The collection of projective sets of reals—meaning $\bigcup_{n < \omega} \sum_{n=1}^{n}$ is closed under complementation and projection.

The proof that $\Delta_1^1 = \mathcal{B}$ uses some facts about how borel sets project and interact with Σ_1^1 sets.

— 22 C • 7.	Corollary -
All borel se	ts are Δ_1^1 .

Proof .:.

Clearly every closed set is in Δ_1^1 just by coding: for U a closed set, the set $U \times \mathcal{N}$ is also closed as a continuous preimage of U under $\langle x, y \rangle \mapsto x$. This set clearly has $\mathfrak{p}(U \times \mathcal{N}) = U$ so that $U \in \Delta_1^1$. By Result 22 C • 5, Δ_1^1 is a σ -algebra and therefore contains the borel σ -algebra.

To show the converse of this, that Δ_1^1 consists *only* of borel sets, we need a separation property.

Proof .:.

For $U, V \in \sum_{1}^{1}$, say $U \mathscr{B}$ -sep V iff U and V are separated by a borel set. Firstly note that \mathscr{B} -sep respects unions.

Claim 1 If $V_n \mathcal{B}$ -sep U_m for all $n, m < \omega$, then $\bigcup_{n < \omega} V_n \mathcal{B}$ -sep $\bigcup_{n < \omega} U_n$.

Proof .:.

For each $n, m < \omega$, let $B_{n,m}$ be a borel set with $V_n \subseteq B_{n,m} \subseteq \mathcal{N} \setminus U_m$. Consider $\bigcup_{n < \omega} \bigcap_{m < \omega} B_{n,m}$, which is clearly borel and satisfies the following:

$$V_n \subseteq \bigcap_{m < \omega} B_{n,m} \subseteq \bigcap_{m < \omega} \mathcal{N} \setminus U_m = \mathcal{N} \setminus \bigcup_{m < \omega} U_m \quad \text{for each } n < \omega$$
$$\therefore \bigcup_{n < \omega} V_n \subseteq \bigcup_{n < \omega} \bigcap_{m < \omega} B_{n,m} = B \subseteq \mathcal{N} \setminus \bigcup_{m < \omega} U_m \dashv$$

To reduce the amount of sequences in play, we will opt to use coding in our projections: $pA = \{x \in \mathcal{N} : \exists y \ (x * y \in A)\}$ for any $A \subseteq \mathcal{N}$, where x * y places x on the evens and y on the odds so for any $x \in \mathcal{N}$ we can say $x = x_{\text{even}} * x_{\text{odd}}$ where $x_{\text{even}} = \langle x(2n) : n < \omega \rangle$ and $x_{\text{odd}} = \langle x(2n+1) : n < \omega \rangle$.

Since $X, Y \in \sum_{1}^{1}$ are projections of closed sets and closed sets are the sets of branches of trees, let $X = \mathfrak{p}[T]$ and $Y = \mathfrak{p}[S]$ for T and S trees over ω . For $\tau \in \omega^{<\omega}$ write $T \upharpoonright \tau$ for $\{t \in T : \tau \leq t \lor t \leq \tau\}$ and X_{τ} for $\mathfrak{p}[T \upharpoonright \tau]$ and similarly for $S \upharpoonright \tau, Y_{\tau}$. We have the following properties that are easy to confirm: for $\tau \in {}^{<\omega}\omega, x \in \mathcal{N}$, and $n < \omega$;

1.
$$T_{\emptyset} = T$$
 and $T_{\tau} = \bigcup_{n < \omega} T_{\tau^{\frown}(n)}$.

2.
$$X_{\emptyset} = X$$
 and $X_{\tau} = \bigcup_{n < \omega} X_{\tau \frown \langle n \rangle}$.

3. $\tau_{\text{even}} \triangleleft x$ for every $x \in X_{\tau}$.

And the above similarly hold for S and Y. Suppose towards a contradiction that $X = \bigcup_{n < \omega} X_{\langle n \rangle}$ \mathcal{B} -sep $\bigcup_{n < \omega} Y_{\langle n \rangle} = Y$ is false. By Claim 1, there must be some $n_0, m_0 < \omega$ where $X_{\langle n_0 \rangle}$ \mathcal{B} -sep $Y_{\langle m_0 \rangle}$ is false. Since $X_{\langle n_0 \rangle} = \bigcup_{n < \omega} X_{\langle n_0, n \rangle}$ and similarly for Y, the same argument yields $n_1, m_1 < \omega$ where $X_{\langle n_0, n_1 \rangle}$ \mathcal{B} -sep $Y_{\langle m_0, m_1 \rangle}$ is false. So by (dependent) choice, we can construct reals $n = \langle n_k : k < \omega \rangle$ and $m = \langle m_k : k < \omega \rangle$ where for each $k < \omega, X_{n \upharpoonright k} \mathcal{B}$ -sep $Y_{m \upharpoonright k}$ is false. This gives a contradiction.

To see this, clearly $n \in [T]$ and $m \in [S]$ yield $x = n_{even} \in X$ and $y = m_{even} \in Y$. Since X and Y are disjoint, $x \neq y$ so for some sufficiently large k, $\mathcal{N}_{x \upharpoonright k} \cap \mathcal{N}_{y \upharpoonright k} = \emptyset$. By (3), $X_{n \upharpoonright 2k} \subseteq \mathcal{N}_{x \upharpoonright k}$ and similarly for Y, meaning that $X_{n \upharpoonright 2k} \subseteq \mathcal{N}_{x \upharpoonright k} \subseteq \mathcal{N} \setminus \mathcal{N}_{y \upharpoonright k} \subseteq \mathcal{N} \setminus Y_{m \upharpoonright 2k}$ and so $\mathcal{N}_{x \upharpoonright k} \in \Sigma_1^0$ separates $X_{n \upharpoonright 2k}$ and $Y_{m \upharpoonright 2k}$, a contradiction to the conclusion above that $X_{n \upharpoonright 2k} \ \mathcal{B}$ -sep $Y_{m \upharpoonright 2k}$ is false. \dashv

22C•9. Corollary —

All Δ_1^1 sets are borel. Hence $\Delta_1^1 = \mathcal{B}$.

Proof .:.

Let $X \in \Delta_1^1$. Since $X, \mathcal{N} \setminus X \in \Sigma_1^1$ are disjoint, there is some borel B with $X \subseteq B \subseteq \mathcal{N} \setminus (\mathcal{N} \setminus X) = X$, meaning X = B is borel. \dashv

The proof of The \sum_{1}^{1} -Separation Principle (22 C • 8) serves as an introduction trees will play in the projective hierarchy. To finish off the basic properties of the projective hierarchy, we also get a closure property under continuous preimages, and in fact a much stronger property.

— 22 C • 10. Definition —

Let $\mathcal{M}, \mathcal{M}'$ be two topologies. A function $f : \mathcal{M} \to \mathcal{M}'$ is *borel* iff for any $Y \in \mathcal{B}^{\mathcal{M}'}, f^{-1} "Y \in \mathcal{B}^{\mathcal{M}}$.

One can easily see that all continuous functions are therefore borel, but there are many borel functions that are not continuous. Still, we have no increase in complexity when taking preimages.

— 22 C • 11. Result -

Let $f : \mathcal{N} \to \mathcal{N}$ be borel and $0 < n < \omega$. Therefore $A \in \sum_{n=1}^{n} \mathbb{I}$ implies $f^{-1} A \in \sum_{n=1}^{n} \mathbb{I}$, and similarly for the other projective pointclasses.

Proof .:.

Proceed by induction on *n*. For n = 1, let $A = \mathfrak{p}B$ for $B \in \prod_{i=1}^{0} B \subseteq \mathcal{N} \times \mathcal{N}$. Consider the set $B' = \{\langle x, y \rangle \in \mathcal{N} \times \mathcal{N} : \langle f(x), y \rangle \in B\}$ which is borel since *f* is borel. Note that then $f^{-1} A = \mathfrak{p}B'$. To see that $\mathfrak{p}B' \in \Sigma_{1}^{1}$, note that $B' \in \Delta_{1}^{1} \subseteq \Sigma_{1}^{1}$, and Σ_{1}^{1} is closed under projections.

For n + 1, $A = \mathfrak{p}B$ for $B \in \prod_{n=1}^{n}$, $B \subseteq \mathcal{N} \times \mathcal{N}$ yields similarly that $B' = \{\langle x, y \rangle \in \mathcal{N} \times \mathcal{N} : \langle f(x), y \rangle \in B\}$ as $\prod_{n=1}^{n}$ inductively. Hence the projection of this $\mathfrak{p}B' = f^{-1} \colon A \in \sum_{n=1}^{n}$. The analogous property for $\prod_{n=1}^{n}$ holds because preimages play nicely with complements, and the result holds for $\Delta_{n=1}^{n}$ just because it holds for the other pointclasses.

In fact, we actually get that each $\sum_{n=1}^{1}$ is closed under borel images in addition to preimages, although this fact is unproven here. Instead, we now prove the long overdue facts about the containments shown in Figure 22 C • 4.

- 22 C • 12. Corollary Let $n < \omega$. Therefore, $\Delta_n^1 \subseteq \Sigma_n^1 \subseteq \Delta_{n+1}^1$, and similarly for \prod_n^1 .

Proof .:.

We always have $\Delta_n^1 \subseteq \Sigma_n^1$ by definition of Δ_n^1 . For the other containment, proceed by induction on *n*. For n = 0, we have $\prod_{i=0}^{1} = \prod_{i=0}^{0} \subseteq \Sigma_1^1$: any closed set *U* is clearly the projection of the closed set $U \times \mathcal{N}$. This implies $\Sigma_0^1 = \Sigma_0^1 \subseteq \Sigma_1^1$ since by Result 22 C • 5, Σ_1^1 is closed under countable unions: $\prod_{i=0}^{0} \subseteq \Sigma_1^1$ implies $\Sigma_2^0 \subseteq \Sigma_1^1$, and thus $\Sigma_1^0 \subseteq \Sigma_1^1$. But if $X \in \Sigma_1^1$ then $\mathcal{N} \setminus X \in \prod_{i=0}^{1}$ so $\prod_{i=0}^{0} \subseteq \Sigma_1^1$ implies $\Sigma_0^1 \subseteq \prod_{i=0}^{1}$ and so $\Sigma_0^1 \subseteq \Delta_1^1$.

For n > 0, let $A \in \prod_{n=1}^{1}$. Clearly $A \times \mathcal{N} \in \prod_{n=1}^{1}$ as the continuous (and hence borel) preimage of A under $\langle x, y \rangle \mapsto x$, showing $A = \mathfrak{p}(A \times \mathcal{N}) \in \sum_{n=1}^{1}$. This shows $\prod_{n=1}^{1} \subseteq \sum_{n=1}^{1}$ and thus $\sum_{n=1}^{1} \subseteq \prod_{n=1}^{1}$. Inductively, $\prod_{n=1}^{1} \subseteq \Delta_{n}^{1} \subseteq \prod_{n=1}^{1}$ so taking projections of these yields $\sum_{n=1}^{1} \subseteq \sum_{n=1}^{1}$ and therefore $\sum_{n=1}^{1} \subseteq \Delta_{n+1}^{1}$. \dashv

To show that all of the containments of Figure 22 C • 4 are actually strict, we need to make use of universal sets, just as we did with the borel hierarchy. Recall that a set U in a pointclass Γ is Γ -universal iff every $A \in \Gamma$ is $\{x : \langle x, r \rangle \in U\}$ for some $r \in \mathcal{N}$. The projective pointclasses Σ_n^1 and Π_n^1 , $n < \omega$, have universal sets similar to before, and these show the strict inequalities.

- 22C·13. Theorem

Let $n < \omega$. Therefore, there is a $\sum_{n=1}^{1} \frac{1}{2}$ -universal set, and similarly a $\prod_{n=1}^{1} \frac{1}{2}$ -universal set.

Proof .:.

Proceed by induction on *n*. For n = 0, there is a $\sum_{0}^{1} = \sum_{1}^{0}$ -universal set by Theorem 22 A • 13. For n + 1, let inductively $W \subseteq \mathcal{N}^{3}$ be $\prod_{n=1}^{1}$ -universal, i.e. for every $B \in \prod_{n=1}^{1}$ with $B \subseteq \mathcal{N} \times \mathcal{N}$, there is some $r \in \mathcal{N}$ where $B = W_r = \{\langle x, y \rangle \in \mathcal{N} \times \mathcal{N} : \langle r, x, y \rangle \in W\}$. Consider

 $U = \mathfrak{p}W = \{ \langle r, x \rangle \in \mathcal{N} \times \mathcal{N} : \exists y \in \mathcal{N} \ (\langle r, x, y \rangle \in W) \}.$

We will have that is \sum_{n+1}^{1} -universal. To see this, clearly $U \in \sum_{n+1}^{1}$ as the projection of a $\prod_{n=1}^{1}$ -set. Now for any \sum_{n+1}^{1} -set $A \subseteq \mathcal{N}$, we have $A = \mathfrak{p}B$ for some $\prod_{n=1}^{1}$ -set $B \subseteq \mathcal{N} \times \mathcal{N}$. As W is $\prod_{n=1}^{1}$ -universal, there is some $r \in \mathcal{N}$

where $A = \mathfrak{p}W_r = U_r$. Hence U is \sum_{n+1}^{1} -universal. Also the complement $\mathcal{N} \times \mathcal{N} \setminus U$ is \prod_{n+1}^{1} -universal. \dashv

– 22 C•14. Corollary —

All of the containments of the projective hierarchy are strict: $\Delta_n^1 \subsetneq \Sigma_n^1 \subsetneq \Delta_{n+1}^1$, and similarly for Π_n^1 .

Proof .:.

Firstly, for $n < \omega$ arbitrary, let U be $\sum_{n=1}^{1}$ -universal. Therefore, $U \notin \prod_{n=1}^{1}$. To see this, if U were $\prod_{n=1}^{1}$, then $\mathcal{N} \times \mathcal{N} \setminus U$ would be $\sum_{n=1}^{1}$, and therefore $D = \{x \in \mathcal{N} : \langle x, x \rangle \notin U\}$, the preimage of $\mathcal{N} \times \mathcal{N} \setminus U$ under the continuous map $x \mapsto \langle x, x \rangle$, would also be $\sum_{n=1}^{1}$. As a $\sum_{n=1}^{1}$ -universal set, there must then be some $r \in \mathcal{N}$ where $D = U_r$. Note that $r \in D$ implies $\langle r, r \rangle \notin U$ and so $r \notin U_r = D$, a contradiction. Similarly, $r \notin D$ implies $\langle r, r \rangle \in U$ and so $r \in U_r = D$, another contradiction. Hence $U \notin \prod_{n=1}^{1}$.

This shows $\Delta_n^1 \subsetneq \Sigma_{n+1}^1$ and similarly for Π_n^1 , since the Σ_n^1 and Π_n^1 -universal sets are in $\Sigma_n^1 \setminus \Pi_n^1 = \Sigma_n^1 \setminus \Delta_n^1$ and $\Pi_n^1 \setminus \Delta_n^1$ respectively.

We also immediately get $\sum_{n=1}^{1} \subseteq \Delta_{n+1}^{1}$ and similarly $\prod_{n=1}^{1} \subseteq \Delta_{n+1}^{1}$, since any $\prod_{n=1}^{1}$ -universal set is in Δ_{n+1}^{1} by Corollary 22 C • 12, but isn't in $\sum_{n=1}^{1}$ by the argument above. Similarly, any $\sum_{n=1}^{1}$ -universal set is in $\Delta_{n+1}^{1} \setminus \prod_{n=1}^{1}$.

Let's now reframe some of what we've done so far. Regarding $P \subseteq \mathcal{N}$ as a predicate where P(x) stands for membership $x \in P$, we can write $\neg P$ for $\mathcal{N} \setminus P$. Similarly, we can write $\exists^{\mathcal{N}} P$ for $\mathfrak{p}_{\mathcal{N}} P$ both satisfying

 $x \in \exists^{\mathcal{N}} P \quad \text{iff} \quad \exists y \in \mathcal{N} \ P(x, y).$

We also write $\forall^{\mathcal{N}} P$ for $\neg \exists^{\mathcal{N}} \neg P$, and similarly for other spaces. And of course, for $P, Q \subseteq \mathcal{N}$, we can write $P \land Q$ for $P \cap Q$, and similarly for \lor and \cup . Similarly, for $\Gamma \subseteq \mathcal{P}(\mathcal{N})$, we can write $\exists \Gamma$ for $\{\exists X : X \in \Gamma\}$ and so on. In this way, the various closure properties of Result 22 A • 6 and Result 22 C • 5 can be restated as follows: for $\alpha < \omega_1$ and $n < \omega$,

$\neg \Delta^0_{lpha} = \Delta^0_{lpha}$
${{f \Pi}}^0_lpha\wedge {{f \Pi}}^0_lpha={{f \Pi}}^0_lpha$
$\bigwedge \Pi^{0}_{lpha} = \Pi^{0}_{lpha}$
$i < \omega$
$\bigvee {\mathbf{ \Pi}}{}^{0}_{\alpha} = {\mathbf{ \Sigma}}{}^{0}_{\alpha+1}$
$i < \omega$
$\neg \underline{\mathbf{A}}_n^1 = \underline{\mathbf{A}}_n^1$
$\underline{\Pi}_n^1 \wedge \underline{\Pi}_n^1 = \underline{\Pi}_n^1$
$\underline{\Pi}_n^1 \vee \underline{\Pi}_n^1 = \underline{\Pi}_n^1$
$\forall^{\mathcal{N}} \Pi^0_n = \Pi^0_n$
$\exists^{\mathcal{N}} \prod_{n=1}^{1} = \sum_{n=1}^{1}$

In this way, we may regard sets as relations and predicates and so consider more logical notation. This puts closure under continuous preimages in a slightly different light.

 $\begin{array}{|c|c|c|} \hline & 22 \, \mathbb{C} \cdot 15. \quad \text{Corollary} \\ \hline & \text{Let} \, \alpha < \omega_1. \text{ Suppose } R \subseteq \mathcal{N}^{k+1} \text{ is a } \Sigma^0_{\alpha} \text{-relation and } f : \mathcal{N} \to \mathcal{N} \text{ is continuous. Therefore } \{\langle x_0, \vec{x} \rangle : R(f(x_0), \vec{x}) \} \\ \hline & \text{ is a } \Sigma^0_{\alpha} \text{-relation.} \end{array}$

This motivates calling closure under continuous preimages instead closure under continuous *substitutions*. This also makes it clear why continuous functions are defined in terms of preimages: if P(x) is a predicate defining a simple (i.e. open) set, then the predicate $R(x) \leftrightarrow P(f(x))$ should also define a simple (i.e. open) set.

For the sake of a picture summarizing the containments of the boldface hierarchies, we have the following figure.



22 C • 16. Figure: The borel and projective hierarchies

Section 23. Properties of Sets of Reals

The properties we will consider for now mostly come from older notions when topology was still a major subject studied in its own right. Now topology has essentially become split into many fields where pure topology more or less evolved into the study of independence results around topological statements, and other parts of topology became more frequently used in a variety of places frequently as a framework to study other things in, e.g. algebraic topology, geometry, analysis, and so forth.

We study three fundamental properties each related to different aspects of the study of the real numbers. Each of these properties holds for all of the borel sets, and the question becomes where do they fail? It turns out that the answer is undecidable for each property we investigate. There is some hope, however, since as large cardinal hypotheses give explanations as to *why* these statements are undecidable: they hold of more of the projective hierarchy when we have fairly large large cardinal hypotheses. We will see later that these are all related fundamentally to questions of the determinacy of certain games. Hence the assumptions we have about the larger areas of the universe can have fairly concrete consequences down at the level of the real numbers.

§ 23 A. Perfect sets and trees

We first consider a somewhat natural looking property, being a "perfect" set, and investigate what sets have a subset like this. It turns out that we can prove it happens for all borel sets and the beginning of the projective hierarchy and the question suddenly becomes: does it fail for *any* projective set? This question is unanswerable in ZFC alone, and so marks another point of interest for set theory. In studying these ideas, we will need a great amount of technology about trees and even cardinals, helping connect the almost pure topology described thus far with set theory as we've investigated in previous chapters.

The proof of Theorem 21 B • 7—and subsequently Corollary 21 B • 8—hinged on the idea that we can just keep splitting simple (i.e. open) sets into disjoint sets that are also simple. The result is not just an embedding of \mathcal{N} , but also a way of divvying up the space into isolated points and entire areas that look like \mathcal{N} . This motivates the following concept.

- 23 A \cdot 1. Definition Let \mathcal{M} a be polish space, and let $X \subseteq \mathcal{M}$.

- An element $x \in X$ is *isolated* iff there is some open B around x and $B \cap X = \{x\}$.
- X is *perfect* iff X is closed and contains no isolated points.

For example, over \mathbb{R} , $[0, 1] \cup \{2, 3\}$ is not perfect, because 2 (and also 3) is isolated. On the other hand, [0, 1] is perfect. For a more complicated example, the cantor set^v is perfect.

^vThe cantor set is defined in stages: $C_0 = [0, 1]$, and for C_n already defined, $C_{n+1} = \{x/3 : x \in C_n\} \cup \{(x+2)/3 : x \in C_n\}$; and we define the cantor set to be $\bigcap_{n < \omega} C_n$. If this isn't clear, what's happening is that we are removing the middle thirds of the intervals making up C_n to form C_{n+1} . The cantor set is then the limit of this process.

There are a number of equivalent characterizations of a subset being perfect, mostly a result of on equivalent definitions of $\mathbf{\Pi}_{1}^{0,\mathcal{M}}$ for polish \mathcal{M} . For example, X is closed iff X contains all its limit points, meaning X is perfect iff X is exactly the set of all of its limits of non-constant sequences in $^{\omega}X$.

It should be obvious that no countable polish space has *any* perfect subsets, precisely because all points are isolated. \mathcal{N} , on the other hand, has lots by the following well-known theorem.

— 23 A • 2. Theorem (Cantor–Bendixson Theorem)

Every closed subset of \mathcal{N} is either countable or contains a perfect subset. In fact, any closed set can be uniquely written in the form $X \cup Y$ where X is perfect and Y is countable.

Proof .:.

Let $A \subseteq \mathcal{N}$ be closed. If A is perfect, then we're done. Otherwise let Y_0 be the set of isolated points of A. By separability, Y_0 is countable. As with Corollary 21 B • 8, we can continually remove the isolated points and get a countable set Y such that $X = A \setminus Y$ is closed with no isolated points. To see this, define

$$A_0 = A \qquad A_{\alpha+1} = A_{\alpha} \setminus Y_{\alpha} = A_{\alpha} \setminus \{x \in A_{\alpha} : x \text{ is isolated}\} \qquad A_{\alpha} = \bigcap_{\xi < \alpha} A_{\xi} \text{ for limit } \xi.$$

Inductively, A_0 is closed and $A_{\alpha+1}$ remains closed if A_{α} is closed. To see this, for $B = \{\tau \in {}^{<\omega} : \mathcal{N}_{\tau} \cap Y_{\alpha} = \emptyset\}$,

$$\mathcal{N} \setminus A_{\alpha+1} = (\mathcal{N} \setminus A_{\alpha}) \cap \bigcup_{y \in Y_{\alpha}} (\mathcal{N} \setminus \{y\}) = (\mathcal{N} \setminus A_{\alpha}) \cap \bigcup_{\tau \in B} \mathcal{N}_{\tau} \in \Sigma_{1}^{0}$$

Limit stages take intersections so they remain closed inductively. By the same reasoning in Corollary 21 B • 8, we get $A_{\alpha} = A_{\alpha+1}$ for some $\alpha < \omega_1$ and thus A_{α} is closed with no isolated points. Assuming $A_{\alpha} \neq \emptyset$, this means $A_{\alpha} \subseteq A$ is perfect.

This motivates the following concept.

- 23A·3. Definition Let \mathcal{M} be polish. For $X \subseteq \mathcal{M}$, X has the *perfect set property* iff $|X| \leq \aleph_0$ or X contains a perfect subset. For $\Gamma \subseteq \mathcal{O}(\mathcal{M})$, Γ has the *perfect set property* iff every $X \in \Gamma$ has the perfect set property.

As a result, Cantor-Bendixson Theorem (23 A • 2) says Π_1^0 has the perfect set property. It's also easy to see that all open sets have the perfect set property since each cone \mathcal{N}_{τ} is perfect and any open set contains a cone. Indeed, any perfect set is really just a copy of cantor space.

- 23A·4. Lemma Let \mathcal{M} be a polish space. Let $f : \mathcal{C} \to \mathcal{M}$ be continuous injection as a map from \mathcal{C} , the cantor space $\mathcal{C} = {}^{\omega}2$. Therefore im f is perfect in \mathcal{M} .

Proof .:.

It should be clear that im f has no isolated points, since if $\{f(x)\} \subseteq \mathcal{M}$ is open in im f, then $\{x\}$ would be open in \mathcal{C} , which isn't true as \mathcal{C} has no isolated points (this requires f to be injective). So it suffices to show $C = \operatorname{im} f$ is closed in \mathcal{M} .

Suppose $\vec{y} = \langle y_i : i < \omega \rangle \in {}^{\omega}C$ converges in \mathcal{M} to $y \in \mathcal{M}$. Without loss of generality, $y_i \neq y_j$ for $i \neq j$. We must show $y \in C$. Write $y_i = f(x_i)$ for $x_i \in \tilde{C}$ so that $x_i \neq x_j$ for $i \neq j$. If $\vec{x} = \langle x_i : i < \omega \rangle$ converges to some $x \in C$, then

$$y = \lim_{i \to \infty} f(x_i) = f\left(\lim_{i \to \infty} x_i\right) = f(x)$$

by continuity. So it suffices to show \vec{x} has a convergent subsequence. Suppose not. Therefore, each tail $X_N = \{x_i : N < i < \omega\}$ for $N < \omega$ is a closed subset of C. We also have $C = \bigcup_{n < \omega} C \setminus X_n$ so by compactness (Result 21 B • 10) there is some $N < \omega$ where $C = \bigcup_{n < N} C \setminus X_n$. But $x_N \in C \setminus \bigcup_{n < N} (C \setminus X_n) = \bigcap_{n < N} X_n$, a contradiction. Hence \vec{x} converges to some $x \in C$.

Note that not every perfect set will be a copy of C: continuous images of compact sets like C are compact but there

are non-compact spaces like \aleph that are perfect. Nevertheless, we at least get that every perfect set *contains* a copy of C.

23 A • 5. Lemma

Let \mathcal{M} be a polish space. Let $P \subseteq \mathcal{M}$ be perfect. Therefore there is a closed (and therefore perfect) subset $C \subseteq P$ that (with the inherited topology) is homeomorphic to \mathcal{C} .

Proof .:.

If P is perfect, then it's uncountable in particular. Since perfect sets are closed, there is a continuous injection $f : C \to P$ by Theorem 21 B • 7. Lemma 23 A • 4 tells us $C = \operatorname{im} f$ is perfect. As a result, the inherited topology **C** on C is polish, and therefore f is a continuous bijection. So to show f is a homeomorphism, it suffices to show that $f^{"}U$ is open in **C** whenever U is open in \widetilde{C} , and in fact we only need to show the result for when U is a cone $C_{\tau} = \{x \in C : \tau \lhd x\}$. But this follows from how we originally defined f in Theorem 21 B • 7: $f^{"}C_{\tau} = C \cap \mathcal{M}_{\tau}$ where \mathcal{M}_{τ} was constructed to be open.

The following is the first consequence of the benefit of being able to change the topology and make any borel set clopen.

— 23 A • 6. Corollary –

Every borel subset of a polish space \mathcal{M} has the perfect set property.

Proof .:.

Let $X \subseteq \mathcal{M}$ be borel. By Theorem 22 B•3, there is a topology \mathcal{M}' where X is clopen and $\Sigma_1^{0,\mathcal{M}} \subseteq \Sigma_1^{0,\mathcal{M}'}$. Hence there is a perfect (in \mathcal{M}') subset $P \subseteq X$ which, without loss of generality by Lemma 23 A•5, is (with the inherited topology) homeomorphic to \mathcal{L} by some injective $f : \mathcal{C} \to P$ that is continuous as a map from \mathcal{L} to \mathcal{M}' . Since $\Sigma_1^{0,\mathcal{M}} \subseteq \Sigma_1^{0,\mathcal{M}'}$, the preimages of $\Sigma_1^{0,\mathcal{M}}$ -sets are still open in \mathcal{L} so that f is continuous as a map from \mathcal{L} to \mathcal{M} . By Lemma 23 A•4, P is perfect in \mathcal{M} .

Note that the perfect set property tells us that we won't find a counterexample to CH with such a set.

– 23 A • 7. Result

Suppose X has the perfect set property. Therefore $|X| \leq \aleph_0$ or $|X| = 2^{\aleph_0}$. In particular, X does not witness $\neg CH$.

Proof .:.

If X is uncountable, then the perfect set $Y \subseteq X$ is closed with no isolated points. Inheriting the topology from \mathcal{N} , Y itself is then a polish space with no isolated points and so Theorem 21 B • 7 implies $|Y| \ge 2^{\aleph_0}$. Any polish space has size at most 2^{\aleph_0} , giving $2^{\aleph_0} \le |Y| \le |X| \le 2^{\aleph_0}$ and so equality. \dashv

As a result, any witness to the failure of CH must be *very* complicated, certainly more complicated than any borel set. In fact, we will show that all \sum_{1}^{1} -sets have the perfect set property, but we cannot go beyond this in ZFC alone: measurable cardinals establish the perfect set property for \sum_{2}^{1} -sets, and larger cardinal hypotheses entail the perfect set property for more of the projective hierarchy, but it's consistent relative to ZFC that there is a Δ_{2}^{1} -set without the perfect set property, and $L \models$ "there's a \prod_{1}^{1} -set without the perfect set property". The existence of a set without the perfect set property is provable from ZFC alone (requiring AC), although the resulting set isn't projective.

— 23 A • 8. Result

There is a set $X \subseteq \mathcal{N}$ without the perfect set property.

Proof .:.

There are \aleph_0 -many cones \mathcal{N}_{τ} for $\tau \in {}^{<\omega}\omega$. Open sets of \mathcal{N} have the form $\bigcup_{\tau \in A} \mathcal{N}_{\tau}$ for $A \subseteq {}^{<\omega}2$, and since there are $|\mathcal{P}({}^{<\omega}2)| = 2^{\aleph_0}$ such A, there are 2^{\aleph_0} open sets. It's not difficult to see that there are then (at most) 2^{\aleph_0} perfect sets. So let $\{P_{\alpha} : \alpha < 2^{\aleph_0}\}$ enumerate the (non-empty) perfect sets. Now for $\alpha < 2^{\aleph_0}$, recursively take distinct

$$x_{\alpha}, y_{\alpha} \in P_{\alpha} \setminus \{x_{\xi}, y_{\xi} : \xi < \alpha\}$$

Therefore $B = \{x_{\alpha} : \alpha < 2^{\aleph_0}\} \neq \emptyset$ is distinct from every perfect set because $y_{\alpha} \in P_{\alpha} \setminus B$ so B can't be perfect. Moreover, this also shows no subset of B can be perfect either. B also isn't countable since $|B| = 2^{\aleph_0}$. So B doesn't have the perfect set property.

But the issue is that such a set is very complicated, so much so that it can't in general be placed in our hierarchies because it relies so much on AC. We are more interested in complexity (and later its connection with definability), and so we are interested in which of our pointclasses have the perfect set property. To do this, we want an alternative characterization of perfect subsets in terms of trees.

- 23 A • 9. Definition -

A tree *T* over ω is *perfect* iff it always eventually splits, i.e. for every $\tau \in T$ there are $\sigma_0, \sigma_1 \in T$ with $\tau \leq \sigma_0$ and $\tau \leq \sigma_1$ but $\sigma_0 \not\leq \sigma_1$ and $\sigma_1 \not\leq \sigma_0$.

- 23 A • 10. Result A set $X \subseteq \mathcal{N}$ is perfect iff X = [T] for some perfect tree T over ω .

Proof .:.

 (\rightarrow) Since X is closed, using the tree of approximations from Result 22 • 4,

 $T = \{ \tau \in {}^{<\omega}\omega : \exists x \in X \ (\tau \leq x) \},\$

we get that [T] = X. Moreover, it's not difficult to see that T is perfect just because X has no isolated points. To see this, if T isn't perfect, then there is some $\tau \in {}^{<\omega}\omega$ where all extensions of τ in T are comparable. In particular, there is some $x \in X$ where $\mathcal{N}_{\tau} \cap X = \mathcal{N}_{\tau} \cap [T] = \{x\}$, meaning x is isolated in X, contradicting that X is perfect.

(←) Since X = [T], it follows from Result 22 • 4 that X is closed, so it suffices to show X has no isolated points. Without loss of generality, just by considering the new tree where we remove them, T has no finite branches. So suppose $x \in X$ is isolated. Therefore for some $\tau \triangleleft x$, $\mathcal{N}_{\tau} \cap X = \{x\}$. But then $\tau \in T$ has no incompatible extensions, i.e. every $\sigma_0, \sigma_1 \in T$ with $\tau \triangleleft \sigma_0, \sigma_1$ has $\sigma_0 \triangleleft \sigma_1$ or vice versa, because otherwise we extend σ_0 and σ_1 to elements $y_0, y_1 \in [T]$ where then $y_0, y_1 \in \mathcal{N}_{\tau} \cap X \neq \{x\}$, a contradiction. Hence X has no isolated points and is therefore perfect. \dashv

Our main goal is now to show the following theorem:

- 23A • 11. Theorem - Σ_1^1 has the perfect set property.

To show this, we must expand the kinds of trees we're looking at.

- 23A · 12. Definition Let $X \subseteq \mathcal{N}$ and $\alpha \in \text{Ord.}$ We call $X \alpha$ -suslin iff $X = \mathfrak{p}_{\mathcal{N}}[T]$ for some tree T over $\omega \times \alpha$ (regarding $T \subseteq \bigcup_{n < \omega} {}^{n}\omega \times {}^{n}\alpha$ so that $[T] \subseteq \mathcal{N} \times {}^{\omega}\alpha$).

It's not difficult to see that if X is α -suslin, then X is also β -suslin for every $\beta > \alpha$.

Technically speaking, the above definition doesn't really make sense. In particular, for T a tree on $\omega \times \alpha$ Definition 22 • 3 would say that $T \subseteq {}^{<\omega}(\omega \times \alpha)$. This would imply that the above notation $\mathfrak{p}_{\mathcal{N}}[T]$ makes no sense as elements of T are of the form $\langle \langle n_i, \xi_i \rangle \in \omega \times \alpha : i < N \rangle$ for some $N < \omega$ and hence elements of [T] are in ${}^{\omega}(\omega \times \alpha) \neq {}^{\omega}\omega \times {}^{\omega}\alpha$. To remedy this problem, we will instead identify such sequences of pairs instead with the pair of sequences: $\langle \langle n_i \in \omega : i < N \rangle$, $\langle \xi_i \in \alpha : i < N \rangle \rangle \in {}^N \omega \times {}^N \alpha$. Thus we regard T as a subset of $\bigcup_{n < \omega} {}^n \omega \times {}^n \alpha$ closed under entry-wise initial segments: $\langle \sigma, \tau \rangle \in T$ implies $\langle \sigma \upharpoonright n, \tau \upharpoonright n \rangle \in T$ for every $n < \omega$. When we do this, the notation $\mathfrak{p}[T]$ now makes sense: before $[T] \subseteq {}^{\omega}(\omega \times \alpha)$, but now $[T] \subseteq {}^{\omega}\omega \times {}^{\omega}\alpha = \mathcal{N} \times {}^{\omega}\alpha$. We also commonly identify branches as pairs, because any branch of such a T is of the form $\langle \langle x \upharpoonright n, y \upharpoonright n \rangle \in {}^{<\omega}\omega \times {}^{<\omega}\alpha : n < \omega \rangle$ for some $x \in \mathcal{N}$ and $y \in {}^{\omega}\alpha$, and it's easier to identify this with just $\langle x, y \rangle \in \mathcal{N} \times {}^{\omega}\alpha$.

One result of the axiom of choice is that we can regard every set of reals as α -suslin for some α . This then forms a natural way of categorizing sets of reals according to which α they are α -suslin. In particular, as Result 22 • 4 tells us Π_1^0 -sets are the branches of trees on ω (and so through coding, also trees on $\omega \times \omega$), Σ_1^1 consists of precisely the

 \aleph_0 -suslin sets.

23 A • **13**. Corollary Let $X \subseteq \mathcal{N}$. Therefore $X \in \sum_{i=1}^{1}$ iff X is \aleph_0 -suslin.

— 23 A • 14. Result

Let $X \subseteq \mathcal{N}$. Therefore, X is |X|-suslin. In particular, every subset of \mathcal{N} is 2^{\aleph_0} -suslin.

Proof .:.

Let $f : X \to |X|$ be a bijection. Consider the tree (which is visually more like a bunch of separate lines sharing only the point $\langle \emptyset, \emptyset \rangle$)

 $T = \{ \langle x \upharpoonright n, \operatorname{const}_{f(x)} \upharpoonright n \rangle \in {}^{n}\omega \times {}^{n}|X| : n < \omega \land x \in X \}.$

It's not difficult to see that T is a tree over $\omega \times |X|$. We clearly have $X \subseteq \mathfrak{p}[T]$. The other direction is what makes use of f: any branch of T is of the form $\langle x, \operatorname{const}_{\alpha} \rangle$ for some $x \in \mathcal{N}$ and $\alpha < |X|$ which requires $\alpha = f(x)$. Thus $X = \mathfrak{p}[T]$ is |X|-suslin.

It's a good exercise to note why this *doesn't* work if we naively consider the tree $T' = \{\langle x \upharpoonright n, \text{const}_0 \upharpoonright n \rangle : n < \omega\}$ building up to X. Often times, we are not interested in such general results as Result 23 A • 14, because they don't give *explicit* ways of forming the tree: they rely on this bijection which ignores the complexity of X. Hence we are more interested in the correspondence between complexity and which α s the sets are α -suslin for. We also may place definability restrictions on the sorts of trees we care about.

In L or any other model of CH, it follows that all sets of reals are \aleph_1 -suslin. Therefore we cannot identify the \aleph_1 -suslin sets with any particular pointclass in our hierarchies. Nevertheless, we *can* get partial results that establish that various sets are \aleph_1 -suslin in general. For example, we have the following.

23 A • 15. Result Let $X \subseteq \mathcal{N}$. Therefore $X \in \prod_{1}^{1}$ implies X is \aleph_1 -suslin.

Proof .:.

If $X \in \prod_{i=1}^{1}$, then $\mathcal{N} \setminus X$ is \aleph_0 -suslin, meaning $\mathcal{N} \setminus X = \mathfrak{p}[T]$ for some tree T over $\omega \times \omega$. In particular, $x \in X$ iff there is no infinite branch $\langle x, y \rangle \in \mathcal{N}^2$ of [T]. In particular, if we turn the tree upside down, and restrict our attention to x, we get a well-founded partial order:

$$T_x = \{ \tau \in {}^n \omega : n < \omega \land \langle x \upharpoonright n, \tau \rangle \in T \},\$$

ordered by $\tau \preccurlyeq \sigma$ iff $\sigma \preccurlyeq \tau$. There can be no infinite branch $y \in \mathcal{N}$ of $\langle T_x, \preccurlyeq \rangle$ since otherwise $\langle x, y \rangle \in [T]$ so that $x \in \mathfrak{p}[T] = \mathcal{N} \setminus X$, a contradiction. Stated in another way, $\langle T_x, \preccurlyeq \rangle$ is well-founded. There is then a rank function rank : $T_x \rightarrow \text{Ord}$, i.e. a function where $y \prec z \rightarrow \text{rank}(y) < \text{rank}(z)$ for all $y, z \in T_x$. Since T (and therefore T_x) is countable, we can assume rank : $T_x \rightarrow \omega_1$.

The point of considering T_x is to now build a tree S_T over $\omega \times \omega_1$ of approximations to such rank functions. We define the *shoenfield tree* for T by

 $S_T = \left\{ \langle x, R \rangle \in {}^{<\omega}\omega \times {}^{<\omega}\omega_1 : \forall \tau, \sigma < \mathrm{lh}(x) \ [\tau \preccurlyeq \sigma \land \langle x \upharpoonright \mathrm{lh}(\tau), \tau \rangle \in T \to R(\tau) < R(\sigma)] \right\}.$

By identifying ${}^{<\omega}\omega$ with ω through coding (with $\operatorname{code}(\tau \upharpoonright n) \leq \operatorname{code}(\tau)$ for $\tau \in {}^{<\omega}\omega$ and $n \in \omega$), we don't run into any issues, and may see that $\mathfrak{p}[S] = \mathcal{N} \setminus \mathfrak{p}[T] = X$. Explicitly, $x \in X$ iff $\langle T_x, \leq t \rangle$ is illfounded, which is equivalent to there being a rank function rank : $T_x \to \omega_1$, meaning $\langle x, \operatorname{rank} \rangle \in [S_T]$, i.e. $x \in \mathfrak{p}[S_T]$. \dashv

The tree used for the above proof is very important as we will later study the importance of figuring out which trees have infinite branches through the identification of rank functions on their "upside down" versions.

23A • 16. Definition Let T be a tree over $\omega \times \omega$. The *shoenfield tree* of T, denoted here by S_T , is the tree $S_T = \{ \langle x, R \rangle \in {}^{<\omega}\omega \times {}^{<\omega}\omega_1 : \forall \tau, \sigma < \ln(x) \ [\sigma \lhd \tau \land \langle x \upharpoonright \ln(\tau), \tau \rangle \in T \to R(\tau) < R(\sigma)] \}.$

Firstly, a few remarks about the shoenfield tree. Note that $\sigma \triangleleft \tau \land \langle x \upharpoonright h(\tau), \tau \rangle \in T$ implies $\langle x \upharpoonright h(\sigma), \sigma \rangle \in T$

since T as a tree is closed under entrywise initial segments. So it makes sense to talk about $R(\sigma)$. One might also wonder why we need all of ω_1 . After all, if T has no infinite branches, then any rank function rank : $T \to Ord$ on $\langle T, \triangleleft \rangle$ is actually a function into ω . Presumably, we then just reverse the order and get a rank function for $\langle T, \triangleright \rangle$, rank : $T \to \omega$. The issue with this is that while T may have no infinite branches, the height of T may still be infinite in a complicated way. For example, consider the figure below.



23 A • 17. Figure: A tree T with $[T] = \emptyset$ and no rank function on \hat{T} into ω

In Figure 23 A • 17, there can be no rank function on \hat{T} into ω , precisely because the endpoint x_0 of the tree must have a rank at least as large as any given (finite) branch length. So we would be required to set rank $(x_0) = \omega$ and thus have rank : $T \rightarrow \omega + 1$. We can similarly construct more complicated trees requiring larger (countable) ordinals.

We also then get a structural consequence for $\prod_{i=1}^{1}$ -sets.

— 23 A • 18. Corollary ——

Every $\prod_{i=1}^{1}$ -set is the union of \aleph_1 -many borel sets.

Proof .:.

Let $X \in \prod_{1}^{1}$ so that $\mathcal{N} \setminus X = \mathfrak{p}[T]$ for some T over $\omega \times \omega$. It follows that $X = \mathfrak{p}[S_T]$ where S_T is the shoenfield tree for T as in the proof of Result 23 A • 15. We may restrict our attention from $S_T \subseteq {}^{<\omega}\omega \times {}^{<\omega}\omega_1$ to $S_T \upharpoonright \alpha = S_T \cap ({}^{<\omega}\omega \times {}^{<\omega}\alpha)$ for $\alpha < \omega_1$. In particular, $S_T = \bigcup_{\alpha < \omega_1} S_T \upharpoonright \alpha$. Since T is countable, any rank function for T is bounded in ω_1 so that $[S_T] = \bigcup_{\alpha < \omega_1} [S_T \upharpoonright \alpha]$ and therefore $X = \mathfrak{p}[S_T] = \bigcup_{\alpha < \omega_1} \mathfrak{p}[S_T \upharpoonright \alpha]$. Since each $\alpha < \omega_1$ is countable, we may find a tree S_α over $\omega \times \omega$ with $\mathfrak{p}[S_\alpha] = \mathfrak{p}[S_T \upharpoonright \alpha]$. Thus $\mathfrak{p}[S \upharpoonright \alpha] \in \Sigma_1^1$ with $\mathfrak{p}[S \upharpoonright \alpha] \cap \mathfrak{p}[T] = \emptyset$ for each $\alpha < \omega_1$. Since $\mathfrak{p}[T]$ is also Σ_1^1 , by The Σ_1^1 -Separation Principle (22 C • 8) for each $\alpha < \omega_1$ there is a borel B_α with $\mathfrak{p}[S \upharpoonright \alpha] \subseteq B_\alpha \subseteq \mathcal{N} \setminus \mathfrak{p}[T] = X$. It then follows that $X = \bigcup_{\alpha < \omega_1} B_\alpha$. \dashv

− 23 A • 19. Corollary —

Every \prod_{1}^{1} -set is either countable, of size \aleph_1 , or of size 2^{\aleph_0} .

Proof .:.

We know from Corollary 23 A • 6 that every borel set has the perfect set property which, by Result 23 A • 7, is either countable or of size continuum. As the \aleph_1 -union of borel sets, any given $X \in \prod_{1}^{1}$ has $X = \bigcup_{\alpha < \omega_1} B_{\alpha}$ for borel B_{α} and therefore $|X| \le \aleph_1 \cdot \sup_{\alpha < \omega_1} |B_{\alpha}| \le \aleph_1 \cdot \max\{|B_{\alpha}| : \alpha < \omega_1\}$. Since each $|B_{\alpha}| \in \{n : n < \omega\} \cup \{\aleph_0, 2^{\aleph_0}\}$, the result holds.

We are getting somewhat offtrack with our investigation of \aleph_1 -suslin sets. Let us now think about *why* κ -suslin sets are generally important when thinking about the perfect set property.

23 A • 20. Theorem

Let $\kappa \geq \aleph_0$ be a cardinal. Let $X \subseteq \mathcal{N}$ be κ -suslin with $|X| > \kappa$. Therefore X has the perfect set property.

Proof .:.

It suffices to show X has a perfect subset, and we do this through an argument similar to Result 23 A \cdot 10 and Cantor–Bendixson Theorem (23 A \cdot 2). Firstly, for any given tree, we will remove the parts that yield isolated

points. For T is a tree on $\omega \times \kappa$, assume without loss of generality that T has no finite branches. We set

 $\langle \tau, r \rangle \in \text{prune}(T)$ iff $\exists \langle x_0, R_0 \rangle, \langle x_1, R_1 \rangle \in [T] (x_0 \neq x_1 \land \tau \lhd x_0, x_1 \land r \lhd R_0, R_1)$

It's not difficult to see that prune(T) is a tree as well. Of course, in the process of pruning our tree to remove these isolated branches, we might invariably *introduce* isolated branches because we accidentally removed their neighbors. Hence, we merely keep pruning: for T a tree on $\omega \times \kappa$, define recursively

$$T_0 = T$$

$$T_{\alpha+1} = \text{prune}(T_{\alpha})$$

$$T_{\gamma} = \bigcap_{\alpha < \gamma} T_{\alpha}, \text{ for limit } \gamma$$

We stop the process if $T_{\alpha+1} = T_{\alpha}$ and then define T^* to be T_{α} . Since $|T| \le \kappa$, we can't remove κ^+ elements of T and so this process must eventually stop by some $\alpha < \kappa^+$. With all of this, let $X = \mathfrak{p}[T]$ for $T \kappa$ -suslin. It's not immediately clear whether this process eventually removes everything from T, i.e. whether $T^* = \emptyset$.

- Suppose T^{*} ≠ Ø. Since prune(T^{*}) = T^{*} ≠ Ø, every ⟨τ, r⟩ ∈ T^{*} has two branches, meaning distinct x₀, x₁ ∈ N and R₀, R₁ ∈ ^ωκ where τ ⊲ x₀, x₁ and r ⊲ R₀, R₁. But by ignoring everything else, this means we can find a perfect tree T' over ω where then [T'] ⊆ p[T^{*}] ⊆ X. Explicitly, define by recursion a correspondence between elements of ^{<ω}2 and elements of T^{*}:
 - \emptyset corresponds to $\langle \tau_{\emptyset}, r_{\emptyset} \rangle = \langle \emptyset, \emptyset \rangle \in T^*;$
 - For $\sigma \in {}^{<\omega}\omega$ corresponding to $\langle \tau_{\sigma}, r_{\sigma} \rangle \in T^*$, set $\langle \tau_{\sigma \frown \langle 0 \rangle}, r_{\sigma \frown \langle 0 \rangle} \rangle$ and $\langle \tau_{\sigma \frown \langle 1 \rangle}, r_{\sigma \frown \langle 0 \rangle} \rangle$ to be two incompatible extensions of $\langle \tau_{\sigma}, r_{\sigma} \rangle$ in T^* .

If we consider the tree $T' = \{\rho : \exists \sigma \in {}^{<\omega} 2 \ (\rho \leq \tau_{\sigma})\}$, we get that $[T'] \subseteq \mathfrak{p}[T^*]$ since any branch of T' is of the form $\bigcup_{\sigma \lhd x} \tau_{\sigma}$ for some (any) $x \in \mathcal{N}$ which then has $(\bigcup_{\sigma \lhd x} \tau_{\sigma}, \bigcup_{\sigma \lhd x} r_{\sigma}) \in [T^*]$. Since T' is perfect by construction, [T'] is a perfect subset of X.

• Suppose $T^* = \emptyset$. This will contradict that $|X| > \kappa$. To see this, for each $\langle x, R \rangle \in [T]$, there must then be a maximal $\xi < \alpha$ where $\langle x, R \rangle \in [T_{\xi}]$ but $\langle x, R \rangle \notin [T_{\xi+1}]$. Each where then we removed some initial segment: there is an $n < \omega$ where $\langle x \upharpoonright n, R \upharpoonright n \rangle \in T_{\xi+1}$ but $\langle x \upharpoonright (n+1), R \upharpoonright (n+1) \rangle \notin T_{\xi+1}$. But then, looking back at the definition of prune (T_{ξ}) , there is only one branch extending $\langle x \upharpoonright n, R \upharpoonright n \rangle$ in T_{ξ} . As a result, each branch in $[T_{\xi}] \setminus [T_{\xi+1}]$ can be identified with a single element of $T_{\xi} \setminus T_{\xi+1}$. Since each T_{ξ} has size $\leq \kappa$, it follows that $|[T_{\xi}] \setminus [T_{\xi+1}]| \leq \kappa$ and therefore $X = \bigcup_{\xi < \alpha} \mathfrak{p}([T_{\xi}] \setminus [T_{\xi+1}])$ has size at most $|\alpha| \cdot \kappa = \kappa$, contradicting that $|X| > \kappa$.

This provides a proof of Theorem 23 A • 11, that \sum_{1}^{1} has the perfect set property, and therefore gives an alternative proof that the borel sets have the perfect set property.

- 23A·21. Corollary All \aleph_0 -suslin sets have the perfect set property, i.e. $\sum_{n=1}^{1}$ -sets have the perfect set property.

§23 B. Lebesgue measure

Usually the first experience students have with borel sets and σ -algebras occurs in an analysis course with the goal of defining the lebesgue integral—named after Henri Lebesgue—and generally studying measure spaces. This leads to a number of other (sub)fields of mathematics like ergodic theory. For our purposes, we are interested in the connection between the projective hierarchy and lebesgue measure, wondering which projective pointclasses^{vi} consist of lebesgue measurable sets.

Unfortunately, presenting lebesgue measure on \mathcal{N} is tedious and opaque, to say the least, compared to the usual presentation on \mathbb{R} . So A few words should be said on why it ultimately makes no difference. Firstly, we have the following very important theorem that unfortunately has a very long proof (if we include the relevant lemmas) of little interest to us. Curious readers can read the details in classic books like [18] (Theorem 15.6) or [23] (Theorem 1G.4). The basic

vi All the borel sets are lebesgue measurable almost immediately and so all that remains are the projective pointclasses.

idea is a more topological version of Cantor-Bernstein (5 C \cdot 4) using injections from \mathcal{N} and from the other given polish space.

Let \mathcal{M} be an uncountable polish space. Therefore there is a bijection $f: \mathcal{N} \to \mathcal{M}$ such that $f'' X \subseteq \mathcal{M}$ is borel iff $X \subset \mathcal{N}$ is borel.

Although f might not be continuous, meaning $f^{-1}X$ might not be in \sum_{α}^{0} for $X \in \sum_{\alpha}^{0,\mathcal{M}}$, we still have a way of associating the borel sets of \mathcal{M} with the borel sets of \mathcal{N} . In this way, the borel sets of any two polish spaces "behave" in the same sort of way, and subsequently their projective hierarchies do too. And defining something like lebesgue measure on $[0, 1] \subseteq \mathbb{R}$ with the standard topology defines a similar notion on \mathcal{N} . More precisely, we have the following non-trivial theorem in [18] (Theorem 17.41).

23 B • 2. Theorem (Measure Isomorphism Theorem)

Let \mathcal{M} be an uncountable polish space. Suppose $\mu_{\mathcal{M}}$ is a non-trivial probability measure on \mathcal{M} over the borel sets of \mathcal{M} . Therefore there is a bijection $f:[0,1] \to \mathcal{M}$ such that, writing μ for lebesgue measure (restricted to subsets of [0, 1]);

- $f''X \subseteq \mathcal{M}$ is borel iff $X \subseteq [0, 1]$ is borel; and
- $\mu_{\mathcal{M}}(f''X) = \mu(X)$ for all borel $X \subseteq [0, 1]$.
- In particular, $X \subseteq [0, 1]$ is μ -null (i.e. contained in a borel set of μ -measure 0) iff f''X is $\mu_{\mathcal{M}}$ -null.

In short, the above theorem tells us that all borel probability spaces on uncountable polish spaces are borel isomorphic.vii Questions about which projective sets are lebesgue measurable will have the same answer across all (uncountable) polish spaces because the projective pointclasses are all closed under images and preimages of borel functions.viii

Note that we will be using some common properties about \mathbb{R} without reference. Most of these can be found in any basic analysis book, and indeed most calculus books. Write \mathbb{R}^+ for $[0,\infty) = \{x \in \mathbb{R} : 0 \le x\}$. We frequently think of ∞ as an object where $x \in \mathbb{R} \leftrightarrow x < \infty$ so that sup $\mathbb{R} = \sup \mathbb{N} = \infty$, for example.

- If x : ω → ℝ⁺, then the sum ∑_{n=0}[∞] x(n) doesn't depend on the order: ∑_{n<ω} x(n) makes sense. Moreover, ∑_{n<ω} x(n) = sup_{N<ω} ∑_{n<N} x(n).
 Similarly if x, y : ω → ℝ⁺, then ∑_{n<ω} x(n) + y(n) = ∑_{n<ω} x(n) + ∑_{n<ω} y(n).
- If for every $\varepsilon > 0, r \le x + \varepsilon$ then $r \le x$.
- If sup $X < \infty$ for $X \subseteq \mathbb{R}$, then for every $\varepsilon > 0$ there is some $x \in X$ with sup $X \leq x + \varepsilon$, and similarly for $\inf X$.

Let us first define what measures and outer-measures are before talking about lebesgue measure on R.

- 23 B • 3. Definition

An *outer-measure* on a set \mathcal{M} is a function $\mu^* : \mathcal{P}(\mathcal{M}) \to \mathbb{R}^+ \cup \{\infty\}$ such that

• $0 = \mu^*(\emptyset) \le \mu^*(X) \le \mu^*(Y)$ for all $X \subseteq Y \subseteq \mathcal{M}$; and

• (sub-additivity) $\mu^*(\bigcup_{n < \omega} X_n) \le \sum_{n < \omega} \mu^*(X_n)$ where each $X_n \subseteq \mathcal{M}$.

A measure on \mathcal{M} over a σ -algebra $\Sigma \subset \mathcal{P}(\mathcal{M})$ is a function $\mu : \Sigma \to \mathbb{R}^+ \cup \{\infty\}$ such that

• $0 = \mu(\emptyset) \le \mu(X) \le \mu(Y)$ for all $X \subseteq Y \subseteq \mathcal{M}$; and

• (\aleph_0 -additivity) $\mu(\bigcup_{n < \omega} X_n) = \sum_{n < \omega} \mu(X_n)$ where the X_n s are disjoint subsets of \mathcal{M} .

We call μ a *probability* measure iff $\mu(\mathcal{M}) = 1$. Call a subset $X \subseteq \mathcal{M} \mu^*$ -null iff $\mu^*(X) = 0$.

Some simple examples of measures are the following:

• (Trivial) For any set X, const_0 $\upharpoonright \mathcal{P}(X)$ —the constant 0 function—is a measure on X over $\mathcal{P}(X)$.

viiIt's also possible to transform any given measure on a polish space into a probability measure in a way that preserves which sets are null and so which sets are measurable.

viiiNote that it's consistent relative to large cardinals (in particular, the existence of a "real-valued measurable cardinal" $< 2^{\infty_0}$) that there's a measure defined on all of $\mathcal{P}(\mathbb{R})$. Such a measure won't be translation invariant of course by the proof that there exist non-lebesgue measurable sets. But moreover this measure isn't an issue for Measure Isomorphism Theorem (23 B • 2): the measures will disagree on weird, non-borel, non-null sets.

- (Trivial) For any $X \neq \emptyset$ and $x \in X$, $\mu(Y) = 1$ iff $x \in Y$ defines a probability measure on X over $\mathcal{P}(X)$. This corresponds to a principal ultrafilter.
- (Non-trivial) More generally, if U is a countably complete ultrafilter over X, then there is a measure on X defined by μ(Y) = 1 iff Y ∈ U and otherwise μ(Y) = 0. In particular, a measurable cardinal gives such a measure over 𝒫(κ).
- (Trivial) The *counting measure* defined by $\mu(Y) = |Y|$ (writing $\mu(Y) = \infty$ for $|Y| = \aleph_0$) is a measure on $\{Y \subseteq X : |Y| \le \aleph_0\}$ for any X.

All of the measures labelled "trivial" have something in common.

— 23 B•4. Definition	
We call a measure on a	set X over a σ -algebra Σ non-trivial iff $\mu(X) \neq 0$ and $\mu(\{x\}) = 0$ for every $x \in X$.

This then allows us to interpret Measure Isomorphism Theorem (23 B \cdot 2) so long as we define lebesgue measure. To do this, we really need to define the σ -algebra of sets we will be working with, which comes just as the largest σ -algebra that works.

The important point is that outer-measures allow us to consider more subsets than measures do. On the other hand, measures have more nice properties than outer-measures. For example, for μ a measure, $\mu(A \setminus B) = \mu(A) - \mu(B)$ whenever $B \subseteq A$. Similarly, $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$, as expected from basic probability. Outer-measures, on the other hand, might not have this property and can have sets A and X where $\mu^*(A) \neq \mu^*(A \cap X) + \mu^*(A \setminus X)$. Luckily, the collection of X where this doesn't happen *does* form σ -algebra that the outer-measure is actually a measure on.

- 23 B • 5. Definition Let μ^* be an outer measure on \mathcal{M} . Define $X \subseteq \mathcal{M}$ to be μ^* -measurable iff for all $A \subseteq \mathcal{M}$, $\mu^*(A) = \mu^*(A \cap X) + \mu^*(A \setminus X)$.

- 23 B • 6. Result

Let μ^* be an outer measure on \mathcal{M} . Therefore the collection of μ^* -measurable sets is a σ -algebra $\Sigma \subseteq \mathcal{P}(\mathbb{R})$ such that $\mu^* \upharpoonright \Sigma$ is a measure on \mathcal{M} .

Proof .:.

There are two parts to this:

- 1. that the collection Σ of μ^* -measurable sets is indeed a σ -algebra; and
- 2. $\mu^* \upharpoonright \Sigma$ is a measure on \mathcal{M} .

The first of these is relatively straightforward although quite tedious.

- 1. Clearly Σ is closed under complements. Thus it suffices to that Σ is closed under countable unions. To do this, we first show closure under finite unions. Suppose $X, Y \in \Sigma$. For any $A \subseteq M$, firstly note that
 - $(A \cap (X \cup Y)) \cap Y = (A \cap X \cap Y) \cup ((A \cap Y) \setminus X) = (A \cap X \cap Y) \cup ((A \setminus X) \cap Y)$; and

•
$$(A \cap (X \cup Y)) \setminus Y = (A \cap X) \setminus Y$$

Therefore, when we break apart $\mu^*(A)$, we get

 $\mu^*(A) = \mu^*(A \cap X) + \mu^*(A \setminus X)$ = $\left[\mu^*(A \cap X \cap Y) + \mu^*((A \setminus X) \cap Y)\right] + \mu^*((A \cap X) \setminus Y) + \mu^*(A \setminus (X \cup Y))$ = $\left[\mu^*((A \cap (X \cup Y)) \cap Y) + \mu^*((A \cap (X \cup Y)) \setminus Y)\right] + \mu^*(A \setminus (X \cup Y))$ = $\mu^*(A \cap (X \cup Y)) + \mu^*(A \setminus (X \cup Y)).$

This shows closure under finite unions which implies $\mu^*(A \cap (X \cup Y)) = \mu^*(A \cap X) + \mu^*(A \cap Y)$ for disjoint $X, Y \in \Sigma$ and any $A \subseteq \mathcal{M}$. This also generalizes to finite unions of disjoint sets.

For infinite unions, let $X_n \in \Sigma$ for $n < \omega$. Note that instead considering $X_n \setminus \bigcup_{m < n} X_m \in \Sigma$ yields the same union but now all the sets are disjoint, so we assume this. Write $X_{< N}$ for $\bigcup_{n < N} X_n$ for $N \le \omega$ and similarly for $X_{>N}$. By the finite union case, $\mu^*(A \cap X_{< N}) = \sum_{n < N} \mu^*(A \cap X_n)$ for any $N < \omega$ and

PROPERTIES OF SETS OF REALS

$$A \subseteq \mathcal{M}. \text{ It's also not hard to see that } A \setminus X_{<\omega} \subseteq A \setminus X_{
$$\mu^*(A) = \mu^*(A \cap X_{
$$\therefore \mu^*(A) \ge \sum_{n < \omega} \mu^*(A \cap X_n) + \mu^*(A \setminus X_{<\omega}) \ge \mu^*(A \cap X_{<\omega}) + \mu^*(A \setminus X_{<\omega}) \ge \mu^*(A).$$$$$$

This tells us that $X_{<\omega} \in \Sigma$ and therefore Σ is a σ -algebra.

2. That $\mu^*(\emptyset) = 0$ and $\mu^*(X) \le \mu^*(Y)$ holds by μ^* being an outer-measure. So it suffices to show $\mu^*(\bigcup_{n < \omega} X_n) = \sum_{n < \omega} \mu^*(X_n)$ whenever $X_n \in \Sigma$ has $X_n \cap X_m = \emptyset$ for $n, m < \omega$. One inequality holds by sub-additivity, and for the other, note that

$$\mu^*\left(\bigcup_{n<\omega} X_n\right) \ge \sup_{N<\omega} \mu^*\left(\bigcup_{n< N} X_n\right) = \sup_{N<\omega} \sum_{n< N} \mu^*(X_n) = \sum_{n<\omega} \mu^*(X_n). \quad \exists$$

We also get a kind of converse: every measure defines an outer measure.

– 23 B • 7. Result

Let μ be a measure on a set \mathcal{M} over a σ -algebra Σ . Define $\mu^*(X) = \inf\{\mu(Y) : X \subseteq Y \in \Sigma\}$. Therefore μ^* is an outer-measure on \mathcal{M} with $\mu^* \upharpoonright \Sigma = \mu$. Moreover, all null sets are μ^* -measurable.

Proof .:.

That $\mu^* \upharpoonright \Sigma = \mu$ is immediate, since $\mu(X) \le \mu(Y)$ for $X, Y \in \Sigma$ with $X \subseteq Y \in \Sigma$, meaning $\mu(X) \le \mu^*(X)$, and obviously $X \in \Sigma$ yields $\mu^*(X) \le \mu(X)$.

That $\mu^*(X) \leq \mu^*(Y)$ for $X \subseteq Y$ is immediate: if $X \subseteq Y$ then $\{Z \in \Sigma : Y \subseteq Z\} \subseteq \{Z \in \Sigma : X \subseteq Z\}$ and therefore taking the infimum yields $\mu^*(X) \leq \mu^*(Y)$. It's also clear that $\mu^*(\emptyset) = 0$ since μ is a measure. Thus it suffices to check sub-additivity. Let $X_n \subseteq \mathcal{M}$ for $n < \omega$. For each $\varepsilon > 0$, let $Y_{n,\varepsilon} \in \Sigma$ be such that $\mu(Y_{n,\varepsilon})$ is within $\varepsilon/2^{n+1}$ of $\mu^*(X_n)$, i.e. $\mu(Y_{n,\varepsilon}) \leq \mu^*(X_n) + \varepsilon/2^{n+1}$. So using the \aleph_0 -additivity and hence sub-additivity of μ , for each $\varepsilon > 0$,

$$\mu^*\left(\bigcup_{n<\omega}X_n\right) \le \mu\left(\bigcup_{n<\omega}Y_{n,\varepsilon}\right) \le \sum_{n<\omega}\mu(Y_{n,\varepsilon}) \le \sum_{n<\omega}\mu^*(X_n) + \sum_{n<\omega}\frac{\varepsilon}{2^{n+1}} = \sum_{n<\omega}\mu^*(X_n) + \varepsilon.$$

Taking the infimum over such $\varepsilon > 0$ gives the desired inequality.

To see that $\mu^*(X) = 0$ implies X is μ^* -measurable, for any $A \subseteq \mathcal{M}, \mu^*(A) \leq \mu^*(A \cap X) + \mu^*(A \setminus X)$. Since $\mu^*(A \cap X) \leq \mu^*(X) = 0$, it follows that $\mu^*(A) \leq \mu^*(A \setminus X) \leq \mu^*(A)$ and therefore we have equality: $\mu^*(A) = \mu^*(A \cap X) + \mu^*(A \setminus X)$ and therefore X is μ^* -measurable.

This outer-measure has some nice properties that allows us to expand any given measure to one that's "complete" in the sense that if $\mu(X) = 0$ and $Y \subseteq X$ then $\mu(Y) = 0$. It also gives an alternative and conceptually better characterization of a set being measurable.

– 23 B•8. Theorem

Let μ be a measure on a set \mathcal{M} over a σ -algebra Σ . Let μ^* be the outer-measure derived from μ , i.e. $\mu^*(X) = \inf\{\mu(Y) : X \subseteq Y \in \Sigma\}$. Let Λ be the σ -algebra of μ^* -measurable sets. Therefore $\Lambda \supseteq \Sigma$ is the largest σ -algebra where $\mu^* \upharpoonright \Lambda$ is a measure on \mathcal{M} .

Proof .:.

Result 23 B • 7 tells us that μ^* is indeed an outer-measure, and Result 23 B • 6 tells us that Λ is indeed a σ -algebra where $\mu^* \upharpoonright \Lambda$ is a measure with $\Sigma \subseteq \Lambda$. So it suffices to show Λ is the largest such: let $\Lambda \subseteq \Omega$ be a σ -algebra where $\mu^* \upharpoonright \Omega$ is a measure. Let $X \in \Omega$.

Since $\Sigma \subseteq \Lambda \subseteq \Omega$, for any $U \in \Sigma$, $U \setminus X$, $U \cap X \in \Omega$ and therefore $\mu^*(U) = \mu^*(U \cap X) + \mu^*(U \setminus X)$. In particular, for each $\varepsilon > 0$, let $U_{\varepsilon} \in \Sigma$ such that $X \subseteq U_{\varepsilon}$ and $\mu^*(U_{\varepsilon})$ is within ε of $\mu^*(X)$, implying $\mu^*(U_{\varepsilon} \setminus X) < \varepsilon$. As a result, $\mu^*(\bigcap_{n < \omega} U_{1/n} \setminus X) = \inf\{\mu^*(U_{1/n} \setminus X) : n < \omega\} = 0$ and therefore $\bigcap_{n < \omega} U_{1/n} \setminus X \in \Lambda$. As a σ -algebra, $\bigcap_{n < \omega} U_{1/n} \in \Sigma \subseteq \Lambda$ and therefore the relative complement $X \in \Lambda$, implying $\Omega \subseteq \Lambda$.

In particular, if we can ensure that we have an outer-measure defined on the borel sets of a polish space \mathcal{M} (with \mathcal{M} having measure 1), then by Measure Isomorphism Theorem (23 B • 2), there is a unique σ -algebra of so-called lebesgue measurable sets on \mathcal{M} . But doing this is fairly easy if we just define the outer-measure over the basic open sets, and then expand to an outer-measure. For \mathbb{R} , lebesgue outer-measure is typically defined as follows.

23 B • 9. Definition For $a < b \in \mathbb{R}$, define $\mu^*([a, b]) = b - a$. For general $X \subseteq \mathbb{R}$, define *lebesgue outer-measure* μ^* by $\mu^*(X) = \inf \left\{ \sum_{I \in \mathcal{I}} \sup(I) - \inf(I) : \mathcal{I} \text{ is a countable collection of closed intervals } A \subseteq \bigcup \mathcal{I} \right\}.$ Call a subset $X \subseteq \mathbb{R}$ *lebesgue measurable* iff it is μ^* -measurable. We often call \mathcal{I} a *cover* for X whenever $X \subseteq \bigcup \mathcal{I}$.

It's not difficult to see that the two definitions for μ^* on intervals coincide. This gives an outer-measure where all basic open sets are pretty easily seen to be μ^* -measurable. Hence all borel sets are μ^* -measurable and so μ^* is a measure called *lebesgue measure* over the *lebesgue measurable* sets. In showing this, we must actually confirm that the name "lebesgue outer-measure" isn't stupid.

- 23 B • 10. Result

Lebesgue outer-measure on \mathbb{R} is a non-trivial outer-measure. Moreover, for each $X \subseteq \mathbb{R}$ and $r \in \mathbb{R}$, $\mu^*(X) = \mu^*(\{x + r : x \in X\})$. Furthermore, every interval (i.e. every basic open set) is lebesgue measurable, and therefore all borel sets are lebesgue measurable.

Proof .:.

Let μ^* denote lebesgue outer-measure. That $\mu^*(X) \le \mu^*(Y)$ for $X \subseteq Y$ is clear. Similarly $\mu(\emptyset) \le \mu^*(\{r\}) \le \inf\{\mu^*([r-1/n, r+1/n]) : 0 < n \in \omega\} = \inf\{2/n : 0 < n < \omega\} = 0$ yields that μ^* is non-trivial. So we must check sub-additivity and translation invariance (that $\mu^*(X) = \mu^*(X + r)$ where $X + r = \{x + r : x \in X\}$).

• Clearly if $\sum_{n < \omega} \mu^*(X_n)$ is infinite, then the desired inequality holds. So assume it's finite and therefore we can find covers for each X_n within ε of the outer-measure of X_n for each $n < \omega$. More precisely, for $\varepsilon > 0$ and $n < \omega$ let X_n be covered by intervals $X_n \subseteq \bigcup_{m < \omega} I_{n,m}$ that are close to $\mu^*(X_n)$ in the sense that $\sum_{m < \omega} \mu^*(I_{n,m}) \le \mu^*(X_n) + \frac{\varepsilon}{2^n}$. Therefore $\{I_{n,m} : n, m < \omega\}$ is a covering of $\bigcup_{n < \omega} X_n$ where

$$\sum_{m<\omega}\mu^*(I_{n,m})=\sum_{n<\omega}\sum_{m<\omega}\mu^*(I_{n,m})\leq\sum_{n<\omega}\left(\mu^*(X_n)+\frac{\varepsilon}{2^n}\right)\leq\sum_{n<\omega}\mu^*(X_n)+2\varepsilon.$$

The infimum over such ε yields $\mu^*(\bigcup_{n < \omega} X_n) \leq \sum_{n < \omega} \mu^*(X_n)$.

- For translation invariance, we get the desired inqualities just by shifting over the relevant covers: $X \subseteq \bigcup_{n < N \le \omega} I_n$ iff $X + r \subseteq \bigcup_{n < N} (I_n + r)$ with $\mu^*(I_n + r) = \sup(I_n + r) \inf(I_n + r) = \sup(I_n) + r (\inf(I_n) + r) = \sup(I_n) \inf(I_n) = \mu^*(I_n)$. Therefore all of the relevant sums are the same and so $\mu^*(X) = \mu^*(X + r)$.
- Note that μ*([a, b)) = μ*((a, b)) = μ*([a, b]) = a b so it makes no difference whether we use closed, open, or half-open intervals. Moreover, if I and J are intervals, μ*(J ∩ I) + μ*(J \ I) = μ*(J). So let I be an interval and A ⊆ ℝ. Let I_ε be a countable collection of closed intervals with ∑_{J∈I_ε} μ*(J) within ε of μ*(A) and A ⊆ ∪ I_ε. Note that {J ∩ I : J ∈ I_ε} covers A ∩ I and {J \ I : J ∈ I_ε} covers A \ I and moreover

$$\mu^*(A) \le \mu^*(A \cap I) + \mu^*(A \setminus I) \le \sum_{J \in \mathcal{I}_{\mathcal{E}}} \mu^*(J \cap I) + \mu^*(J \setminus I) = \sum_{J \in \mathcal{I}_{\mathcal{E}}} \mu^*(J) \le \mu^*(A) + \varepsilon. \quad \exists$$

— 23 B•11. Lemma

If $X \subseteq \mathbb{R}$ is lebesgue measurable, then

- 1. $\mu(X) = \inf\{\mu(C) : X \subseteq C \land C \text{ is open}\}.$
- 2. $\mu(X) = \sup\{\mu(C) < \infty : C \subseteq X \land C \text{ is closed}\}.$

Proof .:.

Write I(X) for $\inf\{\mu^*(C) : X \subseteq C \land C \text{ is open}\}$ and S(X) for $\sup\{\mu^*(C) < \infty : C \subseteq X \land C \text{ is closed}\}$.

1. We clearly have $\mu^*(X) \leq I(X)$ since $X \subseteq C$ implies $\mu^*(X) \leq \mu^*(C)$. For the other inequality, for each

 $\varepsilon > 0$, let $\{I_{\varepsilon,n} : n < \omega\}$ be a set of intervals covering X with measure within ε of $\mu^*(X)$. Without loss of generality, we can ensure the $I_{\varepsilon,n}$ s are all open intervals: if they're all closed, consider $J_{\varepsilon,n}$ to be the open interval from $\inf(I_{\varepsilon,n}) - \frac{\varepsilon}{2^{n+2}}$ to $\sup(I_{\varepsilon,n}) + \frac{\varepsilon}{2^{n+2}}$ so that $\mu^*(J_{\varepsilon,n}) = \mu^*(I_{\varepsilon,n}) + \frac{\varepsilon}{2^{n+1}}$. Therefore taking the infimum over all $\varepsilon > 0$ gives the result:

$$\mu^*(X) \le I(X) \le \sum_{n < \omega} \mu^*(J_{\varepsilon,n}) = \sum_{n < \omega} \mu^*(I_{\varepsilon,n}) + \sum_{n < \omega} \frac{\varepsilon}{2^{n+1}} = \sum_{n < \omega} \mu^*(I_{\varepsilon,n}) + \varepsilon \le \mu^*(X) + 2\varepsilon.$$

2. Clearly $\mu^*(X) \ge S(X)$ since any $C \subseteq X$ has $\mu^*(X) \ge \mu^*(C)$. For the other inequality, note that $[-n, n] \setminus X$ is lebesgue measurable for each $n \in \omega$, and so by (1),

$$2n - \mu^*(X \cap [-n, n]) = \mu^*([-n, n] \setminus X) = I([-n, n] \setminus X)$$

= $\inf\{\mu^*([-n, n] \setminus ([-n, n] \setminus C)) : [-n, n] \setminus C \subseteq X \wedge [-n, n] \setminus C \text{ is open}\}$
= $\inf\{2n - \mu^*(C) : C \subseteq X \cap [-n, n] \wedge C \text{ is closed}\}$
= $2n - \sup\{\mu^*(C) < \infty : C \subseteq X \cap [-n, n] \wedge C \text{ is closed}\}$
= $2n - S(X \cap [-n, n]).$

Therefore $\mu^*(X) = \sup_{n < \omega} \mu^*(X \cap [-n, n]) = \sup_{n < \omega} S(X \cap [-n, n]) = S(X).$

 \neg

– 23 B•12. Lemma -

A set $X \subseteq \mathbb{R}$ lebesgue measurable iff there is a borel (in particular $\sum_{n=1}^{\infty} B \subseteq X$ where $X \setminus B$ is null.

Proof .:.

(\leftarrow) Suppose there is a borel *B* where $\mu(X \setminus B) = 0$. Let $A \subseteq \mathbb{R}$ be arbitrary. Since all borel sets are lebesgue measurable,

 $\mu^*(A \cap X) = \mu^*(A \cap X \cap B) + \mu^*((A \cap X) \setminus B) = \mu^*(A \cap B),$ and also $\mu^*(A \setminus X) \le \mu^*(A \setminus B)$. It follows that $\mu^*(A) \le \mu^*(A \cap X) + \mu^*(A \setminus X) \le \mu^*(A \cap B) + \mu^*(A \setminus B) = \mu^*(A).$

(→) Suppose X is lebesgue measurable. By Lemma 23 B • 11, $X = \sup\{\mu(C) < \infty : C \subseteq X \land C \text{ is closed}\}$. In particular, for $n < \omega$, taking a closed $C_n \subseteq X$ where $\mu(C_n) \le \mu(X \cap [-n,n]) + 1/n$ yields that $B = \bigcup_{n < \omega} C_n \subseteq X$ is borel with $\mu(B) = \lim_{n \to \infty} \mu(C_n) = \mu(X)$ and therefore $\mu(X \setminus B) = 0$. \dashv

The next result is slightly subtle. The point is that we can always find a kind of "minimal" (modulo null sets) measurable superset of any given non-measurable set X: A is measurable and $\mu^*(A) = \mu^*(X)$. This does not, however, imply $\mu^*(A \setminus X)$ is null, since X may not be measurable. Instead the best we can do is show that any measurable set between the two $X \subseteq A' \subseteq A$ has $\mu^*(A \setminus A') = 0$.

- 23 B • 13. Lemma For every $X \subseteq \mathbb{R}$ there is a measurable $A \supseteq X$ where every measurable A' with $X \subseteq A' \subseteq A$ has that $A \setminus A'$ is null.

Proof .:.

By Lemma 23 B • 11, for each $\varepsilon > 0$ let U_{ε} be open such that $X \subseteq U_{\varepsilon}$ and $\mu^*(U_{\varepsilon})$ is within ε of $\mu^*(X)$. It follows that $A = \bigcap_{n < \omega} U_{1/n}$ is borel and hence measurable with $\mu^*(A) = \mu^*(X)$. For any $X \subseteq A' \subseteq A$ with A' measurable, it follows that $\mu^*(A') = \mu^*(A)$ and therefore $\mu^*(A) = \mu^*(A \cap A') + \mu^*(A \setminus A') = \mu^*(A)$ implies $\mu^*(A \setminus A') = 0$.

As a result, using Measure Isomorphism Theorem (23 B • 2)—noting that the preimage of a borel set by a borel function is still borel—we can identify the lebesgue measurable sets as just those with measure 0 difference between a borel set. Now to use Measure Isomorphism Theorem (23 B • 2), we need to actually have an outer-measure on \mathcal{N} where the basic open sets are all measurable. But doing this isn't so difficult. Indeed, for any borel bijection $f : \mathcal{N} \to [0, 1]$ we can just define the measure $\mu_f(X) = \mu(f^{"}X)$. But even explicitly, this isn't too bad. Furthermore, once we have fixed this notion, we are free to identify the resulting (probability) measure as "lebesgue measure", even if it technically differs from the measure on \mathbb{R} or [0, 1]. - 23 B • 14. Definition

We define lebesgue outer-measure μ^* on $\mathcal N$ as follows.

- For $\tau \in {}^{<\omega}\omega$, define $\mu^*(\mathcal{N}_{\tau}) = \prod_{n < \mathrm{lh}(\tau)} \frac{1}{2^{\tau(n)+1}}$.
- For $B \subseteq {}^{<\omega}\omega$ such that $\{\mathcal{N}_{\tau} : \tau \in B\}$ is a collection of disjoint sets, define $\mu^*(\bigcup_{\tau \in B} \mathcal{N}_{\tau}) = \sum_{\tau \in B} \mu^*(\mathcal{N}_{\tau})$.
- For arbitrary $X \subseteq \mathcal{N}$, define $\mu^*(X) = \inf\{\mu^*(A) : X \subseteq A \land A \text{ is open}\}.$

It's not difficult to see that the different definitions are compatible for open and basic open subsets.

– 23 B•15. Corollary –

Lebesgue outer-measure on \mathcal{N} is a non-trivial outer-measure where all borel sets are measurable, \mathcal{N} has measure 1, and the family of lebesgue measurable sets is $\{B \cup N : B \text{ is borel } \land N \text{ is null}\}$.

Proof .:.

That μ^* is an outer-measure is easy from Result 23 B • 10: $0 = \mu^*(\emptyset) \le \mu^*(X) \le \mu^*(Y)$ for $X \subseteq Y$ is clear. Sub-additivity uses the same proof as Result 23 B • 10: let $X_n \subseteq \mathcal{M}$ for $n < \omega$. For $\varepsilon > 0$, let $\mathcal{I}_{\varepsilon,n}$ be an open cover for X_n within $\varepsilon/2^{n+1}$ in measure: $\sum_{I \in \mathcal{I}_{\varepsilon,n}} \mu^*(I) \le \mu^*(X_n) + \frac{\varepsilon}{2^{n+1}}$. It follows that $\bigcup_{n < \omega} \mathcal{I}_{\varepsilon,n}$ is a covering of $\bigcup_{n < \omega} X_n$ with

$$\mu^*\left(\bigcup_{n<\omega}X_n\right)\leq \sum_{n<\omega,I\in\mathcal{I}_{\varepsilon,n}}\mu^*(I)=\sum_{n<\omega}\sum_{I\in\mathcal{I}_{\varepsilon,n}}\mu^*(I)\leq \sum_{n<\omega}\left(\mu^*(X_n)+\frac{\varepsilon}{2^{n+1}}\right)\leq \sum_{n<\omega}\mu^*(X_n)+\varepsilon.$$

The infimum over $\varepsilon > 0$ then yields sub-additivity. In particular, μ^* is an outer-measure.

To see that all borel sets are measurable, it suffices to show that all basic open sets are measurable. Let $\tau \in {}^{<\omega}\omega$ and $A \subseteq \mathcal{N}$ be arbitrary. For $\varepsilon > 0$, let $U \subseteq \mathcal{N}$ be open such that $\mu^*(U_{\varepsilon}) \leq \mu^*(A) + \varepsilon$. Note that $\mathcal{N}_{\tau} \cap U_{\varepsilon}$ and $U_{\varepsilon} \setminus \mathcal{N}_{\tau}$ are disjoint open sets (since \mathcal{N}_{τ} is clopen). It's easy to confirm that any disjoint open sets U, V have $\mu^*(U \cup V) = \mu^*(U) + \mu^*(V)$, so it follows that

$$\mu^*(A) \le \mu^*(A \cap \mathcal{N}_{\tau}) + \mu^*(A \setminus \mathcal{N}_{\tau}) \le \mu^*(\mathcal{N}_{\tau} \cap U_{\varepsilon}) + \mu^*(U_{\varepsilon} \setminus \mathcal{N}_{\tau}) = \mu^*(U_{\varepsilon}) \le \mu^*(A) + \varepsilon$$

Taking the infimum over all $\varepsilon > 0$ yields $\mu^*(A) = \mu^*(A \cap \mathcal{N}_{\tau}) + \mu^*(A \setminus \mathcal{N}_{\tau})$ and so all basic open and therefore borel sets are lebesgue measurable.

Moreover by Measure Isomorphism Theorem (23 B • 2), there is a borel bijection $f : \mathcal{N} \to [0, 1]$ where $\mu^*(X) = \mu(f^*X)$ for all borel $X \subseteq \mathcal{N}$ where μ denotes lebesgue measure on $[0, 1] \subseteq \mathbb{R}$. In particular, $f^*X = B \cup N$ for some borel B and $\mu(N) = 0$. We know N is contained in a borel set N' of μ -measure 0, and therefore $X = f^{-1*}B \cup f^{-1*}N$ with the preimages $f^{-1*}B$ and $f^{-1*}N'$ are both borel, and $f^{-1*}N \subseteq f^{-1}N'$ has μ^* -measure 0 so $f^{-1*}N$ is μ^* -null.

With the concept of the lebesgue measurable sets more-or-less nailed down for \mathcal{N} , we can more easily investigate two major questions:

- What (projective) pointleasses consist of lebesgue measurable sets? And
- What properties do lebesgue measurable sets have?

As should be expected, the answers to these questions are undecidable from ZFC alone. In particular, all $\sum_{n=1}^{1}$ -sets are lebesgue measurable. With stronger large cardinal hypotheses we get that more and more sets are lebesgue measurable: it's consistent for all $\sum_{n=1}^{1}$ -sets to be lebesgue measurable, the existence of an inaccessible cardinal yields the consistency of all $\sum_{n=1}^{1}$ -sets being lebesgue measurable, and generally the existence of $n < \omega$ woodin cardinals implies all $\sum_{n=1}^{1}$ -sets are lebesgue measurable. In particular, with infinitely many woodin cardinals an axiom PD (projective determinacy) holds and *all* projective sets are lebesgue measurable.

Nevertheless, proving that \sum_{1}^{1} -sets are lebesgue measurable is the best we can do, because it's consistent that there is a Δ_{2}^{1} -set which is not lebesgue measurable. In particular, this holds in L. There's a subtlety to this statement, since L may not contain all real numbers nor sets of intervals and thus might get the outer-measure of a set wrong and hence whether it's lebesgue measurable.

Regardless of what happens with the projective hierarchy, under AC there is always a non-lebesgue measurable set.
The issue is that because this relies on a well-ordering of \mathcal{N} or \mathbb{R} , there's ostensibly no way to place such a set in our hierarchy.^{ix} It is consistent relative to the existence of an inaccessible cardinal that *all* sets of reals (even the projective ones) are lebesgue measurable, but AC is false in any model of this.

- 23 B • 16. Result There is a set $X \subseteq \mathcal{N}$ that is not lebesgue measurable.

Proof .:.

It suffices to show a subset of \mathbb{R} that isn't lebesgue measurable as we just translate the result to \mathcal{N} . Consider the equivalence relation $x \approx y$ iff $x - y \in \mathbb{Q}$. Let Vit $\subseteq [0, 1]$ be a set of representatives of the \approx -equivalence classes, i.e. $\forall x \in \mathbb{R} \exists y \in \text{Vit} (x \approx y)$ and $\forall x, y \in \text{Vit} (x \neq y \rightarrow x \not\approx y)$. Suppose Vit (a *Vitali* set, named after Giuseppe Vitali) were lebesgue measurable. Note that $[0, 1] \subseteq \bigcup_{q \in \mathbb{Q} \cap [0, 1]} \text{Vit} + q \subseteq [0, 2]$ and therefore by \aleph_0 -additivity

$$1 = \mu^*([0,1]) \le \sum_{q \in \mathbb{Q} \cap [0,1]} \mu^*(\operatorname{Vit} + q) \le \mu^*([0,2]) = 2,$$

where μ^* is lebesgue outer-measure. By translational invariance, $\mu^*(\text{Vit} + q) = \mu^*(\text{Vit})$ and therefore $1 \leq \sum_{q \in \mathbb{Q} \cap [0,1]} \mu^*(\text{Vit}) \leq 2$. But if $\mu^*(\text{Vit}) \neq 0$ then $\sum_{q \in \mathbb{Q} \cap [0,1]} \mu^*(\text{Vit}) = \infty$, a contradiction; and if $\mu^*(\text{Vit}) = 0$ then $\sum_{q \in \mathbb{Q} \cap [0,1]} \mu^*(\text{Vit}) = 0$, another contradiction. Hence Vit can't be lebesgue measurable.

To show that all \sum_{1}^{1} -sets are lebesgue measurable, we need another way at arriving at \sum_{1}^{1} -sets. We do this through the so-called *suslin* operation \mathcal{A}^{x} . It turns out that the \sum_{1}^{1} -sets are applications of operation \mathcal{A} to collections of closed sets and that the collection of lebesgue measurable sets is closed under operation \mathcal{A} , meaning \sum_{1}^{1} -sets are all lebesgue measurable since closed sets are.

- 23 B • 17. Definition Consider a set $X = \{X_{\tau} \subseteq \mathcal{M} : \tau \in {}^{<\omega}\omega\}$ of subsets of a set \mathcal{M} indexed by ${}^{<\omega}\omega$. Define the suslin operation applied to X as $\mathcal{A}X = \bigcup_{x \in \mathcal{N}} \bigcap_{n \in \omega} X_{x \upharpoonright n}$.

– 23 B•18. Lemma ——

A set is Σ_1^1 iff it is $\mathcal{A}X$ for some $X = \{X_\tau : \tau \in {}^{<\omega}\omega\} \subseteq \Pi_1^0$.

Proof .:.

(\leftarrow) Let $X = \{X_{\tau} : \tau \in {}^{<\omega}\omega\}$ be a collection of closed sets. There are then trees $\{T_{\tau} \subseteq {}^{<\omega}\omega : \tau \in {}^{<\omega}\omega\}$ where $X_{\tau} = [T_{\tau}]$. Note that $x \in AX$ iff there is some $y \in \mathcal{N}$ where $x \in \bigcap_{n < \omega} X_{y \mid n}$ iff $x \in {}_{n < \omega}[T_{y \mid n}]$. In particular, the following is a tree over $\omega \times \omega$

$$T = \left\{ \langle \tau, \sigma \rangle \in {}^{<\omega}\omega \times {}^{<\omega}\omega : \mathrm{lh}(\tau) = \mathrm{lh}(\sigma) \wedge \tau \in \bigcap_{n < \mathrm{lh}(\sigma)} T_{\sigma \upharpoonright n} \right\}$$

It follows that for $x, y \in \mathcal{N}, \langle x, y \rangle \in [T]$ iff $x \in [T_{y \upharpoonright n}]$ for every $n < \omega$. In particular, $\mathcal{A}X = \mathfrak{p}[T]$ is Σ_1^1 .

 (\rightarrow) Let $A \in \sum_{i=1}^{1}$ as witnessed by $A = \mathfrak{p}[T]$ for some tree T over $\omega \times \omega$. We therefore have $x \in A$ iff $\exists y \in \mathcal{N} \ \forall n < \omega \ (\langle x \upharpoonright n, y \upharpoonright n \rangle \in T)$. For each $\sigma \in {}^{<\omega}\omega$ consider the tree over ω

 $T_{\sigma} = \left\{ \tau \in {}^{<\omega}\omega : \exists \sigma' \ \left((\sigma \triangleleft \sigma' \lor \sigma' \triangleleft \sigma) \land \langle \tau, \sigma' \rangle \in T \right) \right\}.$

It follows that each $[T_{\sigma}]$ is closed. Moreover if $A = \mathcal{A}\{[T_{\sigma}] : \sigma \in {}^{<\omega}\omega\}$. To see this, if $x \in A$ then there is some $y \in \mathcal{N}$ where $\langle x, y \rangle \in [T]$ so that for every $n < \omega, x \upharpoonright n \in T_{y \upharpoonright n}$ (and therefore $x \upharpoonright n \in T_{y \upharpoonright m}$ for all $n, m < \omega$). This implies $x \in [T_{y \upharpoonright n}]$ for all $n < \omega$ and therefore $x \in \mathcal{A}\{[T_{\sigma}] : \sigma \in {}^{<\omega}\omega\}$. The converse holds in the same way.

Thus we only need to show that the lebesgue measurable sets are closed under the suslin operation A. The reader may

^{ix}There's no explicit way to well-order \mathcal{N} under ZFC. The best explanation that can be given at this point is that if we could get a projective well-ordering, we'd get one over just ZF, but ZF + \neg AC is consistent, and the usual model for demonstrating this has no well-ordering for \mathcal{N} in particular. For another argument, the skeptical reader is dared to just try. See what happens. I dare you.

^xAs far as I know, Mikhail Suslin (Михаил Суслин) proposed the notation ' \mathcal{A} ', which is written in honor of Pavel Aleksandrov (Павел Александров) who eventually took to calling the sets given by this—i.e. the $\sum_{i=1}^{1}$ -sets— \mathcal{A} -sets and from here the name "analytic" seems to be derived and used in part due to Nikolai Luzin (Николай Лузин). I am not a fan of the term "analytic" because it can be confused with "the analytical hierarchy", which is a related, but totally distinct concept.

note that in fact $\sum_{i=1}^{1}$ is closed under the suslin operation \mathcal{A} as well, though this is unnecessary for showing the lebesgue measurability for $\sum_{i=1}^{1}$ -sets.

- 23 B • 19. Lemma
Let
$$X = \{X_{\tau} : \tau \in {}^{<\omega}\omega\}$$
 be a set of lebesgue measurable sets. Therefore AX is lebesgue measurable.

Proof .:.

Without loss of generality, we can assume $X_{\tau} \subseteq X_{\sigma}$ whenever $\sigma \leq \tau$ just by replacing X_{τ} with $\bigcap_{n < \ln(\tau)} X_{\tau \upharpoonright n}$, because doing yields the same result when applying the suslin operation \mathcal{A} . Write μ^* for lebesgue outer-measure on \mathcal{N} . For each $\tau \in {}^{<\omega}\omega$, define

$$A_{\geq \tau} = \bigcup_{x \in \mathcal{N}_{\tau}} \bigcap_{n \in \omega} X_{x \upharpoonright n}.$$

In particular, $A_{\geq \emptyset} = AX$. By our adjustment at the start of the proof, $A_{\geq \sigma} \subseteq X_{\sigma}$. By Lemma 23 B • 13, for each $\tau \in {}^{<\omega}\omega$ there are measurable sets $B_{\geq \tau}$ where (just by intersecting with X_{τ}) $A_{\geq \tau} \subseteq B_{\geq \tau} \subseteq X_{\tau}$ and such that $B_{\geq \tau} \setminus B$ is null for any measurable B with $A_{\geq \tau} \subseteq B \subseteq B_{\geq \tau}$. As a result, for $\sigma \leq \tau$ and $n < \omega$, we have all the following containments:

 $X_{\tau} \subseteq X_{\sigma}$ $A_{\geq \tau} \subseteq A_{\geq \sigma}$ $B_{\geq \tau} \subseteq B_{\geq \sigma} \subseteq X_{\sigma}$ and $\forall Y$ measurable $(A_{\geq \tau} \subseteq Y \subseteq B_{\geq \tau} \to \mu^*(B_{\geq \tau} \setminus Y) = 0)$. Since $A_{\geq \tau} = \bigcup_{n < \omega} A_{\geq \tau^{\frown}(n)}$, by analogy $N_{\tau} = B_{\geq \tau} \setminus \bigcup_{n < \omega} B_{\geq \tau^{\frown}(n)}$ is null. By countable additivity the union $\bigcup_{\tau \in <\omega} N_{\tau}$ is null. Define

$$B = B_{\geq \emptyset} \setminus \bigcup_{\tau \in {}^{<\omega} \omega} N_{\tau}$$

so that *B* is measurable. We will see that $B \subseteq AX$, implying $AX \setminus B \subseteq \bigcup_{\tau \in {}^{<\omega} \omega} N_{\tau}$ is null and therefore AX is measurable. To see that $B \subseteq AX$, let $x \in B$ so that $x \in B_{\geq \emptyset} \setminus N_{\emptyset}$, i.e. $x \in B_{\geq \langle n \rangle}$ for some $n \in \omega$. So we construct a $y \in \mathcal{N}$ recursively: take y(0) to be such an *n*, and define y(n + 1) = m to be such that $x \in B_{\geq y \mid n} \setminus N_{y \mid n}$. Such an *m* exists because $x \in B_{\geq y \mid n} \setminus N_{y \mid n}$. In particular, such a *y* witnesses that

$$x \in \bigcap_{n < \omega} B_{\geq y \restriction n} \subseteq \bigcap_{n < \omega} X_{y \restriction n} \subseteq \mathcal{A}X.$$

– 23 B • 20. Corollary -

All $\sum_{n=1}^{1}$ -sets are lebesgue measurable.

Proof .:.

All Π_1^0 -sets are lebesgue measurable so by Lemma 23 B • 19, for any $X = \{X_\tau : \tau \in {}^{<\omega}\omega\} \subseteq \Pi_1^0, AX$ is lebesgue measurable. By Lemma 23 B • 18, every Σ_1^1 -set has this form and therefore is lebesgue measurable. \dashv

§23 C. The baire property

The last idea we will consider for now is the *baire property*, named after René-Louis Baire just as with baire space. Unlike lebesgue measure, the baire property involves topology more than any analysis of \mathbb{R} , and the ideas can be stated generally for topological spaces. The idea involves the topological notion of "category", distinct from the algebraic sense as in category theory. Category talks about how a set is built up from "small" sets. What sets are small isn't determined by a sense of measure, where we look at how the set can be covered, but instead by how "dense" the set is. The motivating result is the following result on \mathbb{R} .

— 23 C • 1. Theorem (The Baire Category Theorem) —

For $n < \omega$, let $D_n \in \sum_{i=1}^{n} be a \text{ dense subset of } \mathcal{N}$. Therefore $\bigcap_{n < \omega} D_n \neq \emptyset$. In particular, \mathcal{N} is not the countable union of the complements of open, dense sets.

Proof .:.

It's easy to see that a dense subset of \mathcal{N} is any set $D \subseteq \mathcal{N}$ such that for any $\tau \in {}^{<\omega}\omega$, there is an $x \in \mathcal{N}$ where $\tau \triangleleft x \in D$. Firstly, note that the intersection of any two dense open sets D, D' of \mathcal{N} is dense and open: that $D \cap D'$ is open is clear. For density, suppose $\tau \in {}^{<\omega}\omega$. There is then an element $x \in \mathcal{N}_{\tau} \cap D$. As an open set, there is then a τ' with $\tau \triangleleft \tau'$ and $\mathcal{N}_{\tau'} \subseteq D$ so that there is then an element $x' \in \mathcal{N}_{\tau'} \cap D'$. It follows that $\tau \triangleleft x' \in D \cap D'$ and therefore $D \cap D'$ is dense as τ was arbitrary.

So define by recursion the sequence $\tau_n \leq \tau_{n+1}$ for $n \in \omega$ in the same way. Set $\tau_0 < x$ for any $x \in D_0$. For τ_n already defined, let $x \in \bigcap_{i < n+1} D_i \cap D_{i+1}$ which is open and dense by the argument above. In particular, there is some τ such that $\tau_n \leq \tau$ and $\mathcal{N}_{\tau} \subseteq \bigcap_{i \leq n+1} D_i$, and we set τ_{n+1} to be such a τ . It follows that $\bigcup_{n < \omega} \tau_n \in \mathcal{N}$ is an element of $\bigcap_{n < \omega} D_n$.

One might think the only open dense subsets of \mathcal{N} would be \mathcal{N} itself. Ostensibly, (rephrasing things in terms of \mathbb{R}) one could take an open interval around each point and seemingly cover everything because the set is dense. But this isn't the case: consider $\mathcal{N} \setminus \{x\}$ for any $x \in \mathcal{N}$, or $(-\infty, x) \cup (x, \infty)$ for any $x \in \mathbb{R}$. The only way for an open set dense to necessarily be the whole space is for every element to be isolated, meaning the space must be discrete.

Let us introduce some very common concepts in topology that we will use throughout this subsection.

- 23 C • 2. Definition Let \mathcal{M} be a topological space. For any $X \subseteq \mathcal{M}$, define the following:

• The closure of X is $cl(X) = \bigcap \{C : X \subseteq C \text{ is closed} \}.$

• The *interior* of X is $int(X) = \bigcup \{ U \subseteq X : U \text{ is open} \} = \mathcal{M} \setminus cl(\mathcal{M} \setminus X).$

• The boundary of X is $\partial X = \operatorname{cl}(X) \setminus \operatorname{int}(X) = \operatorname{cl}(X) \cap \operatorname{cl}(\mathcal{M} \setminus X)$.

We have some very easy properties about these new concepts. In particular,

- int(X) is always open, and $cl(X) = \mathcal{M} \setminus int(\mathcal{M} \setminus X)$ is always closed.
- $\operatorname{int}(X) \subseteq X \subseteq \operatorname{cl}(X) = \operatorname{int}(X) \cup \partial X$.
- If X is open, int(X) = X; and if X is closed, cl(X) = X so that $\partial X \subseteq X$ in this case.
- Similarly, X is open iff $\partial X \cap X = \emptyset$.
- $\partial X \subseteq X$ implies $cl(X) \subseteq X$ and therefore X is closed.

In general, there's no connection between ∂X and X, and we might have $\partial X \cap X$ range anywhere from \emptyset to X itself.

– 23 C • 3. Definition —

- Let \mathcal{M} be a topological space.
 - A set $N \subseteq \mathcal{M}$ is nowhere dense iff $\mathcal{M} \setminus N$ contains a dense and open set.
 - A set $M \subseteq \mathcal{M}$ is *meagre* iff M is the countable union of nowhere dense sets.
 - A set $X \subseteq \mathcal{M}$ has the *baire property* iff there is some $U \in \sum_{i=1}^{0,\mathcal{M}}$ where $X \bigtriangleup U = (X \setminus U) \cup (U \setminus X)$ is meagre.

Thus The Baire Category Theorem (23 C • 1) is saying that \mathcal{N} is not meagre. It should be clear that \mathcal{N} has the baire property since \mathcal{N} itself is open and $\mathcal{N} \bigtriangleup \mathcal{N} = \emptyset$ is clearly meagre. Similarly, any set of isolated points is nowhere dense and therefore meagre. Every open set and every meagre set has the baire property. In fact, the sets with the baire property form a σ -algebra which then contains all borel sets. Before establishing this, however, let's prove some basic facts about meagre and nowhere dense sets.

— 23 C • 4. Lemma

Let \mathcal{M} be a topological space. Therefore, for $X, Y \subseteq \mathcal{M}$;

- 1. If X is nowhere dense and $Y \subseteq X$ then Y is nowhere dense.
- 2. In particular, if X is meagre and $Y \subseteq X$, then Y is meagre.
- 3. The finite intersection of open dense sets is open dense. In particular, the union of two nowhere dense sets is nowhere dense.
- 4. The countable union of meagre sets is meagre.

5. For any open set U, $cl(U) \setminus U$ is nowhere dense.

Proof .:.

- 1. Suppose X is nowhere dense so that $\mathcal{M} \setminus X$ contains a dense and open set $D \subseteq \mathcal{M} \setminus X$. It's clear that $D \subseteq \mathcal{M} \setminus X \subseteq \mathcal{M} \setminus Y$ so that Y is also nowhere dense.
- 2. If X is meagre, then write $X = \bigcup_{n < \omega} X_n$ where each X_n is nowhere dense. Hence $Y \subseteq X$ has $Y = \bigcup_{n < \omega} X_n \cap Y$ where $X_n \cap Y \subseteq X_n$ is also nowhere dense by (1).
- 3. Let D_X and D_Y be open dense sets. Therefore $D_X \cap D_Y$ is open. For density, let $U \subseteq \mathcal{M}$ be open so that $D_X \cap U$ is also open (since D_X is) and non-empty (since D_X is dense). By the density of D_Y , $D_Y \cap D_X \cap U \neq \emptyset$. Since U was arbitrary, $D_Y \cap D_X$ is dense. In particular, if X, and Y are nowhere dense with $D_X \cap X = D_Y \cap Y = \emptyset$, then $(D_X \cap D_Y) \cap (X \cup Y) = \emptyset$ with $D_X \cap D_Y$ open and dense. In other words, $X \cup Y$ is nowhere dense.
- 4. Suppose $X_n = \bigcup_{m < \omega} X_{n,m}$ is meagre for all $n < \omega$ where each $X_{n,m}$ is nowhere dense. Therefore $\bigcup_{n < \omega} X_n = \bigcup_{n,m < \omega} X_{n,m}$ is the countable union of nowhere dense sets and is therefore meagre.
- 5. The complement of $cl(U) \setminus U$ is just $U \cup (\mathcal{M} \setminus cl(U))$. It's clear that cl(U), as the intersection of closed sets, is closed and therefore $U \cup (\mathcal{M} \setminus cl(U))$ is open. It's also dense because any counterexample is a non-empty open $V \subseteq \mathcal{M}$ with $V \cap U = \emptyset$ and $V \cap (\mathcal{M} \setminus cl(U)) = \emptyset$, i.e. $V \subseteq cl(U) \setminus U$. But then $cl(U) \setminus V$ is a closed set with $U \subseteq cl(U) \setminus V$. But $V \neq \emptyset$ then contradicts the definition of cl(U). So no such V can exist and therefore $U \cup (\mathcal{M} \setminus cl(U))$ is open and dense, i.e. $cl(U) \setminus U$ is nowhere dense. \dashv

- 23 C • 5. Result

Let \mathcal{M} be a topological space. Therefore $\{X \subseteq \mathcal{M} : X \text{ has the baire property}\}$ is a σ -algebra containing all open sets, and hence every borel set has the baire property.

Proof .:.

All open sets have the baire property. Now suppose X_n for $n < \omega$ has the baire property as witnessed by an open U_n with meagre $X_n \Delta U_n$. Note that $\bigcup_{n < \omega} U_n$ is open with

$$\left(\bigcup_{n<\omega}X_n\right) \bigtriangleup \left(\bigcup_{n<\omega}U_n\right) \subseteq \bigcup_{n<\omega}(X_n \bigtriangleup U_n),$$

which is the countable union of meagre sets. By Lemma 23 C • 4 (4) and (2), $(\bigcup_{n < \omega} X_n) \bigtriangleup (\bigcup_{n < \omega} U_n)$ is meagre and therefore $\bigcup_{n < \omega} X_n$ has the baire property. Therefore it suffices to show the set of subsets with the baire property is closed under complements.

Suppose X has the baire property. To show that $\mathcal{M} \setminus X$ does too, let U be open with $X \bigtriangleup U$ meagre. Note that the closure of $cl(U) \setminus U$ is nowhere dense by Lemma 23 C • 4 (5). In particular,

$$(\mathcal{M} \setminus X) \bigtriangleup (\mathcal{M} \setminus \mathrm{cl}(U)) = X \bigtriangleup \mathrm{cl}(U) = ((X \bigtriangleup U) \setminus (X \cap \mathrm{cl}(U))) \cup (\mathrm{cl}(U) \setminus U).$$

Since $X \triangle U$ is meagre, the subset $(X \triangle U) \setminus (X \cap cl(U))$ is too. Since $cl(U) \setminus U$ is meagre, and the union of two meagre sets is also meagre, it follows that $\mathcal{M} \setminus X$ is meagre. \dashv

To explain our terminology a little, a nowhere dense set N is clearly not dense since otherwise the complement $\mathcal{M} \setminus N$ would contain an open dense set $D \subseteq \mathcal{M} \setminus N$ whose complement $\mathcal{M} \setminus D \supseteq N$ is a *closed* set that is still dense because it contains a dense set N. But the only closed dense set is the whole space itself, meaning $\mathcal{M} \setminus D = \mathcal{M}$ requires D to be empty and therefore not dense, a contradiction.

More than just not being dense, however, a nowhere dense set isn't dense in *any* (open) subset's inherited topology. This then motivates the name.

— 23 C • 6. Result -

Let \mathcal{M} be a topological space. Therefore $X \subseteq \mathcal{M}$ is nowhere dense iff for every open set $U \subseteq \mathcal{M}, X \cap U$ is not dense in the inherited topology on U.

Proof .:.

- (\rightarrow) Suppose X is nowhere dense but $X \cap U$ is dense in an open set U, meaning $X \cap V \neq \emptyset$ for every open $V \subseteq U$. Let $D \subseteq \mathcal{M} \setminus X$ be an open dense subset of \mathcal{M} so that $D \cap U$ is also open and dense in U. Since the complement $U \setminus D$ is closed (in the inherited topology) and contains a dense set, it follows that $U \setminus D = U$ contradicting the density of D in U.
- (\leftarrow) Suppose $X \cap U$ is not dense for every open set $U \subseteq \mathcal{M}$. Consider D as the interior of $\mathcal{M} \setminus X$, i.e. $D = \bigcup \{U \subseteq \mathcal{M} \setminus X : U \text{ is open}\}$. It follows that $D \subseteq \mathcal{M} \setminus X$ is open. Suppose D is not dense. This means for some $U \subseteq \mathcal{M}, U \cap D = \emptyset$ so that $D \cup U \subseteq \mathcal{M} \setminus X$ is a larger open set than D, contradicting the definition of D.

The name "meagre" invokes a sense a smallness. In a space where there are no isolated points, every set can be written as the union of some collection of nowhere dense sets (considering unions of singletons). Since nowhere dense sets are closed under finite unions, meagre sets mark the first new stage, being the result of countable unions. As a result, meagre sets are sometimes referred to as of the "first category". One can similarly define second cateogry just as the sets which aren't of the first category. The analysis of meagre sets and nowhere dense sets in general is what is sometimes meant by the study of "category" in topology and is why it appears in the name of The Baire Category Theorem $(23 \text{ C} \cdot 1)$.



23 C • 7. Figure: Containments between the categorical and measure-theoretic properties

Now despite meagre sets and lebesgue measure 0 sets both being a sense of "smallness", the two notions do not coincide at all. In particular, we can partition \mathcal{N} itself into the union of a meagre set $M \in \sum_{2}^{0}$ and a lebesgue null set $N \in \prod_{2}^{0}$. Since \mathcal{N} isn't meagre, N can't be meagre; and since \mathcal{N} has measure one, M isn't lebesgue null. Indeed, it's possible to have a nowhere dense set with positive measure. Figure 23 C • 7 gives the containments provable in ZFC between these notions. In particular, we give short descriptions for non-empty sets in each region above, demonstrating all strict containments in a non-trivial way. A reader interested in these sets can look at [6] and elsewhere for their definitions and properties:

- let B be any bernstein set (cf. Result 23 A 8);
- let Non be the set of non-normal numbers in [0, 1];
- let C be the cantor set and FC the fat cantor set; and
- let NBC be any non-borel subset of C (there are $2^{2^{\aleph_0}}$ subsets of C and only 2^{\aleph_0} -many borel ones).

$a. (B \cap Non) \cup [1, 2];$	к. ℝ;
3. NBC \cup [1, 2];	λ . An $N \in \prod_{i=1}^{0}$ where $\mathbb{R} = M \cup N$
$Y. \ \mathbf{B} \cup \mathbf{Non};$	with M meagre and N null;
$\delta. \ [0,1] \cap \mathbf{B} \setminus \mathrm{Non}$	μ. Q ;
E. $FC \cap B$;	v. C;
ζ . FC \ NBC;	ξ. NBC $∪$ \mathbb{Q} ;
ן. NBC $\cup \mathbb{Q} \cup [1, 2];$	o. NBC;
Э. FC;	<i>π</i> . Non;
1. FC $\cup \mathbb{Q}$;	ρ . B \cap Non.

Nevertheless, the two notions *act* similarly at a low level and often arguments regarding one can be easily translated into the other just by replacing "measure 0" with "meagre" or vice versa. In particular, with an analog of Lemma 23 B • 13, we can follow the proof of Lemma 23 B • 19 to get that all $\sum_{i=1}^{1}$ -sets have the baire property.

- 23C·8. Lemma For every $X \subseteq \mathcal{N}$ there is a $B \supseteq X$ with the baire property such that for any B' with the baire property and $X \subseteq B' \subseteq B, B \setminus B'$ is meagre.

Proof .:.

Define $Y = \bigcap \{\mathcal{N} \setminus \mathcal{N}_{\tau} : \tau \in {}^{<\omega} \omega \land \mathcal{N}_{\tau} \cap X \text{ is meagre} \}$. Perhaps a confusing definition, Y is sort-of the closure of X modulo meagre sets: whenever $\mathcal{N}_{\tau} \cap X$ is meagre, we remove \mathcal{N}_{τ} , constantly chipping away until we arive at Y. Note that Y is closed as the intersection of closed sets. We don't necessarily have that $X \subseteq Y$, but we at least have that $X \setminus Y$ is meagre, since $X \setminus Y = \bigcup \{X \cap \mathcal{N}_{\tau} : \tau \in {}^{<\omega} \omega \land \mathcal{N}_{\tau} \cap X \text{ is meagre} \}$ is the countable union of meagre sets.

Consider $B = X \cup Y = (X \setminus Y) \bigtriangleup Y$. Note that B has the baire property: $X \setminus Y$ is meagre and Y is closed so that $\mathcal{N} \setminus B = (\mathcal{N} \setminus Y) \bigtriangleup (X \setminus Y)$ which has the baire property so by Result 23 C • 5, B does too.

To show the "minimality" of B, let $X \subseteq B' \subseteq B$ where B' has the baire property. We want to show that $B \setminus B'$ is meagre. We know $B \setminus B'$ has the baire property so there $B \setminus B' = U \bigtriangleup M$ for some open U and meagre M so that $U = (B \setminus B') \bigtriangleup M \subseteq (B \setminus X) \cup M$. In particular, $U \cap X \subseteq M$ is meagre.

To see that $U \setminus X \subseteq M$, proceed as follows. Since $\emptyset \setminus X \subseteq M$ trivially, suppose U is non-empty. As an open set there is some $\tau \in {}^{<\omega}\omega$ where $\mathcal{N}_{\tau} \subseteq U \subseteq (B \setminus X) \cup M$. Since $\mathcal{N}_{\tau} \cap X \subseteq U \cap X$ is meagre, it follows that $Y \subseteq \mathcal{N} \setminus \mathcal{N}_{\tau}$ and hence $\mathcal{N}_{\tau} \subseteq \mathcal{N} \setminus Y$. As the union of τ with $\mathcal{N}_{\tau} \subseteq U$, it follows that $U \subseteq \mathcal{N} \setminus Y$. As a result, $(U \setminus X) \cap (B \setminus X) = (U \setminus X) \cap Y \subseteq (\mathcal{N} \setminus (Y \cup X)) \cap Y = \emptyset$. Since $U \subseteq (B \setminus X) \cup M$, $U \subseteq M$ is meagre and therefore $B \setminus B' \subseteq U \cup M$ is meagre.

The above proof was not particularly similar to Lemma 23 B \cdot 13, but nevertheless, we may practically copy and paste the proof of Lemma 23 B \cdot 19 to get that the subsets with the baire property are closed under the suslin operation A, merely replacing "measurable" with "has the baire property" and "null" with "meagre".

- 23C•9. Lemma Let $X = \{X_{\tau} : \tau \in {}^{<\omega}\omega\}$ be such that each X_{τ} has the baire property. Therefore $\mathcal{A}X$ has the baire property.

Proof .:.

Write BP(Z) for "Z has the baire property". Without loss of generality, we can assume $X_{\tau} \subseteq X_{\sigma}$ whenever $\sigma \leq \tau$ just by replacing X_{τ} with $\bigcap_{n < \ln(\tau)} X_{\tau \upharpoonright n}$, because doing yields the same result when applying the suslin operation \mathcal{A} . For each $\tau \in {}^{<\omega}\omega$, define

$$A_{\geq \tau} = \bigcup_{x \in \mathcal{N}_{\tau}} \bigcap_{n \in \omega} X_{x \upharpoonright n}.$$

In particular, $A_{\geq \emptyset} = AX$. By our adjustment at the start of the proof, $A_{\geq \sigma} \subseteq X_{\sigma}$. By Lemma 23 C • 8, for each $\tau \in {}^{<\omega}\omega$ there are sets $B_{\geq \tau}$ such that

PROPERTIES OF SETS OF REALS

Ch IV §23 C

- $BP(B_{\geq \tau});$
- $B_{\geq \tau} \setminus B$ is meagre for any B with BP(B); and
- (just by intersecting with X_{τ}) $A_{\geq \tau} \subseteq B_{\geq \tau} \subseteq X_{\tau}$.

As a result, for $\sigma \leq \tau$ and $n < \omega$, we have all the following:

- 1. $X_{\tau} \subseteq X_{\sigma};$
- 2. $A_{\geq \tau} \subseteq A_{\geq \sigma}$;
- 3. $B_{\geq \tau} \subseteq B_{\geq \sigma} \subseteq X_{\sigma}$; and
- 4. $\forall Y (\mathsf{BP}(Y) \land A_{\geq \tau} \subseteq Y \subseteq B_{\geq \tau} \to B_{\geq \tau} \setminus Y \text{ is meagre}).$

Since $A_{\geq \tau} = \bigcup_{n < \omega} A_{\geq \tau^{\frown}(n)}$, one can easily check that $N_{\tau} = B_{\geq \tau} \setminus \bigcup_{n < \omega} B_{\geq \tau^{\frown}(n)}$ is meagre. By countable additivity the union $\bigcup_{\tau \in {}^{<\omega}\omega} N_{\tau}$ is meagre. Define

$$B = B_{\geq \emptyset} \setminus \bigcup_{\tau \in {}^{<\omega} \omega} N_{\tau}$$

so that BP(B). We will see that $B \subseteq AX$, implying $AX \setminus B \subseteq \bigcup_{\tau \in {}^{<\omega}\omega} N_{\tau}$ is meagre and therefore BP(AX). To see that $B \subseteq AX$, let $x \in B$ so that $x \in B_{\geq \emptyset} \setminus N_{\emptyset}$, i.e. $x \in B_{\geq \langle n \rangle}$ for some $n \in \omega$. So we construct a $y \in \mathcal{N}$ recursively: take y(0) to be such an n, and define y(n + 1) = m to be such that $x \in B_{\geq y \upharpoonright n^{\frown}(m)}$. Such an m exists because $x \in B_{> v \upharpoonright n} \setminus N_{v \upharpoonright n}$. In particular, such a y witnesses that

$$x \in \bigcap_{n < \omega} B_{\geq y \upharpoonright n} \subseteq \bigcap_{n < \omega} X_{y \upharpoonright n} \subseteq \mathcal{A}X.$$

- 23 C • 10. Corollary Every $\sum_{n=1}^{1}$ -set has the baire property.

Proof .:.

All open sets have the baire property, and by Result 23 C • 5, all closed sets do too. The closure of \prod_{1}^{0} under the suslin operation \mathcal{A} yields all \sum_{1}^{1} -sets, and since the collection of subsets with the baire property is closed under this, all \sum_{1}^{1} -sets have the baire property.

Of course, not every set has the baire property as the same non-lebesgue measurable set from Result 23 B • 16 shows.

Proof .:.

Let Vit be the set of equivalence classes under the equivalence relation $x \approx y$ iff $x - y \in \mathbb{Q}$. For $A \subseteq \mathbb{R}$ and $x \in \mathbb{R}$, write A + x for $\{a + x : a \in A\}$. Again, $\mathbb{R} = \bigcup_{q \in \mathbb{Q}} \text{Vit} + q$ is the disjoint union of countably many sets which can't all be meagre by The Baire Category Theorem (23 C • 1). But any Vit + q is meagre iff Vit is, and so Vit isn't meagre.

If Vit has the baire property, then there's an open set $U \subseteq \mathbb{R}$ and a meagre set M where $U \bigtriangleup M = \text{Vit}$. Take $(a, b) \subseteq U$ for $a < b \in \mathbb{R}$. When translating things around, note that $(a, b) + r \cap (a, b)$ is an open interval whenever the intersection is non-empty (i.e. for r < |b-a|), and similarly, $M \cup M + r$ will be meagre. We have that Vit $+ r \cap \text{Vit} = \emptyset$ and yet for sufficiently small r, as an open interval subtracting a meager set,

$$\emptyset \neq ((a,b) + r \cap (a,b)) \setminus (M + r \cup M) \subseteq \text{Vit} + r \cap \text{Vit},$$

a contradiction. Hence Vit can't have the baire property.

 \dashv

Again, it's not clear that such a set can be placed in our hierarchy. So the question becomes can we do better than $\sum_{i=1}^{1}$ sets? As is usual, we cannot: L believes there's a Δ_{2}^{1} -set without the baire property (and that is not measurable). We can still prove from ZFC that there are sets without the baire property. This is partly because the baire property is related to the determinacy of certain games, and the axiom of determinacy is incompatible with AC. But assuming the weaker PD—which is comptible with AC assuming the consistency of some large cardinal axioms—it follows that all projective sets have the baire property.

Section 24. The Lightface Hierarchies

We will define lightface variants of the borel and projective hierarchies. In general, we say a pointclass is *boldface* iff it's closed under continuous preimages. And so clearly by Result 22 A • 7 and Result 22 C • 11, all the borel and projective pointclasses are boldface. Moreover, for any $\Gamma \subseteq \mathcal{O}(\mathcal{N})$, we can associate a boldface variant $\underline{\Gamma} \subseteq \mathcal{O}(\mathcal{N})$ just by taking the closure under continuous preimages. We really have no need for this added generality, but an equivalent characterization does provide motivation. In particular, we usually think of Γ as being some notion of definability and $\underline{\Gamma} \subseteq \mathcal{O}(\mathcal{N})$ as being the corresponding notion where we allow parameters: $\underline{\Gamma}$ can be thought of as $\bigcup_{x \in \mathcal{N}} \Gamma(x)$, where $\Gamma(x)$ allows x as a parameter; each $X \in \Gamma(x)$ being the preimage of some $Y \in \Gamma$ under the map $y \mapsto \langle x, y \rangle$ or perhaps $y \mapsto \operatorname{code}(x, y) = x * y$.

Much of this section relies on familiarity with computability theory. An overiew can be found in Appendix A. Moreover, these ideas on computability and the *effective* theory of these spaces are further explored in Appendix B.

What exactly is the connection between the borel hierarchy and computability? The traditional understanding tells us that being computable, computably enumerable, or more generally Δ_n^N , Σ_n^N , or Π_n^N for $n < \omega$ is a property of subsets of ω —i.e. of real numbers—not of *sets* of real numbers like being Δ_n^0 , Σ_n^0 , or Π_n^0 . Nevertheless, there is still a notion of computability that generalizes to \mathcal{N} and other polish spaces, giving a computable analogue of their topologies.

Let us state first the corresponding concepts with the boldface hierarchies.

$x \in \sum_{n=1}^{\infty} $	corresponds to	$x \in \Sigma_n^0$
$x \in \sum_{n=1}^{n} \sum_{n=1}^{n}$	corresponds to	$x \in \Sigma_n^1$
continuity	corresponds to	computability
borel	corresponds to	hyperarithmetical
$\bigcup_{n<\omega} \sum_{n=0}^{\infty} \sum_{n=0}^$	corresponds to	arithmetical, i.e. $\bigcup_{n < \omega} \Sigma_n^0$
projective	corresponds to	analytical
F : A		I THE STATE IN THE REPORT OF A DESCRIPTION



We will only explore the beginning of the hyperarithmetical hierarchy, which is just to say Σ_{α}^{0} for $\alpha < \omega$, also called the arithmetical hierarchy. This is partly because the definitions are easier to work with, but also because we have no need for the rest of the hierarchy. Indeed, we rarely have the need to go beyond Σ_{2}^{0} . The curious reader can further explore the hierarchy in Appendix B.

§ 24 A. Generalizing computability

Rather than generalize computability immediately, we instead generalize computably enumerable sets: Σ_1^0 sets in the standard notation from computability theory. Recall that such sets have a semi-decidable procedure for membership. In particular, for $X \subseteq \omega$ with $X \in \Sigma_1^0$, there is some (computable) relation $R \subseteq \omega^2$ that satisfies

$$x \in X$$
 iff $\exists y \in \omega R(x, y)$.

When generalizing this, note that every $X \subseteq \omega$ is open, since any countable polish space is discrete. So we should be looking to find simple open sets as our Σ_1^0 -sets. First consider \mathcal{N} and its basic open sets, which take the form \mathcal{N}_{τ} for some $\tau \in {}^{<\omega}\omega$, which may be regarded essentially as an element of ω just by coding. There is clearly an *effective* test for membership in \mathcal{N}_{τ} because for $x \in \mathcal{N}$, we just check that $x(n) = \tau(n)$ for each $n < \ln(\tau) < \omega$; and if this holds then $x \in \mathcal{N}_{\tau}$ and otherwise $x \notin \mathcal{N}_{\tau}$. So membership in these basic open sets should be considered "computable". Going beyond this, note that open sets are unions of these basic open sets, and so

$$x \in \bigcup_{n < \omega} \mathcal{N}_{\tau_n}$$
 iff $\exists n < \omega \ (x \in \mathcal{N}_{\tau_n}).$

If membership in \mathcal{N}_{τ_n} is seen as "computable", and $n \mapsto \tau_n$ is computable, then it's natural to consider the above as defining a Σ_1^0 -set as with subsets of ω . And the above idea clearly is the same as with ω if we consider the basic open sets to just be singletons, hinting that this is the correct idea to have.

— 24 A • 1. Result

For $m \in \omega$, let $N_m \subseteq \omega$ be the basic open set $\{m\}$ (as the canonical topology on ω is discrete). Let $\emptyset \neq X \subseteq \omega$. Therefore X is Σ_1^0 in the computability sense iff $X = \bigcup_{n < \omega} N_{f(n)}$ for some computable function $f : \omega \to \omega$.

Proof .:.

A non-empty subset of ω is Σ_1^0 iff it's the image of some computable function f so that im $f = \bigcup_{n < \omega} N_{f(n)}$.

We basically are assuming that the basic open sets are *simple* in some sense. As countable polish spaces are discrete, we let singletons be the basic open sets, which are certainly simple.

— 24 A • 2. Definition -

Let \mathcal{M} be a polish space with basic open sets $\{\mathcal{M}_n : n < \omega\}$. Define $X \subseteq \mathcal{M}$ to be $\Sigma_1^{0,\mathcal{M}}$ iff $X = \emptyset$ or else $X = \bigcup_{n < \omega} \mathcal{M}_{f(n)}$ for some $f : \omega \to \omega$ that is computable. We write just Σ_1^0 for $\Sigma_1^{0,\mathcal{N}}$.

For product spaces $\mathcal{M} \times \mathcal{W}$ where \mathcal{W} has basic open sets $\{\mathcal{W}_n : n < \omega\}$, the basic open sets of the product space will be rectangles: $\{\mathcal{M}_{n_0} \times \mathcal{W}_{n_1} : n_0, n_1 < \omega\}$.

Naturally, $\Sigma_1^{0,\mathcal{M}}$ sets depends on the presentation and basic sets we've chosen.^{xi} That said, we will mostly be focused on \mathcal{N}, ω , and their products, whose basic open sets are somewhat canonical: cones and singletons respectively. Result 24 A • 1 then shows this is the same as the traditional definition for computability: there's no confusion between Σ_1^0 in the computability sense and $\Sigma_1^{0,\omega}$ in the polish space sense. By standard facts about computation over ω , we also get seom equivalent characterizations of Σ_1^0 -sets.

— 24 A • 3. Corollary

Let \mathcal{M} be polish with basis $\{\mathcal{M}_n : n < \omega\}$. Let $X \subseteq \mathcal{M}$. Therefore the following are equivalent.

- 1. X is $\Sigma_1^{0,\mathcal{M}}$.
- 2. $X = \bigcup_{n \in B} \mathcal{M}_n$ for some $\Sigma_1^{0,\omega}$ -set $B \subseteq \omega$.
- 3. $X = \bigcup_{n < \omega} \mathcal{M}_{f(n)}$ for some computable *partial* function $f : \omega \rightharpoonup \omega$.

Proof .:.

(1) \leftrightarrow (2) If X is empty, we're done: $B = \emptyset$. Otherwise for $f : \omega \to \omega$ witnessing $X \in \Sigma_1^{0,\mathcal{M}}$, im(f) is $\Sigma_1^{0,\omega}$ and satisfies $X = \bigcup_{n \in \text{im}(f)} \mathcal{N}_n$. Similarly, if $X = \bigcup_{n \in B} \mathcal{M}_n$ for some $B \in \Sigma_1^{0,\omega}$, then B = im(f)for some computable $f : \omega \to \omega$ where then $X = \bigcup_{n < \omega} \mathcal{M}_{f(n)}$.

(2) \leftrightarrow (3) Since the images of partial functions are $\Sigma_1^{0,\omega}$, we clearly get (3) \rightarrow (2). So suppose (2) holds: let *B* be $\Sigma_1^{0,\omega}$, i.e. $x \in B$ iff $\exists k \in \omega \ R(x,k)$ where $R \subseteq \omega^2$ is computable. We can then consider the partial function defined by $f(n) = k_0$ for the least $k = \text{code}(k_0, k_1)$ where $R(k_0, k_1)$ and $\forall m < n \ (f(m) < k_0)$. This has *f* as computable with im f = B and therefore (3) holds.

^{xi}Since there are only countably many computable functions $f : \omega \to \omega$, this yields only countably many $\Sigma_1^{0,\mathcal{M}}$ -sets for any choice of (countably many) basic open sets. Since there are uncountably many open sets, if we then add any other set $M \in \Sigma_1^{0,\mathcal{M}} \setminus \Sigma_1^{0,\mathcal{M}}$ to our list of basic open sets, we get a new notion of $\Sigma_1^{0,\mathcal{M}}$ differing from the previous one: M is $\Sigma_1^{0,\mathcal{M}}$ in the new sense, but not the old. This issue can be partially avoided through the use of the open balls as neighborhoods and a fixed countable dense subset of \mathcal{M} . This added terminology, however, will serve us little, since we may simply deal with \mathcal{N} and its products with itself and ω where such things are dealt with by proxy with initial segments. The only point where this comes up is just in the very rarely referenced properties of the basis. In particular, using this open ball approach, $\{\mathcal{M}_n : n < \omega\} \cup \{\emptyset\}$ is closed under finite intersections. This will be useful in showing $\Sigma_1^{0,\mathcal{M}}$ is closed under finite intersections, for example. But as this holds for the usual spaces (\mathcal{N} and its products with itself and ω), we don't take this approach in general until necessary later.

— 24 A • 4. Corollary

For $X \subseteq \mathcal{N}, X \in \Sigma_1^0$ iff there is a computable set B such that $\forall x \in \mathcal{N} \ (x \in X \Leftrightarrow \exists n < \omega \ (x \upharpoonright n \in B)).$ This also easily generalizes to products: for $X \subseteq {}^{\omega}\mathcal{N}, X \in \Sigma_1^{0,\omega}\mathcal{N}$ iff there is a computable B such that $\langle x_i : i < \omega \rangle \in X \Leftrightarrow \exists m < \omega \exists n < \omega \ (\langle x_i \upharpoonright n : i < m \rangle \in B).$

Proof .:.

$$X \in \Sigma_1^0 \text{ iff } X = \bigcup_{\tau \in R} \mathcal{N}_{\tau} \text{ for some } R \in \Sigma_1^{0,\omega} \text{ by Corollary 24 A} \cdot 3. \text{ Thus}$$
$$x \in X \Leftrightarrow \exists \tau \in R \ (x \in \mathcal{N}_{\tau}) \Leftrightarrow \exists \tau \in R \ (\tau \lhd x)$$
$$\Leftrightarrow \exists n < \omega \ \exists \tau \in R \ (x \upharpoonright n = \tau) \Leftrightarrow \exists n < \omega \ (x \upharpoonright n \in R).$$

As $\Sigma_1^{0,\omega}$ -relations have the form $\vec{x} \in Y$ iff $\exists m < \omega \ P(m, \vec{x})$ for some computable P, we have

 $x \in X \leftrightarrow \exists n < \omega \; \exists m < \omega \; (\langle x \upharpoonright n, m \rangle \in P).$

Clearly we can absorb this extra information into *n* by coding it onto the length: say $\langle \sigma, m \rangle \in P'$ iff $\exists n < h(\sigma)$ $(n = \text{code}(n_0, n_1) \land \langle \sigma \upharpoonright n_0, n_1 \rangle \in P)$. Then we have

$$x \in X \Leftrightarrow \exists n < \omega \ (x \upharpoonright n \in P')$$

To see that the two are equivalent, if $x \in X$, then for some $n_0 < \omega$ and some $n_1 < \omega$, $\langle x \upharpoonright n_0, n_1 \rangle \in P$. In particular, for $n = \operatorname{code}(n_0, n_1), x \upharpoonright n \in P'$. Similarly, if $x \upharpoonright n \in P'$ then there is some $n' = \operatorname{code}(n_0, n_1) < n$ where then $\langle (x \upharpoonright n) \upharpoonright n_0, n_1 \rangle = \langle x \upharpoonright n_0, n_1 \rangle \in P$ and thus $x \in X$. This proves the (\rightarrow) direction. For the converse, if $x \upharpoonright n \in B$ then $x \in \mathcal{N}_{x \upharpoonright n} \subseteq \bigcup_{\tau \in B} \mathcal{N}_{\tau} = X$.

We can also relativize Σ_1^0 just as in computability with oracles and in definability with parameters. Recall that for $A \in \mathcal{N}$, a subset $X \subseteq \omega$ is $\Sigma_1^0(A)$ iff we have some computable $R \subseteq \omega^3$ where

 $x \in X \quad \text{iff} \quad \exists y \in \omega \; \exists n \in \omega \; R(x, y, A \upharpoonright n).$

To generalize this in the same way as $\Sigma_1^{0,\omega}$ was generalized to $\Sigma_1^{0,\mathcal{N}}$, we can just consider such As as parameters.

— 24 A • 5. Definition

Let \mathcal{M} be polish with basis $\{\mathcal{M}_n : n < \omega\}$. Let $A \subseteq \mathcal{M}$ and $X \subseteq \mathcal{M}$. Define X to be $\Sigma_1^{0,\mathcal{M}}(A)$ iff there is some $R \in \Sigma_1^{0,\mathcal{M} \times \mathcal{M}}$ and some $\vec{a} \in A^{<\omega}$ such that $X = \{x \in \mathcal{M} : R(x, \vec{a})\}$.

Commonly, for $A = \{a\}$, we just write $\Sigma_1^{0,\mathcal{M}}(a)$ rather than $\Sigma_1^{0,\mathcal{M}}(\{a\})$. But a result of this new definition is another way of referring to all the open sets of bairs space (and its products).

24A•6. Corollary $\Sigma_1^0 = \bigcup_{A \in \mathcal{N}} \Sigma_1^0(A)$. In other words, $\Sigma_1^0 = \Sigma_1^0(\mathcal{N})$.

Proof .:.

If X is open, then $X = \bigcup_{\tau \in B} \mathcal{N}_{\tau}$ for some $B \subseteq {}^{<\omega}\omega$ where then $X \in \Sigma_1^0(B)$, regarding $B \in \mathcal{N}$ after coding. And clearly, for every $A \subseteq \mathcal{N}$, since every element of $\Sigma_1^0(A)$ is a union of basic open sets, $\Sigma_1^0(A) \subseteq \Sigma_1^0$. \dashv

It's natural to think of this Σ_1^0 as generating a kind of computable topology on \mathcal{N} . This especially makes sense given that the standard topology on ω is just the discrete topology where all sets are open. The arithmetical hierarchy on ω , however, is very rich compared to the standard topology. We get similar closure properties for Σ_1^0 as with Σ_1^0 , as we will see, strengthening this connection. It is ultimately this idea that leads to an entire lightface version of the borel hierarchy, called the hyperarithmetical hierarchy, so-called because it extends the arithmetical hierarchy of Σ_n^0 , Π_n^0 , Δ_n^0 for $n < \omega$ to Σ_α^0 , Π_α^0 , Δ_α^0 for $\alpha < \omega_1^{CK}$, where ω_1^{CK} is some particular countable ordinal (larger than ω). The precise definition isn't important for us now.

The point for us is that Definition 24 A • 2 determines what it means for a subset $X \subseteq \mathcal{N}$ to be Σ_1^0 , and this gives a notion of a subset of \mathcal{N} being computable as well as a notion of a function $f : \mathcal{N} \to \mathcal{N}$ being computable. This

generalization is motivated from the result on ω that a (partial) function $f: \omega \to \omega$ is computable iff $f \subseteq \omega \times \omega$ as a relation is $\Sigma_1^{0,\omega \times \omega}$.

The obvious generalization, saying that $f : \mathcal{M} \to \mathcal{M}$ is computable iff it's a $\Sigma_1^{0,\mathcal{M}\times\mathcal{M}}$ -relation, isn't the proper one, because there are no such functions if \mathcal{M} has no isolated points: $f \in \Sigma_1^{0,\mathcal{M}\times\mathcal{M}}$ implies f is open. In particular, for $\langle x, y \rangle \in f$, there is some open rectangle $\mathcal{M}_n \times \mathcal{M}_m \subseteq f$ with $\langle x, y \rangle \in \mathcal{M}_n \times \mathcal{M}_m$. Since $\mathcal{M}_m \neq \{y\}$, then there is some $y' \neq y$ with $\langle x, y' \rangle \in \mathcal{M}_n \times \mathcal{M}_m \subseteq f$, meaning f isn't a function.

So instead we consider f as computable when dealing with the basic open sets: where $x \in \mathcal{M}$ is mapped in terms of the basic open sets of the target space, which we deal with in terms of ω . In particular, f is computable iff whether $f(x) \in \mathcal{M}_n$ or not is $\Sigma_1^{0,\mathcal{M}\times\omega}$ for $x \in \mathcal{M}$ and $n \in \omega$.

Let \mathcal{M} and \mathcal{W} be polish with bases $\{\mathcal{M}_n : n < \omega\}$ and $\{\mathcal{W}_m : m < \omega\}$. Let $A \subseteq \mathcal{M}$. Let $f : \mathcal{M} \to \mathcal{W}$ be a function. Define f to be A-computable iff the neighborhood relation $f(x) \in \mathcal{W}_{\sigma}$ is $\Sigma_1^{0,\mathcal{M}\times\omega}(\langle A, 0 \rangle)$:

$$\mathrm{NG}_f = \{ \langle x, \sigma \rangle \in \mathcal{M} \times \omega : f(x) \in \mathcal{W}_\sigma \} \in \Sigma_1^{0, \mathcal{M} \times \omega}(\langle A, 0 \rangle).$$

For $\mathcal{M} = \mathcal{W} = \omega$, the basic open sets are singletons and so $NG_f = f$. Thus f is computable iff f as a relation is $\Sigma_1^{0,\omega\times\omega}$, in line with what we know from computability on ω . But in what sense is does this make f "computable" in an intuitive sense? The idea is that we can approximate f(x) from below—from basic open sets around f(x)—in an effective way, as the result of a genuinely computable function. In particular, a more algorithmic characterization for \mathcal{N} (and its products with itself and ω) is the following. The basic idea being that on \mathcal{N} , the neighborhood graph simply represents the relation $\sigma \triangleleft f(x)$ for $\langle x, \sigma \rangle \in \mathcal{N} \times \omega$. Hence if this is $\Sigma_1^{0,\mathcal{N}\times\omega}$, we should be able to find f(x)(n), $n < \omega$, just by searching along the computable union for a suitably large σ .

There are a lot of different ways to state computability for \mathcal{N} as with ω . The two most popular revolve around this idea stated in slightly different, equivalent ways. First, we give a little bit of background on computability theory: a program can be coded just by a single number $e \in \omega$ which can then be decoded into the function $[\![e]\!] : \omega \to \omega$. Programs can also involve more functions assumed to be given. The roles of these given functions are *oracles* in that the program asks for their value and it is given with no further computation involved. In the program $e \in \omega$, these are merely syntactic and must be interpreted to give meaning to $[\![e]\!]$. For $\vec{x} \in {}^{<\omega}\mathcal{N}$, the function computed by e using \vec{x} as its oracles will be $[\![e]\!]_{\vec{x}}$. In this way, $x \mapsto [\![e]\!]_x \in \mathcal{N}$ is "computable" in that it is computable *uniformly* or *by the same algorithm*: given $x \in \mathcal{N}$ we can then calculate every value of $[\![e]\!]_x \in \mathcal{N}$ with e and x.

- 24 A • 8. Theorem

A function $f : \mathcal{N} \to \mathcal{N}$ is computable iff there is some $e \in \omega$ where $f = x \mapsto \llbracket e \rrbracket_{x,A}$. Informally, f is computable iff there is an *x*-computable algorithm that computes f(x) and moreover, this algorithm is uniform across all $x \in \mathcal{N}$. This also easily generalizes to *A*-computability for $A \subseteq \mathcal{N}$ and also products of \mathcal{N} with itself and ω .

This is an easy consequence of the following, equivalent form where we merely compute initial segments of f in a uniform way, i.e. by a computable \hat{f} with two arguments: one for the oracle, and the other for the place to evaluate the computed real.

— 24 A • 9. Result -

For $A \subseteq \mathcal{N}$, a function $f : \mathcal{N} \to \mathcal{N}$ is A-computable iff there is some $\vec{a} \in A^{<\omega}$ and $\oplus \vec{a}$ -computable $\hat{f} : {}^{<\omega}\omega \times \omega \to \omega$ such that for every $x \in \mathcal{N}$ and $n \in \omega$,

 $\exists m < \omega \ (\hat{f}(x \upharpoonright m, n) = f(x)(n)).$

And this easily generalizes to products of \mathcal{N} with itself and ω .

Proof .:.

We take $A = \emptyset$, as the proof easily generalizes. We may reduce to a finite $\vec{a} \in A^{<\omega}$ —taking the computable join $\oplus \vec{a}$ of them—because the parameters that make f A-computable use only finitely many parameters from A. The basic open subsets of $\mathcal{N} \times {}^{<\omega}\omega$ are of the form $\mathcal{N}_{\tau} \times \{\sigma\}$ for $\tau, \sigma \in {}^{<\omega}\omega$.

 (\rightarrow) If NG_f is $\Sigma_1^{0,\mathcal{N}\times\omega}$, then there is some computable $g: \omega \to {}^{<\omega}\omega \times {}^{<\omega}\omega$ (where we write $g(n) = \langle g_0(n), g_1(n) \rangle$) such that NG_f = $\bigcup_{n < \omega} \mathcal{N}_{g_0(n)} \times \{g_1(n)\}$. For $x \in \mathcal{N}$ and $\sigma \in {}^{<\omega}\omega$, recall that NG_f (x, σ) iff $\sigma \triangleleft f(x)$. So given $x \in \mathcal{N}$, to calculate f(x)(n), we just search for some sufficiently long $\sigma \triangleleft f(x)$, i.e. some sufficiently long σ with $\langle x, \sigma \rangle \in NG_f$: define $\hat{f}(\tau, n)$ as follows: find the least *m* such that

1. $lh(g_1(m)) > n;$

2. $g_0(m) \triangleleft \tau$, meaning $\forall k < \operatorname{lh}(g_0(m)) (g_0(m)(k) = \tau(k));$

and then output $g_1(m)(n)$. \hat{f} is then clearly computable with $\hat{f}(x \upharpoonright m, n) = f(x)(n)$.

(←) Let \hat{f} : ${}^{<\omega}\omega \times \omega \to \omega$ be as in the statement. By Corollary 24 A • 3, it suffices to show there is a $\Sigma_1^{0,\omega}$ -relation $B \subseteq {}^{<\omega}\omega \times {}^{<\omega}\omega$ with NG_f = $\bigcup_{\langle \tau,\sigma \rangle \in B} \mathcal{N}_{\tau} \times \{\sigma\}$. But this is obvious: just write

 $\langle \tau, \sigma \rangle \in B$ iff $\exists \tau' \forall k < \ln(\sigma) \ (\sigma(k) = \hat{f}(\tau^{\frown}\tau', k))$

So that *B* is of the form $\exists \tau' R(\tau, \tau', \sigma)$ for some computable relation *R*, meaning *B* is $\Sigma_1^{0,\omega}$. It should be clear that this *B* works. To see this, if $\langle x, \sigma \rangle \in \operatorname{NG}_f$, then $\sigma \triangleleft f(x)$, in which case, there is some $m < \omega$ with $\hat{f}(x \upharpoonright m, n) = f(x)(n) = \sigma(n)$ for all $n < \operatorname{lh}(\sigma)$ so that $x \upharpoonright m = \tau$ with $\tau' = \emptyset$ witnesses $\langle \tau, \sigma \rangle \in B$. This shows $\operatorname{NG}_f \subseteq \bigcup_{\langle \tau, \sigma \rangle \in B} \mathcal{N}_\tau \times \{\sigma\}$. The converse is clear: $\langle \tau, \sigma \rangle \in B$ as witnessed by $\tau' \in {}^{<\omega}\omega$ has any $x \in \mathcal{N}_{\tau \frown \tau'}$ satisfy $f(x) \upharpoonright \operatorname{lh}(\sigma) = \sigma$ where then $\langle x, \sigma \rangle \in \operatorname{NG}_f$.

This algorithmic approach can sometimes be a little easier to intuitively see than confirming a neighborhood graph is Σ_1^0 . It also further motivates why such functions are worthy of the description "computable".

Computability will be very important for two reasons: its absoluteness between transitive models of sufficiently large fragments of set theory, and the role it plays here analogous to continuity in topology. The following result should raise eyebrows relating the two.

24A•10. Lemma Let $f : \mathcal{N} \to \mathcal{N}$ be a function. Therefore, f is continuous iff there is some $A \in \mathcal{N}$ such that f is A-computable.

Proof .:.

 (\rightarrow) Suppose f is continuous. This means if f(x) = y, for every initial segment $\sigma \triangleleft y$, there is an initial segment $\tau \triangleleft x$ with $f^{-1}"\mathcal{N}_{\sigma} \subseteq \mathcal{N}_{\tau}$. So if we consider $A = \{\langle \tau, \sigma \rangle \in {}^{<\omega}\omega \times {}^{<\omega}\omega : f^{-1}"\mathcal{N}_{\sigma} \subseteq \mathcal{N}_{\tau}\}$, then f is A-computable. In particular, given any x, to define f(x)(n), we consider the least $m = \operatorname{code}(\tau, \sigma, k)$ such that

- τ is the code of $x \upharpoonright n$;
- σ is the code of a sequence of length > k; and
- $\langle \tau, \sigma \rangle \in A$.

Then we can notice $f(x)(n) = \sigma(n)$ for appropriate *n*, and the above procedure is computable from *A* using an initial segment of *x*.

(\leftarrow) Suppose f is A-computable by \hat{f} : ${}^{<\omega}\omega \times \omega \to \omega$: for each $n < \omega$, $\hat{f}(x \upharpoonright m, n) = f(x)(n)$ for sufficiently large m. For $x \in \mathcal{N}$ arbitrary, suppose $f(x) = y \in \mathcal{N}$. If $n < \omega$, we can compute every value of $y \upharpoonright n$ using just some sufficiently large value of m, meaning $\mathcal{N}_{x \upharpoonright m} \subseteq f^{-1} :: \mathcal{N}_{y \upharpoonright n}$, implying f is continuous.

As a result, we should expect closure properties of $\Sigma_1^0(A)$ under not *all* continuous functions, but only the *A*-computable functions.

24A•11. Lemma Let \mathcal{M} and \mathcal{W} be polish with bases $\{\mathcal{M}_n : n < \omega\}$ and $\{\mathcal{W}_n : n < \omega\}$. Let $f : \mathcal{M} \to \mathcal{W}$ be computable. Therefore if $X \in \Sigma_1^{0,\mathcal{W}}$ then f^{-1} " $X \in \Sigma_1^{0,\mathcal{M}}$.

 \dashv

Proof .:.

Recall that $NG_f = \{\langle x, m \rangle \in \mathcal{M} \times \omega : f(x) \in \mathcal{W}_m\}$. Firstly, we need a way of translating between the basic open sets in a computable way. We don't need to worry about empty preimages, since trivially $\emptyset \in \Sigma_n^{0,\mathcal{M}}$.

— Claim 1 For *m* where f^{-1} " $\mathcal{W}_m \neq \emptyset$, there is a computable $h : \omega \times \omega \to \omega$ such that f^{-1} " $\mathcal{W}_m = \bigcup_{k < \omega} \mathcal{M}_{h(m,k)}$.

Proof .:.

Since f is computable, $NG_f \in \Sigma^{0, \mathcal{M} \times \omega}$, meaning for some computable $p_0, p_1 : \omega \to \omega$, $NG_f = \bigcup_{k < \omega} \mathcal{M}_{p_0(k)} \times \{p_1(k)\}$. Hence $f(x) \in \mathcal{W}_m$ iff there is some $k < \omega$ such that $x \in \mathcal{M}_{p_0(k)}$ and $m = p_1(k)$. So set

$$h(m,k) = \begin{cases} p_0(k) & \text{if } p_1(k) = m \\ p_0(k') & \text{for some fixed } k' \text{ with } p_1(k') = m \text{ otherwise.} \end{cases}$$

Then $f(x) \in W_m$ iff $x \in \bigcup_{k < \omega} \mathcal{M}_{h(m,k)}$.

This provides a basis to define a function transforming arithmetical sets according to how they're built up. In particular, proceed by induction on *n* to show the result. For n = 1, let $X \in \Sigma_1^{0,\mathcal{W}}$ have $f^{-1} X \neq \emptyset$. Therefore $X = \bigcup_{m \le \omega} W_{g(m)}$ for some computable $g : \omega \to \omega$ where then

$$f^{-1}"X = \bigcup_{m < \omega} f^{-1}"\mathcal{W}_{g(m)} = \bigcup_{m,k < \omega} \mathcal{M}_{h(g(m),k)} \in \Sigma_1^{0,\mathcal{M}}.$$

We also get some expected properties of the class of computable functions.

- 24 A • 12. Result

The class of computable functions between polish spaces is closed under composition, substitutions (i.e. f defined by $f(\vec{x}) = \langle f_0(\vec{x}), \dots, f_n(\vec{x}) \rangle$ for f_0, \dots, f_n computable) and contains

• all projections: $p_i(\vec{x}) = x_i$; and

• all (in the usual computability sense) computable $f: \omega \to \omega$.

Proof ...

That all computable $f: \omega \to \omega$ are computable in the sense of Definition 24 A • 7 follows just from facts about these functions as in Appendix A. Let \mathcal{M}_i for $i < N < \omega$ be polish spaces with bases $\{(\mathcal{M}_i)_n : n < \omega\}$ for i < N. Let $p_i : \prod_{n < N} \mathcal{M}_n \to \mathcal{M}_i$ be the projection map for i < N: $p_i(\vec{x}) = x_i$. Therefore p_i is computable since the neighborhood graph has $p_i(\vec{x}) \in (\mathcal{M}_i)_n$ iff $x_i \in (\mathcal{M}_i)_n$:

$$\mathrm{NG}_{p_i} = \bigcup_{n,\mathrm{code}(\vec{k}) < \omega} \left(\left(\prod_{j < i} (\mathcal{M}_j)_{k_j} \right) \times (\mathcal{M}_i)_n \times \left(\prod_{i < j < N} (\mathcal{M}_j)_{k_j} \right) \times \{n\} \right).$$

— Claim 1

Let \mathcal{M}, \mathcal{W} , and \mathcal{U} be polish with bases $\{\mathcal{M}_n : n < \omega\}, \{\mathcal{W}_n : n < \omega\}, \{\mathcal{U}_n : n < \omega\}$ respectively. Suppose $f : \mathcal{M} \to \mathcal{W}$ and $g : \mathcal{U} \to \mathcal{U}$ is computable. Therefore $F : \mathcal{M} \times \mathcal{U} \to \mathcal{W} \times \mathcal{U}$ is computable defined by $F(x, n) = \langle f(x), g(n) \rangle$.

Proof .:.

For $m, k < \omega$, let the basic open set of $\mathcal{W} \times \mathcal{U}$ indexed by $\operatorname{code}(m, k)$ just be the rectangle $\mathcal{W}_m \times \mathcal{U}_k$. If $\operatorname{NG}_f = \bigcup_{n < \omega} \mathcal{M}_{f_0(n)} \times \{f_1(n)\}$ for computable $f_0, f_1 : \omega \to \omega$ and similarly $\operatorname{NG}_g = \bigcup_{n < \omega} \mathcal{U}_{g_0(n)} \times \{g_1(n)\}$, then we have

$$\mathrm{NG}_F = \bigcup_{n < \omega} (\mathcal{M}_{f_0(n)} \times \mathcal{U}_{g_0(n)}) \times \{ \mathrm{code}(f_1(n), g_1(n)) \}.$$

We get closure under substitutions just by repeating applications of Claim 1 and composition. So it suffices to

 \neg

show closure under compositions. Suppose $f : \mathcal{M} \to \mathcal{W}$ and $g : \mathcal{W} \to \mathcal{U}$ is computable where \mathcal{M}, \mathcal{W} , and \mathcal{U} are polish with bases $\{\mathcal{M}_n : n < \omega\}, \{\mathcal{W}_n : n < \omega\}$, and $\{\mathcal{U}_n : n < \omega\}$ respectively as expected. Therefore,

$$g \circ f(x) \in \mathcal{U}_n$$
 iff $\langle f(x), n \rangle \in \mathrm{NG}_g \in \Sigma_1^{0, W \times \omega}$ iff $\langle x, n \rangle \in F^{-1}$ "NG_g

where $F : \mathcal{M} \times \omega \to \mathcal{W} \times \omega$ is defined by $F(x, n) = \langle f(x), n \rangle$, which is computable by Claim 1. By Lemma 24 A • 11, NG_{g o f} = F^{-1} "NG_g $\in \Sigma_1^{0, \mathcal{M} \times \omega}$ so that the $g \circ f$ composition is computable. \dashv

As in Theorem 21 B • 5 and Theorem 21 B • 7, we can find computable injections and surjections with \mathcal{N} , so long as the basic open sets we're working with look sufficiently nice. We will develop a term for such bases later.

- 24A • 13. Theorem - Let \mathcal{M} be polish with metric d. Let $\{\mu_i : i < \omega\}$ be a dense subset of \mathcal{M} yielding basic open sets as open balls:

$$\mathcal{M}_{n_0,n_1,n_2} = \left\{ x \in \mathcal{M} : d(x,\mu_{n_0}) < \frac{n_1}{n_2 + 1} \right\}$$

Suppose the relations $D_{\leq} \subseteq \omega^4$, defined by by $d(\mu_i, \mu_j) < \frac{n_1}{n_2+1}$, and similarly $D_{\leq} \subseteq \omega^4$ are computable. Therefore, there is a computable $f : \mathcal{N} \to \mathcal{M}$ that is surjective.

Proof .:.

Regard $\mathcal{M}_{n_0,n_1,n_2}$ as instead $\mathcal{M}_{q(\tau)} = \mathcal{M}(\tau)$ for some $\tau \in {}^{<\omega}\omega$ and computable $q : {}^{<\omega}\omega \to \omega$. And we will then define f(x) to be the unique element of $\bigcap_{n<\omega} \mathcal{M}(x \upharpoonright n)$. So now we must define this $q : {}^{<\omega}\omega \to \omega^3$ (writing $q(\tau) = \langle q_0(\tau), q_1(\tau), q_2(\tau) \rangle$). We will always have $q_1(\tau) = 1$ and $q_2(\tau) = 2^{\ln(\tau)+1}$ for $\tau \neq \emptyset$.

Without loss of generality, assume the metric d has d(x, y) < 1 for all $x, y \in \mathcal{M}$ (e.g. take instead $d' = \frac{d}{1+d}$). Let $q(\emptyset) = \langle 0, 1, 0 \rangle$ so that $\mathcal{M}(q(\emptyset)) = \mathcal{M}_{0,1,0} = \mathcal{M}$. For $n < \omega$, define $q_0(\langle n \rangle) = n$. For $q(\tau)$ already defined for $\tau \neq \emptyset$, define $q_0(\tau \cap \langle n \rangle)$ be the *n*th least *i* such that $D_{<}(i, q_0(\tau), 1, 2^{\ln(\tau)+2})$ if there is one, otherwise leave $q_0(\tau \cap \langle n \rangle) = q_0(\tau)$. It should be clear that q is computable in the usual sense. Now define $\mathcal{M}(\tau)$ to be the basic open set $\mathcal{M}_{q(\tau)}$, and therefore

- For any $x \in \mathcal{N}$, $\langle \mu_{q(x \upharpoonright n)} : n < \omega \rangle$ is cauchy and therefore converges in \mathcal{M} .
- For any $y \in \mathcal{M}$, there is an $x \in \mathcal{N}$ with $\langle \mu_{q(x \upharpoonright n)} : n < \omega \rangle$ converging to y.

The first of these is immediate by definition: for all τ , $d(\mu_{q_0(\tau)}, \mu_{q_0(\tau^{\frown}(n))}) < \frac{1}{2^{\ln(\tau)+2}+1}$. For the second, we just choose any cauchy $\langle \mu_{y_i} : i < \omega \rangle$ converging to y, and then construct such an $x = \langle x(n) : n < \omega \rangle$: set $x(0) = y_0$, and for x(n+1), we enumerate the $i < \omega$ such that μ_i such that $d(\mu_i, q(x \upharpoonright n+1)) < \frac{1}{2^{n+2}+1}$ and set x(n+1) to be the least i that is one of the μ_{y_n} s. The resulting $x \in \mathcal{N}$ works.

In particular, for $\mathcal{M}(\tau) = \mathcal{M}_{q(\tau)}$, for $x \in \mathcal{N}$, taking f(x) as the unique element of $\bigcap_{n < \omega} \mathcal{M}(x \upharpoonright n)$ yields that $f : \mathcal{N} \to \mathcal{M}$ is well-defined and surjective. Moreover,

$$\langle x, \langle n_0, n_1, n_2 \rangle \rangle \in \mathrm{NG}_f$$
 iff $f(x) \in \mathcal{M}_{n_0, n_1, n_2}$ iff $\exists m < \omega \ D_<(q(x \upharpoonright m), n_0, n_1, n_2)$,
which is $\Sigma_1^{0, \mathcal{N} \times \omega^3}$, meaning $\mathrm{NG}_f \in \Sigma_1^{0, \mathcal{N} \times \omega^3}$ and so f is computable.

Computable functions also give way to computable relations by way of whether their characteristic functions are computable.

- 24A • 14. Definition Let \mathcal{M} be polish with basis $\{\mathcal{M}_n : n < \omega\}$. Let $X \subseteq \mathcal{M}$. The *characteristic function* of X is $\chi_X : \mathcal{M} \to 2$ defined by $\chi_X(x) = 1$ iff $x \in X$, and otherwise $\chi_X(x) = 0$. For $A \subseteq \mathcal{M}$, we call X A-computable iff χ_X is A-computable.

An alternative characterization is the following, analogous to that for ω .

$-24A \cdot 15$. Theorem

Let \mathcal{M} be polish with basis $\{\mathcal{M}_n : n < \omega\}$. Let $X \subseteq \mathcal{M}$. Therefore X is A-computable iff $X \in \Sigma_1^{0,\mathcal{M}}(A)$ and $\neg X \in \Sigma_1^{0,\mathcal{M}}(A)$.

Proof .:.

We consider only the case where $A = \emptyset$, as the general case follows from this just by introducing parameters. Suppose $X, \neg X \in \Sigma_1^{0,\mathcal{M}}$. Write $X = \bigcup_{n < \omega} \mathcal{M}_{f(n)}$ and $\neg X = \bigcup_{n < \omega} \mathcal{M}_{g(n)}$ for $f, g : \omega \to \omega$. Therefore χ_X is computable, and so X is, since

$$\mathrm{NG}_{\chi_X} = \chi_X = \bigcup_{n < \omega} \mathcal{M}_{f(n)} \times \{1\} \cup \bigcup_{n < \omega} \mathcal{M}_{g(n)} \times \{0\} = \bigcup_{n < \omega} \mathcal{M}_{f \ast g(n)} \times \{\chi_{\mathrm{even}}(n)\} \in \Sigma_1^{0, \mathcal{M} \times \omega}.$$

Now suppose X is computable by χ_X with $NG_{\chi_X} \in \Sigma_1^{0,\mathcal{M}\times\omega}$. Therefore $X \in \Sigma_1^{0,\mathcal{M}}(A)$ as the preimage of NG_{χ_X} under the map $x \mapsto \langle x, 1 \rangle$, and similarly $\neg X \in \Sigma_1^{0,\mathcal{M}}(A)$ as the preimage under the map $x \mapsto \langle x, 0 \rangle$. \dashv

We can also generalize Σ_1^0 and its closure under computable preimages to further along other levels in another hierarchy, being the lightface variant of the borel hierarchy. We will have almost no need of Σ_{α}^0 for large α , and so only the first ω -levels of the hierarchy are presented here, called the *arithmetical* hierarchy, analogous to that on ω defined for computability. But in principle, we may go beyond to explore Σ_{α}^0 for $\alpha < \omega_1$ (and in fact just for $\alpha < \omega_1^{CK} < \omega_1$), the *hyperarithmetical* hierarchy, and this is covered further in Appendix B.

§ 24 B. The arithmetical hierarchy for polish spaces

We will now be concerned with two kinds of projections: projections over ω , corresponding to countable unions in the borel hierarchy; and projections over \mathcal{M} , just as in the projective hierarchy. As such, we introduce the notation \exists^{ω} to denote projection where we eliminate a copy of ω from the product space: for $X \subseteq \mathcal{M} \times \omega$, $x \in \exists^{\omega} X$ iff $\exists n \in \omega \ (\langle x, n \rangle \in X)$ so that $\exists^{\omega} X \subseteq \mathcal{M}$. Another way of writing this is just that $\exists^{\omega} X = \mathfrak{p} X$, but \exists^{ω} makes it more clear what exactly is being projected, or really *removed*: a copy of ω . And we can do similarly for the operations $\exists^{\mathcal{N}}$ or $\exists^{\mathcal{M}}$, and even the \forall versions of these: \forall^{ω} is the operation $\neg \exists^{\omega} \neg$, for example.

— 24 B•1. Definition -

Let \mathcal{M} be polish with basis $\{\mathcal{M}_n : n < \omega\}$. Let $A \subseteq \mathcal{M}$. We define the (*relativized* or A-) arithmetical hierarchy by recursion on $n < \omega$ as follows: for $X \subseteq \mathcal{M}$,

• X is $\Sigma_1^{0,\mathcal{M}}(A)$ iff X is as before in Definition 24 A • 2 and Definition 24 A • 5;

• X is $\Pi_n^{0,\mathcal{M}}(A)$ iff $\mathcal{M} \setminus X$ is $\Sigma_n^{0,\mathcal{M}}(A)$;

- X is $\sum_{n+1}^{0,\mathcal{M}}(A)$ iff $X = \exists^{\omega} Y$ for some $\prod_{n=1}^{0,\mathcal{M}\times\omega}(A)$ -set $Y \subseteq \mathcal{M}\times\omega$; and
- X is $\Delta_n^{0,\mathcal{M}}(A)$ iff X is $\Sigma_n^{0,\mathcal{M}}(A)$ and $\Pi_n^{0,\mathcal{M}}(A)$.

Continuing the fact that this is really the computable analog of the borel hierarchy, we have that Π_1^0 consists of the branches of *computable* trees on ω .

Let $X \subseteq \mathcal{N}$. Therefore $X \in \Pi_1^0$ iff there is some *computable* tree $T \subseteq {}^{<\omega}\omega$ where X = [T].

Proof .:.

Suppose membership in *T* is computable. Therefore $x \in [T]$ iff $\forall n < \omega$ ($x \upharpoonright n \in T$), which is Π_1^0 by Corollary 24 A • 4. Similarly, if $X \in \Pi_1^0$, then there is some computable $R \subseteq {}^{<\omega}\omega$ where $x \in X$ iff $\forall n \in \omega R(x \upharpoonright n)$. So take $T = \{\tau \in R : \forall \sigma \preccurlyeq \tau R(\sigma)\}$ and get that *T* is still computable with [T] = X.

Looking back to Definition 24 B • 1, it follows that A-computable sets are $\Delta_1^{0,\mathcal{M}}(A)$. Moreover,

$$\Sigma_{n+1}^{0,\mathcal{M}}(A) = \exists^{\omega} \Pi_n^{0,\mathcal{M} \times \omega}(A) = \exists^{\omega} \neg \Sigma_n^{0,\mathcal{M} \times \omega}(A)$$

And we can continue this on: any Σ_3^0 -set has the form

$$\begin{aligned} x \in X \quad \text{iff} \quad \exists m_2 \in \omega \neg P_2(x, m_2) \quad \text{for } P_2 \in \Sigma_2^{0, \mathcal{N} \times \omega} \\ \text{iff} \quad \exists m_2 \in \omega \neg \exists m_1 \in \omega \neg P_1(x, m_1, m_2) \quad \text{for } P_1 \in \Sigma_1^{0, \mathcal{N} \times \omega^2} \\ \text{iff} \quad \exists m_2 \in \omega \ \forall m_1 \in \omega \ P_1(x, m_1, m_2) \quad \text{for } P_1 \in \Sigma_1^{0, \mathcal{N} \times \omega^2}. \end{aligned}$$

The analogy between the arithmetical hierarchy and the lévy hierarchy of formulas (with quantification only over ω) should then be clear, although we won't discuss this analogy formally until the end of the section. More than stopping at just Σ_1^0 -relations, for \mathcal{N} , we can actually continue down to *computable* relations over ${}^{<\omega}\omega$ by Corollary 24 A • 4: for $X \in \Sigma_1^0$,

 $x \in X$ iff $\exists m \in \omega \ R(x \upharpoonright m)$ for $R \subseteq {}^{<\omega}\omega$ computable.

So really, the basis for the arithmetical hierarchy is the computable relations over ω . This also gives an alternative way of generating the arithmetical hierarchy, as negations of computable relations are still computable: we start with computable relations, and alternatively add \exists and \forall quantifiers over ω .

As usual, we should establish the expected closure and inclusion properties.



24 B • 3. Figure: The arithmetical hierarchy

First we have the following useful result expanding Lemma 24 A • 11. In particular, coding over products of ω or \mathcal{N} makes no difference.

24 B • 4. Result

Let \mathcal{M} and \mathcal{W} be polish with bases $\{\mathcal{M}_n : n < \omega\}$ and $\{\mathcal{W}_n : n < \omega\}$. Let $0 < n < \omega$ and let $f : \mathcal{M} \to \mathcal{W}$ be computable. Therefore if $X \in \Sigma_n^{0,\mathcal{W}}$ then $f^{-1} X \in \Sigma_n^{0,\mathcal{M}}$, and similarly for the other arithmetical pointclasses.

Proof .:.

Lemma 24 A • 11 establishes the result for n = 1. Note that complements work nicely with preimages, yielding the result for $\Pi_1^{0,W}$, $\Pi_1^{0,\mathcal{M}}$ as well. For n + 1, let $X = \exists^{\omega} Y$ where $Y \in \Pi_n^{0,W\times\omega}$. Rather than f, consider $\hat{f} : \mathcal{M} \times \omega \to \mathcal{W} \times \omega$ defined by $\hat{f}(x,m) = \langle f(x),m \rangle$, which is computable by Result 24 A • 12. So inductively, $\hat{f}^{-1}"Y \in \Pi_n^{0,\mathcal{M}\times\omega}$ and it should be clear that $f^{-1}"X = \exists^{\omega}(\hat{f}^{-1}"Y) \in \Sigma_{n+1}^{0,\mathcal{M}}$.

We now aim to establish the basic inclusion properties giving the usual argyle picture of Figure $24B \cdot 3$. This unfortunately requires some restrictions on what the basic open sets of our polish space are, however, as in Theorem $24A \cdot 13$.

- 24 B • 5. Definition

Let \mathcal{M} be polish with metric d. A computable presentation of \mathcal{M} is a dense set $\{\mu_i : i < \omega\} \subseteq \mathcal{M}$ such that the relations $D_{\leq} \subseteq \omega^4$ and $D_{\leq} \subseteq \omega^4$ are computable, defined by

 $D_{\leq}(i, j, p, q)$ iff $d(\mu_i, \mu_j) \leq p/(q+1)$, and $D_{<}(i, j, p, q)$ iff $d(\mu_i, \mu_i) < p/(q+1)$.

A basis $\{\mathcal{M}_n : n < \omega\}$ is *presented* iff there is some computable presentation $\{\mu_i : i < \omega\}$ of \mathcal{M} such that each \mathcal{M}_n is the open ball around μ_{n_0} of radius $n_1/(n_2 + 1)$, where $n = \operatorname{code}(\langle n_0, n_1, n_2 \rangle)$.

The basic idea here is that our choice of basis isn't too outlandish: it's coming from open balls on dense sets rather than a mix of simple sets and horribly complicated open sets. For example, any polish space \mathcal{M} without isolated points has uncountably many open sets while only countably many $\Sigma_1^{0,\mathcal{M}}$ -sets (for whatever basis $\{\mathcal{M}_n : n < \omega\}$ we have). Hence some open set $U \in \sum_1^{0,\mathcal{M}} \setminus \Sigma_1^{0,\mathcal{M}}$ yields that U is "complicated". But we could have just as easily considered $\{\mathcal{M}_n : n < \omega\} \cup \{U\}$ as our basis, yielding that U is trivially $\Sigma_1^{0,\mathcal{M}}$, against our intuition. This has the negative effect that U might not be at all computably related to to the other basic open sets in terms of codes—especially if we added infinitely many such U—and so things like intersections of basic open sets suddenly difficult to deal with. But so long as our basis is presented, we have the natural properties one would expect, because the basic open sets work well with each other.

It should be clear that the canonical bases of the baire space and its products with itself and ω can be presented in this sense. This is obvious for ω if we take the computable presentation ω with the basis $\{\{n\} : n < \omega\} \cup \{\omega\}$, and similarly for \mathcal{N} . For products, for example with \mathcal{N}^n , we just take μ_i to be $\langle \mu_{i_0}^{\mathcal{N}}, \cdots, \mu_{i_n}^{\mathcal{N}} \rangle$ for $n < \omega$ with $i = \text{code}(\langle i_0, \cdots, i_n \rangle)$. The fine details of D_{\leq} and $D_{<}$ are left to the reader in this example, but it should be clear from the exposition back in Section 21 and Example 21 A • 13 that the maps are computable. This whole topic is brought up only for the following lemma (and its consequences).

- 24B·6. Result Let \mathcal{M} be polish with presented basis $\{\mathcal{M}_n : n < \omega\}, A \in \mathcal{M}$, and $0 < n < \omega$. Therefore $\Delta_n^{0,\mathcal{M}}(A) \subseteq \Sigma_n^{0,\mathcal{M}}(A) \subseteq \Delta_{n+1}^{0,\mathcal{M}}(A)$, and similarly for $\Pi_n^{0,\mathcal{M}}(A)$.

Proof .:.

We prove this for $A = \emptyset$ for the sake of notation. By definition, $\Delta_n^{0,\mathcal{M}} \subseteq \Sigma_n^{0,\mathcal{M}}$ so we must show 1. $\Sigma_n^{0,\mathcal{M}} \subseteq \Pi_{n+1}^{0,\mathcal{M}}$; and 2. $\Sigma_n^{0,\mathcal{M}} \subseteq \Sigma_{n+1}^{0,\mathcal{M}}$. (1) is easy: for $X \in \Sigma_n^{0,\mathcal{M}}$, $X \times \omega \in \Sigma_n^{0,\mathcal{M} \times \omega}$ by Result 24 B • 4 so that $\neg X \times \omega$ is $\Pi_n^{0,\mathcal{M}}$ and thus $\neg X = \exists^{\omega}(\neg X \times \omega) \in \Sigma_{n+1}^{0,\mathcal{M}}$. So $X \in \Pi_{n+1}^{0,\mathcal{M}}$.

Showing (2) is more difficult, and uses the next claim to show the base case of $\Sigma_1^{0,\mathcal{M}} \subseteq \Sigma_2^{0,\mathcal{M}}$.

There is a computable $h: \omega^2 \to \omega$ where $\mathcal{M} \setminus \mathcal{M}_n = \bigcup_{k < \omega} \mathcal{M}_{h(n,k)}$ if $\mathcal{M} \setminus \mathcal{M}_n \neq \emptyset$.

Proof ...

- Claim 1 -----

Let $\{\mu_i : i < \omega\}$ be a computable presentation and assume $\mathcal{M} \setminus \mathcal{M}_n \neq \emptyset$. Note that $x \notin \mathcal{M}_n$ iff $d(x, \mu_{n_0}) \ge \frac{n_1}{n_2+1}$. Since the μ_i s are dense, by the triangle inequality, this is equivalent to the existence of $i, c_1, c_2 < \omega$ where

a.
$$d(x, \mu_i) < \frac{c_1}{c_2+1}$$
, i.e. $x \in \mathcal{M}_{\text{code}(i,c_1,c_2)}$;
b. $d(\mu_i, \mu_{n_0}) \ge \frac{n_1}{n_2+1} + \frac{2c_1}{c_2+1}$, i.e. $\neg D_{\le}(i, n_0, n_1 \cdot (c_2+1) + 2 \cdot c_1 \cdot (n_2+1), (n_2+1)(c_2+1) - 1)$.

So $\mathcal{M}\setminus\mathcal{M}_n$ is the union of $\mathcal{M}_{\text{code}(i,c_1,c_2)}$ where (b) holds. Note that (b) holding of $i, c_1, c_2 < \omega$ is computable. Hence we can enumerate them: h(n, 0) is the least (code of a) triple where (b) holds, and h(n, k + 1) is the least (code of a) triple $\tau < k$ with $\tau > h(n, k)$ (setting h(n, k + 1) = h(n, k) if there is no such triple). Then h is computable with $\mathcal{M} \setminus \mathcal{M}_n = \bigcup_{k < \omega} \mathcal{M}_{h(n,k)}$.

Now we show (2) by induction on *n*. For n = 1, let $X = \bigcup_{n < \omega} \mathcal{M}_{f(n)}$ for $f : \omega \to \omega$ computable. If $X = \mathcal{M}$, then X is easily $\Sigma_2^{0,\mathcal{M}} : \emptyset \in \Sigma_1^{0,\mathcal{M} \times \omega}$ implies $\mathcal{M} \times \omega \in \Pi_1^{0,\mathcal{M} \times \omega}$ with then $\mathcal{M} = \exists^{\omega}(\mathcal{M} \times \omega) \in \Sigma_2^{0,\mathcal{M}}$.

So without loss of generality, assume $X \neq \mathcal{M}$ and therefore $\mathcal{M} \setminus \mathcal{M}_{f(n)} \neq \emptyset$ for each $n < \omega$. Note that

$$\begin{split} X &= \exists^{\omega} \left(\bigcup_{n < \omega} \mathcal{M}_{f(n)} \times \{n\} \right) = \mathfrak{p} \left(\mathcal{M} \times \omega \setminus \left(\mathcal{M} \times \omega \setminus \bigcup_{n < \omega} \mathcal{M}_{f(n)} \times \{n\} \right) \right) \\ &= \exists^{\omega} \left(\mathcal{M} \times \omega \setminus \bigcup_{n < \omega} (\mathcal{M} \setminus \mathcal{M}_{f(n)}) \times \{n\} \right). \end{split}$$

Applying Claim 1, we have that $\bigcup_{n < \omega} (\mathcal{M} \setminus \mathcal{M}_{f(n)}) \times \{n\} = \bigcup_{n,k < \omega} \mathcal{M}_{h(f(n),k)} \times \{n\} \in \Sigma_{1}^{0,\mathcal{M} \times \omega}$ and therefore $X = \exists^{\omega} \left(\underbrace{\mathcal{M} \times \omega \setminus \bigcup_{n,k < \omega} \mathcal{M}_{h(f(n),k)} \times \{n\}}_{\prod_{1}^{0,\mathcal{M} \times \omega}} \right) \in \Sigma_{2}^{0,\mathcal{M}}.$

This shows the base case of $\Sigma_1^{0,\mathcal{M}} \subseteq \Sigma_2^{0,\mathcal{M}}$. For the inductive case n > 1, let $X \in \Sigma_n^{0,\mathcal{M}}$ so that $X = \exists^{\omega} Y$ for some $Y \in \prod_{n=1}^{0,\mathcal{M} \times \omega}$. Inductively, $\Sigma_{n-1}^{0,\mathcal{M}} \subseteq \Sigma_n^{0,\mathcal{M}}$ and therefore $\neg \Sigma_{n-1}^{0,\mathcal{M}} \subseteq \neg \Sigma_n^{0,\mathcal{M}}$, in other words $\prod_{n=1}^{0,\mathcal{M}} \subseteq \prod_{n=1}^{0,\mathcal{M}}$. In particular, $Y \in \prod_n^{0,\mathcal{M} \times \omega}$ and therefore $X = \exists^{\omega} Y \in \Sigma_{n+1}^{0,\mathcal{M}}$.

In particular, through a computable coding and a computable decoding, we can in essence identify the arithmetical hierarchy on products of baire space with itself and ω as the one just on baire space.

Now some basic closure properties can be established. Mostly these just come the fact that the constructions for the usual borel pointclasses can be done in a computable manner, or absorbed into defining formulas through a simple coding.

- 24 B • 7. Lemma Let \mathcal{M} be polish with presented basis $\{\mathcal{M}_n : n < \omega\}$. Let $0 < n < \omega_1$ and $A \in \mathcal{M}$. Therefore 1. $\Sigma_n^{0,\mathcal{M}}(A)$ is closed under $\cup, \cap, \exists^{\omega}$, and A-computable preimages; 2. $\Pi_n^{0,\mathcal{M}}(A)$ is closed under $\cup, \cap, \forall^{\omega}$, and A-computable preimages; and 3. $\Delta_n^{0,\mathcal{M}}(A)$ is closed under \cup, \cap, \neg , and A-computable preimages. Proof .:.

We only prove the result for $A = \emptyset$ for the sake of notation. The computable preimages were proven in Result 24 B • 4. Let $\{\mu_i : i < \omega\} \subset \mathcal{M}$ be a computable presentation.

 For n = 1, this is immediate: for X = Ø, this is trivial. For X ≠ Ø, if X ∈ Σ₁^{0,M×ω}, then for some computable f₀, f₁: ω → ω, we have X = ⋃_{n<ω} M_{f₀(n)} × {f₁(n)} so that ∃^ωX = ⋃_{n<ω} M_{g(n)} for g defined as follows: g(0) is the least m with ∃n ∈ ω (M_m × {n} ⊆ X); and recursively g(n + 1) = 1. $f_0(m_0)$ for the least $m = \text{code}(m_0, m_1) < n$ such that $f_1(m_0) = m_1$ and $\forall k < m_0 (f_1(m_0) > g(k))$ (set g(n + 1) = g(0) if there is no such m). Thus g is computable and witnesses that $\exists^{\omega} X \in \Sigma_1^{0,\mathcal{M}}$.

For n > 1, let X be $\sum_{n=0}^{0,\mathcal{M}\times\omega}$ so that $X = \exists^{\omega} Y$ for some $Y \in \prod_{n=1}^{0,\mathcal{M}\times\omega\times\omega}$ so by (de)coding with Result 24 B • 4, $Y' = \{\langle x, \operatorname{code}(\tau) \rangle : \langle x, \tau \rangle \in Y\}$ is $\prod_{n=1}^{0,\mathcal{M}\times\omega}$ with $\exists^{\omega} X = \exists^{\omega} Y' \in \sum_{n=0}^{0,\mathcal{M}}$.

- For finite unions, proceed by induction. This is immediate for $\Sigma_1^{0,\mathcal{M}}$, as $X = \bigcup_{n < \omega} \mathcal{M}_{f(n)}$ and $Y = \bigcup_{n < \omega} \mathcal{M}_{g(n)}$ have $X \cup Y = \bigcup_{n < \omega} \mathcal{M}_{h(n)}$ where h = f * g. Now suppose $X = \exists^{\omega} X'$ and $Y = \exists^{\omega} Y'$ where $X', Y' \in \prod_n^{0,\mathcal{M} \times \omega}$. Inductively, $X' \cup Y' \in \prod_n^{0,\mathcal{M} \times \omega}$ so that $X \cup Y = \exists^{\omega} (X' \cup Y') \in \Sigma_n^{0,\mathcal{M}}$.
- For finite intersections, we again proceed by induction. Suppose $X = \bigcup_{n \le \omega} \mathcal{M}_{f(n)} \in \Sigma_1^0$ and $Y = \bigcup_{n < \omega} \mathcal{M}_{g(n)} \in \Sigma_1^0 \text{ for computable } f, g : \omega \to {}^{<\omega}\omega.$

- Claim 1 —

There is a computable $h : \omega^3 \to \omega$ where h(n, m, k) = 0 iff $\mathcal{M}_n \cap \mathcal{M}_m = \emptyset$, and otherwise $\mathcal{M}_n \cap \mathcal{M}_m = \bigcup_{k < \omega} \mathcal{M}_{h(n,m,k)-1}.$

Proof .:. Given an h' that satisfies $\mathcal{M}_n \cap \mathcal{M}_m = \bigcup_{k < \omega} \mathcal{M}_{h'(n,m,k)}$, we merely define h(n,m,k) by $h(n,m,k) = \begin{cases} h'(n,m,k) + 1 & \text{if } \mathcal{M}_n \cap \mathcal{M}_m \neq \emptyset\\ 0 & \text{otherwise.} \end{cases}$ Note that $\mathcal{M}_n \cap \mathcal{M}_m = \emptyset$ is computable from n and $m: \mathcal{M}_n \cap \mathcal{M}_m = \emptyset$ iff $d(\mu_{n_0}, \mu_{m_0}) < \frac{n_1}{n_2+1} + \frac{m_1}{m_2+1}$, i.e. $D_<(n_0, m_0, n_1 \cdot (m_2 + 1) + m_1 \cdot (n_2 + 1), (n_2 + 1) \cdot (m_2 + 1))$. So now we work towards defining such an h': assume $\mathcal{M}_n \cap \mathcal{M}_m \neq \emptyset$. Note that each \mathcal{M}_k is the open ball of radius $k_1/(k_2 + 1)$ around μ_{k_0} where $k = \text{code}(\langle k_0, k_1, k_2 \rangle)$. In particular, $x \in \mathcal{M}_n \cap \mathcal{M}_m$ iff $d(x, \mu_{n_0}) < \frac{n_1}{n_2+1} \wedge d(x, \mu_{m_1}) < \frac{m_1}{m_2+1}$. By the density of the μ_i s, this happens iff there is some μ_i within $\frac{c_1}{c_2+1}$ of x and within $\frac{n_1}{n_2+1} - \frac{c_1}{c_2+1}$ of μ_{n_0} , and similarly for μ_{m_0} . So $x \in \mathcal{M}_n \cap \mathcal{M}_m$ iff there are some $i, c_1, c_2 \in \omega$ where a. $x \in \mathcal{M}_{\operatorname{code}(i,c_1,c_2)};$ b. $c_1 \cdot (n_2 + 1) < n_1 \cdot (c_2 + 1)$ and $c_1 \cdot (m_2 + 1) < m_1 \cdot (c_2 + 1)$; c. $D_{\leq}(i, n_0, n_1 \cdot (c_2 + 1) - c_1 \cdot (n_2 + 1), (n_2 + 1) \cdot (c_2 + 1) - 1)$; and d. $D_{\leq}(i, m_0, m_1 \cdot (c_2 + 1) - c_1 \cdot (m_2 + 1), (m_2 + 1) \cdot (c_2 + 1) - 1).$ Hence $\mathcal{M}_n \cap \mathcal{M}_m$ is the union of $\mathcal{M}_{code(i,c_1,c_2)}$ for the i, c_1, c_2 where (b)–(d) hold. So define by recursion h'(n, m, 0) to be the least (code of a) triple where (b)-(d) hold; and h'(n, m, k + 1)to be the least such (code of a) triple $\tau < k$ with $\tau > h'(n,m,k)$ (set h'(n,m,k+1) =h'(n,m,k) if there are no such triples). This h' is clearly computable (and total) with $\mathcal{M}_n \cap$ $\mathcal{M}_m = \bigcup_{k < \omega} \mathcal{M}_{h'(n,m,k)}.$ Thus $X \cap Y = \bigcup_{(n,m,k) \in P} \mathcal{M}_{h(f(n),g(m),k)}$ where $P = \{(n,m,k) : h(f(n),g(m),k) \neq 0\}$ which is computable and clearly gives $X \cap Y$ as $\Sigma_1^{0,\mathcal{M}}$.

For n > 1, suppose $X = \exists^{\omega} X'$ and $Y = \exists^{\omega} Y'$ for $X', Y' \in \prod_{n=1}^{0,\mathcal{M}\times\omega}$. By Result 24 B • 4, consider $X'' = \{\langle x, \langle \tau_0, \tau_1 \rangle \rangle \in \mathcal{M} \times \omega^2 : \langle x, \tau_0 \rangle \in X'\} \in \prod_{n=1}^{0,\mathcal{M}\times\omega^2}$ $Y'' = \{\langle x, \langle \tau_0, \tau_1 \rangle \rangle \in \mathcal{M} \times \omega^2 : \langle x, \tau_1 \rangle \in Y'\} \in \prod_{n=1}^{0,\mathcal{M}\times\omega^2}$. Inductively, $Z = X'' \cap Y'' \in \prod_{n=1}^{0,\mathcal{M}\times\omega^2}$ and it's not difficult to see that $\exists^{\omega} Z = X \cap Y$, as desired.

- 2. As complements of $\Sigma_n^{0,\mathcal{M}}$ -sets, this is clear: closure under finite unions follows from $\Sigma_n^{0,\mathcal{M}}$'s closure under finite intersections; and closure under \forall^{ω} follows from Σ_n^{0} 's closure under \exists^{ω} .
- 3. Closure under finite unions and intersections follows just from the fact that both $\Sigma_{\alpha}^{0,\mathcal{M}}$ and $\Pi_{\alpha}^{0,\mathcal{M}}$ are closed under these. For complements, $X \in \Delta_{\alpha}^{0,\mathcal{M}}$ implies $X \in \Sigma_{\alpha}^{0,\mathcal{M}}$ so that $\mathcal{M} \setminus X \in \Pi_{\alpha}^{0,\mathcal{M}}$. But we also have from $X \in \Delta_{\alpha}^{0,\mathcal{M}}$ that $X \in \Pi_{\alpha}^{0,\mathcal{M}}$, which tells us $\mathcal{N} \setminus X \in \Sigma_{\alpha}^{0,\mathcal{M}}$ and hence $\mathcal{N} \setminus X \in \Delta_{\alpha}^{0,\mathcal{M}}$.

As a result of Theorem 24 A • 15 and $\Sigma_1^{0,\mathcal{M}}$'s closure under \exists^{ω} , we have an alternative characterization of the arithmetical hierarchy: we start with Δ_1^0 -relations and then add \exists^{ω} and \forall^{ω} quantifiers to get the Σ_n^0 and Π_n^0 pointclasses.

— 24 B · 8. Corollary

- Let \mathcal{M} be polish with presented basis $\{\mathcal{M}_n : n < \omega\}$. Let $A \subseteq \mathcal{M}$ and $X \subseteq \mathcal{M}$. Let $0 < n < \omega$. Therefore,
 - X is $\Sigma_n^{0,\mathcal{M}}(A)$ iff there is some $R \subseteq \Delta_1^{0,\mathcal{M} \times \omega^n}(A)$ where $X = \exists^{\omega} \forall^{\omega} \exists^{\omega} \cdots Q^{\omega} R$, alternating *n* quantifiers where *Q* is ' \exists ' if *n* is odd and ' \forall ' if *n* is even.
 - X is $\Pi_n^{0,\mathcal{M}}(A)$ iff there is some $R \subseteq \Delta_1^{0,\mathcal{M} \times \omega^n}(A)$ where $X = \forall^{\omega} \exists^{\omega} \forall^{\omega} \cdots Q^{\omega} R$, alternating *n* quantifiers where Q is ' \forall ' if *n* is odd and ' \exists ' if *n* is even.

Proof .:.

Proceed by induction on *n* with $A = \emptyset$ for the sake of notation. For n = 1, if X has the form $\exists^{\omega} R$ for $R \in \Delta_1^{0,\mathcal{M}\times\omega} \subseteq \Sigma_1^{0,\mathcal{M}\times\omega}$, then X is $\Sigma_1^{0,\mathcal{M}}$ by Lemma 24 B • 7. If X is $\Sigma_1^{0,\mathcal{M}}$, we must show that there is a computable $R \subseteq \mathcal{M} \times \omega$ where $x \in X$ iff $\exists n \in \omega R(x, n)$. Since $X = \bigcup_{n < \omega} \mathcal{M}_{f(n)}$ for some computable $f : \omega \to \omega$, define $R = \bigcup_{n < \omega} \mathcal{M}_{f(n)} \times \{n\}$. So clearly $\exists^{\omega} R = X$, and therefore it suffices to show the following claim.

- Claim 1 -----R is computable.

Proof .:.

By Theorem 24 A • 15, it suffices to show $R, \neg R \in \Sigma_1^{0,\mathcal{M}}$. Clearly $R \in \Sigma_1^{0,\mathcal{M}}$ and $\neg R$ is $\Sigma_1^{0,\mathcal{M}}$ since \mathcal{M} 's basis is presented: let $\{\mu_i : i < \omega\}$ be a computable presentation of \mathcal{M} . Therefore, for any $n = \operatorname{code}(n_0, n_1, n_2) < \omega, x \notin \mathcal{M}_n$ iff the distance between x and n_0 is at least $\frac{n_1}{n_2+1}$. By the density of the presentation, this happens iff there is some $i < \omega$ and $c_1, c_2 < \omega$ where x is within $\frac{c_1}{c_2+1}$ of μ_i , and μ_i is not within $\frac{c_1}{c_2+1} + \frac{n_1}{n_2+1}$, i.e. for

$$B = \{ \langle i, c_1, c_2 \rangle \in \omega^3 : \neg D_{<}(i, n_0, c_1 \cdot (n_2 + 1) + n_1 \cdot (c_2 + 1), (c_2 + 1) \cdot (n_2 + 1) - 1) \},\$$

we get $\mathcal{M} \setminus \mathcal{M}_n = \bigcup_{(i,c_1,c_2) \in B} \mathcal{M}_{\text{code}(i,c_1,c_2)}$. So there is some $h : \omega^2 \to \omega$ where $\neg \mathcal{M}_n = \bigcup_{k < \omega} \mathcal{M}_{h(n,k)}$. In particular,

$$\neg R = \mathcal{M} \times \omega \setminus \bigcup_{n < \omega} \mathcal{M}_{f(n)} \times \{n\} = \bigcup_{n < \omega} (\mathcal{M} \setminus \mathcal{M}_{f(n)}) \times \{n\} = \bigcup_{n, k < \omega} \mathcal{M}_{h(f(n), k)} \times \{n\} \in \Sigma_1^{0, \mathcal{M} \times \omega}. \quad \exists$$

For n + 1, any $X \in \sum_{n+1}^{0,\mathcal{M}}$ has $X = \exists^{\omega} Y$ for some $Y \in \prod_{n=1}^{0,\mathcal{M}\times\omega}$ where inductively $Y = \forall^{\omega}\exists^{\omega}\cdots R$ for some $R \in \Delta_{1}^{0,\mathcal{M}\times\omega\times\omega^{n}} = \Delta_{1}^{0,\mathcal{M}\times\omega^{n+1}}$ and therefore $X = \exists^{\omega}\forall^{\omega}\exists^{\omega}\cdots R$ as in the problem statement. Through negations, it's easy to show the result for $\prod_{n=1}^{0,\mathcal{M}}$ from the result on $\sum_{n=1}^{0,\mathcal{M}}$.

We now aim to show that all of the levels of these hierarchies are distinct in the same sort of way as with the borel hierarchy in Theorem 22 A • 13, and with the arithmetical hierarchy on ω regarding computability. Recall the notion of a universal set from Definition 22 A • 12. The fact that the pointclass in question there is presented as "boldface" and a subset of $\mathcal{P}(\mathcal{N})$ (rather than the powerset of an arbitrary polish space) is irrelevant: the idea is just that we have parametrized every element of the pointclass Γ by a real number. In the case of $\Sigma_n^0 s$, we actually are doing this relative to a natural number. Rather than modify the definition and show that the two work out equivalently, we will just identify the open set $\{e\} \subseteq \omega$ with the cone $\mathcal{N}_{\{e\}} \subseteq \mathcal{N}$.

- 24 B • 9. Theorem

For each $n < \omega$, there is a $\sum_{n=1}^{\infty} n^{-1}$ universal set and a $\prod_{n=1}^{\infty} n^{-1}$ universal set.

Proof .:.

Mostly this just mimics the proof of Theorem 22 A • 13 in a computable way. Again, it suffices to give just Σ_n^0 -universal sets, since if U is such a set, $\neg U$ is Π_n^0 -universal.

For n = 1, we just consider the computable map $f(e, n) = \llbracket e \rrbracket(n)$ where $\llbracket e \rrbracket$ is the computable function computed by the program (coded by) $e \in \omega$. Take $U = \bigcup_{e,n < \omega} \mathcal{N}_{(e)} \times \mathcal{N}_{f(e,n)}$ so that $U \in \Sigma_1^{0,\mathcal{N} \times \omega}$. Hence for $r = \langle r_n : n < \omega \rangle \in \mathcal{N}$, the section $U_r = \{x \in \mathcal{N} : \langle r, x \rangle \in U\} = \bigcup_{n < \omega} \mathcal{N}_{f(r_0,n)} = \bigcup_{n < \omega} \mathcal{N}_{\llbracket r_0 \rrbracket(n)}$. Any $X \in \Sigma_1^0$ is $\bigcup_{n < \omega} \mathcal{N}_{\llbracket e \rrbracket(n)}$ for some e and therefore $X = U_{\langle e \rangle \frown r}$ for any $r \in \mathcal{N}$.

For n + 1, let U be $\prod_{n=1}^{0}$ -universal. Using a computable coding and decoding with Result 24 B • 4, we can instead regard $U \subseteq \mathcal{N}^2 \times \omega$ as $\prod_{n=1}^{0,\mathcal{N}\times\omega}$ -universal, in which case $\mathfrak{p}_{\mathcal{N}\times\mathcal{N}}U$ is Σ_{n+1}^0 -universal. \dashv

- **24B**•10. Corollary For $0 < n < \omega$, $\Delta_n^0 \subsetneq \Sigma_n^0 \subsetneq \Delta_{n+1}^0$, and similarly for Π_n^0 .

Proof .:.

Let $U \subseteq \mathcal{N} \times \mathcal{N}$ be Σ_n^0 -universal as per Theorem 24 B • 9. By coding, consider instead $U' = \{x * y : \langle x, y \rangle \in U\} \in \Sigma_n^0$. Therefore $U' \in \Sigma_n^0 \setminus \Pi_n^0 = \Sigma_n^0 \setminus \Delta_n^0$. To see this, otherwise we'd have $\neg U' \in \Sigma_n^0$ as well as $D = \{x \in \mathcal{N} : x * x \in \neg U'\}$ as the computable preimage of $\neg U'$ under the map $x \mapsto x * x$. Therefore $D = U_r$ for some $r \in \mathcal{N}$ where then $r \in U_r$ iff $r \in D$ iff $r * r \in \neg U'$ iff $r \notin U_r$, a contradiction. So there can be no such r, meaning $\neg U' \notin \Sigma_n^0$ i.e. $U' \notin \Pi_n^0$. This shows $\Delta_n^0 \subsetneq \Sigma_n^0$.

To see that $\Sigma_n^0 \subsetneq \Delta_{n+1}^0$, just note that $\neg U' \in \Pi_n^0 \subseteq \Delta_{n+1}^0$ but $\neg U' \notin \Sigma_n^0$ for the reasons as above. Hence $\neg U' \in \Delta_{n+1}^0 \setminus \Sigma_n^0$.

These can also be easily relativized to $A \subseteq \mathcal{N}$ just by using parameters over the slightly larger spaces $\mathcal{N} \times \mathcal{N}$ (or via coding).

§ 24 C. The analytical hierarchy

The analytical hierarchy is incredibly important, as much as the projective hierarchy is, given that it is a refinement of it. So we give a definition analogous to that of the projective hierarchy. Given that the arithmetical hierarchy corresponds to adding quantifiers over ω , the analytical hierarchy corresponds to adding quantifiers over the (potentially) larger polish space. In particular, we will use $\exists^{\mathcal{M}}$ defined by

$$\exists^{\mathcal{M}} X = \{ x : \exists y \in \mathcal{M} \ (\langle x, y \rangle \in X) \},\$$

and similarly $\forall^{\mathcal{M}}$ is just $\neg \exists^{\mathcal{M}} \neg$. We could also do this just with coding rather than actual order pairs, as we will later show.

24C • 1. Definition
Let M be polish with basis {M_n : n < ω}. We form the *relativized analytical hierarchy* as follows: for X ⊆ M, A ⊆ M, and n < ω,
X is Σ^{1,M}₀(A) iff X is Σ^{0,M}₁(A);
X is Π^{1,M}_n(A) iff N \ X is Σ^{1,M}_n(A);
X is Σ^{1,M}_{n+1}(A) iff X = ∃^MY for some Y ∈ Π^{1,M×M}_n(A); and
X is Δ^{1,M}_{n+1}(A) iff X is both Σ^{1,M}_n(A) and Π^{1,M}_n(A).
These Σ^{1,M}_n(A), Π^{1,M}_n(A), Δ^{1,M}_n(A) are the A-analytical pointclasses, and the sets in them A-analytical. We write just Σ^{1,M}_n for A = Ø and just Σ¹_n for M = N.

One gets the expected properties from the notation and definition, similar to Definition 22 C • 3 and Figure 22 C • 4.

$$\begin{split} \Sigma_1^0 &= \Sigma_0^1 & \Sigma_{\omega_1^{CK}}^0 = \Delta_1^1 & \Sigma_1^1 & \Sigma_2^1 \\ & \ddots & \ddots \\ \Pi_1^0 &= \Pi_0^1 & \Pi_1^1 & \Pi_2^1 & \Pi_2^1 \end{split}$$

24 C • 2. Figure: The analytical hierarchy

Here $\sum_{\omega_1}^{0} \sum_{\omega_1}^{0} \sum_{\omega_1}^{0}$

We can do the same analysis as with Corollary 24 B • 8 to get that any $\Sigma_1^{1,\mathcal{M}}$ -relation is just of the form $\exists^{\mathcal{M}}\forall^{\omega} R$ for $R \in \Delta_1^{0,\mathcal{M}\times\mathcal{M}\times\omega}$, and so on.

— 24 C • 3. Corollary

Let \mathcal{M} be polish with presented basis $\{\mathcal{M}_n : n < \omega\}$. Let $A \subseteq \mathcal{M}$ and $X \subseteq \mathcal{M}$. Let $0 < n < \omega$. Therefore,

- X is $\Sigma_n^{1,\mathcal{M}}(A)$ iff there is some A-computable $R \subseteq \mathcal{M}^{n+1} \times \omega$ where $X = \exists^{\mathcal{M}} \forall^{\mathcal{M}} \exists^{\mathcal{M}} \cdots Q^{\omega} R$, alternating n quantifiers where Q is ' \exists ' if n is odd and ' \forall ' if n is even.
- X is $\Pi_n^{1,\mathcal{M}}(A)$ iff there is some A-computable $R \subseteq \mathcal{M}^{n+1} \times \omega$ where $X = \forall \mathcal{M} \exists \mathcal{M} \forall \mathcal{M} \cdots Q^{\omega} R$, alternating *n* quantifiers where Q is '\delta' if *n* is odd and '\exp if *n* is even.

Proof .:.

Proceed by induction on *n* with $A = \emptyset$ for the sake of notation. For n = 1, if X has the form $\exists^{\mathcal{M}} \forall^{\omega} R$ for $R \in \Delta_1^{0,\mathcal{M}^2 \times \omega} \subseteq \Pi_1^{0,\mathcal{M}^2 \times \omega}$, then X is $\Sigma_1^{1,\mathcal{M}}$ by Lemma 24 B•7 (2). This shows one direction. Furthermore, if X is $\Sigma_1^{0,\mathcal{M}}$, then $X = \exists^{\mathcal{M}} Y$ for some $Y \in \Pi_1^{0,\mathcal{M} \times \mathcal{M}}$ which, by Corollary 24 C•3, takes the form $\forall^{\omega} R$ for some computable $R \subseteq \mathcal{M}^2 \times \omega$ where then $X = \exists^{\mathcal{M}} \forall^{\omega} R$, as desired.

For n + 1, any $X \in \sum_{n+1}^{1,\mathcal{M}}$ has $X = \exists^{\mathcal{M}} Y$ for some $Y \in \prod_{n=1}^{1,\mathcal{M}\times\omega}$ where inductively $Y = \forall^{\mathcal{M}}\exists^{\mathcal{M}}\cdots R$ for some computable R and therefore $X = \exists^{\mathcal{M}}\forall^{\mathcal{M}}\exists^{\mathcal{M}}\cdots R$ as in the problem statement. Through negations, it's easy to show the result for $\prod_{n=1}^{1,\mathcal{M}}$ from the result on $\sum_{n=1}^{1,\mathcal{M}}$.

Other polish spaces will be practically irrelevant when working with \mathcal{N} , as we will see later. As a result, we will focus just on \mathcal{N} and ω for the remainder of the section, allowing ourselves to code products freely. We now show the following expected closure properties of the analytical pointclasses.

- 24 C • 4. Result

- Let $A \subseteq \mathcal{N}$ and $0 < n < \omega$. Therefore,
 - $\Sigma_n^1(A)$ is closed under $\cap, \cup, \exists^{\omega}, \forall^{\omega}, \text{ and } \exists^{\mathcal{N}};$
 - $\Pi_n^1(A)$ is closed under $\cap, \cup, \exists^{\omega}, \forall^{\omega}, \text{ and } \forall^{\mathcal{N}};$ and
 - $\Delta_n^1(A)$ is closed under $\cap, \cup, \exists^{\omega}, \forall^{\omega}, \text{ and } \neg$.

Proof ...

Note that computable relations—being $\Delta_1^{0,\omega}$ —are closed under \lor , \land , and \neg . Proceed by induction on *n* for all of these pointclasses simultaneously, assuming for simplicity that $A = \emptyset$.

• For n = 1, an $X \subseteq \mathcal{N}$ is $\Sigma_1^1(A)$ iff $X = \exists^{\mathcal{N}} Y$ for some $Y \in \Pi_1^0$, meaning

 $x \in X$ iff $\exists y \in \mathcal{M} \ \forall n \in \omega \ R(x \upharpoonright n, y \upharpoonright n)$

for some computable relation R. So if $X_0, X_1 \in \Sigma_1^1$, are given by computable relations R_0 and R_1 respectively as above, then

$$\begin{split} x \in X_0 \land X_1 & \text{iff} \quad \exists y_0 \in \mathcal{M} \ \exists y_1 \in \mathcal{M} \ \forall n \in \omega \ (R_0(x \upharpoonright n, y_0 \upharpoonright n) \land R_1(x \upharpoonright n, y_1 \upharpoonright n)) \\ & \text{iff} \quad \exists y \in \mathcal{M} \ \forall n \in \omega \ (y = y_0 \ast y_1 \land R_0(x \upharpoonright n, y_0 \upharpoonright n) \land R_1(x \upharpoonright n, y_1 \upharpoonright n)) \\ x \in X_0 \lor X_1 & \text{iff} \quad \exists y \in \mathcal{M} \ \forall n \in \omega \ (R_0(x \upharpoonright n, y \upharpoonright n)) \lor R_1(x \upharpoonright n, y \upharpoonright n)) \\ x \in \exists^{\omega} X_0 & \text{iff} \quad \exists m \in \omega \ \exists y \in \mathcal{M} \ \forall n \in \omega \ R_0(x \upharpoonright n, y \upharpoonright n, m) \\ & \text{iff} \quad \exists y \in \mathcal{M} \ \forall n \in \omega \ (y = \langle y_0 \rangle^\frown y' \land R_0(x \upharpoonright n, y', y_0)) \\ x \in \forall^{\omega} X_0 & \text{iff} \quad \forall m \in \omega \ \exists y \in \mathcal{M} \ \forall n \in \omega \ R_0(x \upharpoonright n, y \upharpoonright n, m) \\ & \text{iff} \quad \exists y \in \mathcal{M} \ \forall n \in \omega \ (y = \langle y_i \in \mathcal{M} : i < \omega \rangle \land n = \text{code}(n_0, n_1) \land R_0(x \upharpoonright n_0, y_n \upharpoonright n_0, n_1)). \end{split}$$

Here, the formulas after the quantifiers are all computable. In some sense, we should be a bit careful, since $y = \langle y_i : i < \omega \rangle$ isn't computable over ω . But it's not difficult to see that we're just leaving out the relevant restrictions: the computable map $\langle n_0, n_1 \rangle \mapsto m$ where $y \upharpoonright m$ contains the info of $y_{n_1} \upharpoonright n_0$.

For n + 1, the inductive results for Π_n^1 yields the same results for $\Sigma_{n+1}^1 = \exists^{\mathcal{N}} \Pi_n^1$ by similar ideas as above: $\Sigma_{n+1}^1 \vee \Sigma_{n+1}^1 \subseteq \Sigma_{n+1}^{1}$, $\exists^{\omega} \Sigma_{n+1}^1 \subseteq \Sigma_{n+1}^{1}$, and $\exists^{\mathcal{N}} \Sigma_{n+1}^1 \subseteq \Sigma_{n+1}^{1}$ are easy by any computable

coding of pairs. Conjunctions and disjunctions follow nicely by coding as well: $R, R' \in \prod_n^1$ yields $\exists y \ R(x, y) \land \exists y \ R'(x, y) \quad \text{iff} \quad \exists y_0 \ \exists y_1(R(x, y_0) \land R(x, y_1)),$ which is $\sum_{n+1}^1 . \ \forall^{\omega} \sum_{n+1}^1 \subseteq \sum_{n+1}^1$ by the same reasoning above: for $R \in \prod_n^1,$ $x \in \forall^{\omega} \exists^{\mathcal{N}} R \quad \text{iff} \quad \forall m \in \omega \ \exists y \in \mathcal{N} \ R(x, y, m)$ $\text{iff} \quad \exists y \in \mathcal{N} \ \forall m \in \omega \ (y = \langle y_i \in \mathcal{N} : i < \omega \rangle \land R(x, y_m, m)).$

- These follow from the results on Σ_n^1 .
- All but negation follows from the results on Σ_n^1 and Π_n^1 . For negation, $X \in \Sigma_n^1$ iff $\neg X \in \Pi_n^1$, and similarly $X \in \Pi_n^1$ iff $\neg X \in \Sigma_n^1$. Hence $X \in \Delta_n^1$ implies $X, \neg X \in \Sigma_n^1 \cap \Pi_n^1 = \Delta_n^1$.

We also have the expected containments.

- **24C·5.** Result Let $n < \omega$. Therefore $\Delta_n^1 \subseteq \Sigma_n^1 \subseteq \Delta_{n+1}^1$, and similarly for Π_n^1 .

Proof .:.

Proceed by induction on n. We always have $\Delta_n^1 \subseteq \Sigma_n^1$ just by definition. So we need to show

1. $\Sigma_n^1 \subseteq \Pi_{n+1}^1$; and 2. $\Sigma_n^1 \subseteq \Sigma_{n+1}^1$.

For (1), we easily have that if $X \in \Sigma_n^1$, then $X \times \mathcal{M} \in \Sigma_n^{1,\mathcal{M}\times\mathcal{M}}$ and so $\forall^{\mathcal{M}}X \times \mathcal{M} = X \in \prod_{n+1}^1$. For (2), we just add a dummy quantifier at the beginning and rely on coding: any $X \in \Sigma_n^1$ has $X = \exists^{\mathcal{N}}\forall^{\mathcal{M}} \cdots P^{\mathcal{M}}Q^{\omega}R$ for some computable R, where $\{P, Q\} = \{`\exists', `\forall'\}$ and they continue the alternating pattern. As a result,

 $x \in X \quad \text{iff} \quad \exists y_0 \in \mathcal{N} \ \forall y_1 \in \mathcal{N} \ \cdots \ P \ y_n \in \mathcal{N} \ Qm \in \omega \ R(x, \vec{y}, m)$ $\text{iff} \quad \exists y_0 \in \mathcal{N} \ \forall y_1 \in \mathcal{N} \ \cdots \ P \ y_n \in \mathcal{N} \ Qz \in \mathcal{N} \ Pk \in \omega \ R(x, \vec{y}, z_0),$

since computable relations are closed under computable substitution (by Result 24 B • 4), it follows that $R' = \{\langle x, \vec{y}, z, k \rangle \in \mathcal{N} \times \mathcal{N}^n \times \mathcal{N} \times \omega : R(x, \vec{y}, z_0)\}$ is also computable with clearly $X = \exists^{\mathcal{N}} \forall^{\mathcal{N}} \cdots P^{\mathcal{N}} Q^{\mathcal{N}} P^{\omega} R'$, and thus $X \in \Sigma_{n+1}^1$.

Firstly, note that, just as with Σ_1^1 , there are many different equivalent definitions for Σ_1^1 that we will end up proving. This suggests Σ_1^1 is quite canonical, or at least a natural idea to consider.

24C • 6. Result
Let X ⊆ N. Therefore the following are equivalent.
1. X = {x : ∃y ∈ N (x * y ∈ Y)} for some Y ∈ Π₁⁰...
2. X = ∃^NY for some Y ∈ Π₁^{0,N×N}.
3. X = ∃^MY for some polish M with presented basis {M_n : n < ω} and some Y ∈ Π₁^{0,N×M}.
4. X = ∃^NY for some arithmetical Y ⊆ N × N.
5. X = im f for some computable f : N → N.

Proof ...

(1) ↔ (2) Suppose (1) holds with Y ∈ Π⁰₁. Consider Y' = {⟨x, y⟩ : x * y ∈ Y}, the computable preimage of ⟨x, y⟩ ↦ x * y so that Y' is also Π⁰₁ by Result 24 B • 4, and clearly X = pY'. Similarly, if (2) holds, Y' = {x * y : ⟨x, y⟩ ∈ Y} is the continuous preimage of x * y ↦ ⟨x, y⟩ so that A = {x : ∃y (x * y ∈ Y')}. For the remainder of the proof, we therefore identify Π⁰₁, w×N and Π⁰₁, and similarly for the other arithmetical pointclasses (partially just for the sake of notation).

(2) \leftrightarrow (3) One direction is easy as $\mathcal{N} = \mathcal{M}$ shows (2) \rightarrow (3). So suppose (3) holds, working towards (2). By

Theorem 24 A • 13, there is a computable surjection $f : \mathcal{N} \to \mathcal{M}$. So define $F : \mathcal{N} \times \mathcal{N} \to \mathcal{N} \times \mathcal{M}$ by $F(x, y) = \langle x, f(y) \rangle$ and set $Y' = F^{-1} "Y \in \Pi_1^{0, \mathcal{N} \times \mathcal{N}}$. We then have $X = \exists^{\mathcal{M}} Y = \exists^{\mathcal{N}} Y' \in \Sigma_1^0$, i.e. (2) holds.

- (2) \leftrightarrow (4) One direction is easy as every $\Pi_1^{0,\mathcal{N}\times\mathcal{N}}$ -set is arithmetical. So assume (4). Since $\exists^{\mathcal{N}}\Sigma_1^1$ (using $\exists^{\mathcal{N}}$ in the sense of (1)) is contained in Σ_1^1 , it suffices to show $\bigcup_{n<\omega}\Sigma_n^0 \subseteq \Sigma_1^1$ as then $\exists^{\mathcal{N}}\bigcup_{n<\omega}\Sigma_n^0 \subseteq \exists^{\mathcal{N}}\Sigma_1^1 = \Sigma_1^1$. To do this, it suffices to show all computable relations are Σ_1^1 , and Σ_1^1 is closed under \exists^{ω} and \forall^{ω} . By Corollary 24 C 3, this generates the arithmetical hierarchy, showing that all arithmetical sets are Σ_1^1 . Clearly all computable relations are Σ_1^1 as they are all Π_1^0 and any $Y \in \Pi_1^0$ has $Y \times \mathcal{N} \in \Pi_1^0$ where then $\exists^{\mathcal{N}}(Y \times \mathcal{N}) = Y \in \Sigma_1^1$. Result 24 C 4 implies Σ_1^1 is closed under \forall^{ω} and \exists^{ω} , and so the arithmetical hierarchy is contained in Σ_1^1 which is closed under $\exists^{\mathcal{N}}$. Hence (2) holds.
- (2) \rightarrow (5) Let $X = \exists^{\mathcal{N}} Y$ for some $Y \in \Pi_1^0$. Knowing a bit about metric spaces, since $\Pi_1^{0,\mathcal{N}^2} \subseteq \Pi_1^0$ implies Y is a closed subset of $\mathcal{N} \times \mathcal{N}$, Y (with the restricted topology) is complete and hence also a polish space. So by Theorem 24 A • 13, there is a computable surjection $f : \mathcal{N} \to Y$ where then the projection map $p_1(x, y) = x$ has $p_1 \circ f : \mathcal{N} \to \mathcal{N}$ as computable with $\operatorname{im}(p_1 \circ f) = \exists^{\mathcal{N}} Y = X$.
- (5) \rightarrow (4) If f is computable then the neighborhood graph of f as a relation is $\Sigma_1^{0, N \times \omega}$. Hence f as a relation is Π_2^0 : f(x) = y iff $\forall m < \omega \ (y \upharpoonright m \lhd f(x))$. Since $y \upharpoonright m \lhd f(x)$ is Σ_1^0 , it follows that f(x) = y is Π_2^0 so that Y = f implies $X = \text{im } f = \exists^{\mathcal{N}} Y$ for $Y \in \Pi_2^0$. \dashv

We also get that the above are equivalent with $X = \exists^{\mathcal{N}} Y$ for some hyperarithmetical $Y \subseteq \mathcal{N} \times \mathcal{N}$, but this will not be proven here.

– 24 C • 7. Result -

For $n < \omega$, Σ_n^1 , Π_n^1 , and Δ_n^1 are all closed under computable preimages, meaning that if \mathcal{M} is polish with presented basis $\{\mathcal{M}_n : n < \omega\}$ and $f : \mathcal{N} \to \mathcal{M}$ is computable with $X \in \Sigma_n^{1,\mathcal{M}}$, then $f^{-1} X \in \Sigma_n^1$, and similarly for the other pointclasses.

Proof .:.

Proceed by induction on *n*. The case of n = 0 was shown in Lemma 24 B • 7. For n + 1, suppose $X = \exists^{\mathcal{M}} Y$ for $Y \in \prod_{n}^{1,\mathcal{M}\times\mathcal{M}}$. Therefore $f(x) \in X$ iff $\exists y \in \mathcal{M}$ ($\langle f(x), y \rangle \in Y$). Inductively, $\langle f(x), y \rangle \in Y$ defines a $\prod_{n}^{1,\mathcal{N}\times\mathcal{M}}$ -relation Y' and therefore $f^{-1}"X = \exists^{\mathcal{M}} Y' \in \Sigma_{n+1}^{1}$ by Result 24 C • 6 (3). The result for \prod_{n+1}^{1} follows easily by taking negations, and the result for Δ_{n+1}^{1} follows from the results on Σ_{n+1}^{1} and \prod_{n+1}^{1} .

This shows that computable coding and decoding does no harm, as expected, and so as indicated above, we will work merely with \mathcal{N} and freely consider, e.g. Σ_n^1 instead of more technically proper $\Sigma_n^{1,\mathcal{N}\times\mathcal{N}}$.

We also get a connection between the analytical and projective hierarchies, just as with the borel hierarchy and the (hyper)arithmetical hierarchy.

- 210.0	Pocult
Z4 C * 0.	Result
For $n < \omega$,	$\sum_{n=1}^{1} = \bigcup_{A \in \mathcal{N}} \Sigma_{n}^{1}(A)$; and similarly for $\prod_{n=1}^{1}$ and Π_{n}^{1} . Equivalently, $\sum_{n=1}^{1} \Sigma_{n}^{1}(\mathcal{N})$.

Proof .:.

Proceed by induction on *n*. For n = 0, this follows from the result on the borel hierarchy, Corollary 24 A • 6. For n + 1, if $X \in \prod_{n=1}^{n} \mathbb{I}$ then inductively $X \in \prod_{n=1}^{1} (A)$ for some $A \in \mathcal{N}$ so that $\exists^{\mathcal{N}} X \in \sum_{n=1}^{1} (A)$ and thus $\sum_{n=1}^{n} \mathbb{I} = \bigcup_{A \in \mathcal{N}} \sum_{n=1}^{1} (A)$.

Given that the result holds for $\sum_{n=1}^{1}$, it's clear that it holds for $\prod_{n=1}^{1}$, since

$$\prod_{n=1}^{1} = \neg \sum_{n=1}^{1} = \neg \bigcup_{A \in \mathcal{N}} \Sigma_{n}^{1}(A) = \bigcup_{A \in \mathcal{N}} \neg \Sigma_{n}^{1}(A) = \bigcup_{A \in \mathcal{N}} \Pi_{n}^{1}(A). \quad \dashv$$

We now aim to show that all of the levels of the analytical hierarchy are distinct in the same sort of way as with arithmetical, borel, and projective hierarchies in Theorem $24 B \cdot 9$, Theorem $22 A \cdot 13$, and Theorem $22 C \cdot 13$.

– 24 C • 9. Theorem

For each $n < \omega$, there is a $\sum_{n=1}^{1}$ -universal set and a $\prod_{n=1}^{1}$ -universal set.

Proof .:.

Mostly this just mimics the proof of Theorem 22 C • 13 in a computable way. Again, it suffices to give just Σ_n^1 -universal sets, since if U is such a set, $\neg U$ is Π_n^1 -universal.

For n = 1, there is a $\Sigma_0^1 = \Sigma_1^0$ -universal set by Theorem 24 B•9. For n + 1, let inductively $W \subseteq \mathcal{N}^3$ be Π_n^1 -universal, i.e. for every $B \in \Pi_n^1$ with $B \subseteq \mathcal{N} \times \mathcal{N}$, there is some $r \in \mathcal{N}$ where $B = W_r = \{\langle x, y \rangle \in \mathcal{N}^2 : \langle r, x, y \rangle \in W\}$. Consider then

 $U = \exists^{\mathcal{N}} W = \{ \langle r, x \rangle \in \mathcal{N}^2 : \exists y \in \mathcal{N} \ (\langle r, x, y \rangle \in W) \}.$

It's not difficult to see that $U \in \Sigma_{n+1}^1$ and is Σ_{n+1}^1 -universal. To see the latter, any $X = \exists^{\mathcal{N}} Y$ for $Y \in \prod_n^1$ has an $r \in \mathcal{N}$ where

$$X = \exists^{\mathcal{N}} W_r = \{x \in \mathcal{N} : \exists y \in \mathcal{N} (\langle r, x, y \rangle \in W)\} = \{x \in \mathcal{N} : \langle r, x \rangle \in U\} = U_r.$$

- **24 C**•10. Corollary For $n < \omega$, $\Delta_n^1 \subsetneq \Sigma_n^1 \subsetneq \Delta_{n+1}^1$, and similarly for Π_n^1 .

Proof .:.

Let $U \subseteq \mathcal{N} \times \mathcal{N}$ be Σ_n^1 -universal as per Theorem 24 C • 9. Therefore $U \in \Sigma_n^1 \setminus \Pi_n^1 = \Sigma_n^1 \setminus \Delta_n^1$. To see this, otherwise we'd have $\neg U \in \Sigma_n^1$ as well as $D = \{x \in \mathcal{N} : \langle x, x \rangle \in \neg U\}$ as the computable preimage of $\neg U$ under the map $x \mapsto \langle x, x \rangle$. Therefore $D = U_r$ for some $r \in \mathcal{N}$ where then $r \in U_r$ iff $r \in D$ iff $\langle r, r \rangle \in \neg U$ iff $r \notin U_r$, a contradiction. So there can be no such r, meaning $\neg U \notin \Sigma_n^1$ i.e. $U \notin \Pi_n^1$. This shows $\Delta_n^1 \subsetneq \Sigma_n^1$.

To see that $\Sigma_n^1 \subsetneq \Delta_{n+1}^1$, just note that $\neg U \in \Pi_n^1 \subseteq \Delta_{n+1}^1$ but $\neg U \notin \Sigma_n^1$ for the reasons as above. Hence $\neg U \in \Delta_{n+1}^1 \setminus \Sigma_n^1$.

§ 24 D. The lightface hierarchies and definability

Thus far our discussions have been a mix of topology and computability, neither of which is particularly reminiscent of the "descriptive" part of descriptive set theory. One may recall, however, that the arithmetical hierarchy on ω from computability theory lines up with the definable relations on ω under the levy hierarchy: $\Sigma_n^{0,\omega} = \Sigma_n^N$. In generalizing this, we need either to introduce a second-order extension of the first-order model $\mathbf{N} = \langle \omega, 0, 1, +, \cdot \rangle$ —and thus introduce second-order logic as well—or a more complicated first-order structure. We will go with the former.

Second-order logic just allows additional quantifiers over the powerset of the model's universe. The syntax is practically identical to first-order logic, except that there are now two types of variables: ones ranging over the universe of the model, and ones ranging over the *powerset* of the universe of the model.

— 24 D•1. Definition

Let σ be a signature. The variables of SOL(σ) are $\{v_n : n < \omega\} \cup \{P_n^i : n < \omega\}$. We call the v_n individual variables and the P_n^i predicate variables with arity *i*. Terms are built up in the same way as before just on individual variables. Define recursively φ to be a SOL(σ)-formula iff

• φ is *atomic*, i.e.

- "x = y" for individual variables x, y;
- $-\varphi$ is " $R(\vec{x})$ " for $R \in \sigma$ a relation symbol and \vec{x} terms (that match the arity of R);
- φ is " $P_n^i(\vec{x})$ " where \vec{x} is a sequence of *i* terms;

- φ is " $\neg \psi$ ", $\psi \rightarrow \theta$, $\psi \land \theta$, etc. for ψ , θ SOL(σ)-formulas; or
- φ is " $\forall X \psi$ " or " $\exists X \psi$ " for ψ a SOL(σ)-formula and *X* any variable.

We don't care about the proof theory of second-order logic, since it isn't complete in the way first-order logic is.^{xii} So we can jump right into the semantics of second-order logic. Mostly the only difference from first-order logic is how the new variables are interpreted. Note that this requires knowledge of the powerset of the model's universe.

- 24D.2. Definition Let σ be a signature. A SOL(σ)-model is defined in the same way as with FOL(σ)-models. Let M be a SOL(σ)-model, φ a SOL(σ)-sentence. Let $\vec{m} \in M^{<\omega}$ and $P_0 \in \bigcup_{n < \omega} \mathcal{O}(M^n)$. We define the interpretation of φ in the same way as with first-order logic with the following additions: $M \models "P_0(\vec{m})"$ if and only if $\vec{m} \in P_0$, $M \models "\forall P_n^i \varphi(P_n^i)"$ if and only if $M \models \varphi(P)$ for every $P \subseteq M^i$ $M \models "\forall P_n^i \varphi(P_n^i)"$ if and only if $M \models \varphi(P)$ for some $P \subseteq M^i$. To distinguish the two, for M a FOL(σ)-model, the SOL(σ)-extension is denoted M^1 .

Formally, there's no difference between the FOL-model M and the SOL-model M^1 . Really the notation marks the distinction between the use of " \models " rather than the structures.

Note that with this added language, the definable sets are no longer just subsets of $M^{<\omega}$ but of the more complicated $\mathcal{P}(M^{<\omega})^{<\omega} \times M^{<\omega}$, since predicate variables range over $\mathcal{P}(M^{<\omega})$ while individual variables range over M.

In the case of N, we write \mathcal{N} for N¹. Moreover, instead of $P \in \mathcal{P}(\omega)$, we usually regard them as elements of $\omega \omega$, which can be done as follows.

- 24 D • 3. Example Write \mathcal{N} for $N^1 = \langle \omega, 0, 1, +, \cdot \rangle^1$. \mathcal{N} and ω are SOL-definable over \mathcal{N} .

Proof .:.

 ω is definable by the formula $v_0 = v_0$: $\{v_0 \in \omega : \mathcal{N} \models "v_0 = v_0"\} = \omega$. For \mathcal{N} , we just rely on the fact that being a function is definable: $\mathcal{N} = \{P \subseteq \omega^2 : \mathcal{N} \models "\forall v_0 \exists ! v_1 P(v_0, v_1)"\}$.

There is then an ambiguity between the use of \mathcal{N} as a model and as the baire space. The reader will need to use context clues to determine the meaning, but for the rest of the section we will focus on the model.

In practice, we will often just use " x_i " to range over \mathcal{N} and " m_i " to range over ω in \mathcal{N} . In fact, we will often mark them in that we write " $\forall x \in \mathcal{N}$ " or " $\exists m \in \omega$ ". In other words, we use " x_i " for binary predicate variables that are functions, and " m_i " for individual variables. One can then think of their use as shorthand for the definitions in Example 24 D • 3. So for the most part, rather than $P_n^2(t_0, t_1)$ for P_n^2 a variable ranging over \mathcal{N} , we write $x_n(t_0) = t_1$. In other words, we identify x_n ranging over \mathcal{N} with P_n^2 as the relation $P_n^2(a, b) \leftrightarrow x_n(a) = b$. But this association is merely meta-theoretic.

With this new language, we also can consider another hierarchy! This time an extension of the lévy hierarchy with the addition of our new variable types.

— 24 D•4. Definition -

Let σ be a signature. For **M** a model and $\varphi(\vec{v}, \vec{P})$ a SOL(σ)-formula, write

 $\varphi(\mathsf{M}) = \{ \langle \vec{m}, \vec{P} \rangle \in M^{<\omega} \times \mathcal{P}(M^{<\omega})^{<\omega} : \mathsf{M}^1 \vDash "\varphi(\vec{m}, \vec{P})" \}.$

For the language SOL({0, 1, +, ·}), we call a quantifier *bounded* iff its of the (shorthand) form " $\exists x < y$ " or " $\forall x < y$ " where "x < y" itself is also shorthand for " $\exists z \ (x + z = y)$ ", and x, y, z are individual variables. For $n < \omega$ and a

 x^{ii} This holds just from some knowledge of computability theory and Gödel's theorems. In particular, full induction is SOL-definable with a single formula. As a result, **N** is uniquely described by the second-order analog of PA. If completeness held for second-order logic's proof system, then the consequences of PA would be exactly the SOL-theory of **N**. But no proof system for arithmetic can be intelligible (i.e. computable), complete, and sound. Hence second-order logic's proof system couldn't computable (i.e. one can't tell whether a sequence of formulas is a proof or not). But any reasonable definition of a proof system for second-order logic *will* be computable, a contradiction.

 $SOL(\{0, 1, +, \cdot\})$ -formula φ , define recursively

- φ is Σ₀⁰ = Π₀⁰ iff all quantifiers in φ are bounded;
 φ is Σ_{n+1}⁰ iff φ is of the form "∃m ∈ ω ψ" for m an individual variable and ψ a Π_n⁰-formula;
- φ is Π_n^0 iff φ is of the form " $\neg \psi$ " for ψ a Σ_n^0 -formula;

For $A \subseteq \mathcal{N}$, define

$$SOL_{\Sigma_n^0}(A) = \{\varphi(\mathcal{N}) : \varphi \text{ is } \Sigma_n^0 \text{ with parameters in } A\}$$

and similarly for $SOL_{\Pi_n^0}(A)$, where we then define $SOL_{\Delta_n^0}(A) = SOL_{\Sigma_n^0}(A) \cap SOL_{\Pi_n^0}(A)$.

We also can define Σ_n^1 variants, and we will do so later. For now, let's focus on this hierarchy, which just corresponds to the formulas that quantify only over ω . It should be obvious just by adding unnecessary quantifiers (although we will prove this) that $\operatorname{FOL}_{\Sigma_n^0}(A) \subseteq \operatorname{FOL}_{\Delta_{n+1}^0}(A) \subseteq \operatorname{FOL}_{\Sigma_{n+1}^0}(A)$ for $n < \omega$ and $A \subseteq \mathcal{N}$, and similarly for Π_n^0 . Hence we get the usual argyle picture. Note that, of course, not every formula is equivalent to a formula in the hierarchy





above: $\exists P_0^1 \varphi(x, P_0^1)$, a priori, cannot be placed in this hierarchy. But of the formulas using only quantification over ω , through coding, each is equivalent over Th(\mathcal{N}) to a formula in this hierarchy, and so every relation of \mathcal{N} SOLdefinable with only quantification over ω is somewhere in the hierarchy of $SOL_{\Sigma_n^0}$, $SOL_{\Pi_n^0}$, and $SOL_{\Delta_n^0}$ for $n < \omega$. For example, equality between real numbers is Π_1^0 -definable: $x_0 = x_1$ iff $\forall n \in \omega$ ($x_0(n) = x_1(n)$), itself shorthand for the cumbersome formula

And now the reader should be convinced (if not already) that fully writing out formulas in the language should be avoided. That said, we will still often identify $\mathcal{P}(\omega)$ with \mathcal{N} for the sake of simplicity. Alternatively, one may just consider the subsets of $\mathcal{N}^{<\omega} \times \omega^{<\omega}$ when interpreting SOL-formulas. Whatever is more understandable to the reader is recommended.

We get many expected closure properties of these classes by coding, showing some of the above facts.

– 24 D•6. Lemma — Let $0 < n < \omega$ and $A \subseteq \mathcal{N}$. Write $\exists^{\omega} X = \{ \langle \vec{x}, \vec{m} \rangle \in \mathcal{N}^{<\omega} \times \omega^{<\omega} : \exists n \in \omega \; (\langle \vec{x}, \vec{m}, n \rangle \in X) \}$ for $X \subseteq \mathcal{N}^{<\omega} \times \omega^{<\omega} \times \omega$, and similarly for the other quantifiers we write. Therefore,

- 1. SOL_{Σ_{n}^{0}}(A) is closed under $\exists^{\omega}, \wedge, \vee$, bounded quantification, and A-computable substitution (i.e. A-computable preimäges).
- 2. SOL_{$\Pi_n^0(A)$} is closed under $\forall^{\omega}, \wedge, \vee$, bounded quantification, and A-computable substitution.
- 3. SOL_{$\Delta_n^0(A)$} is closed under \neg , \land , \lor , bounded quantification, and A-computable substitution.

Moreover, $SOL_{\Delta_n^0}(A) \subseteq SOL_{\Sigma_n^0}(A) \subseteq SOL_{\Delta_{n+1}^0}(A)$, and similarly for $SOL_{\Pi_n^0}(A)$.

Proof .:.

We consider only the case where $A = \emptyset$ as this trivially implies the result for $A \neq \emptyset$. Note that Σ_0^0 -formulas are trivially closed under bounded quantification, conjunction, disjunction, and negations. First we show the relevant containments.

That $\operatorname{SOL}_{\Delta_n^0} \subseteq \operatorname{SOL}_{\Sigma_n^0}$ is immediate just by definition. That $\operatorname{SOL}_{\Sigma_n^0} \subseteq \operatorname{SOL}_{\Pi_{n+1}^{0}}$ follows just by adding an unnecessary quantifier after all the others: for $x \in X$ defined by the Σ_n^0 -formula $\varphi(x)$, then " $\forall y \in \omega \varphi(x)$ " is Π_{n+1}^0 and still defines X. To show $\operatorname{SOL}_{\Sigma_n^0} \subseteq \operatorname{SOL}_{\Sigma_{n+1}^0}$, we add these unnecessary quantifiers *before* the first one. It's not difficult to see that any Σ_n^0 -formula is of the form

$$"\exists m_1 \in \omega \ \forall m_1 \in \omega \ \cdots \ Q_n m_n \in \omega \ \varphi(\vec{x}, \vec{m}) "$$

where φ is Σ_0^0 , and Q_n is \forall if *n* is even and \exists if *n* is odd. So adding an unnecessary quantifier to φ yields the Σ_{n+1}^0 -formula

" $\exists m_1 \in \omega \ \forall m_1 \in \omega \ \cdots \ Q_n m_n \in \omega \ Q_{n+1} k \in \omega \ \varphi(\vec{x}, \vec{m})$ "

equivalent to the one above, where k doesn't appear in φ . Thus $SOL_{\Sigma_n^0} \subseteq SOL_{\Delta_{n+1}^0}$, and similarly for Π_n^0 .

1. We can deal with \exists^{ω} with coding pairs: proceed by induction on $n < \omega$. Let $X \subseteq \mathcal{N}^a \times \omega^b$ for $a, b < \omega$ with $\langle \vec{x}, \vec{m} \rangle \in X$ defined by the Σ_n^0 -formula " $\exists k \in \omega \varphi(\vec{x}, k, \vec{m})$ " where φ is Π_{n-1}^0 . Note that then $\langle \vec{x}, m_1, \cdots, m_b \rangle \in \exists^{\omega} X$ is then defined by the formula

$$\langle \vec{x}, m_1, \cdots, m_b \rangle \in \exists^{\omega} X \quad \text{iff} \quad \underbrace{\mathcal{N}}_{\sim} \models ``\exists m_0 \in \omega \exists k \in \omega \ \varphi(\vec{x}, k, m_0, \cdots, m_b)'' \\ \text{iff} \quad \underbrace{\mathcal{N}}_{\sim} \models ``\exists y \in \omega \ \exists m_0 < y \ \exists k < y \ \underline{\varphi(\vec{x}, k, m_0, \cdots, m_b)}'' \\ \underbrace{\Pi_{n-1}^0}_{\Pi_{n-1}^0} \\ \underline{\Pi_{n-1}^0}_{\mathbb{P}^0}$$

Closure under intersections and unions can also be done easily by the same idea: for Π_{n-1}^0 -formulas $\varphi(\vec{x}, m_0, \dots, m_b)$ and $\psi(\vec{x}, m_0, \dots, m_b)$,

$$\begin{split} & \underbrace{\mathcal{N}} \vDash ``\exists m_0 \in \omega \ \varphi(\vec{x}, m_0, \cdots, m_b) \land \exists m_0 \in \omega \ \psi(\vec{x}, m_0, \cdots, m_b) " \\ & \text{iff} \quad \underbrace{\mathcal{N}}_{\mathcal{N}} \vDash ``\exists y \in \omega \ \exists m_0 < y \ \exists m'_0 < y \ (\varphi(\vec{x}, m_0, \cdots, m_b) \land \psi(\vec{x}, m'_0, m_1, \cdots, m_b)) ", \text{ and} \\ & \underbrace{\mathcal{N}}_{\mathcal{N}} \vDash ``\exists m_0 \in \omega \ \varphi(\vec{x}, \vec{m}) \lor \exists m_0 \in \omega \ \psi(\vec{x}, \vec{m}) " \\ & \text{iff} \quad \underbrace{\mathcal{N}}_{\mathcal{N}} \vDash ``\exists m_0 \in \omega \ (\varphi(\vec{x}, \vec{m}) \lor \psi(\vec{x}, \vec{m})) ". \end{split}$$

For bounded quantification, let $X \subseteq \mathcal{N}^a \times \omega^b$ for $a, b < \omega$ be defined by the Σ_n^0 -formula " $\exists k \in \omega \varphi(\vec{x}, k, \vec{m})$ ", where φ is Π_{n-1}^0 . For existential bounded quantification, we can freely switch the quantifiers; and for universal quantification, we can use a Σ_0^0 -definable coding of pairs:

$$\begin{split} & \underset{\sim}{\mathcal{N}} \vDash ``\exists m'_0 < m_0 \; \exists k \in \omega \; \varphi(\vec{x}, k, m'_0, m_1, \cdots, m_b)" \\ & \text{iff} \quad & \underset{\sim}{\mathcal{N}} \vDash ``\exists k \in \omega \; \exists m'_0 < m_0 \; \varphi(\vec{x}, k, m'_0, m_1, \cdots, m_b)" \end{split}$$

$$\mathscr{N} \vDash "\forall m'_0 < m_0 \; \exists k \in \omega \; \varphi(\vec{x}, k, m'_0, m_1, \cdots, m_b)"$$

iff $\mathcal{M} \models ``\exists k' \in \omega (\underbrace{k' \text{ codes sequences of pairs } \langle i, k_i \rangle \text{ for } i < m_0}_{\Sigma_0^0 = \Pi_0^0} \land \underbrace{\forall i < m_0 \varphi(\vec{x}, k_i, i, m_1, \cdots, m_b)}_{\text{inductively } \Pi_{n-1}^0})''$

$$\overline{\Sigma_0}$$

Computable substitutions holds by some facts from computability theory: every such function is Δ_1 -definable over **N** so that translating the Σ_1 and Π_1 -formulas over to \mathcal{N} by relativizing quantifiers to ω yields Σ_1^0 and Π_1^0 -definitions over \mathcal{N} :

$$\underbrace{\mathscr{N}}_{\Sigma_{n}^{0}} \vDash \operatorname{\exists} k \in \omega \; \exists k_{0} < k \; \exists k_{1} < k \; (\underbrace{f(k_{0}) = k_{1}}_{\Delta_{1}^{0}} \land \underbrace{\varphi(\vec{x}, k_{1}, \vec{m})}_{\Pi_{n-1}^{0}})^{"}}_{\Sigma_{n}^{0}}$$

2. These facts follow easily from the facts on $SOL_{\Sigma_n^0}$ just by pushing the '¬' symbol through the quantifiers in the definitions.

3. These facts follow easily from the facts on $SOL_{\Sigma_{\mu}^{0}}$ and $SOL_{\Pi_{\mu}^{0}}$ as usual.

 \dashv

The strict inequalities, as usual, are quite difficult to show. But we are more-or-less unconcerned about them. The main purpose for this added hierarchy of formulas is that we will show that the arithmetical hierarchy on the baire space is (with some slight modifications) this hierarchy of definable relations over \mathcal{N} ; connecting topology, computability, and description.

We are now prepared to show that the arithmetical hierarchy on the baire space lines up with elementary relations of \mathcal{N} that use quantifiers only over ω . In particular, we have $\bigcup_{n < \omega} \Sigma_n^0 = \mathcal{P}(\mathcal{N}) \cap \bigcup_{n < \omega} \text{SOL}_{\Sigma_n^0}$, and in fact SOL_{Σ_n^0} $\cap \mathcal{P}(\mathcal{N}) = \Sigma_n^0$, and similarly for the other pointclasses.

- **24D**•7. Lemma Let $A \subseteq \mathcal{N}$ and $a, b < \omega$. Write $\mathcal{M} = \mathcal{N}^a \times \omega^b$. Therefore $\operatorname{SOL}_{\Sigma_0^0} \cap \mathcal{P}(\mathcal{M}) \subseteq \Delta_1^{0,\mathcal{M}}(A)$.

Proof .:.

As we're considering only elements of \mathcal{N} , we modify our language slightly by considering $x \in \mathcal{N}$ as functions rather than as relations.

Without loss of generality, assume $A = \emptyset$, as the general case follows immediately from this. The x_i then range over \mathcal{N} and the m_i range over ω . Let $X \in \text{SOL}_{\Sigma_0^0}$ as witnessed by the formula $\varphi(\vec{x}, \vec{m})$. Without loss of generality, $X \neq \emptyset$. Proceed by induction on φ to show that $X \in \Sigma_1^{0,\mathcal{M}}$. The Σ_0^0 -formulas are the least collection of formulas containing the atomic formulas, and closed under conjunctions, negations, and bounded quantification. Since $\Delta_1^{0,\mathcal{M}}$ -sets are closed under intersections, complements, and bounded quantification, it suffices to show that the atomic formulas define $\Delta_1^{0,\mathcal{M}}$ -sets.

To show this, what matters are the terms that can be built up in this language SOL($\{0, 1, +, \cdot\}$). But the only atomic formulas are equality, and evaluation of reals: for equality, $t_0(\vec{m}) = t_1(\vec{m})$ is a statement involving only natural numbers and computable functions. So $X' = \{\vec{m} \in \omega^b : t_0(\vec{m}) = t_1(\vec{m})\} \in \Delta_1^{0,\omega^b}$ and it's easy to see that then $X = \mathcal{N}^a \times X' \in \Delta_1^{0,\mathcal{M}}$.

The other atomic formulas are " $P_n^i(\vec{t}(\vec{m}))$ " for $n, i < \omega$. Consider for simplicity (and as we're focused on $\mathcal{N} \subseteq \mathcal{O}(\omega^2)$) binary relations: " $P_n(t_0(\vec{m}), t_1(\vec{m}))$ " taken as evaluation $x_n(t_0(\vec{m})) = t_1(\vec{m})$. This us just one evaluation the real x_n : as t_0 and t_1 are computable, the defined set is just

$$X = \bigcup_{\langle \tau, \vec{m} \rangle \in A} \mathcal{N}_{\tau} \times \mathcal{N}^{a-1} \times \{ \vec{m} \} \in \Sigma_{1}^{0, \mathcal{M}}$$

where A is the set of all $\langle \tau, \vec{m} \rangle \in ({}^{<\omega}\omega)^2$ where τ is of appropriate length such that $\tau(t_0(\vec{m}) = t_1(\vec{m}))$, which is computable. It's not difficult to see that $\neg P_n(t_0(\vec{m}), t_1(\vec{m}))$ also defines a Δ_1^0 -set, since we just consider

$$\neg X = \bigcup_{\langle \tau, \vec{m} \rangle \in B} \mathcal{N}_{\tau} \times \mathcal{N}^{a-1} \times \{ \vec{m} \} \in \Sigma_1^{0, \mathcal{M}}$$

where *B* is the set of all $\langle \tau, \vec{m} \rangle \in ({}^{<\omega}\omega)^2$ where τ is of appropriate length such that $\tau(t_0(\vec{m})) \neq t_1(\vec{m})$, which is again computable.

Corollary 24 C • 3 tells us that Δ_1^0 is the starting point for the arithmetical hierarchy, so it's good that all Σ_0^0 -definable relations are Δ_1^0 , as then $\Sigma_1^0 = \exists^{\omega} \Delta_1^0 \supseteq \exists^{\omega} SOL_{\Sigma_0^0} = SOL_{\Sigma_1^0}$, and so on. So the result holds for the rest of the hierarchy easily.

- 24D·8. Theorem Let $a, b < \omega$ and set $\mathcal{M} = \mathcal{N}^a \times \omega^b$. Therefore, for each $0 < n < \omega$, $\Sigma_n^{0,\mathcal{M}}(A) = \text{SOL}_{\Sigma_n^0}(A) \cap \mathcal{P}(\mathcal{M})$, and similarly for $\prod_{n=1}^{n} M$ and $\Delta_n^{0,\mathcal{M}}$.

Proof .:.

Proceed by induction on n > 0, starting with $A = \emptyset$, as the result for $A \neq \emptyset$ follows immediately from this case. Note that the basic closure properties for the $SOL_{\Sigma_n^0}s$ as in Lemma 24 D • 6 also hold for $SOL_{\Sigma_n^0}$ by the same reasoning: when eliminating " $x_i = x_j$ ", the Σ_0^0 -definable relations still trivially have the relevant closure properties. We begin with the base case of n = 1. Write \mathcal{M} for $\mathcal{N}^a \times \omega^b$ for the sake of notation.

(⊆) Suppose $X = \emptyset$. Clearly this is SOL_{Σ_1^0} by the formula $m_0 \neq m_0$. So suppose $X = \bigcup_{n < \omega} \mathcal{N}_{f(n)}$ for some computable $f : \omega \to {}^{<\omega}\omega$. Since f is computable, the relation f(x) = y is Σ_1 -definable over $\mathbf{N} = \langle \omega, 0, 1, +, \cdot \rangle$. In particular, using the same definition (modified to mark all variables as in ω), we get that f is SOL_{Σ_1^0}. So by Lemma 24 D • 6, the following witnesses that X is SOL_{Σ_1^0}:

$$x \in X$$
 iff $\mathcal{N} \models \exists n \in \omega \forall k < \ln(f(n)) (x(k) = f(n)(k))$

A similar idea applies to \mathcal{M} .

 (\supseteq) By Lemma 24 D•7, SOL_{Σ_0^0} $\cap \mathcal{P}(\mathcal{M} \times \omega) \subseteq \Delta_1^{0,\mathcal{M}}$, and therefore

$$\mathrm{SOL}_{\Sigma_1^0} \cap \mathfrak{P}(\mathcal{M}) = \exists^{\omega} \mathrm{SOL}_{\Sigma_0^0} \cap \mathfrak{P}(\mathcal{M}) \subseteq \exists^{\omega} \Delta_1^{0, \mathcal{M} \times \omega} \subseteq \Sigma_1^{0, \mathcal{M}}.$$

This establishes the base case of n = 1, and the result for $\prod_{n=1}^{0,M}$ is, as always, immediate as well the inductive case n + 1:

This has a number of nice consequences, the easiest of which is the following.

- **24D**•9. Corollary
For
$$A \subseteq \mathcal{N}, \bigcup_{n < \omega} \Sigma_n^0(A) = \mathcal{P}(\mathcal{N}) \cap \bigcup_{n < \omega} \operatorname{SOL}_{\Sigma_n^0}(A).$$

This also allows a slightly more easy to confirm characterization of the arithmetical and analytical pointclasses. As stated above, we may start from the computable relations over ω and then generalize this to \mathcal{N} by taking initial segments as before in Corollary 24 A • 4: Π_1^0 relations take the form $x \in X \leftrightarrow \forall n < \omega$ ($x \upharpoonright n \in R$). Similarly Σ_3^0 -relations take the form $x \in X \leftrightarrow \exists m_3 \forall m_2 \exists m_1 (\langle x \upharpoonright m_1, m_2, m_3 \rangle \in R)$ for some computable R, and so on throughout the arithmetical hierarchy.

It also allows us to more easily see some of the closure properties of Σ_n^0 and $\text{SOL}_{\Sigma_n^0}$. And thus $\bigcup_{n < \omega} \sum_n^0$ are just the FOLp-definable subsets of \mathcal{N} (with quantifiers only over ω). Moreover, we can generalize this to the analytical and projective hierarchies by allowing quantifiers over \mathcal{N} .

24D • 10. Definition
For n < ω and a SOL({0, 1, +, ·})-formula φ, define recursively
φ is Σ₀¹ iff φ is Σ₀¹;
φ is Σ_{n+1}¹ iff φ is of the form "∃x ∈ N ψ" for x a predicate variable and ψ a Π_n¹-formula;
φ is Π_n¹ iff φ is of the form "¬ψ" for ψ a Σ_n¹-formula.
For A ⊆ N, define SOL_{Σ_n¹}(A), SOL_{Π_n¹}(A), and SOL_{Δ_n¹}(A) as with SOL_{Σ_n⁰}(A) and so on.

Again, not every formula is equivalent to a formula in this extended hierarchy. For example, if φ involves quantification over variables that are not binary, then it's not equivalent to any formula in the hierarchy. But this is the only obstruction: for any definable subset of \mathcal{N} , there is a Σ_n^1 -definition for some n. The proof of this actually follows just from equating SOL $_{\Sigma_n^1}$ with Σ_n^1 .

- 24D·11. Theorem For $A \subseteq \mathcal{N}$, $0 < n < \omega$, and $a, b < \omega$, write $\mathcal{M} = \mathcal{N}^a \times \omega^b$. Therefore $SOL_{\Sigma_n^1}(A) \cap \mathcal{P}(\mathcal{M}) = \Sigma_n^{1,\mathcal{M}}(A)$, and similarly for Π_n^1 and Δ_n^1 .

Proof .:.

Proceed by induction on *n*. For n = 1, Theorem 24 D • 8 implies

 $\Sigma_n^1(A) = \exists^{\mathscr{N}} \Pi_1^0(A) = \exists^{\mathscr{N}} \operatorname{SOL}_{\Pi_1^0}(A) = \operatorname{SOL}_{\Sigma_n^1}(A).$

For n + 1, the result holding on $\prod_{n=1}^{1}(A)$ implies the result for $\sum_{n=1}^{1}(A)$ as, by the same idea as above,

$$\Sigma_{n+1}^1(A) = \exists^{\mathcal{N}} \Pi_n^1(A) = \exists^{\mathcal{N}} \operatorname{SOL}_{\Pi_n^1}(A) = \operatorname{SOL}_{\Sigma_{n+1}^1}(A).$$

The result for the other pointclasses follows easily by taking negations for Π_n^{n+1} and intersections for Δ_n^1 . \dashv

By the closure properties of the analytical pointclasses under \exists^{ω} and \forall^{ω} , we then get that the analytical hierarchy encompasses more than just the relations definable using a bunch of quantifiers over \mathcal{N} and then one quantifier over ω : it encompasses *all* definable relations of \mathcal{N} over \mathcal{N} .

— 24 D•12. Corollary -

For $A \subseteq \mathcal{N}$, $X \subseteq \mathcal{N}$ is SOL-definable over \mathcal{N} iff $X \in \bigcup_{n < \omega} \Sigma_n^1$. And so $X \subseteq \mathcal{N}$ is SOL-definable over \mathcal{N} iff $X \in \bigcup_{n < \omega} \Sigma_n^1(\mathcal{N})$.

Proof .:.

The \leftarrow direction was proven in Theorem 24 D • 11. So let $X = \varphi(\mathcal{N})$ for φ a SOL({0, 1, +, ·})-formula. Without loss of generality, we can put φ is prenex normal form so that φ has a bunch of quantifiers out in front of a Σ_0^0 formula ψ . Clearly $\psi(\mathcal{N}) \in \Sigma_1^1$, and so it suffices to show inductively that $\bigcup_{n < \omega} \Sigma_n^1$ is closed under $\exists^{\omega}, \exists^{\mathcal{N}},$ \forall^{ω} , and $\forall^{\mathcal{N}}$. And this was proven in Result 24 C • 4. The boldface result holds since $\Sigma_n^1(\mathcal{N}) = \Sigma_n^1$. \dashv

Hence the arithmetical and analytical hierarchies combined form a natural way of categorizing the definable relations of \mathcal{N} . Unfortunately, SOL-truth is not absolute because it relies on information about the powerset, which isn't absolute. But for (transitive) models (of enough set theory) that agree on the powerset of a model's universe, SOL-truth *is* absolute and can be witnessed in the same way as with FOL-truth: bounding all quantifiers over the the powerset of the model's universe to arrive at a Σ_0^0 -formula with parameters in both models. Hence to know what's true about \mathcal{N} , we merely need to know what reals exist, and vice versa. This is the basis for many absoluteness results: merely showing that L has the reals to demonstrate the result.

More than just truth about N, we can instead consider truth about V, and really merely $V_{\omega} = H_{\aleph_0} = HF$, the hereditarily finite sets: sets that are finite, and all of their elements are finite, and so on. The reader is encouraged to read Subsection 7 C for some of the basic facts about these sets. Nevertheless, in this way, the study of \mathcal{N} will inevitably require facts about the larger set theoretic universe.

- 24D•13. Definition Write $HF = \langle HF, \in \rangle$ for H_{\aleph_0} , the *hereditarily finite* sets. We also occasionally write $HC = \langle HC, \in \rangle$ for H_{\aleph_1} , the *hereditarily countable* sets.

This is ultimately just a result of the fact that HF and membership in HF are definable (in a coded way) over **N**. In particular, since $|\text{HF}| = \aleph_0$, we can consider the coded relation $\{\text{code}(n, m) : n = \text{code}(x) \land m = \text{code}(y) \land x \in y \in \text{HF}\}$. So it suffices to show that this is definable over **N** as then we may decode this relation.

There is a coding of HF into ω such that code "HF = ω and the coded relation $\hat{\epsilon} = \{\langle n, m \rangle \in \omega^2 : \exists x, y \in \text{HF} (n = \text{code}(x) \land m = \text{code}(y) \land x \in y)\} \subseteq \omega$ is Σ_0 -definable over $N = \langle \omega, 0, 1, +, \cdot \rangle$. Proof .:.

There are several ways to do this. The easiest way conceptually is just to note that any element of HF can be written down in our set builder notation: $\emptyset = \{\}$ and $\{\emptyset, 2\} = \{\{\}, \{\{\}, \{\{\}\}\}\}\}$, for example. Then we identify HF with coded finite strings of this sort and identify elements easily by how many unpaired parantheses exist to the left and right. This conceptual approach isn't exactly the easiest to formalize however. Another approach is through well-ordering $V_{\omega} = \bigcup_{n < \omega} V_n$. In particular, we view $n \in \omega$ as a string of 0s and 1 (i.e. in binary), and then the position of the 1s lists out the (coded) elements of lower rank:

$$n \in m$$
 iff $1 = \left\lfloor \frac{m}{2^n} \right\rfloor \mod 2$

iff
$$\exists q_0 < m \ \exists r < m \ (m = q_0 \cdot 2^n + r \land r < 2 \land \exists q_1 < m \ (q_0 = 2 \cdot q_1 + 1)).$$

The above shows that $n \in m$ is Σ_0 -definable over **N**. From this, we recursively define

 $\operatorname{decode}(m) = \{\operatorname{decode}(n) : n \in m\}.$

This makes sense since $n \in m$ implies n < m. Now we have the crucial claim that this decoding is truly a means of coding HF in N.

– Claim 1 –

decode : $\omega \to HF$ is an isomorphism between **HF** and $\langle \omega, \hat{\epsilon} \rangle$.

Proof .:.

It's obvious that decode respects membership:

 $n \in m \leftrightarrow \operatorname{decode}(n) \in \{\operatorname{decode}(n) : n \in m\} = \operatorname{decode}(m).$

So it suffices to show injectivity and surjectivity. For injectivity, we need to show $m \mapsto \operatorname{pred}_{\hat{e}}(m)$ is injective. Note that $n \in m$ iff the *n*th digit of *m* (written in binary) is 1. Hence the \hat{e} -predecessors of *m* tell us where the 1s are in the binary expansion with the other digits being 0s, which then uniquely determine *m*. This shows decode is injective by a simple induction on *m*: the least *m* with $\exists m' (\operatorname{decode}(m) = \operatorname{decode}(m'))$ yields that

decode "pred_{$\hat{\in}$}(m) = decode(m) = decode(m') = decode "pred_{$\hat{\in}$}(m'),

which means some $n \in m$ (and thus n < m) has decode(n) = decode(n') for some $n' \in m'$, contradicting minimality of m.

To show decode is surjective, proceed by induction on rank. For $x = \emptyset$, one may see that decode(0) = x. For $x \in V_{n+1}$, inductively $V_n \subseteq im(decode)$ and therefore we may consider $x' = \{n \in \omega : decode(n) \in x\}$ and then $m = \sum_{n \in x'} 2^n$. It's not difficult to check decode(m) = x.

As a result, $code = decode^{-1}$ works as in the statement of the result.

 \dashv

We also get the reverse: that membership in ω , the operations +, \cdot , and constants 0, 1 are definable in a simple way over HF.

- 24D • 15. Lemma The following relations are Σ_1 and Π_1 -definable over HF: 1. $x \in \omega$; 2. $x, y \in \omega \land x + y = z$, and $x, y \in \omega \land x \cdot y = z$; 3. x = 0, and x = 1;

Proof .:.

- 1. As a transitive set, $\operatorname{Ord}^{\mathsf{HF}} = \operatorname{Ord} \cap \operatorname{HF}$. As the set of all hereditarily finite sets, $\omega \subseteq \operatorname{HF}$, and clearly $\omega \notin \operatorname{HF}$ as it's infinite. Therefore $\operatorname{Ord} \cap \operatorname{HF} = \omega$ so $x \in \omega$ iff $\mathsf{HF} \models "x \in \operatorname{Ord}"$. Since being an ordinal is Σ_0 -definable, this gives the result.
- 2. x + y = z (or $x \cdot y = z$) iff there is *some* or *any* function f obeying the usual definition of + (or \cdot) with x, y in the domain with f(x, y) = z. Obeying the usual definition with $x, y \in \text{dom}(f)$ is Σ_0 -definable, which means x + y = z is both Σ_1 (when talking about the existence of such a function) and Π_1 (when talking about all such functions with $x, y \in \text{dom}(f)$) definable.

3. x = 0 iff $\forall y \in x (y \neq y)$, which is Σ_0 -definable, and x = 1 iff $\forall y \in x (y = \emptyset) \land \exists y \in x (y = y)$, which is Σ_0 -definable.

As a result, we can give an alternative characterization of the arithmetical hierarchy on ω through first-order logic, and on \mathcal{N} through second-order logic. This gives another connection between the lévy hierarchy and the topology of \mathcal{N} .

- 24D • 16. Result Let $0 < n < \omega$ and $X, A \subseteq \omega$. Therefore $X \in \Sigma_n^{0,\omega}(A)$ iff there is a Σ_n FOL($\{\in, A\}$)-formula φ such that $x \in X$ iff $\mathsf{HF} \models \varphi(x)$.

Proof .:.

Work with $A = \emptyset$ for the sake of notation. Really we prove the result on all product spaces ω^m for $m < \omega$.

(\leftarrow) We have this result for **N** in place of **HF**, so it suffices to show X is Σ_n -definable over **N**. Since (the coded version of) \in^{HF} is Σ_0 -definable, we just replace \in with its computable definition. Therefore the resulting formula φ^{HF} is still Σ_n . More explicitly, for a FOL(\in)-formula φ , we define the FOL($\{0, 1, +, \cdot\}$)-formula φ^{HF} such that

$$\mathsf{HF} \vDash \varphi(\vec{x}) \quad \text{iff} \quad \mathsf{N} \vDash \varphi^{\mathsf{HF}}(\mathsf{code}(\vec{x})) \quad \text{iff} \quad \mathsf{N} \vDash ``\exists \vec{z} \ (\underbrace{\vec{z} = \mathsf{code}(\vec{x})}_{\Sigma_1} \land \underbrace{\varphi^{\mathsf{HF}}(\vec{z})}_{\Sigma_n})''.$$

Assuming φ^{HF} is Σ_n whenever φ is, it follows that " $\exists \vec{z} \ (\vec{z} = \text{code}(\vec{x}) \land \varphi^{\text{HF}}(\vec{z}))$ " is a Σ_n -definition for X over **N**, as the above indicates. It's not difficult to show that the map $n \mapsto \text{code}(n)$ —i.e. regarding $n \in \omega$ as an element of HF and finding an $m \in \omega$ such that decode(m) = n—is computable and therefore Σ_1 -definable. In particular, one may define code(0) = 0 and $\text{code}(n + 1) = 2^{\text{code}(n)} + \text{code}(n)$. By closure of Σ_1 -definable sets under conjunctions and existential quantification, we get that it's Σ_n . We now define φ^{HF} recursively:

- "x = y"^{HF} is "x = y".
- "x ∈ y"^{HF} is "x < y ∧ x ∈ y". (This ensures bounded quantifers "∃x ∈ y"(···) are transformed into bounded quantifiers "∃x < y (x ∈ y ∧ ···)". And because x ∈ y implies x < y, this causes no problems in terms of the logical equivalence.)
- Unsurprisingly " $\psi \wedge \theta$ "^{HF} is " ψ ^{HF} $\wedge \theta$ ^{HF}", and " $\neg \psi$ "^{HF} is " $\neg \psi$ ^{HF}", and " $\exists x \psi$ "^{HF} is " $\exists z \psi$ ^{HF}".

It's easy to then see that if φ is Σ_n then φ^{HF} is also Σ_n and the above reasoning yields a Σ_n -definition for X from this. Computability theory proves Σ_n -definable relations are $\Sigma_n^{0,\omega}$.

- (\rightarrow) If X is $\Sigma_n^{0,\omega}$ then by some knowledge of computability theory, there is a Σ_n -formula defining X over $\mathbf{N} = \langle \omega, 0, 1, +, \cdot \rangle$. So we replace $+, \cdot, 0$, and 1 with their defining Σ_1 or Π_1 -formulas over HF and then relativize all quantifiers to the class $\omega \subseteq$ HF. More explicitly, proceed by induction on *n*. If φ is Σ_0 already, then (using the complexity results of Lemma 24 D 15) we
 - 1. replace every bounded quantifier " $\exists x < y \ (\dots)$ " with " $\exists x \in y \ (\dots)$ ";
 - 2. replace every occurrence of 0 and 1 with their Σ_0 -definitions as in Lemma 24 D 15;
 - 3. for every free variable x, append " $\land x \in \omega$ " to the end of the formula; and
 - replace every occurrence of + and ⋅ with their Σ₁ or Π₁-definition. (We do this just by putting φ is prenex normal form and writing the quantifier free portion in disjunctive normal form, then using the Π₁-definition if there is a ¬ in front of the original relation and otherwise using the Σ₁-definition.)

The resulting formula φ^{ω} is then Σ_1 and $\mathsf{HF} \models \varphi(x)$ iff $\mathsf{N} \models \varphi^{\omega}(x)$. This then tells us X is $\Sigma_1^{0,\omega}$, and establishes the base case.

So suppose φ is Σ_{n+1} . Therefore φ is of the form " $\exists y \neg \psi$ " for a Σ_n -formula ψ . So inductively we consider φ^{ω} as " $\exists y \ (y \in \omega \land \neg \psi^{\omega})$ ", and we get that this Σ_{n+1} -formula defines X over HF. \dashv

- 24D • 17. Theorem Let $0 < n < \omega$ and $X, A \subseteq \mathcal{N}$. Therefore $X \in \Sigma_n^0(A)$ iff there is a Σ_n^0 SOL($\{\in\} \cup A$)-formula φ such that $x \in X$ iff $\mathsf{HF}^1 \models \varphi(x)$.

Proof Sketch .:.

Here we use the same idea as with Result $24 D \cdot 16$, but now using Theorem $24 D \cdot 8$ to identify Σ_n^0 -definable subsets of \mathcal{N} with Σ_n^0 instead of identifying Σ_n -definable subsets of **N** with $\Sigma_n^{0,\omega}$. To give a sketch, we relativize quantifiers to either ω or the code of HF in a computable way, and exchange \in , +, \cdot , 0, and 1 in a computable way. The result changes φ without increasing complexity and defines X over the other space. \dashv

And we get a similar result when generalizing to the analytical hierarchy.

— 24 D•18. Theorem

Let $n < \omega$ and $X, A \subseteq \mathcal{N}$. Therefore $X \in \Sigma_n^1(A)$ iff there is a Σ_n^1 SOL($\{\in\} \cup A$)-formula φ such that $x \in X$ iff $\mathsf{HF}^1 \models \varphi(x)$.

Proof Sketch .:.

We again use the same idea as with Result 24 D • 16 and Theorem 24 D • 18, but now using Theorem 24 D • 11 to identify Σ_n^1 -definable subsets of \mathcal{N} with Σ_n^1 instead of identifying Σ_n -definable subsets of N with $\Sigma_n^{0,\omega}$. To give a sketch, we relativize quantifiers to either ω or the code of HF in a computable way, and exchange \in , +, \cdot , 0, and 1 in a computable way. The result changes φ without increasing complexity and defines X over the other space. \dashv

The main takeaway from this is that we don't lose or add anything by considering concepts more easily definable with set theory. We also get a related result for $HC = H_{\aleph_1}$, the hereditarily countable sets, not needing the extra mechanisms of second-order logic. We are not yet prepared to prove this theorem yet, however, as we need some results on absoluteness properties of these analytical sets and relations. We will return to this idea in a later section.

- 24D·19. Theorem Let $n < \omega$ and $X, A \subseteq \mathcal{N}$. Therefore $X \in \Sigma_{n+1}^{1}(A)$ iff there is a Σ_{n} FOL($\{\in\} \cup A$)-formula φ such that $x \in X$ iff $\mathsf{HC} \models \varphi(x)$.

Given that any countable structure has an isomorphic copy in HC, what we can express about arbitrary countable structures is reflected in the complexity of the analytical hierarchy. Moreover, since large cardinal properties often entail the existence of certain countable structures, we get interesting absoluteness properties about things as concrete as \mathbb{R} conditional on the abstract and unknowable existence of certain large cardinals.

Section 25. Properties of the Lightface Hierarchies

As the lightface pointclasses are contained in their boldface counterparts, many properties of the arithmetical and analytical pointclasses are inherited from the borel and projective pointclasses. In particular, all Σ_1^1 -sets have the perfect set property, the baire property, and are lebesgue measurable, as established in Section 23. Our goal here will be to establish Σ_1^1 as the limit of these as a result of some properties of **L**, and mostly the result of its Δ_2^1 -definable well-ordering of \mathcal{N}^L . Moreover, we will look at properties that are useful for dealing with the limitations of computability. This involves new concepts and new techniques that become the subject of study in their own right.

§ 25 A. Complexity and absoluteness

When studying absoluteness and complexity, we first looked at the levy hierarchy for FOL(\in)-formulas, concluding that concepts defined by Σ_0 -formulas were absolute and getting partial absoluteness results for Σ_1 and Π_1 -formulas. We cannot generally go further than this, however, as "V = L" is Π_2 -definable (by " $\forall x \exists \alpha \ (\alpha \in \text{Ord} \land x \in L_{\alpha})$ " where " $x \in L_{\alpha}$ " is Σ_0 -definable) but certainly isn't absolute since $L \models$ "V = L" but it's consistent that $V \neq L$, i.e. $V \models$ "V $\neq L$ ". That said, the results at the end of Section 24—and in particular at the end of Subsection 24 D—suggest a natural albeit more limited direction to head. Rather than deal with how things are defined over the entire universe, we may ask questions just of HF or HC and see what absoluteness results we can get there. The result combines the potential of coding countable structures in \mathcal{N} with the absoluteness of computability to easily get a large number of results simply by examining how simple the formulas are.

We already have the absoluteness (between transitive models of ZF - P) of all arithmetical formulas because of Result 24 D • 16 noting that the satisfaction relation for first-order logic is absolute and $HF \subseteq L$.

— 25 A • 1. Result

All arithmetical relations are absolute between transitive models of ZF - P.

Proof .:.

Note that HF is in all transitive models of ZF - P. In particular, FOL-satisfaction for HF is absolute. Since all arithmetical relations are FOL-definable over HF, they are absolute.

We do not, however, have the absoluteness of the satisfaction relation for second-order logic because given any model **A**, to determine satisfaction we require knowledge about $\mathcal{P}(A)$, which is merely Π_1 -definable and hence not generally absolute. Hence we cannot get the absoluteness of Σ_1^1 -sentences so easily. Instead, we use the results from the boldface pointclasses.

Note that since a lightface pointclass is contained in the boldface variant, the analytical hierarchy inherits the properties of the boldface pointclasses from Section 23. In particular, all Σ_1^1 -sets (and so all arithmetical sets too) have the perfect set property, have the baire property, and are lebesgue measurable. Moreover, all Σ_1^1 -sets are \aleph_0 -suslin while Π_1^1 -sets are \aleph_1 -suslin. We can say a little more than this, however. The idea is that Π_1^0 -sets are the branches of arbitrary trees while Π_1^0 -sets are the branches of *computable* trees. Hence Σ_1^1 consists *precisely* of \aleph_0 -suslin sets while Σ_1^1 consists of those given by computable trees, and a similar result holds for Π_1^1 .

First we modify the characterization of Σ_1^1 -sets as $\mathfrak{p}[T]$ for a computable tree T over $\omega \times \omega$. Instead, we do away with T and instead work directly with the subtrees we defined in all the relevant proofs (mostly in Subsection 23 A). Mostly this means we decompose T into its sections: for $x \in \mathcal{N}$ defining

 $T_x = \{ \sigma \in {}^n \omega : n < \omega \land \langle x \upharpoonright n, \sigma \rangle \in T \}.$

We can then reform T just by considering $\{\langle \tau, \sigma \rangle \in {}^{n}\omega \times {}^{n}\omega : n < \omega \land \exists x \in \mathcal{N} \ (\tau \lhd x \land \sigma \in T_{x})\}$.

— 25 A · 2. Lemma (Σ_1^1 Normal Form) –

A set $X \subseteq \mathcal{N}$ is Σ_1^1 iff there is a computable map $x \mapsto T_x$ where T_x is a computable tree over ω and $X = \{x \in \mathcal{N} : [T_x] \neq \emptyset\}.$

Moreover, if φ is a Σ_1^1 -formula, the definition of $x \mapsto T_x$ (for $X = \{x \in \mathcal{N} : \varphi(x)\}$) definable from φ .

Proof .:.

Let X be defined by the Σ_1^1 -formula φ , meaning φ has the form

 $x \in X$ iff $\varphi(x)$ iff $\exists y \in \mathcal{N} \forall n \in \omega R(x \upharpoonright n, y \upharpoonright n)$,

where *R* is some computable relation. So for each $x \in \mathcal{N}$, define $T_x = \{\sigma \in {}^{<\omega}\omega : R(x \upharpoonright h(\sigma), \sigma)\}$.

(→) It's clear that $x \in X$ iff $[T_x] \neq \emptyset$: $x \in X$ iff there's a $y \in \mathcal{N}$ with $\forall n < \omega \ R(x \upharpoonright n, y \upharpoonright n)$ iff $\langle x, y \rangle \in [T]$ iff $y \in [T_x]$.

To see that the map f defined by $f(x) = T_x$ is computable, one can easily use Theorem 24 A • 8: the (characteristic function of the) relation $R(x \upharpoonright h(\sigma), \sigma)$ uses x as an oracle in the same algorithm across all x that computes (the characteristic function of) T_x .

(\leftarrow) If $x \mapsto T_x$ is computable, then the characteristic function of T_x is uniformly x-computable, meaning there is a computable $R \subseteq {}^{<\omega}\omega \times {}^{<\omega}\omega$ where $\sigma \in T_x$ iff some sufficiently large initial segment of x is used in a certain computable computation: $\exists N < \omega \ \forall n \ge N \ R(x \upharpoonright n, \sigma)$. We then set

 $T = \{ \langle \tau, \sigma \rangle : \exists \tau' \in {}^{<\omega} \omega \ ((\tau \lessdot \tau' \lor \tau' \vartriangleleft \tau) \land R(\tau', \sigma) \land \operatorname{lh}(\tau) = \operatorname{lh}(\sigma) \}.$

By hypothesis, $x \in X$ iff $[T_x] \neq \emptyset$ iff there is a $y \in \mathcal{N}$ with $\forall n < \omega \ R(x \upharpoonright n, y \upharpoonright n)$ iff $\langle x, y \rangle \in [T]$ iff $x \in \mathfrak{p}[T]$. As T is a computable tree, X is Σ_1^1 .

The definition of $x \mapsto T_x$ is pretty easily definable from φ . More precisely, $\varphi(x)$ is just " $\exists y \in \mathcal{N} \ \forall n \in \omega \ \psi(x \upharpoonright n, y \upharpoonright n)$ " for some $\Sigma_0 \ \text{FOL}(\{0, 1, +, \cdot\})$ -formula ψ . The formula defining $x \mapsto T_x$ just depends on this fixed ψ .

– 25 A•3. Corollary –

A set $X \subseteq \mathcal{N}$ is Π_1^1 iff there is a computable map $x \mapsto T_x$ where T_x is a computable tree over ω and $\{X = \{x \in \mathcal{N} : [T_x] = \emptyset\}$.

This gives us Σ_1^1 -absoluteness as follows, first noting some easy results about computability.

- 25 A•4. Lemma

Every transitive model of ZF contains every computable tree over ω . Moreover, every computable map $f : \mathcal{N} \to \mathcal{P}({}^{<\omega}\omega)$ such that f(x) is always a computable tree over ω has a unique $f' \in M$ where $f' = f \upharpoonright (\mathcal{N} \cap M)$.

Proof .:.

Every computable relation is given by a program $e \in \omega$ following simple rules of computation. ZF is far more than enough to ensure M can do this. Similarly, if $f : \mathcal{N} \to \mathcal{P}({}^{<\omega}\omega)$ as in the statement is computable, then there is a program e where $f = (x \mapsto [\![e]\!]_x)$. But as an x-computable function, the absoluteness of computation yields $[\![e]\!]_x^V = [\![e]\!]_x^M$ so long as $x \in M$. In particular, $f' = (x \mapsto [\![e]\!]_x)^M \subseteq f$ and f' is defined on all of $\mathcal{N}^M = \mathcal{N} \cap M$.

– 25 A • 5. Result (Mostowski Absoluteness) –

Every Σ_1^1 -relation (and hence every Π_1^1 -relation) is absolute between transitive models of ZF.

Proof .:.

Let $\mathbf{M} \models \mathsf{ZF}$ be an arbitrary transitive model and $\varphi \neq \Sigma_1^1$ -formula. By Σ_1^1 Normal Form (25 A • 2), there is a computable map $x \mapsto T_x$ (definable from φ) such that $\mathsf{ZF} \models ``\varphi(x) \leftrightarrow [T_x] \neq \emptyset$ ". Lemma 25 A • 4 tells us that $T_x^{\mathsf{M}} = T_x^{\mathsf{V}}$ for each $x \in \mathcal{N} \cap \mathsf{M}$. Therefore, it suffices to show $\mathsf{V} \models ``[T_x] \neq \emptyset$ " iff $\mathsf{M} \models ``[T_x] \neq \emptyset$ " for each $x \in \mathcal{N} \cap \mathsf{M}$.
PROPERTIES OF THE LIGHTFACE HIERARCHIES

Note that any tree S (ordered by \triangleleft) has $[S] = \emptyset$ iff the upside down version $\langle S, \rhd \rangle$ is ill-founded. Since well-foundedness is absolute between transitive models of ZF, if $S \in M$, $\mathbf{V} \models "[S] \neq \emptyset$ " iff $\mathbf{M} \models "[S] \neq \emptyset$ ". By Lemma 25 A • 4, so each $T_x \in M$ for $x \in \mathcal{N} \cap M$ and so, as desired,

$$\varphi(x)$$
 iff $[T_x] \neq \emptyset$ iff $\mathsf{M} \models "[T_x] \neq \emptyset$ " iff $\mathsf{M} \models "\varphi(x)$ " -

- 25 A • 6. Corollary -

Every Δ_2^1 -relation is absolute between transitive models of ZF.

Proof .:.

By Mostowski Absoluteness (25 A • 5), $\Sigma_2^1 = \exists^{\mathcal{N}} \Pi_1^1$ -relations are upward absolute and $\Pi_2^1 = \forall^{\mathcal{N}} \Sigma_1^1$ -relations are downward absolute. So any Δ_2^1 -relation is both.

We can also relativize these results to allow for parameters. This unsurprisingly restricts our attention to models that actually contain these parameters, but nevertheless allows us to talk about boldface absoluteness. This has a proof identical to Mostowski Absoluteness (25 A • 5) with the added requirement that the transitive $\mathbf{M} \models \mathsf{ZF}$ has $X \subseteq \mathsf{M}$ to ensure that the model \mathbf{M} can talk about the parameters in the defining Σ_1^1 -formula.

— 25 A • 7. Corollary -

For $X \subseteq \mathcal{N}$, every $\Sigma_1^1(X)$ -relation (and hence every $\Delta_2^1(X)$ -relation) is absolute between transitive models of ZF containing X.

Our next goal will be to examine the next level up: Σ_2^1 -relations. This represents the best absoluteness we can get, since the statement " $\mathcal{N} \not\subseteq L$ " can be written in a Σ_3^1 way, and it's consistent that $\mathcal{N} \not\subseteq L$ although certainly $L \models "\mathcal{N} \subseteq L$ ". So the downward absoluteness of Σ_3^1 -relations (and so the upward absoluteness of the Π_3^1 -relation " $\mathcal{N} \subseteq L$ ") doesn't hold in general.^{xiii}

The basic idea behind the proof of Shoenfield involves more trees, but now indexed by countable ordinals, similar to the fact that Π_1^1 -sets are \aleph_1 -suslin. We will also worry about the shoenfield tree S_T for a tree T, as defined in Definition 23 A • 16. For the reader uninterested in returning to that definition, the idea is just that for $T \subseteq ({}^{<\omega}\omega)^3$, S_T is composed of approximations to rank functions: $\langle x, y, R \rangle \in S_T$ iff R is a rank function on (the upside down version of) $T \subseteq \omega \times \omega$ restricted to triples of the form $\langle x \upharpoonright \ln(\sigma), y \upharpoonright \ln(\sigma), \sigma \rangle$ for some σ . The key point about the shoenfield tree S_T is that its infinite branches are then triples $\langle x, y, R \rangle$ where $x, y \in \mathcal{N}$ and R is a rank function on $\langle T_{x,y}, \rhd \rangle$. In this way $x \in \mathfrak{p}[S_T]$ iff $\exists y \in \mathcal{N}$ ($[T_{x,y}] = \emptyset$).

— 25 A • 8. Theorem (Shoenfield Absoluteness) —

Every Σ_2^1 -relation is absolute between transitive models of ZF containing ω_1 .

Proof .:.

Let φ be a Σ_2^1 -formula with $X = \{x \in \mathcal{N} : \varphi(x)\}$. Note that φ has the form

 $\varphi(x)$ iff $\exists y \in \mathcal{N} \ \forall z \in \mathcal{N} \ \exists n < \omega \ R(x \upharpoonright n, y \upharpoonright n, z \upharpoonright n)$

for some computable *R*. We can then consider the tree that essentially builds initial segments of $x, y, z \in \mathcal{N}$ until we find an $n < \omega$ where $R(x \upharpoonright n, y \upharpoonright n, z \upharpoonright n)$. Then we can focus on *z* and then work with the shoenfield tree's slices S_{α} , which is basically the shoenfield tree of approximating rank function but where there's a height limit of α : for $x, y \in \mathcal{N}$ and α , define

$$T = \{ \langle \tau, \sigma, \varsigma \rangle \in (^{n}\omega)^{\mathfrak{s}} : n < \omega \land \forall m \leq n \neg R(\tau \upharpoonright m, \sigma \upharpoonright m, \varsigma \upharpoonright m) \}$$
$$T_{x,y} = \{ \varsigma \in {}^{<\omega}\omega : \langle x \upharpoonright \operatorname{lh}(\varsigma), y \upharpoonright \operatorname{lh}(\varsigma), \varsigma \rangle \in T \}$$

^{xiii}Not all is lost in our investigation into absoluteness, however. There are a great number of *generic* absoluteness (also called *forcing* absoluteness) results investigating the agreement between V and all of its generic extensions used in the technique of *forcing*. Assuming sufficiently large cardinals, we can get generic absoluteness for all analytical relations. For example, Σ_3^1 -generic absoluteness is equivalent to closure under "sharps" [3], which we will investigate later. Moreover, assuming CH and the existence of a proper class of measurable cardinals that are also woodin, Σ_1^2 -absoluteness holds between V and any generic extension that also satisfies CH. Given that CH is also $\Sigma_1^2 \supseteq \bigcup_{n < \omega} \Sigma_n^1$, this is essentially the best we could hope for [8].

PROPERTIES OF THE LIGHTFACE HIERARCHIES

$$= \{ \varsigma \in {}^{<\omega}\omega : \forall m \le \ln(\varsigma) \neg R(x \upharpoonright m, y \upharpoonright m, \varsigma \upharpoonright m) \}$$

$$S_{\alpha} = \{ \langle \tau, \sigma, \rho \rangle \in S_{T} : \operatorname{im}(\rho) \subsetneq \alpha \}$$

$$= \{ \langle \tau, \sigma, \rho \rangle \in ({}^{n}\omega)^{2} \times {}^{n}\alpha : n < \omega \land \forall \varsigma \forall \varsigma' \lhd \varsigma \ (\langle \tau, \sigma, \varsigma \rangle \in T \to \rho \operatorname{(code}(\varsigma)) < \rho \operatorname{(code}(\varsigma'))) \}$$

$$S_{\alpha,x} = \{ \langle \sigma, \rho \rangle \in {}^{<\omega}\omega \times {}^{<\omega}\alpha : \langle x \upharpoonright \ln(\sigma), \sigma, \rho \rangle \in S_{\alpha} \}.$$

Note that, using the above definitions (which rely only on R),

 $\mathsf{ZF} \vdash ``\varphi(x) \leftrightarrow \exists y \in \mathcal{N} ([T_{x,y}] = \emptyset) \leftrightarrow \exists \alpha ([S_{\alpha,x}] \neq \emptyset)`'.$

Note also the following absoluteness results between transitive models of ZF. In particular, if $M \models ZF$ is transitive with $\omega_1 \subseteq M$ (e.g. any inner model), then

- T is computable and hence being T is absolute between such models and $T \in M$.
- The map $\langle x, y \rangle \mapsto T_{x,y}$ is computable and hence the property of being $T_{x,y}$ is absolute between such models containing x, y. Thus each $T_{x,y} \in M$ for $x, y \in M$.
- Being equal to $S_{\alpha}(S_{\alpha,x})$ is absolute between such models containing $\alpha < \omega_1(\alpha, x)$ since being in T is computable. Thus each $S_{\alpha}, S_{\alpha,x} \in M$ for $x \in \mathcal{N} \cap M$ and $\alpha < \omega_1 \subseteq M$.
- A tree having no infinite branches is absolute between such models containing the trees.

In particular, $\mathbf{M} \models "\varphi(x)"$ iff $\mathbf{M} \models "\exists \alpha \in \text{Ord} ([S_{\alpha,x}] \neq \emptyset)"$ iff $\mathbf{V} \models "\exists \alpha < \omega_1^{\mathbf{V}} ([S_{\alpha,x}] \neq \emptyset)"$ iff $\mathbf{V} \models "\varphi(x)"$ for $x \in \mathcal{N} \cap \mathbf{M}$.

In fact, by examining the proof of Shoenfield Absoluteness (25 A • 8), we get the following.

– 25 A•9. Corollary –

If X is Σ_2^1 , then X is \aleph_1 -suslin as witnessed by a tree in L: $X = \mathfrak{p}[T]$ where $T \in L$ and T is a tree over $\omega \times \omega_1$.

Proof .:.

At the risk of repeating the proof of Shoenfield Absoluteness $(25 \text{ A} \cdot 8)$, we will show $X = \mathfrak{p}[S_T]$ for the *T* defined before. More explicitly, let φ a Σ_2^1 -formula defining *X*, φ has the form " $\exists y \in \mathcal{N} \ \forall z \in \mathcal{N} \ \exists n < \omega \ R(x \upharpoonright n, y \upharpoonright n, z \upharpoonright n)$ " and we set

 $T = \{ \langle \tau, \sigma, \varsigma \rangle : \forall m \le \mathrm{lh}(\tau) = \mathrm{lh}(\sigma) = \mathrm{lh}(\varsigma) \neg R(\tau \upharpoonright m, \sigma \upharpoonright m, \varsigma \upharpoonright m) \}.$

As this is computable, $T \in L$ and it's not difficult to see that S_T is constructible from T since $\omega_1 \subseteq L$ and taking just the first component of $[S_T]$ yields $\mathfrak{p}[S_T] = X$.

Note that Corollary 25 A • 9 is actually enough to show Shoenfield Absoluteness (25 A • 8) for inner models of ZF. The idea is that for any $x \in \mathcal{N} \cap L$, if $x \in \mathfrak{p}[S_T]$, then the absoluteness of well-foundedness yields a rank function $R \in L$ with $\langle x, R \rangle \in [S_T]^L$ so that $L \models "\varphi(x)$ ", and the upward absoluteness gives the same for any other inner model of ZF. That being said, Shoenfield Absoluteness (25 A • 8) is slightly stronger than this, giving absoluteness for models that are *sets*, unlike inner models.^{xiv}

Introducing parameters gives similar absoluteness results, as one should expect.

— 25 A • 10. Corollary -

Let $X \subseteq \mathcal{N}$. Every $\Sigma_2^1(X)$ -relation (and hence every $\Delta_3^1(X)$ -relation) is absolute between transitive models of ZF containing ω_1 and X.

The usefulness of shoenfield absoluteness mostly comes from the ability to code statements in a Σ_2^1 -way. One example of this is the following.

— 25 A • 11. Result

Assume there are countable, transitive models of ZFC. Therefore L contains transitive models of ZFC + "V \neq L".

^{xiv}The same general idea works for set models, but it needs to be modified slightly and requires an analysis of where and how the shoenfield tree appears in the L_{α} levels for $\alpha < \omega_1$.

Proof .:.

The technique of *forcing* allows one to transform a countable, transitive model of ZFC into a countable, transitive model of ZFC $+ \neg$ CH (the details of this will be discussed in later chapters, for now just take this as provable in ZFC). Given that ZFC $+ "V = L" \vdash CH$, these models must satisfy "V \neq L". Thus it suffices to show that in L there are such models. In general, the existence of these models is the following statement:

$$\exists M (|M| = \aleph_0 \land \forall x, y (x \in y \in M \to x \in M) \land \mathsf{M} \vDash \mathsf{ZFC} + "\mathsf{V} \neq \mathsf{L}")$$

The existence of such an M then states the existence of a certain real number coding the structure of $\langle M, \in \rangle$ with these properties. In particular, if we have a bijection $f: M \to \omega$, rather than $\langle M, \in \rangle$, we want to consider $E = \{ \operatorname{code}(f(x), f(y)) : x, y \in M \land x \in y \} \subseteq \omega$ and vice versa. Then instead of asking questions about $\langle M, \in \rangle$, we ask questions about $\langle \omega, E \rangle$. Moreover, by taking the transitive collapse, we can translate back to M and hence we merely need to ensure $E \in \mathcal{N}$ (regarded as a subset of $\omega \times \omega$) is well-founded and extensional, something already contained in satisfying ZFC. Hence the existence of a countable, transitive model of ZFC + "V \neq L" is equivalent to the Σ_2^1 -statement

$$\exists E \in \mathcal{N} \ (E \text{ is well-founded} \land \langle \omega, E \rangle \vDash \mathsf{ZFC} + "\mathsf{V} \neq \mathsf{L}")$$

$$\leftrightarrow \exists E \in \mathcal{N} \ (\neg \exists x \in \mathcal{N} \ \forall n < \omega \ (x(n+1) \ E \ x(n)) \land \forall n < \omega \ (\langle \omega, E \rangle \vDash \text{the } n \text{th axiom of } \mathsf{ZFC} + "\mathsf{V} \neq \mathsf{L}")).$$

_	$\neg \Sigma_1^1 = \Pi_1^1$		arithmetical		
		Π_1^1			
		Σ_2^1			

In V, we have the existence of such models and therefore the above Σ_2^1 -sentence is true. By Shoenfield Absoluteness (25 A • 8), it holds in L. Taking the transitive collapse of $\langle \omega, E \rangle$ yields the result.

A corollary of the above proof is the following.

– 25 A•12. Corollary -

Let T be a theory definable over N (e.g. any finite theory) and φ a formula. Therefore the set of $x \subseteq \omega$ where there is a countable, transitive model (which is then in HC) satisfying $T + \varphi(x)$ is Σ_2^1 .

Proof .:.

At the risk of repeating the proof of Result 25 A • 11, consider the statement $\exists E \in \mathcal{N} \ (\exists \exists x \in \mathcal{N} \ \forall n < \omega \ (x(n+1) E x(n)) \land (\omega E) \models "\omega(x)" \land (\omega E)$



This is equivalent to there being a countable, transitive model satisfying $T + "\varphi(x)"$ since any such model, **M**, has a bijection $f : M \to \omega$ which yields $E = \{ \operatorname{code}(f(x), f(y)) : x, y \in M \}$ witnessing (*). Any E witnessing (*) has the transitive collapse of $\langle \omega, E \rangle$ as countable, transitive, and isomorphic to $\langle \omega, E \rangle$ meaning it satisfies $T + "\varphi(\pi(x))"$ where π is the collapsing map. Since $x \subseteq \omega$ is transitive, $\pi(x) = x$ by Corollary 6 C • 2. \dashv

In particular, not every transitive model of ZFC in L is a level of L, in contrast to Condensation (8 B • 3). This shows that the sentence "V = L" is indeed required (or some other requirement) to ensure being a level of L. Although our transitive $M \subseteq L$ with $M \models$ ZFC has $L_{Ord \cap M} = L^M \subseteq M \subseteq L_\beta$ for some β , there's no reason to think that this $\beta = Ord \cap M$ —that $M = V_{Ord \cap M}^M \subseteq L_{Ord \cap M}$ —unless $M \models$ "V = L".

The method of proving Result 25 A • 11 is important mostly for the idea of coding countable structures into ω and then categorizing the properties we want into the analytical hierarchy. So Result 25 A • 11 is really saying that well-foundedness is Π_1^1 -definable. Hence the existence of a certain countable well-founded structure with some Γ -definable properties is $\exists^{\mathcal{N}} (\Pi_1^1 \wedge \Gamma) = \exists^{\mathcal{N}} \Pi_1^1 = \Sigma_2^1$ for $\Gamma \subseteq \Sigma_1^1$. In particular, for arithmetical properties like first-order satisfaction, L has the same things consistent with transitive models as V.

(*)

As explained at the end of Subsection 24 D, there is another characterization of Σ_2^1 -sets, and in fact Σ_n^1 for $n \ge 2$. Proving this isn't that difficult, but relies on the coding idea above and noting that the existence of a countable structure with certain properties implies the existence of such a structure in HC, the hereditarily countable sets: $\text{HC} = \text{H}_{\aleph_1} = \{x : |\text{trcl}(x)| < \aleph_1\}$. First we prove the beginning of the induction, and the rest of the analytical hierarchy follows easily. Note that $\mathcal{N} \subseteq \text{HC}$ is a class of $\text{HC} = \langle \text{HC}, \in \rangle$, defined by $x \in \mathcal{N}$ iff $\text{HC} \models "x$ is a function $\land x \subseteq \omega \times \omega$ ".

- 25A · 13. Lemma -
A set
$$X \subseteq \mathcal{N}$$
 is Σ_2^1 iff X is Σ_1^{HC}

Proof .:.

Suppose $X \in \Sigma_2^1$ as defined by the formula " $\exists y \in \mathcal{N} \varphi(x, y)$ ", where φ is Π_1^1 . Note that Mostowski Absoluteness (25 A • 5) really states absoluteness of φ for transitive models of some *finite* subset of $\Delta \subseteq$ ZF. It follows that $\exists y \in \mathcal{N} \varphi(x, y)$ iff there is a countable, transitive model $\mathbf{M} \models \Delta$ and a $y \in \mathcal{N} \cap M$ with $\mathbf{M} \models$ " $\varphi(x, y)$ " (recall Corollary 7 D • 8, that any finite fragment of ZFC has countable transitive models). Note that this is Σ_1 -definable over **HC**:

$$\exists M \exists y \ \left(\forall z \in M \ \forall x \in z \ (x \in M) \land (\varphi(x, y))^M \land \forall n \in y \ (n \in \omega) \land \bigwedge_{\psi \in \Delta} \psi^M \right).$$

Now suppose φ is a Σ_1 -formula of the form $\exists y \ \psi(x, y)$ where ψ is Σ_0 . Result 7 A • 1 implies that ψ is absolute between transitive sets. In particular, by (\rightarrow) considering $M = \operatorname{trcl}(\{x, y\})$ or (\leftarrow) considering upward absoluteness, for $x \in \mathcal{N} \subseteq \operatorname{HC}$, $\operatorname{HC} \models \exists y \in \mathcal{N} \ \varphi(x, y)$ " iff there is a countable, transitive $M \in \operatorname{HC}$ with $x \in M$ where $\mathbf{M} \models \exists y \in \mathcal{N} \ \psi(x, y)$ ". The latter can be coded in a Σ_1^2 -way by Corollary 25 A • 12, meaning the set Xof $x \in \mathcal{N}$ where this happens is Σ_2^1 .

Note that the above uses AC, mostly because of its reliance on taking skolem hulls for Corollary 7 D • 8. This isn't actually necessary for the (\leftarrow) direction. DC is enough in this case because at the heart of the matter, Σ_0 -formulas like ψ are absolute between all transitive models. So rather than working with skolem hulls to witness them, we can merely take trcl({x, y}) and proceed using DC.

- 25 A • 14. Theorem Let $0 < n < \omega$ and $X \subseteq \mathcal{N}$. Therefore X is Σ_{n+1}^1 iff X is Σ_n^{HC} .

Proof .:.

Lemma 25 A • 13 gives the result for n = 1. Assuming the result for n, we immediately get the corresponding result for Π_{n+1}^1 and Π_n^{HC} just by taking negations in the defining formulas. For the inductive step, note that since $\mathcal{N} \subseteq \text{HC}$, for any formula φ ,

 $\exists^{\mathcal{N}}\{\langle x, y \rangle \in \mathcal{N}^2 : \mathsf{HC} \vDash \varphi(x, y)^{"}\} = \{x \in \mathcal{N} : \mathsf{HC} \vDash \exists y \in \mathcal{N} \ \varphi(x, y)^{"}\}.$

 $(\rightarrow) \text{ Suppose } X \text{ is defined by the } \Sigma_{n+2}^1 \text{-formula "} \exists y \in \mathcal{N} \varphi(x, y) \text{" where } \varphi \text{ is a } \Pi_{n+1}^1 \text{-formula. Inductively,} \\ Y = \{\langle x, y \rangle \in \mathcal{N}^2 : \varphi(x, y)\} \text{ is } \Pi_n^{\mathsf{HC}} \text{ by some formula } \psi \text{ so that } X = \exists^{\mathcal{N}} Y \text{ is } \Sigma_{n+1}^{\mathsf{HC}}.$

(\leftarrow) Suppose X is defined by the Σ_{n+1} -formula " $\exists y \in \mathcal{N} \ \psi(x, y)$ " where ψ is a Π_n -formula. Inductively, $Y = \{\langle x, y \rangle \in \mathcal{N}^2 : \mathsf{HC} \models "\psi(x, y)"\}$ is Π_{n+1}^1 so that $X = \exists^{\mathcal{N}} Y \in \Sigma_{n+2}^1$. \dashv

§ 25 B. L and WO with the lightface pointclasses

Mentioned in previous sections, L has

- a Π_1^1 -set without the perfect set property;
- a Δ_2^1 -set without the baire property; and
- a Δ_2^1 -set that isn't lebesgue measurable.

These mostly come from the fact that L has a Δ_2^1 well-order of \mathcal{N} in addition to other regularity properties. What exactly is this Δ_2^1 well-order? It's precisely the definable well-order of the entire universe of L: the constructibility order $<_L = \bigcup_{\alpha < Ord} <_{L_{\alpha}}$ as in Theorem 8 A • 8 restricted to \mathcal{N} .

25B•1. Lemma

The map
$$\alpha \mapsto L_{\alpha}$$
 is Δ_1^{ZF-P} -definable and hence absolute between transitive models of $ZF - P$

Proof .:.

This is really just a result of examining the previous results of absoluteness in Subsection 7 B. In particular, Corollary 8 A • 4 gives the definition of L_{α} by recursion. The first step is Δ_1^{ZF-P} and the inductive step is too: x = FOLp(y) is $\Delta_1^{\text{ZF-P}}$. (It's $\Pi_1^{\text{ZF-P}}$ as "for all things closed under the operations, x is contained in them". And it's Δ_1^{ZF-P} as "there exists a function that iteratively defines the closure and x is the the image of this".) Theorem 7 B•4 then tells us that the recursion $x = L_{\alpha}$ is also Δ_1^{ZF-P} -definable (for any and all functions L obeying the Σ_1 or Π_1 -definition, the output is *x*).

- 25 B • 2. Theorem \mathcal{N}^{L} is Σ_2^1 . In fact, $<_{\mathsf{L}}$ is a Σ_2^1 well-ordering of \mathcal{N}^{L} . If $\mathcal{N}^{\mathsf{L}} = \mathcal{N}$, then these are Δ_2^1 and so $\mathsf{L} \models$ "there's a Δ_2^1 well-order of \mathcal{N} ".

Proof .:.

To show \mathcal{N}^{L} is Σ_{2}^{1} , we merely need to show \mathcal{N}^{L} is 1-definable over HC by Lemma 25 A • 13. The proof of $L \models GCH$ (Theorem 8 C • 5) tells us that

$$\mathcal{N}^{\mathsf{L}} = \mathcal{N} \cap \mathsf{L} = \mathcal{N} \cap \bigcup_{\alpha < \omega_1} \mathsf{L} \subseteq \mathsf{HC}.$$

Lemma 25 B • 1 tells us that L_{α} is Σ_1 -definable over between transitive models of ZF – P. Theorem 7 C • 8 tells us that $\mathbf{HC} = \mathbf{H}_{\aleph_1} \vDash \mathsf{ZFC} - \mathsf{P}$. As a result, $x \in \mathcal{N}^{\mathsf{L}}$ iff

$$\mathsf{HC} \vDash "\exists \alpha \in \operatorname{Ord}(\underbrace{x \in \mathbf{L}_{\alpha}}_{\Sigma_1})".$$

This shows \mathcal{N}^{L} is Σ_2^1 . To show $<_{\mathsf{L}}$ is Σ_2^1 , we also show that it's Σ_1^{HC} : first-order satisfaction is Σ_0 and $x <_{\mathsf{L}} y$ iff **HC** \models " $\exists \alpha \in Ord (x, y \in L_{\alpha} \land L_{\alpha} \models$ "x < y")"

$$\sum_{i=1}^{n} \sum_{\alpha \in OIU} \left(\underbrace{x, y \in L_{\alpha} \land L_{\alpha} \vdash x < \underline{L} y}_{\Sigma_{1}} \right)$$

This is Σ_1 and hence $x <_L y$ is Σ_2^1 by Lemma 25 A • 13.

Thus the various sets considered in Section 23 by use of AC can be placed in the projective hierarchy of L because L can classify a well-order of \mathcal{N} . One might wonder if because this well-order is Σ_2^1 , does Shoenfield Absoluteness $(25 \text{ A} \cdot 8)$ tell us that this the existence of a well-order is absolute between all transitive models of ZF containing ω_1 ? The answer is no: $<_{L}$ being well-order of \mathcal{N}^{L} is absolute between such models, but $\mathcal{N} = \mathcal{N}^{L}$ isn't absolute and is indeed a Π_3^1 statement: $\mathcal{N} \setminus \mathcal{N}^{\mathsf{L}} \neq \emptyset$ is Σ_2^{HC} and therefore Σ_3^1 :

$$\mathcal{N} \setminus \mathcal{N}^{\mathsf{L}} = \emptyset \quad \text{iff} \quad \mathsf{HC} \models ``\exists x \underbrace{\forall \alpha \ (x \notin \mathsf{L}_{\alpha})}_{\Sigma_2}`'$$

The point is being constructible is absolute and the order of constructibility is absolute, but these need not encompass everything.

- 25 B • 3. Corollary
The relation
$$\{r_n \in \mathcal{N} : n \in \omega\} = \{y \in \mathcal{N} : y <_L x\}$$
 is Σ_2^1 .

Proof .:.

Regard $r \in \mathcal{N}$ as $r' = \{r_n \in \mathcal{N} : n \in \omega\}$ through coding. Since this is a countable set, $r' \subseteq L_{\alpha} \in HC$ for some

 \dashv

PROPERTIES OF THE LIGHTFACE HIERARCHIES

 \dashv

$$\alpha < \omega_1. \text{ So we show again via Lemma 25 A • 13 that the relation is } \Sigma_2^1: r' = \{y \in \mathcal{N} : y <_L x\} \text{ iff}$$
$$\mathsf{HC} \vDash \underbrace{\forall n < \omega \ (r_n <_L x)}_{\Sigma_1} \land \exists \alpha \in \operatorname{Ord} \left(\underbrace{r' \subseteq \mathsf{L}_{\alpha} \land x \in \mathsf{L}_{\alpha} \land \mathsf{L}_{\alpha} \vDash \forall y <_L x \exists n \in \omega \ (y = r_n)^n}_{\Sigma_1}\right)^n \dashv$$

It turns out that the well-order $<_L \subseteq \mathcal{N} \times \mathcal{N}$ both isn't lebesgue measurable and doesn't have the baire property in L. We prove the weaker statement that there are (potentially different) projective sets without these properties in L

- 25B•4. Corollary $L \models$ "there are analytical sets without the baire property and that are non-measurable".

Proof .:.

The set Vit that isn't lebesgue measurable from Result 23 B • 16 was (any) set of equivalence classes under the equivalence relation $x \approx y$ iff $x - y \in \mathbb{Q}$. This set can be defined in L by

$$x \in \text{Vit}$$
 iff $\forall y \in \mathbb{R} (\underbrace{\exists q \in \mathbb{Q} (x - y = q)}_{\Delta_1^1} \rightarrow \underbrace{x \leq_L y}_{\Sigma_2^1})$
 \prod_{3}^{1}

This set also doesn't have the baire property by Result 23 C • 11.

Again, the stronger theorem is as follows. Only a proof sketch is given, since it requires more knowledge about lebesgue measure and meagre sets.^{xv}

- 25 B • 5. Theorem

 $L \models "<_L$ doesn't have the baire property and isn't lebesgue measurable". Hence it's consistent that there are Δ_2^1 -sets that aren't lebesgue measurable and that don't have the baire property.

Proof Sketch .:.

Work in L. For $A \subseteq \mathcal{N} \times \mathcal{N}$ and $x \in \mathcal{N}$, write $A_x = \{y : \langle y, x \rangle \in A\}$. For example, $(\langle x \rangle_x) = \{y : y \langle x \rangle_x \}$ which is countable and therefore both meagre and lebesgue null. Since this holds for each $x \in \mathcal{N}$, it follows that:

- If $<_{\rm L}$ is lesbesgue measurable, it's null; and
- If $<_{L}$ has the baire property, it's meagre.

But similarly, if we consider the complement $A = \mathcal{N}^2 \setminus \langle _L, A = \{\langle x, y \rangle : \langle y, x \rangle \in \langle _L \} \cup \{\langle x, x \rangle : x \in \mathcal{N}\}.$ So the same argument applies (taking the second component slices instead of the first-component slices) to tell us that if A is measurable then it's null and if A has the baire property then it's meagre. But A is measurable iff \langle_L is, and similarly with the baire property. But \mathcal{N}^2 isn't the union of two measure 0 sets, nor the union of two meagre sets. Hence \langle_L can't be lebesgue measurable, nor have the baire property. \dashv

The proof $L \models$ "there's a Π_1^1 -set without the perfect set property" requires much more involved analysis of the lightface pointclasses, so we will not prove it just yet. Nevertheless, we *can* show there's a Σ_2^1 -set without the perfect set property. To do this, we need more information about well-orders. In particular, we need the very useful Boundedness Lemma, telling us that a Σ_1^1 -set of reals X coding ordinals must have the set { $\alpha < \omega_1 : \alpha$ is coded by an element of X } *bounded* in ω_1 , even if X itself is uncountable.

We first introduce some definitions that have been implicitly used in the background above.

^{xv}In particular, it uses Fubini--Tonelli for null sets and an analogue of Fubini--Tonelli for category. Fubini--Tonelli says more or less that $\iint f(x, y) \, dx \, dy = \iint f(x, y) \, dy \, dx$ for a measurable function f. In our case, the measurable function is the characteristic function χ_A where $\chi(x, y) = 1$ if $\langle x, y \rangle \in A$ and otherwise $\chi_A(x, y) = 0$. It follows that $\iint \chi_A(x, y) \, dy \, dx = \mu(A)$. If for all $x, \int \chi_A(x, y) \, dy = \mu(A_x) = 0$, then $\mu(A) = \iint \chi_A(x, y) \, dy \, dx = \int 0 \, dx = 0$. There's an analogous result for the baire property though, as usual, the proof is somewhat different.

25 B • 6. Definition
Let code : ω² → ω be any computable coding. For x ∈ N, define the relation E_x = {⟨n,m⟩ ∈ ω² : x(code(n,m)) = 1}. We then set
WF = {x ∈ N : E_x is well-founded}; and
WO = {x ∈ N : E_x is a well-order} ⊆ WF.
For x ∈ WF, set ||x|| to be the height of E_x.

We've analyzed these before as both being Π_1^1 .

- 25B·7. Corollary WF, WO $\in \Pi_1^1$.

Proof .:.

 $x \in WF$ iff $\forall y \in \mathcal{N} \neg \forall n \in \omega$ $(y(n + 1) E_x y(n))$ which can be decoded further if one desires to the more explicitly Π_1^1 statement

 $\forall y \in \mathcal{N} \neg \forall n \in \omega \ (x(\operatorname{code}(y(n+1), y(n))) = 1).$

Similarly, $x \in WO$ iff $x \in WF \land E_x$ is linear. The statement that a real codes a linear order over ω will be arithmetical and hence Π_1^1 .

This is the easy part. The harder part is showing that we can talk about ||x|| in a simple way. It turns out that we can do this in a Δ_1^1 -way.

- 25B•8. Result For $x, y \in WO$, $||x|| \le ||y||$ is Δ_1^1 . More precisely, there are Σ_1^1 and Π_1^1 relations \le_0 and \le_1 where if $y \in WO$ then $x \le_0 y \Leftrightarrow x \le_1 y \Leftrightarrow x \in WO \land ||x|| \le ||y||$.

Proof .:.

Suppose $y \in WO$. Note that $x \in WO \land ||x|| \le ||y||$ is equivalent to the existence of an order preserving injection from E_x into E_y whenever E_x is a linear order. Such a function is merely a member of baire space and thus $x \in WO \land ||x|| \le ||y||$ iff the following Σ_1^1 -relation holds:

 $x \leq_0 y$ iff x is a linear order $\land \exists r \in \mathcal{N} \forall n, m \in \omega \ (r(n) \neq r(m) \land (n \ E_x \ m \rightarrow r(n) \ E_y \ r(m))).$

To show $x \in WO \land ||x|| \le ||y||$ is Π_1^1 , for $x \in WO$, $||x|| \le ||y||$ iff $||y|| \ne ||x||$, meaning there's no orderpreserving injection from y into x that isn't surjective. In other words, $||x|| \le ||y||$ iff every order-preserving injection from E_y isn't mapping onto an initial segment of E_x . But this is clearly Π_1^1 by the same idea as before:

$$x \leq_1 y \quad \text{iff} \quad x \in WO \land \forall r \in \mathcal{N} \ \neg \exists b \in \omega \ \forall n, m \in \omega \ \left(\begin{array}{c} r(n) \neq r(m) \neq b \land \\ (n \ E_y \ m \to r(n) \ E_x \ r(m) \ E_x \ b) \end{array} \right). \qquad \dashv$$

Note that $||x|| \le ||y||$ isn't itself Δ_1^1 . Really, assuming $y \in WO$, $\{x \in WO : ||x|| \le ||y||\}$ is Δ_1^1 in a uniform way although the relation $R(x, y) \leftrightarrow ||x|| \le ||y||$ isn't Δ_1^1 by the following result. This is partly because \le_0 and \le_1 above might differ significantly if $y \notin WO$. But we will only use the relation if dealing with codes of well-orders anyway.

- 25B•9. Lemma Let $A \subseteq \mathcal{N}$. If A is Π_1^1 then there is a computable $f : \mathcal{N} \to \mathcal{N}$ such that $A = f^{-1}$ "WO. In particular, if A is Π_1^1 , then there is a continuous $f : \mathcal{N} \to \mathcal{N}$ such that $A = f^{-1}$ "WO.

Proof .:.

Consider Σ_1^1 Normal Form (25 A • 2), which tells us A is Π_1^1 iff $A = \{x \in \mathcal{N} : [T_x] = \emptyset\}$ for some computable map $x \mapsto T_x$. Using this computable map, we can define the computable $f : \mathcal{N} \to \mathcal{N}$ by f(x) = y where y is an order on $\langle T, \rhd \rangle$. In other words, we take f(x) to code the relation $\tau E_{f(x)} \sigma$ for $\tau, \sigma \in {}^{<\omega}\omega$ defined as true iff

• $\tau \rhd \sigma$ with $\tau, \sigma \in T_x$; or

• $\tau <_{\text{lex}} \sigma$ and neither $\tau \leq \sigma$ nor $\sigma \leq \tau$ for $\tau, \sigma \in T_x$; or

PROPERTIES OF THE LIGHTFACE HIERARCHIES

• $\tau \notin T_x$ but $\sigma \in T_x$; or

• $\tau, \sigma \notin T_x$ and $\operatorname{code}(\tau) < \operatorname{code}(\sigma) \in \omega$.

This relation essentially has that T_x is well-found iff $E_{f(x)}$ is a well-order where $\tau \leq \sigma$ implies $\sigma E_{f(x)}\tau$. If T_x had any infinite branches, this corresponds to an infinite $E_{f(x)}$ decreasing sequence. Similarly, if $E_{f(x)}$ has an infinite decreasing sequence, it's not due to the elements not in T_x (by the last two conditions) nor to $<_{\text{lex}}$ -decreasing sequences.

Therefore $A = \{x \in \mathcal{N} : [T_x] = \emptyset\} = \{x \in \mathcal{N} : f(x) \in WO\}$. So it suffices to show f is computable, but this is clear because $x \mapsto T_x$ is computable and f is computable from this map. \dashv

We are now in a position to prove the boundedness lemma. There are actually different versions of the boundedness lemma. The version presented below says that \sum_{1}^{1} -subsets of WO are bounded below ω_{1} . The recursive analogue of this is that \sum_{1}^{1} -subsets of WO are bounded below ω_{1}^{CK} . Because we have not introduced ω_{1}^{CK} formally here (just in Appendix B), we only prove the bound of ω_{1} rather than ω_{1}^{CK} .

− 25 B • 10. Theorem (The Boundedness Lemma) Let $X \subseteq$ WO be Σ_1^1 . Therefore sup{ $||x|| : x \in X$ } < ω_1 . In particular, WO = { $x \in \mathcal{N} : ||x|| < \omega_1$ } is not Σ_1^1 .

Proof .:.

Suppose not so that every ordinal $\alpha < \omega_1$ is below the height of some $x \in X$. In particular, WO = $\{y \in \mathcal{N} : \exists x \in X (||y|| \le ||x||)\}$ is $\sum_{i=1}^{1}$ by Result 25 B • 8. Thus it suffices to show WO isn't $\sum_{i=1}^{1}$. But this follows from Lemma 25 B • 9: if WO were $\sum_{i=1}^{1}$, then any $\prod_{i=1}^{1}$ set would be the continuous preimage of a $\sum_{i=1}^{1}$ set and hence $\sum_{i=1}^{1}$, contradicting that $\prod_{i=1}^{1} \neq \sum_{i=1}^{1}$.

One corollary to this is that the relation R(x, y) iff $x, y \in WO$ and $||x|| \le ||y||$ isn't Δ_1^1 (compare with Result 25 B • 8) because otherwise $\mathfrak{p}R \in \Sigma_1^1$ but $\sup\{||x|| : x \in \mathfrak{p}R\} = \omega_1$, contradicting The Boundedness Lemma (25 B • 10).

For the lightface variant, we take ω_1^{CK} to be the set of all recursive ordinals.

- 25B•11. Definition An $\alpha \in \text{Ord is recursive}$ iff there is some computable $R \subseteq \omega \times \omega$ such that $\langle \omega, R \rangle \cong \langle \alpha, \in \rangle$. $\omega_1^{CK} = \sup\{\alpha \in \text{Ord } : \alpha \text{ is recursive}\}$.

It should be clear that every recursive ordinal is countable and as there are only countably many computable relations, $\omega_1^{\text{CK}} < \omega_1$. The only result we need is the intuitively clear result that if α is recursive and $\beta < \alpha$, then β is recursive (just by considering an initial segment of the computable relation for σ).

- 25 B • 12. Theorem (The Lightface Boundedness Lemma) Let $X \subseteq$ WO be Σ_1^1 . Therefore $\sup\{||x|| : x \in X\} < \omega_1^{CK}$.

A corollary of The Boundedness Lemma (25 B \cdot 10) is that L has a Σ_2^1 -set without the perfect set property.

- 25 B • 13. Result

$$L \models :: \Sigma_1^1$$
 doesn't have the perfect set property"

Proof .:.

Argue in L. For each $\alpha < \omega_1$, we have a real number $x \in \mathcal{N}$ coding $\langle \alpha, \in \rangle$. There might be multiple, so for each $\alpha < \omega_1$, let $f(\alpha)$ be the $<_L$ -least $x \in \mathcal{N}$ coding $\langle \alpha, \in \rangle$. To show im f = X is Σ_2^1 , just note that $x \in X$ iff $\underbrace{x \in WO}_{\Pi_1^1} \land \exists r \in \mathcal{N} (\underbrace{\{r_n : n \in \omega\} = \{y \in \mathcal{N} : y <_L x\}}_{\Sigma_2^1} \land \forall n \in \omega (\underbrace{r_n \notin WO}_{\Sigma_1^1} \lor \underbrace{\|r_n\| \neq \|x\|}_{\Delta_1^1})).$

So it suffices to show X doesn't have the perfect set property. Clearly X = im f is uncountable since $f : \omega_1 \rightarrow \omega_1$

X is bijective. To show *X* has no perfect subset, we can actually show *X* has no (uncountable) closed subsets and in fact, no (uncountable) \sum_{1}^{1} -sets at all! This follows from The Boundedness Lemma (25 B • 10), since if $A \subseteq X$ is $\prod_{1}^{0} \subseteq \sum_{1}^{1}$, then sup{ $||x|| : x \in A$ } < ω_1 . Since $x \neq y \in X$ have $||x|| \neq ||y||$, this implies *A* is countable and therefore not perfect.

Again, this can be improved to a Π_1^1 -set without the perfect set property, but this requires more knowledge about uniformization. The last thing we will talk about in this subsection before getting to more general properties of the lightface pointclasses is the how this relates to *relative* constructibility, giving striking theorems that are more precise about how wrong L can be.

- 25B • 14. Definition For $x \in \mathcal{N}$, L[x] is the least inner model of ZFC with x as an element.

We will give a proper introduction to relative constructibility later. For now, we state the following interesting results, appealling to intuition about L[x] which is constructed in much the same way as L, but given access to x as a kind of oracle.

− 25 B • 15. Theorem For every $x \in \mathcal{N}$, $\omega_1^{L[x]} < \omega_1$ iff for every $x \in \mathcal{N}$, $L[x] \models "\omega_1^V$ is (weakly) inaccessible".

Proof .:.

Clearly if $L[x] \models "\omega_1^V$ is inaccessible" then $\omega_1 > \omega_1^{L[x]}$. So the (\leftarrow) direction is clear. Suppose $\omega_1^{L[x]} < \omega_1$. We know ω_1 is still regular in L[x] by downward absoluteness. So ω_1 is (weakly) inaccessible in L[x] iff ω_1 is a limit cardinal.

Suppose not: let $\omega_1 = (\kappa^+)^{L[x]}$ for some κ a cardinal of L[x] which tells us $\kappa < \omega_1$ so that κ is countable in V. Let $y \in \mathcal{N}$ code κ in that $\langle \omega, E_y \rangle \cong \langle \kappa, \epsilon \rangle$. So if we consider $x, y \in L[x * y]$, we can decode y to get a bijection: $L[x * y] \models "|\kappa| = \omega$ " and thus

$$\mathbf{L}[x * y] \vDash ``\omega_1 = |\kappa|^+ \ge (\kappa^+)^{\mathbf{L}[x]} = \omega_1^{\mathbf{V},*},$$

contradicting that $\omega_1^{\mathsf{L}[x*y]} < \omega_1$.

This is striking because of the following theorem, which then relates these topological properties with large cardinal hypotheses: if ω_1^{V} is inaccessible in every $\mathsf{L}[x]$ for $x \in \mathcal{N}$, then $\prod_{i=1}^{1}$ has the prefect set property. It turns out that the three conditions are actually equivalent, as we will show. Thus the stronger statement tells us we can satisfy the hypothesis of Theorem 25 B • 15 just by ensuring $\prod_{i=1}^{1}$ or $\sum_{i=1}^{2}$ has the perfect set property (which clearly L doesn't satisfy by Result 25 B • 13)

- 25B • 16. Theorem For every $x \in \mathcal{N}$, $\omega_1^{L[x]} < \omega_1 \to \Sigma_2^1(x)$ has the perfect set property $\to \Pi_1^1(x)$ has the perfect set property.

Proof .:.

For the sake of notation, take $x = \emptyset$ as the proof easily generalizes to parameters.

• Suppose $\omega_1^L < \omega_1$. Let $X \in \Sigma_2^1$ be uncountable. We must show X has a perfect subset. By Corollary 25 A • 9, $X = \mathfrak{p}[T]$ for some $T \in L$ a tree over $\omega \times \omega_1$. We can continually thin out T as in Theorem 23 A • 20 just by removing isolated branches:

$$T_0 = T$$
 $T_{\alpha+1} = \text{prune}(T_\alpha)$ $T_\gamma = \bigcap_{\alpha < \gamma} T_{\alpha < \gamma}$

where prune(S) just consists of all nodes in S which have incompatible extensions above them (so that their branch isn't isolated). Clearly as $T \in L$, prune $(T) \in L$ as having incompatible extensions is absolute between transitive models. So inductively all members of this sequence are in L. This process stabilizes by some stage α yielding $T^* = T_{\alpha} \in L$.

 \dashv

If X doesn't have a perfect subset, then the reasoning of Theorem 23 A • 20 tells us $T^* = \emptyset$, giving an alternative characterization of $X: x \in X$ iff x is removed from the projection of $[T_\alpha]$ at some stage α . But for each α , the things removed from T_α are uniquely the things in some initial segment's only corresponding branch. Since these branches are unique and L certainly thinks they exist by the absoluteness of well-foundedness, all of these branches are in L and thus $X \subseteq L$. Because $X \subseteq (\mathfrak{p}[T])^L \subseteq \mathfrak{p}[T] = X, X \in L$.

But in L, X is still Σ_2^1 : the Σ_2^1 -relation φ defining X is still absolute

 $\{x \in \mathcal{N} \cap \mathcal{L} : \mathbf{L} \models ``\varphi(x)"\} = X \cap \mathcal{L} = X \in \mathcal{L}.$

Hence applying Corollary 25 A • 9 inside L, we get that X is in fact \aleph_1^L -suslin. Since X still doesn't have a perfect subset, Theorem 23 A • 20 tells us $|X| \le \aleph_1^L < \aleph_1$ and hence X is countable, a contradiction.

• If every $\Sigma_2^1(x)$ -set has the perfect set property, then every $\Pi_1^1 \subseteq \Sigma_2^1$ -set does too.

Theorem 25 B • 15 and Theorem 25 B • 16 together tell us that \sum_{2}^{1} has the perfect set property iff $L[x] \models "\omega_{1}^{V}$ is inaccessible" for all $x \in \mathcal{N}$ and for L = L[0] in particular. Thus we have a connection between large cardinal hypotheses like the consistency of inaccessible cardinals with topological properties of the real numbers. As another example, one can show that the existence of a measurable cardinal implies the inaccessibility of ω_{1}^{V} in each L[x], $x \in \mathcal{N}$, and hence the perfect set property for \sum_{1}^{2} -sets.

§25C. Prewellorders

The motivating concept we will look at will be uniformization, which can be thought of as a choice function with some restrictions on complexity. To study these sets, we will need to look at scales and prewellorderings. Our study of these concepts tell us that in L there's a Π_1^1 -subset without a perfect subset, and we also get the converses to Theorem 25 B • 16. These basic ideas also give other theorems *similar* to The Boundedness Lemma (25 B • 10), like that well-founded \sum_{2}^{1} -relations have length $< \omega_2$, which can be said in the more "impressive" way that $\delta_2^1 \le \omega_2$, to be explained later.

- 25C•1. Definition Let $X \subseteq A \times B$. A *uniformization* is a function $f \subseteq X$ such that dom(f) = dom(X). For Γ a pointclass, Γ -*uniformization* is the statement that every $X \in \Gamma$ has a uniformization $f \in \Gamma$.

In the end, the two pointclasses in our hierarchies that provably have uniformization are Π_1^1 and Σ_2^1 (and their relativizations). These analytical (or projective) pointclasses can be talked about in more general terminology. So we introduce the idea of an *adequate* pointclass, encompassing all the lightface and boldface pointclasses of our hierarchies thus far.

— 25C•2. Definition ——

A pointclass $\Gamma \subseteq \mathfrak{P}(\mathcal{N})$ is *adequate* iff

- Γ contains all computable relations;
- Γ is closed under computable preimages;
- Γ is closed under finite unions and intersections;
- Γ is closed under bounded quantification (over ω).

It's not difficult to see that the borel, arithmetical, projective, and analytical pointclasses (and their relativizations) are all adequate.

The best way to frame the proofs of these results on uniformization is to introduce the concepts of *norms* and *scales*. Norms should be fairly familiar, but scales are a difficult concept to digest, requiring substantial background.

- 25C·3. Definition

A *prewellorder* is a relation \leq that is transitive, total, and well-founded.

This is a *pre*wellorder in the sense that if we "mod out" by the equivalence relation $x \approx y$ iff $x \leq y \leq x$, the result is a well-order. Alternatively, prewellorders are well-orders where we allow clusters of loops, but which don't

PROPERTIES OF THE LIGHTFACE HIERARCHIES

fundamentally change the length of the order if we merely think of these loops as single elements. In particular, because prewellorders are well-founded, we have a rank function on them which gives the length. It's not hard to see that every element of a loop is given the same rank: if $x \approx y$ then by transitivity, $\{z : z < x\} = \{z : z < y\}$ and thus the rank of x is the rank of y.

Frequently these rank functions are thought of as *norms*, a more general kind of function which will in turn define a prewellorder.

- **25C**•**4**. **Definition** A *norm* on a set *X* is a function $\varphi : X \to \text{Ord.}$

— 25 C • 5. Corollary

Every prewellorder has a norm, being its rank function. Moreover, for every norm $\varphi : X \to \text{Ord}$, there is a prewellorder $\leq \subseteq X^2$ defined by $x \leq y$ iff $\varphi(x) \leq \varphi(y)$.

Proof .:.

That a rank function is a norm is obvious. So suppose $\varphi : X \to \text{Ord}$ is a norm, and define \leq as in the statement. \leq is clearly well-founded since if $Y \subseteq X$, $\varphi''Y$ has a minimal element $\varphi(y)$ which implies y is \leq -minimal in Y. That \leq is total and transitive follows from \leq being total and transitive on Ord. \dashv

Note that the norm associated with a prewellorder isn't unique. The rank function will be unique just because it's constructed iteratively, but if we consider for example $\langle \{0, 1\}, < \rangle$, then $\varphi = \{\langle 0, 0 \rangle, \langle 1, 1 \rangle\}$ and $\varphi' = \{\langle 0, 0 \rangle, \langle 1, 4 \rangle\}$ are distinct norms giving the same (pre) well-order: x < y iff $\varphi(x) < \varphi(y)$ iff $\varphi'(x) < \varphi'(y)$.

We've actually already seen norms and prewellorders with WO: $x \mapsto ||x||$ is a norm for this, defining the Δ_1^1 -relation $||x|| \le ||y||$ for $y \in$ WO by Result 25 B • 8. This is important because we may actually generalize this idea to all Π_1^1 -sets. The resulting property is called the prewellordering property for Π_1^1 -sets, often written PWO(Π_1^1).

- **25 C • 6.** Definition - Let Γ ⊆ $\mathcal{P}(\mathcal{N})$ be a pointclass.

X ⊆ N has a Γ-norm iff there's a norm φ : X → Ord such that ≤φ and <φ are both in Γ, defined by x ≤φ y iff x ∈ X ∧ (y ∈ X → φ(x) ≤ φ(y)) x <φ y iff x ∈ X ∧ (y ∈ X → φ(x) < φ(y)).
Γ has the *prewellordering property*, PWO(Γ), iff every X ∈ Γ has a Γ-norm.

One might think that *every* set has a simple to define norm as just the constant 0 function. While it's true that this will be a norm, it may not have the best complexity: this constant function might be fairly complex if X is complex. In particular, if X has a Γ -norm,^{xvi} then $X \in \Gamma$, defined by $x \in X \leftrightarrow x \leq_{\varphi} x$. Nevertheless, the constant 0 function *does* tell us that PWO(Λ) holds whenever Λ is a σ -algebra like Δ_n^1 .

— 25C•7. Corollary -

Let a pointclass Λ be adequate and closed under complements. Therefore PWO(Λ). In particular, PWO($\Delta_n^1(X)$) for each $n < \omega$ and $X \subseteq \mathcal{N}$.

Proof .:.

Let $X \in \Lambda$ and consider $\varphi : X \to \{0\}$, the constant 0 map. Therefore $y \in X \to \varphi(x) = \varphi(y)$ for all $y \in X$. Hence $x \leq_{\varphi} y$ iff $x \in X$ which is in Λ (i.e. $\leq_{\varphi} = X \times \mathcal{N} \in \Lambda$). Similarly, $x <_{\varphi} y$ iff $x \in X \land (y \in X \to \varphi(x) < \varphi(y))$. But no $y \in X$ has $\varphi(x) < \varphi(y)$, meaning $x <_{\varphi} y$ iff $x \in X \land y \notin X$. Given the closure properties of Λ , this is in Λ .

We don't really care about σ -algebras though, because they admit such a trivial norm. Instead, we will consider pointclasses Γ with $\Gamma \neq \neg \Gamma$. To do this, we proceed similarly to Π_1^1 . An alternative characterization of Γ -norms is as follows, similar to comparing (coded) heights of elements of WO in Result 25 B • 8.

^{xvi}and Γ is closed under $A \mapsto \{x : \langle x, x \rangle \in A\}$

-

- 25 C • 8. Result

Let Γ be an adequate pointclass and $X \in \Gamma$. Therefore X has a Γ -norm iff there is a prewellordering \leq of X; a relation $\leq_0 \in \Gamma$; and a relation $\leq_1 \in \neg \Gamma$ where for all $y \in X$,

 $x \in X \land x \preccurlyeq y$ iff $x \preccurlyeq_0 y$ iff $x \preccurlyeq_1 y$.

Proof .:.

Suppose X has a Γ -norm φ . Define $x \preccurlyeq y$ iff $x, y \in X \land \varphi(x) \le \varphi(y)$. Therefore, for all $y \in X$, $x \in X \land x \preccurlyeq y$ iff $x \in X \land \varphi(x) \le \varphi(y)$ iff $x \le_{\varphi} y$, with $\le_{\varphi} \in \Gamma$. So take \le_0 to be \le_{φ} . As a total order, we also have that if $y \in X$, $x \in X \land x \preccurlyeq y$ iff $x \in X \land \varphi(x) \le \varphi(y)$ iff $x \in X \land \varphi(y) \ne \varphi(x)$ iff $y \ne_{\varphi} x$. So take $x \le_1 y$ iff $y \ne_{\varphi} x$ which is then in $\neg \Gamma$.

Alternatively, if there are relations \leq , \leq_0 , and \leq_1 as in the statement, by Corollary 25 C • 5, there is a norm $\varphi : X \to \text{Ord}$ which gives the prewellordering $x \leq y$ iff $\varphi(x) \leq \varphi(y)$. Now we'd like to define, say, $x \leq_{\varphi} y$ iff $x \in X \land (y \in X \to x \leq_0 y)$, but this has the wrong complexity. Instead, note that for $x \in X$, $y \not\leq_1 x$ iff $y \notin X \lor (y \in X \land \varphi(y) > \varphi(x))$. Hence we can use this in place of the conditional $y \in X \to x \leq_0 y$:

$$\begin{array}{ll} x \leq_{\varphi} y & \text{iff} \quad x \in X \land (x \leq_{0} y \lor y \not\leq_{1} x) \\ x <_{\varphi} y & \text{iff} \quad x \in X \land y \not\leq_{1} x. \end{array}$$

And it's not difficult to see that these are both in Γ .

In other words, for $x, y \in X, x \leq y$ is $\Gamma \cap \neg \Gamma$, meaning every initial segment of \leq is in $\Gamma \cap \neg \Gamma$. It's important to note, however, for $y \notin X, \leq_0$ and \leq_1 might differ significantly. The real statement is that for each $y \in X, \{x \in X : x \leq y\}$ is $\Gamma \cap \neg \Gamma$ although \leq itself might not be for one reason or another.^{xvii}

The pointclass that is easiest to show the prewellordering property for is $\Sigma_1^0 = \Sigma_0^1$.

- 25 C • 9. Result (PWO(Σ_0^1)) -For any $X \subseteq \mathcal{N}$, PWO($\Sigma_0^1(X)$)

Proof .:.

Work with $X = \emptyset$, as the proof easily generalizes. For $A \in \Sigma_0^1 = \Sigma_1^0$, $A = \bigcup_{n \in \omega} \mathcal{N}_{f(n)}$ for some computable $f : \omega \to {}^{<\omega}\omega$. So define $\varphi : A \to \text{Ord by}$

 $\varphi(x) = n$ iff *n* is the least such that $x \in \mathcal{N}_{f(n)}$.

It's not hard to see that φ is a Σ_0^1 -norm: each \mathcal{N}_{τ} is Δ_1^0 , and Σ_1^0 is closed under bounded quantification, intersections, and existential quantification over ω .

$$\begin{aligned} x \leq_{\varphi} y & \text{iff} \quad x \in A \land \exists n \in \omega \ \left(x \in \mathcal{N}_{f(n)} \land \forall m < n \ y \notin \mathcal{N}_{f(m)} \right) \\ x <_{\varphi} y & \text{iff} \quad x \in A \land \exists n \in \omega \ \left(x \in \mathcal{N}_{f(n)} \land \forall m \leq n \ y \notin \mathcal{N}_{f(m)} \right). \end{aligned}$$

A harder pointclass to show has the prewellordering property is Π_1^1 . Luckily, with the results from Subsection 25 B, we get it fairly easily because WO has a Π_1^1 -norm $x \mapsto ||x||$.

- 25C·10. Theorem (PWO(Π_1^1)) -

For any $X \subseteq \mathcal{N}$, $\Pi_1^1(X)$ has the prewellordering property.

Proof .:.

Work with $X = \emptyset$ as the proof easily generalizes. Let $A \in \Pi_1^1$ be arbitrary. By Lemma 25 B • 9, there's a computable $f : \mathcal{N} \to \mathcal{N}$ such that $A = f^{-1}$ "WO. WO has a Π_1^1 -norm by Result 25 B • 8 and the equivalence

^{xvii}For example, we know WO is Π_1^1 and $||x|| \le ||y||$ is Δ_1^1 for each $y \in$ WO. But we don't have the relation R(x, y) iff $||x|| \le ||y||$ as Δ_1^1 , however, as then $\mathfrak{p}R \in \Sigma_1^1$ has $\sup\{||x|| : x \in R\} = \omega_1$, contradicting The Boundedness Lemma (25 B • 10).

PROPERTIES OF THE LIGHTFACE HIERARCHIES

Result 25 C • 8. So using $\leq_{\varphi}, <_{\varphi} \in \Pi_1^1$ we get the norm $\varphi \circ f$ on A with $x \leq_{\varphi \circ f} y$ iff $f(x) \leq_{\varphi} f(y)$, and similarly for $<_{\varphi \circ f}$. Given that Π_1^1 is closed under computable substitutions, both of these are in Π_1^1 and so $\varphi \circ f$ is a Π_1^1 -norm on A.

This easily generalizes to Σ_2^1 by the following theorem of Moschovakis.

− 25 C • 11. Theorem Let Γ be an adequate pointclass. Suppose $X \in \Gamma$ has a Γ-norm. Therefore $\exists^{\mathcal{N}} X$ has a $\exists^{\mathcal{N}} \forall^{\mathcal{N}} \Gamma$ -norm. In particular, PWO(Γ) implies PWO($\exists^{\mathcal{N}} \forall^{\mathcal{N}} \Gamma$).

Proof .:.

Since X has a Γ -norm φ , we can define a norm $\psi : \exists^{\mathcal{N}} X \to \text{Ord by } \psi(x) = \min\{\varphi(x, y) : \langle x, y \rangle \in X\}$. It's then clear that this is a $\exists^{\mathcal{N}} \forall^{\mathcal{N}} \Gamma$ -norm:

 $\begin{array}{ll} x \leq_{\psi} y & \text{iff} & \exists y \in \mathcal{N} \ \forall z \in \mathcal{N} \ (\langle x, y \rangle \leq_{\varphi} \langle x, z \rangle) \\ x <_{\psi} y & \text{iff} & \exists y \in \mathcal{N} \ \forall z \in \mathcal{N} \ (y \neq z \rightarrow \langle x, y \rangle <_{\varphi} \langle x, z \rangle). \end{array}$

As a result of $\mathsf{PWO}(\Pi_1^1)$ (25 C • 10), $\mathsf{PWO}(\Pi_1^1)$ implies $\mathsf{PWO}(\Sigma_2^1)$.

```
- 25 C • 12. Corollary (PWO(\Sigma_2^1)) --
```

For any $X \subseteq \mathcal{N}$, $\mathsf{PWO}(\Sigma_2^1(X))$.

Note that although $\Sigma_1^1 \subseteq \Sigma_2^1$ and $\mathsf{PWO}(\Sigma_2^1)$, this does not tell us $\mathsf{PWO}(\Sigma_1^1)$ —i.e. that every Σ_1^1 -set has a Σ_1^1 -norm. It only tells us that every Σ_1^1 -set has a Σ_2^1 -norm. In fact, we can show $\neg \mathsf{PWO}(\Sigma_1^1)$, as we will see later. The following details what's known in ZFC. It turns out that Σ_2^1 seems to be the best we can do in ZFC, although not much is known. What is known is the independence of $\mathsf{PWO}(\Sigma_n^1)$ for odd $n \in \omega$ assuming the consistency of certain large cardinal hypotheses.



25 C • 13. Figure: Known analytical $\Gamma \neq \Delta_n^1$ such that ZFC \models PWO(Γ)

The pointclasses with the prewellordering property in L, form an initial zig-zag followed by a straight line. This shows $PWO(\Sigma_n^1)$ is at least consistent for odd $n \in \omega$ (in addition to even $n \in \omega$).



25 C • 14. Figure: Analytical $\Gamma \neq \Delta_n^1$ such that $L \models PWO(\Gamma)$

On the other hand, assuming projective determinacy continues the zig-zag pattern: $PWO(\Sigma_n^1)$ fails for odd $n \in \omega$ while holding for even n.

This shows the independence of $\mathsf{PWO}(\Sigma_n^1)$ for *odd* $n \in \omega$. The independence of $\mathsf{PWO}(\Sigma_n^1)$ for even n is less clear, but unpublished notes of Leo Harrington states that it's possible for $\neg \mathsf{PWO}(\Sigma_n^1)$ and $\neg \mathsf{PWO}(\Pi_n^1)$ for any particular n > 2.



25 C • 15. Figure: Analytical $\Gamma \neq \Delta_n^1$ such that ZFC + PD \models PWO(Γ)

We should now briefly consider the ideas of reduction and separation before introducing scales, which will tell us that at most one of Σ_n^1 and Π_n^1 have the prewellordering property.

§ 25 D. Reduction and separation

- 25D • 1. Definition A pointclass $\Gamma \subseteq \mathcal{P}(\mathcal{N})$ has the *reduction property* iff for every $X, Y \in \Gamma$, there are disjoint $X_0, Y_0 \in \Gamma$ such that $X_0 \subseteq X, Y_0 \subseteq Y$, and $X \cup Y = X_0 \sqcup Y_0$.

It's not difficult to show that the prewellordering property gives the reduction property.

1 I

Proof .:.

For $X, Y \in \Gamma$, consider the disjoint union $(X \times \{0\}) \cup (Y \times \{1\}) \in \Gamma$. By PWO(Γ), we have a Γ -norm φ on this set. So define

 $\begin{array}{ll} x \in X_0 & \text{iff} \quad x \in X \land \langle x, 0 \rangle \leq_{\varphi} \langle x, 1 \rangle \\ y \in Y_0 & \text{iff} \quad y \in Y \land \langle y, 1 \rangle <_{\varphi} \langle y, 0 \rangle. \end{array}$

As \leq_{φ} and $<_{\varphi}$ are both in Γ , this will be in Γ and clearly $X_0 \subseteq X$, $Y_0 \subseteq Y$, and $X_0 \cap Y_0 = \emptyset$. To show that $X_0 \cup Y_0 = X \cup Y$, if $x \in X \cap Y$, then $\varphi(x, 0) \leq \varphi(x, 1)$ implies $x \in X_0$ and otherwise $x \in Y_0$ so that $x \in X_0 \cup Y_0$ in either case. If $x \in X \setminus Y$, then $\langle x, 1 \rangle \notin (X \times \{0\}) \cup (Y \times \{1\})$ and hence $\langle x, 0 \rangle \leq_{\varphi} \langle x, 1 \rangle$ vacuously, meaning $x \in X_0$. A similar idea holds to show $Y \setminus X \subseteq Y_0$ and therefore $X \cup Y \subseteq X_0 \cup Y_0$. \dashv

The reduction property for an adequate pointclass Γ corresponds to a "separation" property for the dual class $\neg \Gamma$. Note the similarity with The $\sum_{i=1}^{1}$ -Separation Principle (22 C • 8).

- 25D·3. Definition A pointclass Γ has the *separation property* iff for any disjoint $X, Y \in \Gamma$, there's an $X' \in \Gamma \cap \neg \Gamma$ such that $X \subseteq X' \subseteq \mathcal{N} \setminus Y$.

The \sum_{1}^{1} -Separation Principle (22 C • 8) then says \sum_{1}^{1} has the separation property.

- 25 D • 4. Result - Let Γ be an adequate pointclass. Suppose Γ has the reduction property. Therefore $\neg \Gamma$ has the separation property.

Proof .:.

Let $X, Y \in \neg \Gamma$ such that $X \cap Y = \emptyset$. Therefore $(\mathcal{N} \setminus X) \cup (\mathcal{N} \setminus Y) = \mathcal{N}$ with $\mathcal{N} \setminus X, \mathcal{N} \setminus Y \in \Gamma$. By reduction for Γ , there are $X_0 \subseteq \mathcal{N} \setminus X$ and $Y_0 \subseteq \mathcal{N} \setminus Y$ such that $X_0, Y_0 \in \Gamma, X_0 \cup Y_0 = \mathcal{N}$, and $X_0 \cap Y_0 = \emptyset$. But then $\mathcal{N} \setminus X_0 = Y_0 \in \Gamma$ so that $X_0, Y_0 \in \Gamma \cap \neg \Gamma$. Moreover, because $X_0 \subseteq \mathcal{N} \setminus X$, it follows that $X' = \mathcal{N} \setminus X_0 \supseteq \mathcal{N} \setminus (\mathcal{N} \setminus X)$ so that $X \subseteq X'$ and similarly $Y \subseteq Y' = \mathcal{N} \setminus Y_0$. Since X' is disjoint from Y', X' is disjoint from Y: $X \subseteq X' \subseteq \mathcal{N} \setminus Y$.

Given that a pointclass can't have both the reduction property and the separation property, this tells us that at most one of Γ and $\neg \Gamma$ can have the prewellordering property and thus PWO(Π_1^1) implies \neg PWO(Σ_1^1).

– 25 D · 5. Result

Let Γ	be an adequate	pointclass with a I	-universal set.	Therefore Γ	and $\neg \Gamma$	cannot both have the reduction property.	
							1

Proof .:.

Let $U \in \Gamma$ be Γ -universal. Note that we identify x = even(x) * odd(x). Consider

$$X = \{x \in \mathcal{N} : \langle \operatorname{even}(x), x \rangle \in U\}$$

 $Y = \{ y \in \mathcal{N} : \langle \text{odd}(x), x \rangle \in U \}.$

It follows that both $X, Y \in \Gamma$ since $x \mapsto \text{even}(x)$ is computable. By the reduction property for Γ , there are disjoint $X_0 \subseteq X, Y_0 \subseteq Y$ in Γ such that $X \cup Y = X_0 \sqcup Y_0$.

Since $\neg \Gamma$ has the reduction property, Result 25 D • 4 tells us $\neg \neg \Gamma = \Gamma$ has the separation property. So let $X' \in \Gamma \cap \neg \Gamma$ be such that $X_0 \subseteq X' \subseteq \mathcal{N} \setminus Y$. Write $X' = U_r$ and $\mathcal{N} \setminus X' = U_s$ for some $r, s \in \mathcal{N}$ and consider s * r.

• If $s * r \in X' = U_r$ then $s * r \in Y$ by definition, contradicting that $X' \cap Y = \emptyset$.

• If $s * r \notin X'$ then $s * r \in \mathcal{N} \setminus X' = U_s$ implying by definition that $s * r \in X \subseteq X'$, a contradiction. \dashv

25 D•6. Corollary –

For every $X \subseteq \mathcal{N}$, $\Pi_1^0(X)$, $\Sigma_1^1(X)$, and $\Pi_2^1(X)$ do not have the prewellordering property.

Note that every σ -algebra, like any Δ_n^1 , has both the reduction property and the separation property by Result 25 D • 2 and Corollary 25 C • 7. This has the nice side effect that each Δ_n^1 has no Δ_n^1 -universal set, although this is easy enough to prove on its own just by a simple diagonalization argument.

– 25 D•7. Result

Let $\Lambda \subseteq \mathcal{P}(\mathcal{N})$ be adequate and closed under complements. Therefore Λ has the reduction and separation property. In fact, $\mathsf{PWO}(\Lambda)$ holds.

Proof .:.

This of course follows from earlier results, but we can show these both directly as well. For the reduction property, $X, Y \in \Lambda$ implies $X \setminus Y \in \Lambda$. So take $X_0 = X \setminus Y$ and $Y_0 = Y$ which are disjoint subsets in Λ whose union is $X \cup Y$. The separation property is trivial: for any disjoint $X, Y \in \Lambda$, take X' = X because a σ -algebra Λ satisfies $\neg \Lambda \cap \Lambda = \Lambda$.

– 25 D•8. Corollary –

For each $n < \omega$ and $X \subseteq \mathcal{N}$, $\Delta_n^1(X)$ has no $\Delta_n^1(X)$ -universal set.

Proof .:.

This follows from Result 25 D • 5 and Result 25 D • 7: $\Delta_n^1 = \neg \Delta_n^1$ is an adequate pointclass with the reduction property. A more standard proof is still provided. Let $U \in \Delta_n^1$ be universal. Note that $D = \{x \in \mathcal{N} : \langle x, x \rangle \notin U\} \in \neg \Delta_n^1 = \Delta_n^1$. Thus $D = U_r$ for some $r \in \mathcal{N}$, but $r \in D$ iff $\langle r, r \rangle \notin U$ iff $r \notin U_r = D$, a contradiction. \dashv

Most of this is just to say that we shouldn't be thinking of the Δ_n^0 s or Δ_n^1 s when investigating these properties: they almost trivially have them. The real work to be done is with the other (analytical) pointclasses.

§ 25 E. Scales and Uniformization

Let's return to the study of uniformization and norms. The main idea behind this is a certain *sequence* of norms that work nicely with sequences.

- 25 E • 1. Definition

Let $X \subseteq \mathcal{N}$. A scale on X is a sequence $\vec{\varphi} = \langle \varphi_n : n \in \omega \rangle$ such that

- each φ_n is a norm on X;
- for all convergent $x \in {}^{\omega}X$ such that each $\varphi_n \circ x$ is eventually constant, we have
 - $-\lim x \in X$; and
 - (lower semi-continuity) for all $n < \omega$, $\varphi_n(\lim x) \le \lim(\varphi_n \circ x)$.

If $\vec{\varphi}$ satisfies all of the above except lower semi-continuity, $\vec{\varphi}$ is called a *semi-scale*.

As with norms, we are more interested in scales^{xviii} with certain definability restrictions: there are easily scales on any X, but we want to be able to use scales in arguments about pointclasses.

– 25E•2. Corollary ———

Let $X \subseteq \mathcal{N}$. Therefore there is a scale on X.

Proof .:.

For any bijection $f : X \to |X|$, set $\varphi_n = f$ for every $n \in \omega$. Clearly each φ_n is a norm. If $x \in {}^{\omega}X$ is convergent and $f \circ x$ is eventually constant, then x is eventually constant and so $\lim x \in \lim x \subseteq X$. Lower semi-continuity also is easy since it follows that $f(\lim x) = \lim (f \circ x)$.

This is actually a result of $X \subseteq \mathcal{N}$ being |X|-suslin, as the next theorem shows.

- 25 E · 3. Result

For $X \subseteq \mathcal{N}$ and κ an infinite cardinal, the following are equivalent:

1. *X* has a scale $\vec{\varphi}$ where $\varphi_n(x) < \kappa$ for all $n < \omega, x \in X$.

- 2. X has a semi-scale $\vec{\varphi}$ where $\varphi_n(x) < \kappa$ for all $n < \omega, x \in X$.
- 3. X is κ -suslin.

Proof .:.

 $(1) \rightarrow (2)$ Trivial.

 $(2) \rightarrow (3)$ Consider the tree building up sequences of elements in X and their corresponding norms:

 $T = \{ \langle \tau, \rho \rangle \in {}^{<\omega}\omega \times {}^{<\omega}\kappa : \exists x \in X \ (\tau \lhd x \land \rho = \langle \varphi_n(x) : n < \mathrm{lh}(\tau) \rangle) \}.$

It's not difficult to see that this is a tree over $\omega \times \kappa$. Moreover, it's clear $X \subseteq \mathfrak{p}[T]$ since $\langle x, \langle \varphi_n(x) : n < \omega \rangle \rangle \in [T]$. To see that $\mathfrak{p}[T] \subseteq X$, if $x \in \mathfrak{p}[T]$ then for each $n < \omega$, we get $x_n \in X$ with $x \upharpoonright n = x_n \upharpoonright n$ witnessing $x \upharpoonright n \in \mathfrak{p}T$. As they extend each other, $\varphi_k(x_k) = \varphi_k(x_n)$ for every k < n. In particular, $\langle \varphi_n(x_k) : k < \omega \rangle$ is eventually constant for each $n < \omega$ and so as a semi-scale, $\lim x_n = x \in X$.

$$(3) \rightarrow (1)$$
 Let $X = \mathfrak{p}[T]$ where T is a tree over $\omega \times \kappa$. Assume $\operatorname{cof}(\kappa) > \omega$. For $\alpha < \kappa$, write

$$T \upharpoonright \alpha = \{ \langle \tau, \rho \rangle \in T : \rho \in {}^{<\omega} \alpha \}.$$

So when we're building x up as a branch of T, for each $n < \omega$, we have $x \upharpoonright n \in \text{dom}(T \upharpoonright \alpha_n)$ for some $\alpha_n < \kappa$. And this means $\langle x, \sup_{n < \omega} \alpha_n \rangle \in [T]$. In other words, $x \in \mathfrak{p}[T \upharpoonright \alpha_x]$ for some $\alpha_x < \kappa$. So we can order $x \in X$ according to this α_x . In particular, we may order the finite initial segments of the least *branch* lexicographically to get a norm for each $n < \omega$. In particular, let α_x be least such that $x \in \mathfrak{p}[T \upharpoonright \alpha_x]$. Of this, let $s_x \in \omega \alpha_x$ be lexicographically-least such that $\langle x, s_x \rangle \in [T \upharpoonright \alpha_x]$. To keep track of this bound α_x (to ensure lower-semicontinuity), define for $x \in X$,

$$\varphi_n(x) = \operatorname{rank}(\langle \alpha_x \rangle \widehat{\ } s_x \upharpoonright n) \le \alpha_x^n < \kappa,$$

where rank(τ) is the lexicographic rank of τ in the set { $\alpha \cap \sigma \in {}^{n+1}\kappa : im \sigma \subseteq \alpha$ }. It follows that

^{xviii}There's also a separate, more *combinatorial* notion of a scale on a cardinal κ . We will not be interested in this idea, and will only consider scales in the descriptive set theoretic sense.

 $\varphi_n(x) < \kappa$ for every $n < \omega$ and $x \in X$.

 $\langle \varphi_n : n < \omega \rangle$ is also a scale on X. To see this, suppose $\langle x_k \in X : k \in \omega \rangle$ converges to $x \in \mathcal{N}$, with $\langle \varphi_n(x_k) : k \in \omega \rangle$ eventually a constant λ_n for each $n \in \omega$. For sufficiently large k, $\lambda_n = \operatorname{rank}(\langle \alpha_{x_k} \rangle \widehat{\ } s_{x_k} \upharpoonright n)$ and $x_k \upharpoonright n = x \upharpoonright n$. But the rank function is a bijection because the lexicographic order is linear. In other words, for sufficiently large k, $s_{x_k} \upharpoonright n$ is a sequence s_n and α_{x_k} is an ordinal α . Since $s_{x_k} \upharpoonright n \lhd s_{x_k} \upharpoonright m$ for n < m, by looking at even larger k, it follows that $s_n \lhd s_m$ for n < m. In particular, for $s = \bigcup_{n < \omega} s_n \in {}^{\omega}\kappa, \langle x \upharpoonright n, s \upharpoonright n \rangle = \langle x_k \upharpoonright n, s_{x_k} \upharpoonright n \rangle \in T \upharpoonright \alpha_k = T \upharpoonright \alpha$ for sufficiently large k. As a result, $\langle x, s \rangle \in [T \upharpoonright \alpha]$ and hence $x \in \mathfrak{p}[T] = X$. This tells us φ is a semisscale. Lower-semicontinuity is immediate because we kept track of α : any branch $\langle x, t \rangle \in [T \upharpoonright \beta]$ with $\beta < \alpha$ has $\langle \beta \rangle \cap t \upharpoonright n$ lexicographically precede $\langle \alpha \rangle \cap s \upharpoonright n$ whose rank is λ_n .

As a result, suslin representations of sets are more-or-less the same as scales on these sets. Of course, this relies on AC to ensure there's an injection from X into Ord. But there are less *ad hoc* examples of scales. This then motivates the idea of considering κ -suslin sets and scales with certain definability restrictions. In particular, we have the following scale generalized from the usual norm on WO.

There is a scale $\vec{\varphi}$ on WO. Moreover, the following relations on triples $\langle x, y, n \rangle \in \mathcal{N}^2 \times \omega$ are both in Π_1^1 : $x \leq_{\varphi_n} y$ iff $x \in WO \land (y \in WO \rightarrow \varphi_n(x) \leq \varphi_n(y))$ $x <_{\varphi_n} y$ iff $x \in WO \land (y \in WO \rightarrow \varphi_n(x) < \varphi_n(y)).$

Proof .:.

For $x \in WO$, let $E_x \subseteq \omega^2$ is the well-order coded by x. For $n < \omega$, let

 $(E_x)_{< n} = \{ \langle k_0, k_1 \rangle \in E_x : k_1 \ E_x \ n \neq k_1 \},\$

basically the initial segment of E_x that precedes n (and $(E_x)_{< n} = \emptyset$ if $n \notin \text{dom}(E_x) \cup \text{ran}(E_x)$). Note that each $(E_x)_{< n}$ itself is a well-order. So define $\varphi_n : WO \to \omega_1$ by

 $\varphi_n(x) = \operatorname{code}_{\operatorname{lex}}(||E_x||, ||(E_x)_{< n}||),$

where $\operatorname{code}_{\operatorname{lex}} : \omega_1 \times \omega_1 \to \omega_1 \cdot \omega_1$ is the lexicographic rank function. This allows us to both keep track of the original height as well as *how* this height is built up as *n* increases. Depending on the order, $\varphi_0(x)$ might still be infinite, but we always have $\varphi_n(x) \leq \varphi_m(x)$ for $n \leq m$.

To see that $\langle \varphi_n : n \in \omega \rangle$ is a scale on WO, clearly each φ_n is a norm. So suppose $\langle x_n \in WO : n \in \omega \rangle$ be a sequence converging to $x \in \mathcal{N}$ such that for each $n < \omega$, $\langle \varphi_n(x_k) : k \in \omega \rangle$ is eventually a constant value $\operatorname{code}(\lambda, \alpha_n) \in \omega_1$. Note that λ doesn't depend on n since $||E_{x_k}||$ doesn't depend on n. Clearly $\langle ||(E_{x_k})_{< n}|| : k < \omega \rangle$ has eventually constant value α_n .

- To show $x \in WO$, we just need to construct an order-preserving map from $\langle \omega, E_x \rangle$ to $\langle \omega_1, \epsilon \rangle$ (where recall $E_x = \{\langle n, m \rangle \in \omega^2 : x(\operatorname{code}(n, m)) = 1\}$). The map we consider will be $n \mapsto \alpha_n$. If $n \in E_x m$, then $x(\operatorname{code}(n, m)) = 1$. But as the limit, we can determine the initial segment of x up to $\max(\operatorname{code}(m, n), \operatorname{code}(n, m)) < \omega$ so that $x_k(\operatorname{code}(n, m)) = 1$ and (as a well-order) $x_k(\operatorname{code}(m, n)) = 0$ for sufficiently large k. This means if $n \in E_x m$ then for sufficiently large k, $(E_{x_k})_{<n} \subseteq (E_{x_k})_{<m} \ni n$ so in fact $\|(E_{x_k})_{<n}\| < \|(E_{x_k})_{<m}\|$, i.e. $\alpha_n < \alpha_m$. Hence $n \mapsto \alpha_n$ witnesses that E_x is a well-order.
- To show $\varphi_n(x) \leq \operatorname{code}(\lambda, \alpha_n)$ for each $n < \omega$, the same idea above tells us for every $n < \omega$, $(E_x)_{<n} \subseteq (E_{x_k})_{<n}$ for sufficiently large k. This means $||(E_x)_{<n}|| \leq ||(E_{x_k})_{<n}|| = \alpha_n$ and as the supremum of these, over n, $||E_x|| \leq \lambda$. Thus $\varphi_n(x) = \operatorname{code}(||E_x||, ||(E_x)_{<n}||) \leq \operatorname{code}(\lambda, \alpha_n) = \lim_{k \to \infty} \varphi_n(x_k)$, as desired.

It's not difficult to define show the relations \leq_{φ_n} and $<_{\varphi_n}$ are Π_1^1 from PWO(Π_1^1) (25 C • 10). We can easily find a map inputting x and n and outputting a real coding $(E_x)_{< n}$ as follows:

$$f(x,n)(\operatorname{code}(k_0,k_1)) = \begin{cases} 1 & \text{if } x(\operatorname{code}(k_0,k_1)) = 1 \land x(\operatorname{code}(k_1,n)) = 1 \land k_1 \neq n \\ 0 & \text{otherwise.} \end{cases}$$

This is x-computable in a uniform way so that $f : \mathcal{N} \times \omega \to \mathcal{N}$ is computable. Using the the Π_1^1 -norm

 $\phi : WO \to \omega_1$ defined by $\phi(x) = ||x||$ and the Π_1^1 -relations \leq_{ϕ} and $<_{\phi}$, because Π_1^1 is closed under computable substitution, $f(x, n) \leq_{\phi} f(y, n)$ and $f(x, n) <_{\phi} f(y, n)$ are both Π_1^1 . Thus the following are also Π_1^1 :

•
$$x \leq_{\varphi_n} y$$
 iff $x \leq_{\phi} y$ or $(x \leq_{\phi} y \land y \leq_{\phi} x \land f(x, n) \leq_{\phi} f(y, n));$
• $x <_{\varphi_n} y$ iff $x <_{\phi} y$ or $(x \leq_{\phi} y \land y \leq_{\phi} x \land f(x, n) <_{\phi} f(y, n)).$

This motivates the idea of a Π_1^1 -scale just as the map $x \mapsto ||x||$ motivated the idea of a Π_1^1 -norm.

- 25E.5. Definition Let Γ be a pointclass. A Γ -scale is a scale $\vec{\varphi}$ on a set X such that the following relations on triples $\langle x, y, n \rangle$ are in Γ : $x \leq_{\varphi_n} y$ iff $x \in X \land (y \in X \to \varphi_n(x) \leq \varphi_n(y))$ $x <_{\varphi_n} y$ iff $x \in X \land (y \in X \to \varphi_n(x) < \varphi_n(y))$. Γ has the scale property iff every $X \in \Gamma$ has a Γ -scale on X.

So Result 25 E • 4 says WO has a Π_1^1 -scale. As is usual, for adequate Γ , a Γ -scale on X implies $X \in \Gamma$, defined by $x \in X$ iff $x \leq_{\varphi_0} x$. Similarly, it's easy to adapt Result 25 C • 8 into the following corollary.

- 25E·6. Corollary — Let Γ be an adequate pointclass and $X \in \Gamma$. Therefore X has a Γ -scale iff there are relations $S_0 \in \Gamma$; and $S_1 \in \neg \Gamma$, where for all $y \in X$,

 $x \in X \land \varphi_n(x) \le \varphi_n(y)$ iff $S_0(x, y, n)$ iff $S_1(x, y, n)$.

We can also easily use the Π_1^1 -scale on WO to show that Π_1^1 has the scale property.

- 25 E • 7. Corollary (Π_1^1 Scale Property) - For any $X \subseteq \mathcal{N}$, $\Pi_1^1(X)$ has the scale property.

Proof .:.

Work with $X = \emptyset$. As with $\mathsf{PWO}(\Pi_1^1)$ (25 C • 10), for $A \in \Pi_1^1$, Lemma 25 B • 9 tells us there's a computable (and therefore continuous) $f : \mathcal{N} \to \mathcal{N}$ where f^{-1} "WO = A. Result 25 E • 4 tells us there's a Π_1^1 -scale $\vec{\varphi}$ on WO, and it's easy to see from Π_1^1 's closure under computable preimages that $\langle \varphi_n \circ f : n < \omega \rangle$ is a Π_1^1 -scale on A.

The scale property is instrumental in proving Π_1^1 -uniformization—sometimes called Kondô's theorem—which we may now prove. As a reminder, this is the statement that for any $X \in \Pi_1^1$ with $X \subseteq \mathcal{N} \times \mathcal{N}$, there's a Π_1^1 -function $f : \operatorname{dom}(X) \to \operatorname{ran}(X)$.

- 25E • 8. Theorem (Π_1^1 -Uniformization) - For any $X \subseteq \mathcal{N}$, $\Pi_1^1(X)$ -uniformization holds.

Proof .:.

Work with $X = \emptyset$ for simplicity. Let $A \in \Pi_1^1$. For each $x \in \mathcal{N}$, let $A^*x = \{y \in \mathcal{N} : \langle x, y \rangle \in A\}$. The general strategy will find a unique $y \in A^*x$. Uniqueness is relatively easy to establish. The purpose of the scale is to ensure f is Π_1^1 and that dom $(f) = \mathfrak{p}A$.

So let $x \in \mathcal{N}$ be arbitrary. Let $\vec{\varphi}$ be a Π_1^1 -scale on A by Π_1^1 Scale Property (25 E • 7). Define f(x) = y iff $\langle x, y \rangle \in A$ and for all $z \in \mathcal{N}$, and all $n \in \omega$,

1. If $\langle x, z \rangle \in A \land z \upharpoonright n = y \upharpoonright n \land \forall m < n \ (\varphi_m(x, z) = \varphi_m(x, y)) \ (\Sigma_1^1 \text{ using } S_1 \text{ from Corollary } 25 \text{ E} \bullet 6);$ 2. Then y(n) < z(n) or else $y(n) = z(n) \land \varphi_n(x, y) \le \varphi_n(x, z) \ (\Pi_1^1)$

It's not too difficult to show that this indeed uniquely defines a $y \in A^n x$ because if this holds for two $y_1, y_2 \in \mathcal{N}$, then for the least *n* where $y_1(n) \neq y_2(n)$, (1) holds. Suppose $y_1(n) < y_2(n)$, then if we look at (2) applied to $y = y_2$ and $z = y_1$, then $y_2(n) < y_1(n)$ (which is false) or else $y_2(n) = y_1(n)$ (which is also false).

PROPERTIES OF THE LIGHTFACE HIERARCHIES

As a result, f is $\neg \Sigma_1^1 \lor \Pi_1^1 = \Pi_1^1$ with $f \subseteq A$. So it suffices to show the *existence* of such a y for any given $x \in pA$. We use that $\vec{\varphi}$ is a scale and use this to ensure the limit of constructing y has $\langle x, y \rangle \in A$. In particular, recursively define A_n such that $A_0 = A^n x$, and

$$A_0 = A \ x$$

$$A_{2n+1} = \{ y \in A_n : \varphi_n(x, y) = \min\{\varphi_n(x, z) : z \in A_n \} \}$$

$$A_{2n+2} = \{ y \in A_{n+1} : y(n) = \min\{z(n) : z \in A_n \} \}.$$

Any sequence of elements $\langle y_n \in A_n : n \in \omega \rangle$ is necessarily convergent because $y_m \upharpoonright n = y_n \upharpoonright n$ for n < m. Such a sequence also has $\langle \varphi_n(y_k) : k \in \omega \rangle$ as eventually constant and therefore $\lim_{n \to \infty} y_n \in \bigcap_{n < \omega} A_n$, i.e. $y = \lim_{n \to \infty} y_n$ satisfies the definition f(x) = y.

One can generalize this proof to show that if Γ is adequate, closed under \forall^{ω} , and has the scale property, then Γ -uniformization holds. Π_1^1 -uniformization in particular allows us to prove the converses of Theorem 25 B • 16.

- 25 E • 9. Theorem -For $x \in \mathcal{N}$, the following are equivalent:

1. $\omega_1^{L[x]} < \omega_1$.

- 2. $\Sigma_2^1(x)$ has the perfect set property.
- 3. $\Pi_1^1(x)$ has the perfect set property.

Proof .:.

(1) implying (2) is in the proof of Theorem 25 B • 16. (2) implying (3) is trivial, so we need to show (3) implies (1). Let $x \in \mathcal{N}$ and suppose $\omega_1^{L[x]} = \omega_1$. Define the set $X \subseteq \mathcal{N}$ as in Result 25 B • 13, modified to work with L[x]. In particular, for each $\alpha < \omega_1^{L[x]}$, we let $f(\alpha)$ be the $<_{L[x]}$ -least real in WO $\cap L[x]$ coding $\langle \alpha, \in \rangle$. Set $X = f'' \omega_1^{L[x]}$.

All the relevant theorems generalize from L to L[x] to show that;

a. $X \in \Sigma_2^1(x)$ is uncountable (since $\aleph_1 = \aleph_1^{L[x]}$); and

b. X has no uncountable $\sum_{i=1}^{1}$ -subset.

(b) follows from The Boundedness Lemma (25 B • 10): any two elements of X have different height and because any $\sum_{n=1}^{1} 1$ -subset must have height $< \omega_1^{L[x]} = \omega_1$, it must have cardinality $< \aleph_1^{L[x]} = \aleph_1$.

So let $X = \mathfrak{p}Y$ for $Y \in \Pi_1^1(x)$. By Π_1^1 -Uniformization (25 E • 8), there's a $\Pi_1^1(x)$ -function $f \subseteq Y$ with dom(f) = X. Thus $|f| = |X| = \aleph_1$, and f also has no perfect subset. To see this, any closed $g \subseteq f$ has dom $(g) \in \Sigma_1^1$ and dom $(g) \subseteq \text{dom}(f) = X$ so $|g| = |\text{dom}(g)| = \aleph_0$ by (b). \dashv

The uniformization and scale properties also hold for Σ_2^1 , which can be proven just by modifying the proofs for Π_1^1 , but not much is known beyond this because clearly the scale property (used to show uniformity) implies the prewellordering property, and this is often independent.

§ 25 F. Lengths of Definable Prewellorders

The Boundedness Lemma (25 B • 10) tells us that no $\sum_{i=1}^{1}$ -set of (coded) ordinals can reach above ω_1 . But what if instead of a set of ordinals, we consider a single well-order regarded as a relation over \mathcal{N} . Forgetting about definability restrictions, we could reach all the way up to (but not including) $|\mathcal{N}|^+$, which is consistently very large. Rather than work with well-orders, it will be simpler to work with prewellorders.

- 25 F • 1. Definition
Let
$$n < \omega$$
. Define the *projective ordinal*
 $\underline{\delta}_n^1 = \sup\{||R|| : R \in \underline{\Lambda}_n^1 \text{ is a prewellorder}\}.$

We also define the lightface $\delta_n^1(X)$ as the supremum of the (heights of) $\Delta_n^1(X)$ -prewellorders for $X \subseteq \mathcal{N}$.

Note that if $\delta_n^1 = \kappa$, then every Δ_n^1 -prewellorder has height *strictly* less than κ , because otherwise we could append a point at the end to have a larger height without introducing complexity.

For
$$X \subseteq \mathcal{N}$$
, $n < \omega$, if $R \in \Delta_n^1(X)$ is a prewellorder, then $||R|| < \delta_n^1(X)$.

Proof .:.

For $x \in \mathcal{N}$, write $x' = \{\langle n+1, m \rangle : \langle n, m \rangle \in x\}$. If $||R|| = \delta_n^1(X)$ with $R \in \Delta_n^1(X)$, then define $R' = \{\langle x * 0, y * 0 \rangle : \langle x, y \rangle \in R\} \cup \{\langle x * 0, \operatorname{const}_1 \rangle : x \in \operatorname{dom}(R) \cup \operatorname{ran}(R)\}$. This will also be a $\Delta_n^1(X)$ -prewellorder of height $||R'|| = ||R|| + 1 > \delta_n^1(X)$, a contradiction.

Clearly if $any |\delta_n^1(X)| > \aleph_1$, then $\neg CH$. In particular, the following theorem tells us $\mathbf{L} \models "\mathfrak{S}_1^1 < \mathfrak{S}_2^1 = \mathfrak{S}_n^1 = \aleph_2$ " for every n > 1. So we in general cannot proof what precisely these ordinals are. At best, we can place bounds on them. That being said, under certain assumptions, we *can* calculate these.

- **25 F • 3.** Theorem $\delta_1^1 = \omega_1$ while $\delta_1^1 = \omega_1^{CK}$.

Proof .:.

For any infinite $\alpha < \omega_1$, we have a relation on ω coding it: $E = \{\langle n, m \rangle : f(n) \in f(m)\}$ where $f : \omega \to \alpha$ is a bijection. We can translate this to \mathcal{N} just by considering $R = \{\langle x, y \rangle : \langle x(0), y(0) \rangle \in E\} \in \Sigma_1^1(E)$. This shows $\underline{\delta}_1^1 \ge \omega_1$. Assuming $\alpha < \omega_1^{CK}$, then E is computable so that $\Sigma_1^1(E) = \Sigma_1^1$ and therefore $\delta_1^1 \ge \omega_1^{CK}$.

So suppose $R \in \sum_{1}^{1}$ is a prewellorder of height $\geq \omega_{1}$. For every $\alpha < \omega_{1}$, there's therefore an order-preserving map $f : \alpha \to \mathcal{N}$, i.e. if $\gamma < \beta < \alpha$ then $f(\gamma) R f(\beta)$. Using this, we can define WO in a $\Sigma_{1}^{1}(R)$ -way: $x \in WO$ iff x codes a linear order and there's such an order-preserving map from E_{x} to R (i.e. there is a sequence of reals $\langle x_{n} : n \in \omega \rangle$ where for all $n, m \in \omega$, $\langle n, m \rangle \in E_{x}$ implies $\langle x_{n}, x_{m} \rangle \in R$). But The Boundedness Lemma (25 B • 10) tells us WO isn't \sum_{1}^{1} , a contradiction. The same idea applies to ω_{1}^{CK} in place of ω_{1} , just noting that $R \in \Sigma_{1}^{1}$ implies $\Sigma_{1}^{1}(R) = \Sigma_{1}^{1}$.

- 25 F • 4. Corollary CH implies $\delta_1^1 = \aleph_1$, and $\delta_n^1 \le \aleph_2$ for every $1 < n < \omega$.

Proof .:.

Work in L. Theorem 25 F • 3 gives the first equality. For the second, if $\underline{\delta}_2^1 > \aleph_2$, then there's a \aleph_2 -length prewellordering \leq of \mathcal{N} so after modding out by the equivalence relation $x \approx y \leftrightarrow x \leq y \leq x$, we get a wellordering \leq/\approx still of length \aleph_2 , contradicting that $|\mathcal{N}| = \aleph_1$ by CH. Thus $\underline{\delta}_n^1 \leq \aleph_2$ for every $n < \omega$. \dashv

The following theorem, proven independently by Kenneth Kunen and Donald Martin, gives us another way of working with these projective ordinals.

- 25 F • 5. Theorem (Kunen-Martin) If $R \subseteq \mathcal{N}^2$ is a prewellordering and R is κ -suslin, then $||R|| < \kappa^+$.

Proof .:.

Consider the tree *S* of *R*-decreasing sequences in ${}^{<\omega}\mathcal{N}$. Note that because *R* is well-founded, *S* has no infinite branches. Thus we can consider a rank function on the upside down version of $\langle S, \rhd \rangle$ which then means the rank of $\vec{x} \land \langle y \rangle \in S$ is determined by what is above this in *S*, i.e. the *R*-predecessors of *y*. In particular, by a simple induction, rank^{*R*}(*y*) = rank^(S, \triangleright)($\vec{x} \land \langle y \rangle$) whenever $\vec{x} \land \langle y \rangle \in S$. (In particular, $||R|| = \operatorname{rank}^{(S, \triangleright)}(\emptyset) = ||\langle S, \rhd \rangle$

 $\rangle \|.)$

By Result 25 E • 3, let $\vec{\varphi}$ be a scale on R^{-1} with $\varphi_n(x, y) < \kappa$ for every $n < \omega$ and $\langle y, x \rangle \in R$. (We use R^{-1} because finite R^{-1} -increasing sequences are finite R-decreasing sequences and therefore are in S.) For y R x, define

$$\psi_n(x, y) = \operatorname{code}(\langle x(i), y(i), \varphi_i(x, y) : i < n \rangle),$$

where $code(\tau)$ is the lexicographic rank of τ in ${}^{<\omega}(\omega \times \omega \times \kappa)$.

If $\{\vec{v}_n : n < \omega\} \subseteq R^{-1}$, and each $\langle \psi_n(\vec{v}_i) : i < \omega \rangle$ is eventually constant, then $\langle \vec{v}_n : n < \omega \rangle$ converges to some $\vec{v} \in R^{-1}$.

Proof .:.

- Claim 1

If $\langle \psi_n(\vec{v}_i) : i < \omega \rangle$ is eventually constant, then $\langle \varphi_k(\vec{v}_i) : k < \omega \rangle$ is eventually constant and because $\psi_n(\vec{v}_i)$ encodes (and therefore solidifies) more and more of the initial values of \vec{v}_i for larger and larger i, $\langle \vec{v}_n : n < \omega \rangle$ converges to some $\vec{v} \in \mathcal{N}^2$. But $\vec{\varphi}$ being a scale on \mathbb{R}^{-1} then ensures $\vec{v} \in \mathbb{R}^{-1}$.

Now consider the function $f: S \to {}^{<\omega}$ Ord as follows: $f(\emptyset) = \emptyset$, $f(\langle x \rangle) = \emptyset$, and for $\vec{x} = \langle x_0, \cdots, x_{n-1} \rangle \in S$, and $\vec{x} \cap \langle x_n \rangle \in S$,

$$f(\vec{x}^{\wedge}\langle x_n \rangle) = f(\vec{x})^{\wedge} \langle \psi_{n-1}(x_i, x_{i+1}) : i < n \rangle^{\wedge} \langle \psi_{n-2}(x_i, x_{i+1}) : i < n \rangle^{\wedge} \cdots^{\wedge} \langle \psi_0(x_i, x_{i+1}) : i < n \rangle.$$

This way, every $\psi_k(x_i, x_{i+1})$ for i, k < n is coded by $f(\langle x_0, \dots, x_n \rangle)$. The point of this construction is the following two properties we will prove:

1. There is some $\lambda < \kappa^+$ where $f : S \to {}^{<\omega}\lambda$.

- 2. *f* is order-preserving in the sense that $\sigma \triangleleft \tau$ with $\tau, \sigma \in S$ (of length at least 1) implies $f(\sigma) \triangleleft f(\tau)$.
- 3. $(\operatorname{im} f, \rhd)$ is well-founded with therefore $||R|| = ||\langle S, \rhd \rangle|| \le ||\langle \operatorname{im} f, \rhd \rangle|| < \kappa^+$.

(1) is easy enough to see just because each $\varphi_n(x, y) < \kappa$ so that $\psi_n(x, y) < (\omega \cdot \omega \cdot \kappa)^{<\omega}$ (i.e. ordinal exponentiation $(\omega \cdot \omega \cdot \kappa)^{\omega}$) and thus we can regard $f(\tau) < ((\omega \cdot \omega \cdot \kappa)^{<\omega})^{<\omega} < \kappa^+$ for any $\tau \in S$. (2) also isn't difficult to see by the inductive definition of f.

For (3), we first show $\langle \text{im } f, \rhd \rangle$ is well-founded. Suppose not: let $\langle \tau_n \in S : n < \omega \rangle$ yield $\langle f(\tau_n) : n < \omega \rangle$ as an infinite \triangleright -decreasing sequence where (without loss of generality) $\tau_n = \langle \tau_n(k) : k \le n \rangle$ has length n + 1and therefore as an element of S, $\tau_n(k) R^{-1} \tau_n(j)$ for $k < j \le n$. Now because $f(\tau_0) \triangleright f(\tau_1)$, all the coded information from ψ_0 is retained in $f(\tau_1)$. And more generally, the information from ψ_n is retained in $f(\tau_k)$ for k > n. In particular, for any given $n < \omega$, $\langle \psi_n(\tau_i(0), \tau_i(1)) : 1 \le i < \omega \rangle$ is constant. Generalizing this, for any given $k < n < \omega$, $\langle \psi_n(\tau_i(k), \tau_i(k+1)) : k \le i < \omega \rangle$ is constant. But by Claim 1, this means for each $k < \omega$, $\langle \langle \tau_i(k), \tau_i(k+1) \rangle : k \le i < \omega \rangle$ converges to some $\langle \tau(k), \tau(k+1) \rangle \in R^{-1}$. Thus we have an infinite R^{-1} -increasing sequence, contradicting that R is well-founded.

Note that this actually provides another proof that $\delta_1^1 = \aleph_1$ since clearly $\delta_1^1 \ge \aleph_1$, and the fact that $\Delta_1^1 \subseteq \sum_{1}^1$ -sets are \aleph_0 -suslin implies their prewellorders have height $< \delta_1^1 \le \aleph_1$. These two consequences of Kunen–Martin (25 F • 5)—that $\delta_1^1 = \aleph_1$ and $\delta_2^1 \le \aleph_2$ —are more-or-less the only results known in ZFC.

- 25 F · 6. Corollary - $\delta_{22}^1 \leq \omega_2$.

Proof .:.

All \sum_{2}^{1} -sets (and hence Δ_{2}^{1} -sets) are \aleph_{1} -suslin by the proof of Shoenfield Absoluteness (25 A • 8) namely, Corollary 25 A • 9. So by Kunen–Martin (25 F • 5), any $\Delta_{2}^{1} \subseteq \sum_{2}^{1}$ -prewellorder has height at most $\aleph_{1}^{+} = \aleph_{2}$.

In particular, we aren't going to find an order contradicting CH at the level of $\sum_{i=1}^{1}$.

In general, these projective ordinals are not cardinals. That being said, assuming determinacy axioms, they often will

be, and in ZF+AD, every δ_n^1 is actually a regular cardinal. Calculating these however is not exactly easy, and transfering these to results in ZFC isn't exactly possible either since AD is incompatible with AC. The general strategy is instead to look at the consequences of AD and then translate these results to PD and L(\mathbb{R}) (which hasn't yet been defined):

$$\begin{aligned} \mathsf{ZFC} \vdash ``\mathfrak{g}_1^1 &= \aleph_1 "\\ ``\mathfrak{g}_2^1 &\leq \aleph_2 "\\ \mathsf{ZFC} + \mathsf{CH} \vdash ``\mathfrak{g}_1^1 &= \aleph_1 "\\ ``\mathfrak{g}_n^1 &\leq \aleph_2 \text{ for every } n < \omega "\\ \mathsf{ZF} + \mathsf{AD} \vdash ``\mathfrak{g}_n^1 + = (\mathfrak{g}_n^1)^+ \text{ for odd } n"\\ ``\mathfrak{g}_1^1 &= \aleph_1 "\\ ``\mathfrak{g}_2^1 &= \aleph_2 "\\ ``\mathfrak{g}_2^1 &= \aleph_2 "\\ ``\mathfrak{g}_3^1 &= \aleph_{\omega+1} "\\ \vdots\\ \mathsf{ZFC} + \mathsf{PD} \vdash ``\mathfrak{g}_1^1 &= \aleph_1 "\\ ``\mathfrak{g}_2^1 &= \aleph_2^{\mathsf{L}(\mathbb{R})} \leq \aleph_2 "\\ ``\mathfrak{g}_2^1 &= \aleph_{\omega+1}^{\mathsf{L}(\mathbb{R})} \leq \aleph_2 "\\ ``\mathfrak{g}_3^1 &= \aleph_{\omega+1}^{\mathsf{L}(\mathbb{R})} \leq \aleph_3 ". \end{aligned}$$

In ZF + AD, we generally have that δ_n^1 , for odd *n*, is the successor of a *suslin cardinal*^{xix} that has cofinality ω [19]. Determinacy is the right context to study these ideas in, because of the connection between scales and suslin cardinals as per Result 25 E • 3 and Kunen–Martin (25 F • 5), and also noting that stronger determinacy axioms propagate the scale property further in the projective hierarchy. Note that it's currently an open question whether the pattern of $ZF + PD \vdash :\delta_n^1 \leq \aleph_n$ for $n \leq 4$ continues for $n \geq 5$ under PD, even if we make additional assumptions.^{xx}

^{xix}A suslin cardinal is a cardinal κ such that some $X \subseteq \mathcal{N}$ is κ -suslin, but not α -suslin for any $\alpha < \kappa$.

^{xx}Well, it's open with assumptions that don't make the problem trivial like " $|\mathbb{R}| \leq \aleph_4$ " or parts of the conjecture itself like " $\delta_5^1 \leq \aleph_5$ ".

Section 26. Exercises

Exercises

Chapter V. Games and Determinacy*

Section 27. Fundamentals of Games

Determinacy has developed into an *incredibly* rich area of research about fairly concrete questions: what is ξ_2^1 ? What pointclasses have the reduction property? What sets are lebesgue measurable? Answering these questions in ZF making (sometimes modest) determinacy assumptions can help us to understand the situation in $V \models$ ZFC. What's perhaps most remarkable, is that the questions of determinacy can be phrased in a way accessible to most people, mathematicians or not, further motivating just how concrete the questions are. Nevertheless, their connection with large cardinal axioms can help us understand why the questions are hard to answer, and also help provide techniques to study the questions.

We begin with the basic definition of the games in question.

- 27.1. Definition — For our purposes, a *number game* refers to a game between players I and II who take turns playing natural numbers $n_i \in \omega$ for $i < \omega$.

- **I**: n_0 n_2 n_4 ... **II**: n_1 n_3 n_5 ...
- The resulting play is the real $x = \langle n_i : i < \omega \rangle \in \mathcal{N}$.
- On any given turn, we call $x \upharpoonright n = \langle n_i : i < n \rangle$ a *partial play*.
- There is a win-state in that I or II has won based on the resulting play. For simplicity, no ties are allowed: for any resulting play, I wins iff II loses, and a player wins iff that player does not lose.

Of course, this is just the general setup to the games we will be considering: we haven't specified the winning coniditions yet. But given that there are no ties, we can always consider the set of plays where one of the players has won: $A = \{x \in \mathcal{N} : \mathbf{I} \text{ wins}\}$ has \mathbf{I} win iff $x \in A$. This motivates thinking of games where \mathbf{I} and \mathbf{II} take turns are building a real $x \in \mathcal{N}$ where \mathbf{I} tries to ensure $x \in A$ while \mathbf{II} tries to ensure $x \notin A$.

- 27 · 2. Definition

Let $A \subseteq \mathcal{N}$. The game G(A) is the number game where I wins with $x \in \mathcal{N}$ iff the resulting play $x \in A$.

Note that this definition encompasses *all* physical, non-luck based games between two players that can see the other player's decisions. For example, it's not difficult to see that rules don't matter. Here, having "rules" just means that on any given turn, there's only a subset of natural numbers I or II is allowed to play. As a result, we can form a tree T over ω of allowed partial plays according to these rules (say a player loses if they don't have any valid moves).

- 27 • 3. Result

Let T be a tree over ω . Consider the number game G where each partial play must be in T. Therefore, there is a set $A \subseteq \mathcal{N}$ where for any resulting play x in G(A), I wins G(A) iff I wins G with some initial segment of x, and similarly for II.

Proof .:.

Note that because I plays n_0 on the first turn and plays every two-levels thereafter, for any play $x, x \upharpoonright n$ for even $n < \omega$ had II play last. Consider the set

 $A = \{x \in \mathcal{N} : \mathbf{I} \text{ wins in } G \text{ with some initial segment of } x\}$

 $= \{x \in [T] : \mathbf{I} \text{ wins in } G \text{ with } x\} \cup \{x \in \mathcal{N} : \mathbf{II} \text{ broke a rule first}\}.$

Here by "II broke a rule first", we mean there's some $n < \omega$ where $x \upharpoonright n \in T$ but $x \upharpoonright n + 1 \notin T$ with n + 1 even (so that II just played to break a rule). It follows that G(A) is equivalent to G in the sense of the statement. \dashv

We can also encompass finite length games with finitely many options at any given turn just through coding. We give an example rather than work through this informal statement.

With chess, we can label each piece with a number n < 32. As there are 64 possible positions for any given piece, which we can also label with a number. So we can describe the current board with finitely many numbers coding where each piece is on the board. All of this is just to say that a move in chess is just a natural number which describes the new resulting board (such that the move obeys whatever rules of chess there are). So now we have a number game where we want partial plays to be in a certain tree. By Result 27 • 3, this is the same as playing a game G(A) for some $A \subseteq \mathcal{N}$.

The same idea can be applied to other board strategy games like checkers in addition to games like solitaire where I moves cards, and II chooses the revealed cards' values (where II must play according to how the deck of cards was shuffled). Similarly, video games can be seen as number games where each frame is a turn, and the natural number played codes the inputs given on that frame. All this is just to say that it suffices to consider games of the form G(A) for $A \subseteq \mathcal{N}$.

The main purpose of looking at games is figuring out how to win them.

- 27 · 4. Definition

A strategy is a function $\sigma : {}^{<\omega}\omega \to \omega$. A strategy σ is a winning strategy for I (or σ wins for I) in a number game G iff playing according to σ always results in a win for I in G. In other words, σ is a winning strategy for I if for any $x \in \mathcal{N}$ and play of the game of the form

I:
$$\sigma(\emptyset) = \sigma_0$$
 $\sigma(\sigma_0, x(0)) = \sigma_1$ $\sigma(\sigma_0, x(0), \sigma_1)$ $\sigma(\sigma_0, \cdots, x(2))$ \cdots
II: $x(0)$ $x(1)$ $x(2)$ \cdots

Then I wins G with the resulting play, which we denote by $\sigma * x$ so that for every even $n < \omega$, $(\sigma * x)(n) = \sigma(\sigma * x \upharpoonright n)$.

• A number game is *determined* iff one of the players has a winning strategy.

• A set $A \subseteq \mathcal{N}$ is *determined* iff G(A) is determined.

• Γ -determinacy, also written $Det(\Gamma)$, for a pointclass Γ holds iff every $A \in \Gamma$ is determined.

And these definitions work similarly for II: τ is a winning strategy for II iff for all $x \in \mathcal{N}$, $x * \tau$ wins for II.

Now all of this has appealed to the intuitive idea of a game. Formally speaking, questions about G(A) can be recast as questions about A and \mathcal{N} : a winning strategy σ for I in G(A) is just a $\sigma \in {}^{\omega}({}^{<\omega}\omega)$ where $\forall x \in \mathcal{N} \ (\sigma * x \in A)$.

§ 27 A. Determinacy in ZFC

The easiest games to win are those where there are only countably many win-states for I.

- **27** A • 1. **Result** - Let $A \subset \mathcal{N}$ be countable. Therefore **II** wins G(A).

Proof .:.

We give a diagonalization argument. Enumerate $A = \{a_n : n < \omega\}$. Its strategy will be on their *n*th turn (i.e. the 2n + 1st entry of the resulting play) to choose a value different from $a_n(2n + 1)$: for a partial play *p* of length 2n + 1, $\sigma(p) = a_n(2n + 1) + 1$. It follows that for any $x \in \mathcal{N}$ and $n < \omega$, $(x * \sigma)(2n + 1) \neq a_n(2n + 1)$ and thus $x * \sigma \notin \{a_n : n < \omega\} = A$ so that II has won.

This is the easiest sort of game to win: II doesn't even care what I plays and just focuses on diagonalizing against A.

Harder games involve actual strategy where we care what the other player is doing, and where we care what happens after a given partial play. As such, it will be useful to consider the following.

- 27 A • 2. Definition

Let $A \subseteq \mathcal{N}$ and let $p \in {}^{<\omega}\omega$ be a partial play in G(A).

- If p has even length (so $p = \emptyset$ or II just played) say II can force a win at p iff II has a winning strategy in $G(\{x \in \mathcal{N} : p \cap x \in A\})$.
- If p has odd length (so I just played) say II can force a win at p iff there's some $m \in \omega$ where II has a winning strategy in $G(\{x \in \mathcal{N} : p^{\frown} \langle m \rangle^{\frown} x \in A\})$

And these definitions work similarly for I. Basically, II can force a win at p if II has a winning strategy for how to play after p. This is useful especially in showing simple games are determined.

27 A · 3. Theorem (Closed Determinacy) Let $A \subseteq \mathcal{N}$ be closed. Therefore G(A) is determined. In other words, $Det(\Pi_1^0)$.

Proof .:.

Suppose II doesn't have a winning strategy. We will describe a defensive strategy by player I which basically amounts to (at every turn) trying to not lose. In the end, II will not have won by the closure of A so that I wins. The crucial observation is the following, basically saying that if II can always force a win at no matter what I does, then II could have forced a win at their previous turn. We use this for the contrapositive.

Claim 1

Let p be a partial play of odd length (so I just played). Suppose there's a move $m \in \omega$ by player II such that for any move $n \in \omega$ by player I, II can force a win at $p^{\frown}(m, n)$. Therefore II can force a win at p.

Proof .:.

For each $n \in \omega$, let σ_n force a win for **II** at $p^{\frown}(m, n)$. Intuitively, **II**'s strategy is then just to use σ_n if **I** responds with *n*. This strategy will force a win for **II**. In other words, the strategy σ defined by $\sigma(\emptyset) = m$ and $\sigma(\langle m, n \rangle^{\frown} q) = \sigma_n(q)$. forces a win for **II** at *p*.

So what's the strategy that I uses? At the start, because II doesn't have a winning strategy, there's a move $\sigma(\emptyset) \in \omega$ by I such that II can't force a win at $\langle \sigma(\emptyset) \rangle$. But then by Claim 1, no matter what $n_0 \in \omega$ II plays, there's some $\sigma(n_0)$ such that II can't force a win at $\langle \sigma(\emptyset), n_0, \sigma(n_0) \rangle$. So inductively, at every partial play p where I just played according to σ defined thus far, II can't force a win at p and thus by Claim 1, no matter what $n \in \omega$ II plays, there's some $\sigma(p^{-1}\langle n \rangle) \in \omega$ such that II can't force a win at $p^{-1}\langle n, \sigma(p^{-1}\langle p \rangle) \rangle$.

This defines the strategy $\sigma : {}^{<\omega}\omega \to \omega$ for **I**. To see that this wins for **I**, suppose **II** plays $x \in \mathcal{N}$ so the resulting play is $\sigma * x$. Note that **II** can't force a win at any partial play $p \lhd \sigma * x$ by definition. But then $\mathcal{N}_p \cap A \neq \emptyset$ for every $p \lhd \sigma * x$. Hence $\sigma * x \in A$ because A is closed. So σ wins G(A) for **I**.

It's important to note that this *doesn't* say that I always wins closed games. Clearly II wins the game $G(\emptyset)$ where \emptyset is closed. Really, the fact that A is determined says nothing about which player has a winning strategy, just that one of them does. This is especially so with the fact that *open* games are also determined.

```
- 27 A • 4. Result Suppose A and \{x \in \mathcal{N} : \langle n \rangle^{\frown} x \in A\} is determined for every n < \omega. Therefore \mathcal{N} \setminus A is determined.
```

Proof .:.

The idea here is that going from G(A) to $G(\mathcal{N} \setminus A)$, we're switching players, and adding a turn at the beginning. Suppose II wins G(A) with σ . Consider the strategy σ' for I in $G(\mathcal{N} \setminus A)$ defined by $\sigma'(\tau) = \sigma(\langle 0 \rangle \cap \tau)$. It follows that σ' wins for I in $G(\mathcal{N} \setminus A)$ because for any $x \in \mathcal{N}, \sigma' * x = (\langle 0 \rangle \cap x) * \sigma \notin A$.

Suppose I does not win $G(\mathcal{N} \setminus A)$. Thus there's no initial move that forces a win for I. Let $n \in \omega$ be any initial

play by I in $G(\mathcal{N} \setminus A)$. Consider $G(\{x \in \mathcal{N} : \langle n \rangle \cap x \in A\})$. If II wins this, then the above argument tells us I wins, a contradiction. Therefore by determinacy of this game, I wins with a strategy σ_n . Define a strategy σ' for II in $G(\mathcal{N} \setminus A)$ by $\sigma'(\langle n \rangle \cap p) = \sigma_n(p)$, for any $n \in \omega$ and $p \in {}^{<\omega}\omega$. This wins for II in $G(\mathcal{N} \setminus A)$ because for any $x = \langle x(0) \rangle \cap x' \in \mathcal{N}$, $x * \sigma' = \langle x(0) \rangle \cap \sigma_{x(0)} * x'$. Since $\sigma_{x(0)}$ wins for I in $G(\{y \in \mathcal{N} : \langle x(0) \rangle \cap y \in A\})$, it follows this $x * \sigma' \notin A$ so that σ' wins for II in $G(\mathcal{N} \setminus A)$.

In general, we cannot do better than Result 27 A • 4: it's not true in general that if A is determined then $\mathcal{N} \setminus A$ is determined (if there are sets that aren't determined). To see this, let Z be some set that isn't determined, and consider

$$A = \mathcal{N}_{\langle 0 \rangle} \cup \bigcup_{0 < n \in \omega} \{ \langle n \rangle^{\frown} x \in \mathcal{N} : x \notin Z \},$$

Then clearly I wins G(A) just by playing 0 as the first move. But $\mathcal{N} \setminus A$ isn't determined. To see this, any winning strategy for I will have the first move $n_0 \neq 0$ (else I immediately will lose). But then I wins in the game $G(\{x \in \mathcal{N} : \langle n_0 \rangle \widehat{\ } x \notin A\}) = G(Z)$, which is a contradiction. Similarly, if II wins $G(\mathcal{N} \setminus A)$ with σ , then $\sigma'(\tau) = \sigma(\langle 1 \rangle \widehat{\ } \tau)$ wins for I in G(Z).

Nevertheless, given the closure properties of the borel and projective pointclasses, we get open determinacy.

— 27 A • 5. Corollary (Open Determinacy)

Every open $A \subseteq \mathcal{N}$ is determined. In other words, $\mathsf{Det}(\sum_{n=1}^{n})$. In general, $\mathsf{Det}(\sum_{n=1}^{n})$ iff $\mathsf{Det}(\prod_{n=1}^{n})$ for $n < \omega$.

Another corollary to the determinacy of open and closed games is that games of finite length are determined, just because we can consider $\bigcup_{\sigma \in A} \mathcal{N}_{\sigma}$ where $A \subseteq {}^{<\omega}\omega$ is the set of winning positionsⁱ for I in the game. This is open and therefore determined.

We can go much further beyond open and closed determinacy by the remarkable theorem due to Donald Martin. We will not prove this here because the proof is *very* complicated and unnecessary for our purposes.

— 27 A • 6. Theorem (Borel Determinacy)

Every borel $A \subseteq \mathcal{N}$ is determined. In other words, $\mathsf{Det}(\underline{A}_1^1)$.

This just serves as motivation for the idea that simply definable sets of reals should be determined. Going beyond this, however, isn't possible in ZFC alone, as it's consistent that there are projective sets that aren't determined. This is partially a result of the analytical well order of \mathcal{N}^{L} in L, because we have the following theorem showing that not all sets are determined.

- 27 A • 7. Theorem -

- There is a set $A \subseteq \mathcal{N}$ that is not determined.
- In fact, every $B \subseteq \mathcal{N}$ of size $|B| = |\mathcal{N}|$ has a subset $A \subseteq B$ that is not determined.
- As a result, in L, there is an projective (and in fact, analytical) set that is not determined.

Proof .:.

• Since $|^{\omega}({}^{<\omega}\omega)| = |\mathcal{N}|$, there are as many strategies as real numbers. Let $B \subseteq \mathcal{N}$ be given of size $|\mathcal{N}| = \kappa$. Well-order *B* and enumerate the strategies $\{\sigma_{\alpha} : \alpha < \kappa\}$. Define two sequences $\vec{x} = \langle x_{\alpha} : \alpha < \kappa \rangle$ and $\vec{y} = \langle y_{\alpha} : \alpha < \kappa \rangle$ by recursion.

For $\alpha < \kappa$ and $\vec{x} \upharpoonright \alpha$, $\vec{y} \upharpoonright \alpha$ defined, consider $\{\sigma_{\alpha} * r : r \in B\} \setminus \{y_{\xi} : \xi < \alpha\}$, which isn't empty since $|\{r * \sigma_{\alpha} : r \in B\}| = |B| = \kappa > |\alpha| = |\{y_{\xi} : \xi < \alpha\}|.$

So let $x_{\alpha} = r * \sigma_{\alpha}$ for the least $r \in B$ such that $r * \sigma_{\alpha} \notin \{y_{\xi} : \xi < \alpha\}$. Similarly, we can define $y_{\alpha} = \sigma_{\alpha} * r'$ for the least $r' \in B$ such that $\sigma_{\alpha} * r' \notin \{x_{\xi} : \xi \le \alpha\}$. By construction, $x_{\alpha} \neq y_{\beta}$ for $\alpha, \beta < \kappa$.

Define $A = \{x_{\alpha} : \alpha < \kappa\}$. G(A) will not be determined. To see this, suppose I wins with $\sigma = \sigma_{\alpha}$. For any $x \in \mathcal{N}$, write x = even(x) * odd(x). If II plays with $\text{odd}(y_{\alpha})$, then the resulting play is $\sigma_{\alpha} * \text{odd}(y_{\alpha}) = y_{\alpha} \notin I$

ⁱAssuming here that the finite length game has no rules. If it does have rules, we use the same idea as in Result 27 • 3.

A meaning I has lost. Similarly, if II wins with $\sigma = \sigma_{\alpha}$, then the resulting play of even $(x_{\alpha}) * \sigma_{\alpha} = x_{\alpha} \in A$ meaning II has lost.

- To show that L has an analytical set that is not determined, work in L. Consider A above, constructed as a subset of $B = \mathcal{N}$ with the Δ_2^1 -well order from Theorem 25 B 2. We can code strategies as real numbers using some computable coding of ${}^{<\omega}\omega$ into ω , allowing us to define resulting plays in a Δ_1^1 -way: for $x \in \mathcal{N}$ coding σ , write $x \star y$ for $\sigma \star y$ in the sense of Definition 27 4. As a result, $x \in A$ iff x appears in some partial list of our construction $\vec{x} \upharpoonright \alpha$ for $\alpha < |\mathcal{N}| = \aleph_1$. In other words, there are $\alpha \in WO(\Pi_1^1)$ and $\chi, \gamma \in \mathcal{N}$ such that
 - 1. There is an $n < \omega$ such that $\forall k < \omega (\chi(\operatorname{code}(n, k)) = x(k)) (\Delta_1^1)$; and
 - 2. For every $n < \omega$, there is a real χ_n (where $\chi_n(k) = \chi(\operatorname{code}(n, k))$ for $k < \omega$), a real σ coding a strategy, and a real r such that
 - a. $\chi_n = r \star \sigma (\Delta_1^1)$,
 - b. $\forall m \in \omega \ (\alpha(\operatorname{code}(m, n)) = 1 \rightarrow \exists k \in \omega \ (\chi(\operatorname{code}(n, k)) \neq \chi(\operatorname{code}(m, k))) \ (\Delta_1^1), \text{ and }$
 - c. For every $z \in \mathcal{N}$ such that the above occurs for $(z * \sigma)(k)$ in place of $\chi(\operatorname{code}(n, k))$, then $r \leq_{\mathrm{L}} z$ $(\forall^{\mathcal{N}}(\Delta_1^1 \to \Delta_2^1) = \Pi_3^1$; and
 - 3. (1) and (2) hold, switching χ and γ , and replacing " $\alpha(\operatorname{code}(m, n)) = 1 \rightarrow \dots$ " in (2b) with " $\alpha(\operatorname{code}(m, n)) = 1 \lor m = n \rightarrow \dots$ " (Π_1^1).

The same proof as above tells us A isn't determined. Given the above description, we can say A has complexity

$$\exists^{\mathcal{N}} (\Pi_1^1 \wedge \Delta_1^1 \wedge \forall^{\omega} \exists^{\mathcal{N}} (\Delta_1^1 \wedge \Delta_1^1 \wedge \Pi_3^1 \wedge \Pi_3^1)) = \Sigma_6^1$$

 \dashv

Hence there's an analytical set that isn't determined in L.

The set in question above can be categorized as Σ_6^1 in L and hence it's consistent for sets far along in the analytical hierarchy to not be determined. But what about something of lowest complexity not already known by Borel Determinacy (27 A • 6): Σ_1^1 ? It turns out that Σ_1^1 -determinacy is implied by the existence of a measurable cardinal, which we already know is incompatible with L by L Has No Measurable Cardinals (12 D • 4). This is partly due to the fact that a measurable cardinal implies the existence of a certain real 0[#] that codes an elementary embedding $j : L \to L$ in V. In fact, Σ_1^1 -determinacy is equivalent to the existence of 0[#]. More generally, $\Sigma_1^1(x)$ -determinacy for $x \in \mathcal{N}$ is equivalent to the existence of x^{\ddagger} , which codes an elementary embedding $j : L[x] \to L[x]$. We will return to this idea in the next chapter once relative constructibility has been introduced.

We've seen that there are non-determined sets, but it's interesting to note that we needed to use AC in order to do this: we needed to be able to enumerate the strategies and deal with them one by one. But in what sense is choice essential here? Is it possible in ZF alone that there are no undetermined sets of reals?

– 27 A • 8. Definition (Axiom)

(*Determinacy*) AD states: every number game G(A) for $A \subseteq \mathcal{N}$ is determined.

Of course, Theorem 27 A \cdot 7 tells us that AD is incompatible with ZFC. But is ZF + AD consistent? The answer to this is *yes* relative to the existence of sufficiently large cardinals. Of course, this partially begs the question of why we would care to examine a world where AD holds.

Section 28. Determinacy and Pointclass Properties

§ 28 A. Regularity properties from determinacy

One subject of interest is the interaction of AD with the regularity properties of Section 23. In particular, under AD, every set of reals has the perfect set property, the baire property, and is lebesgue measurable. Frequently such results are translatable to the context of ZFC, especially in relation to the projective and analytical hierarchy.

Firstly, note that we still get Closed Determinacy $(27 A \cdot 3)$ and Open Determinacy $(27 A \cdot 5)$ in ZF + DC. The same proof goes through, where the strategy is just to avoid losing. DC is necessary here to choose the strategies in the case that we can ever force a win. Often AD is studied in conjunction with DC, as we will do here. It's actually an open problem whether AD outright implies DC, which would certainly make the stated assumptions much simpler.ⁱⁱ Firstly, recall what exactly DC is saying from Definition 9 B • 6.

– 28 A • 1. Definition (Axiom)

(Dependent Choice) DC states: if a relation has infinite height, then it has an infinite branch, i.e. for every X and R $\forall x \in X \exists y \in X (\langle x, y \rangle \in R) \rightarrow \exists s (s : \omega \rightarrow X \land \forall n \in \omega \langle s(n), s(n+1) \rangle \in R).$

Let's show the first regularity property holds under ZF + DC + AD. The proof of this fact essentially uses a different kind of number game: one in which players play only 0s and 1s. This really ensures PSP(X) for each $X \subseteq C = {}^{\omega}2$, but the result can still be translated to \mathcal{N} by a continuous injection.

— 28 A • 2. Theorem (AD + DC Implies PSP) —

Assume ZF + DC + AD. Therefore PSP(X) for every $X \subseteq \mathcal{N}$.

Proof .:.

We first show the result for \mathcal{C} . Let $X \subseteq \mathcal{C}$ be given. Consider the game G'(X) where I plays finite sequences of 0s and 1s each turn whereas **II** plays only a single 0 or 1 each turn.

I: $s_0 \in {}^{<\omega}2$ s_1 \cdots II: $x_0 \in 2$ x_1 \cdots

The resulting play is $r = s_0 \cap \langle x_0 \rangle \cap s_1 \cap \langle x_1 \rangle \cap \cdots \in C$. We say that I wins iff $r \in X$. This game looks extremely biased towards I, but this is good news for us.

Claim 1

I has a winning strategy for G'(X) iff X contains a perfect subset.

ⁱⁱPart of the reason why this is open is that all the models we have of AD are also models of AD + DC and in fact of an ostensibly stronger (but also unknown whether it's equivalent to AD) hypothesis called AD^+ . One thing we due know due to Kechris is that AD implies $DC_{\mathbb{R}}$ in the model $L(\mathcal{N})$, where $DC_{\mathbb{R}}$ is a restriction of DC to \mathbb{R} (or \mathcal{N}).

DETERMINACY AND POINTCLASS PROPERTIES

Proof .:.

Firstly, suppose that I wins G'(X) with a strategy τ . It follows that X contains a perfect subset. To see this, we find a continuous injection from C into X and use Lemma 23 A • 4. for $x \in {}^{\omega}2$, let f(x) be the resulting play where I uses τ and II plays x: $f(x) = \tau * x \in {}^{\omega}2$. It's clear that f is injective. To see that f is also continuous, note that for any $y, x \upharpoonright n = y \upharpoonright n$ has $f(x) \upharpoonright n = f(y) \upharpoonright n$ (and in fact, probably much more) because

 $f(x) \rhd \tau(\emptyset)^{\frown} \langle x_0 \rangle^{\frown} \tau(x_0)^{\frown} \cdots^{\frown} \langle x_{n-1} \rangle = \tau(\emptyset)^{\frown} \langle y_0 \rangle^{\frown} \tau(y_0)^{\frown} \cdots^{\frown} \langle y_{n-1} \rangle \lhd f(y),$

and the length of that initial segment is at least n. f is thus continuous by Corollary 21 B • 4. So by Lemma 23 A • 4, im f is perfect in \mathcal{C} . Since I wins with τ , $f = x \mapsto \tau * x : \mathcal{C} \to X$ and so im f is a perfect subset of X. Note that the proof of Lemma 23 A • 4 uses DC to show that \mathcal{C} is compact by a use of Kőnig's Lemma on Trees (9 B • 5) (and Theorem 9 B • 7 to get that DC implies countable choice).

Now suppose that X contains a perfect subset P. We can then consider the "tree" resulting from P similarly to Lemma 23 A \cdot 4 by

$$T = \{y \upharpoonright n : y \in \text{im } f \land n \in \omega\}.$$

Since P has no isolated points, at any stage $p \in T$ there is some s such that $p \cap s \cap \langle 0 \rangle$ and $p \cap s \cap \langle 1 \rangle$ are both in T. We can then define a strategy for I by playing according to the tree in that way: at stage p, I plays such an s (as chosen using DC). This strategy clearly wins for I.

- Claim 2 -

II has a winning strategy for G'(X) iff X is countable.

Proof .:.

Now suppose that \mathbf{II} wins G'(X) with a strategy τ . It follows that each $y \in X$ must be *rejected* somewhere, meaning that our play of the game diverges from y where it previously agreed with y. More precisely, a position p (where \mathbf{II} just moved) rejects $y \in C$ iff

• $p \triangleleft y$; and

• for all $s \in {}^{<\omega}2$ with $p \cap s \triangleleft y$, the strategy τ moves away from $y: p \cap s \cap \tau(p, s) \not \lhd y$.

The idea is that each $y \in X$ must be rejected somewhere, but there can only be countably many points rejected, meaning X should be countable. Let's do this rigorously.

Note that at each position p, there's at most one real that's rejected (recall we're working with C rather than \mathcal{N}). To see this, suppose p rejects $y \in C$. We can then construct y from p and τ . Because p rejects y, if I plays $s_0 = \emptyset$, then II playing with τ will move away from $y: \tau(p, s_0) = i \in 2$ has $p^{\frown}\langle i \rangle \not \lhd y$. In other words, $y(\ln(p)) \neq i$ and so $y(\ln(p)) = 1 - i \in 2$. So then we can consider what happens when I plays $s_1 = \tau(p^{\frown}s_0)$. Again, because p rejects y, the play by II with τ will differ from y now at position $\ln(p) + 1$: $y(\ln(p) + 1) = 1 - \tau(p, s_1)$. So we proceed in this way. Recursively define $s_0 = \emptyset$ and $s_{i+1} = s_i^{\frown} \langle 1 - \tau(p, s_i) \rangle$. The result is that $p^{\frown} \bigcup_{n < \omega} s_n = y$.

It's not hard to see that each $y \in X$ is rejected at some position p, since otherwise—by DC to choose I's moves at each stage—there's a play of the game that results in y: there is some play $s \in {}^{<\omega}2$ by I where $p \cap s \cap \tau(p, s) \triangleleft y$ of strictly longer length. If y isn't rejected at this position either, we can continue to lengthen our position until the resulting play is y. Hence $X \subseteq \bigcup_{p \in {}^{<\omega}2} \{y \in C : p \text{ rejects } y\}$ is contained in a countable set. It follows that X is countable.

Now suppose X is countable. Enumerate $X = \{x_n : n \in \omega\}$. It's strategy is to diagonalize: at position p where it's II's nth turn, II plays $1 - x_n(\ln(p)) \neq x_n(\ln(p))$. Since II gets infinitely many turns, the resulting play differs from each x_n eventually, and so the resulting play is not in X and II has won.

Under AD, each game G'(X) is determined by Result 27 • 3, since it's equivalent to a number game with the rules • I plays numbers that are codes for finite binary sequences; • **II** plays 0 or 1.

As a result, X has the perfect set property and thus PSP(X) for each $X \subseteq C$. Now let $Y \subseteq N$ be arbitrary, aiming to show PSP(Y) just by translating it to C. Let $f : N \to C$ be a continuous injection by Theorem 21 B • 7. It follows that $f"Y \subseteq C$ has the perfect set property. So either |f"Y| = |Y| is countable, or there is a perfect subset $P \subseteq f"Y$. In the latter case, consider $f^{-1}"P \subseteq Y$, which must also be closed since P is and f is continuous. $f^{-1}"P$ also must contain no isolated points for the same reason, meaning $f^{-1}"P$ would be a perfect subset of Y.

As a result, ZF + DC + AD implies a version of CH, namely Result 5 E • 3: that every subset of \mathcal{N} has cardinality $\leq \aleph_0$ or cardinality $[\mathcal{N}]_{size}$. Given that AC fails in ZF + DC + AD, it follows that that cardinality of \mathcal{N} is not a cardinal, and in fact, there are no injections from ω_1 into \mathcal{N} .

- 28 A · 3. Corollary — Assume ZF + DC + AD. Therefore there are no injections $f : \omega_1 \to \mathcal{N}$.

Proof .:.

If there were such an injection, then $f''\omega_1 \subseteq \mathcal{N}$ would have size ω_1 . By AD + DC Implies PSP (28 A • 2), $f''\omega_1$ would have size \aleph_0 or would be in bijection with \mathcal{N} . Since the former is by definition impossible, we get that $|\mathcal{N}| = \aleph_1$. But then we can enumerate strategies and real numbers as in Theorem 27 A • 7 to get a set that is not determined, contradicting AD.

This has quite a lot of consequences relating to measure that we will start to consider in the next subsection.

There is also a ZFC-compatible version of AD + DC Implies PSP ($28 \text{ A} \cdot 2$) with the same proof after some tedious consideration about classifying complexity, telling us that under determinacy assumptions, CH would hold for levels of the projective hierarchy.

- 28 A • 4. Corollary Assume ZF + DC. Therefore, for $n \in \omega$, $\text{Det}(\sum_{i=1}^{n})$ implies $\text{PSP}(\sum_{i=1}^{n})$, and similarly for $\prod_{i=1}^{n}$.

Proof .:.

Suppose $X \subseteq \mathcal{N}$ is $\sum_{n=1}^{n}$. Let $f : \mathcal{N} \to C$ be a continuous injection by Theorem 21 B • 7. It follows that f''X is $\sum_{n=1}^{n}$ and G'(f''X) is equivalent to the number game G(B) where $B \subseteq \mathcal{N}$ consists of all $x \in \mathcal{N}$ such that

- 1. For all $n < \omega$, x(2n) is the code of a finite binary sequence τ_n ;
- 2. for all $n < \omega, x(2n + 1) \in 2$;
- 3. $\tau_0^{(x(1))} \tau_1^{(x(3))} \tau_2^{(x(5))} \cdots$ is in f''X.

Since the map taking $x \in B$ to $\tau_0^{\frown} \langle x(1) \rangle^{\frown} \tau_1^{\frown} \langle x(3) \rangle^{\frown} \tau_2^{\frown} \langle x(5) \rangle^{\frown} \cdots$ is continuous by Corollary 21 B•4, taking the preimage of f^*X yields another \sum_n^1 set. (1) and (2) above are clearly Δ_1^1 so that B, being the intersection of these with the preimage of f^*X , is \sum_n^1 just as X is. Thus G(B) and so $G'(f^*X)$ are determined. By the proof of AD+DC Implies PSP (28 A•2), f^*X has the perfect set property and thus X does too: either $|f^*X| = |X| \le \aleph_0$, or there is a perfect subset $P \subseteq f^*X$ where therefore $f^{-1}*P \subseteq X$ is also perfect.

The next regularity concept introduced was that of lebesgue measurability in Subsection 23 B. We established the lebesgue measurability for \sum_{1}^{1} -sets with Corollary 23 B • 20 but could not go beyond due to Theorem 25 B • 5: ZFC + "V = L" implies that there are simply definable sets— Δ_{2}^{1} to be precise—that aren't lebesgue measurable and don't have the baire property. We will see that AD implies that all sets are lebesgue measurable and so AD (or even $\text{Det}(\Sigma_{2}^{1})$) is incompatible with L.

Showing that every set is lebesgue measurable under AD isn't quite as simple a task as it was for the perfect set property. But all three of the proofs that AD implies a regularity property take the following form: first consider a variant game, show that something desirable happens depending on which player wins that game, and then conclude that the game is still determined by AD even though it's not a number game as in Definition $27 \cdot 1$.

DETERMINACY AND POINTCLASS PROPERTIES

The variant game we will consider for lebesgue measurability is the covering game. The basic idea is that II tries to cover a set X with small open sets and I tries to build a point in X that is left uncovered: II has a winning strategy if they can cover X, and otherwise I will try to find a point left uncovered. We need to be careful though, since we need to be able to translate this kind of game into a number game, and so this is where the technical details start to creep in: we need II to only be able to play increasingly smaller sets from a countable list of options. One way to do this is by indexing a certain list of sets and having II play the indices while I builds a real in $[0, 1] \subseteq \mathbb{R}$ as written in binary and considered completely separately from II.

— 28 A • 5. Definition -

Let $\varepsilon > 0$ in \mathbb{R} be given. For $i \in \omega$, let $\langle B_n^i(\varepsilon) : n \in \omega \rangle$ enumerate all sets $B \subseteq \mathbb{R}$ of lebesgue measure $\mu(B) \leq \varepsilon/4^{i+1}$ that are finite unions of closed intervals with rational endpoints.

The definition makes sense of course since there are only $|\mathbb{Q} \times \mathbb{Q}|^{<\omega} = \aleph_0$ -many finite unions of open intervals with rational endpoints. We make use of this definition because we can still correctly calculate lebesgue measure with these sets. The following lemma proves this, but it is unfortunately technical. The reader is recommended to be convinced of the fact and skip the calculation oriented proof of the lemma.

- 28A.6. Lemma (ZF) Suppose $X \subseteq [0, 1] \subseteq \mathbb{R}$. Therefore X is lebesgue measurable with measure 0 iff $0 = \inf \left\{ \sum_{i \in \omega} \mu(B^i) : \text{each } B^i \in \{B_n^i(\varepsilon) : n \in \omega\} \text{ for some } \varepsilon \text{ and } X \subseteq \bigcup_{i \in \omega} B^i \right\}.$ (*)

Proof .:.

Clearly if (*) holds then the lebesgue outer-measure of X is $\mu^*(X) = 0$ and so X is lebesgue measurable by Result 23 B • 7. So suppose X is lebesgue measurable with measure 0.

Let $\mu'(X)$ denote the right-hand side in (*). Suppose $0 < \mu'(X)$. Let $0 < \varepsilon < \mu'(X)$ be arbitrary and let $\{I_i : i \in \omega\}$ be an arbitrary countable collection of closed intervals such that

- 1. $\sum_{i \in \omega} \mu(I_i)$ is less than $\varepsilon/6$ (and therefore below $\mu'(X)$); and
- 2. $X \subseteq \bigcup_{i \in \omega} I_i$.

Each $I_i = [a_i, b_i]$ is contained in a closed interval $[\alpha_i, \beta_i]$ with rational endpoints where $a_i - \alpha_i$ and $\beta_i - b_i$ are arbitrarily small by the density of \mathbb{Q} in \mathbb{R} . In particular, we can have $|a_i - \alpha_i|, |\beta_i - b_1| < \varepsilon/4^{i+2}$. Then X is still covered by $\bigcup_{i \in \omega} [\alpha_i, \beta_i]$ and

$$\sum_{i \in \omega} \mu([\alpha_i, \beta_i]) \le \sum_{i \in \omega} \mu(I_i) + 2\frac{\varepsilon}{4^{i+2}} = \sum_{i \in \omega} (\mu(I_i)) + 2\varepsilon \sum_{i \in \omega} \frac{1}{4^{i+2}}$$
$$\le \sum_{i \in \omega} (\mu(I_i)) + \frac{2\varepsilon}{12} < \frac{\varepsilon}{6} + \frac{\varepsilon}{6} = \frac{\varepsilon}{3} < \mu'(X).$$

We also have that $\sum_{i \in \omega} \mu([\alpha_i, \beta_i])$ is below $\varepsilon/3 = \sum_{n < \omega} \varepsilon/4^{n+1}$. So without loss of generality, let's work with $I_n = [\alpha_n, \beta_n]$, closed intervals with rational endpoints.

Now we can break up $\{I_i : i < \omega\}$ in a way such that we may regard them as $B_n^i(\varepsilon)$ s. Actually doing this is quite tedious.

- Case 1. $\mu(\bigcup_{n < \omega} I_n) > \sum_{j \le i} \varepsilon/4^{j+1}$.
 - Let $m \in \omega$ be the least such that the measure of $\bigcup_{n < m} I_n$ is less than $\sum_{j \le i} \varepsilon/4^{j+1}$ but $\bigcup_{n \le m} I_m$ has a larger measure.
 - Consider the closure of $\bigcup_{n \le m} I_n \setminus \bigcup_{j < i} B^j$ and let B^i be composed of the remaining intervals and initial segments of intervals of this set with as large measure as possible such that $\mu(B^i) \le \varepsilon/4^{i+1}$. (This can be done explicitly with lots more technical detail.)
- Case 2. $\mu(\bigcup_{n<\omega} I_n) \leq \sum_{j\leq i} \varepsilon/4^{j+1}$. This case occurs eventually since $\sum_{j<\omega} \varepsilon/4^{j+1} = \varepsilon/3$ while $\mu(\bigcup_{n<\omega} I_n) \leq \sum_{n<\omega} \mu(I_n) < \varepsilon/6$.
 - Let $m \in \omega$ be such that $\bigcup_{j < i} B^j \subseteq \bigcup_{n < m} I_m$ and $\mu(\bigcup_{n \le m} I_n) \mu(\bigcup_{n < \omega} I_n) < \varepsilon/4^{i+1}$ and then consider $B^i = \bigcup_{n < m} I_n \setminus \bigcup_{j < i} B^j$. Such an *m* exists since $\mu(\bigcup_{n < \omega} I_n) \le \sum_{n < \omega} \mu(I_n)$ converges.

Inductively, $\mu(B^i) \leq \varepsilon/4^{i+1}$ in Case 1. In Case 2, the least i^* for which we enter case 2, will inductively have that $\bigcup_{j < i^*} B^j$ has size $\sum_{j < i^*} \varepsilon/4^{j+1}$ which means the measure of B^{i^*} is $\mu(\bigcup_{n \leq m} I_m) - \mu(\bigcup_{j < i^*} B_j) \leq \sum_{j \leq i^*} \varepsilon/4^{j+1} - \sum_{j < i^*} \varepsilon/4^{j+1} = \varepsilon/4^{i^*+1}$. It's easy to see by induction that all subsequent instances of Case 2, $i > i^*$, similarly obey $\mu(B^i) \leq \varepsilon/4^{i+1}$. It should be clear that $\bigcup_{n < \omega} I_n = \bigcup_{n < \omega} B^n$ so that both cover X. Moreover, $\sum_{n < \omega} \mu(B^n) \leq \varepsilon/3 < \mu'(X)$ even though the B^n s take the proper form to witness $\mu'(X) \leq \varepsilon/3$, a contradiction.

Now we also make use of an old lemma about finding "minimal" measurable sets to give an equivalent characterization of all sets being lebesgue measurable. In particular, we need only deal with sets Y such that every *measurable* subset is null. In the context of ZFC this doesn't imply that Y itself is null since Y might not be measurable. But if we can show in the context of ZF + AD + DC that Y must be measure 0, this would imply every set is measurable.

```
– 28 A • 7. Lemma
```

(ZF) Suppose that for every $X \subseteq \mathbb{R}$, if every measurable subset of X is null, then X is measurable and null. Therefore every subset of \mathbb{R} is lebesgue measurable.

Proof .:.

~ ~

Let $X \subseteq \mathbb{R}$ be arbitrary. Let $X \supseteq Y$ be as in Lemma 23 B • 13: *Y* is measurable and every measurable *A* with $X \subseteq A \subseteq Y$ has $\mu(Y \setminus A) = 0$. It follows that $Y \setminus X$ has measure 0. To see this, let *A* be an arbitrary measurable subset of $Y \setminus X$ and consider $A' = Y \setminus A$ so that $X \subseteq A' \subseteq Y$. Thus $0 = \mu(Y \setminus A') = \mu(A)$. Since $A \subseteq Y \setminus X$ was an arbitrary, measurable subset, $Y \setminus X$ is measurable. Since *Y* is measurable, $X = Y \setminus (Y \setminus X)$ is also measurable.

So really it suffices to work just with subsets that satisfy the hypothesis of Lemma $28 \text{ A} \cdot 7$. So now we can put it all together: the covering game that we now define, the characterization of Lemma $28 \text{ A} \cdot 6$ and using such null sets with Lemma $28 \text{ A} \cdot 7$.

- 28 A • 8. Definition						
Let $X \subseteq [0, 1]$ and $\varepsilon > 0$. The <i>covering game</i> for X, ε is the game $G_{cov}^{\varepsilon}(X)$ that takes the form						
Ŀ	$r_{0} \in \mathcal{I}$	r.	re			
1.	$\lambda_0 \subset \mathcal{I}$	~1	x2			
II:	$B_{n_0}^0(\varepsilon)$	$B_{n_1}^1(\varepsilon)$	$B_{n_2}^2(\varepsilon)$	•••		
where each $n_i \in \omega$ and each $x_i \in 2$. We say I wins iff $x = \sum_{n \in \omega} \frac{x_n}{2^{n+1}} \in X \setminus \bigcup_{i \in \omega} B_{n_i}^i(\varepsilon)$.						

We use this game to show lebesgue measurability as with Lemma 28 A • 7, breaking down into cases depending on which player has a winning strategy.

- 28 A • 9. Theorem (AD + DC Implies Lebesgue Measurability) — Assume ZF + DC + AD. Therefore every $X \subseteq \mathbb{R}$ is lebesgue measurable.

Proof .:.

By Lemma 28 A • 7, it suffices to show that any $X \subseteq [0, 1]$ is null whenever every measurable subset of X is null. So suppose every measurable subset of $X \subseteq [0, 1]$ is null and consider $G_{cov}^{\varepsilon}(X)$. This game is determined due to AD, as it is equivalent to a certain number game where **II** plays $n_i \in \omega$ instead of $B_{n_i}^i(\varepsilon)$.

Claim 1

I does not have a winning strategy in $G_{cov}^{\varepsilon}(X)$ for any ε .

DETERMINACY AND POINTCLASS PROPERTIES

Proof .:.

Let σ be winning and define $f : \mathcal{N} \to \mathbb{R}$ by $f(x) = \operatorname{even}(\sigma * x)$, the play by I where $x = \langle n_m : m < \omega \rangle$ has II playing $\langle B_{n_m}^m(\varepsilon) : m < \omega \rangle$. Since f is calculated term by term from σ , it follows that f is continuous by Corollary 21 B • 4 and $f''\mathcal{N} \subseteq X$. Since \mathcal{N} is $\prod_{i=1}^{0}$, the image is $\sum_{i=1}^{1}$ and therefore lebesgue measurable by Corollary 23 B • 20. Thus $f''\mathcal{N}$ has measure 0. But then by Lemma 28 A • 6, we can cover $f''\mathcal{N}$ by a play $\langle B_{n_i}^i(\varepsilon) : i < \omega \rangle$ by II. So that $\sigma * \langle n_i : i < \omega \rangle \in f''\mathcal{N} \subseteq \bigcup_{i < \omega} B_{n_i}^i(\varepsilon)$ and hence σ wasn't winning. \dashv

Therefore, by DC, for each $0 \neq n < \omega$, II has a winning strategy σ_n for $G_{cov}^{1/n}(X)$. So now we should examine what happens when II wins.

– Claim 2

If II has a winning strategy in $G_{cov}^{\varepsilon}(X)$ then X has lebesgue outer-measure at most $\mu^*(X) \leq \varepsilon$.

Proof .:.

Suppose σ wins for **II**. For each partial play $p \in {}^{<\omega}2$ by **I**, let B(p) be the set **II** plays in response using σ . Since σ wins, every $x \in X$ will be covered by the resulting play by **II**: each $x \in [0, 1]$ can be written in binary so that it can be played by **II** with $\vec{x} = \langle x_n \in 2 : n < \omega \rangle$ and therefore $x \in \bigcup_{p < \vec{x}} B(p)$. Thus

$$X \subseteq \bigcup_{p \in {}^{<\omega} 2} B(p) = \bigcup_{n < \omega} \bigcup_{p \in {}^n 2} B(p),$$

and the measure of $\bigcup_{p \in n_2} B(p)$ is at most $2^n \cdot (\varepsilon/2^{2n}) = \varepsilon/2^n$. Hence the measure of $\bigcup_{p \in <\omega_2} B(p)$ is at most $\sum_{n \in \omega} \varepsilon/2^n = \varepsilon$. So $\mu^*(X) \le \mu^*(\bigcup_{p \in <\omega_2} B(p)) \le \varepsilon$. \dashv

Hence X has outer measure at most 1/n for each $n < \omega$, i.e. X has outer-measure 0. Thus X is lebesgue measurable by Result 23 B • 7. By Lemma 28 A • 7, every subset of [0, 1] is lebesgue measurable. \dashv

As with Corollary 28 A • 4, we get a ZFC-compatible version of this theorem just by showing that for $X \subseteq \mathcal{N}$, the covering game $G_{cov}(X)$ recast as a number game G(A) has A with the same complexity as X whenever X is projective. Doing this explicitly is a little annoying, so we give *most* of a proof.

- 28 A • 10. Corollary -

Assume ZF + DC. Therefore, for $n < \omega$, $\text{Det}(\sum_{n=1}^{n})$ implies every $\sum_{n=1}^{n}$ set is lebesgue measurable, and similarly for $\prod_{n=1}^{n}$.

Proof .:.

Let $X \subseteq [0, 1]$ be $\sum_{n=1}^{1}$. The "minimal" measurable $A \supseteq X$ as in Lemma 23 B • 13 is the countable intersection of open sets and thus is $\Delta_{1}^{1} \subseteq \prod_{n=1}^{1}$. Hence $A \setminus X$ is $\prod_{n=1}^{1}$ and every measurable subset of $A \setminus X$ is null. Note that $\mathsf{Det}(\prod_{n=1}^{1})$ holds by Result 27 A • 4 and $\mathsf{Det}(\sum_{n=1}^{1})$.

Let $f : C \to [0, 1]$ be defined by $f(x) = \sum_{n < \omega} x(n)/2^{n+1}$ which is continuous by Corollary 21 B • 4 and surjective. Hence $f^{-1} X \in \sum_{n=1}^{n} f$ in both \mathcal{N} and \mathcal{C} . To show $A \setminus X$ is Lebesgue measurable, note that the covering game $G_{\text{cov}}^{1/N}(A \setminus X)$ for $N < \omega$ is equivalent to the number game $G(B_N)$ where $B_N \subseteq \mathcal{N}$ is defined by the set of all $x \in \mathcal{N}$ such that

- 1. for every $n < \omega$, x(2n) is either 0 or 1;
- 2. for every $n < \omega$, x(2n + 1) is the code of a sequence $\vec{\tau}_n$ of an even number of elements of $\sqrt[-\omega]{2}$ such that
 - $f(\vec{\tau}_n(m))$ is rational for every $m < \operatorname{dom}(\vec{\tau}_n)$ (a Δ_{1}^{1} -property),
 - $f(\vec{\tau}_n(m)) < f(\vec{\tau}_n(m+1))$ for $m + 1 < \text{dom}(\vec{\tau}_n)$, and
 - $\sum_{2m+1 < \text{dom}(\vec{\tau}_n)} f(\vec{\tau}_n(2m+1)) f(\vec{\tau}_n(2m)) < \frac{1}{N \cdot 4^{n+1}} (a \Delta_1^1 \text{-property});$
- 3. there is some $n < \omega$ such that for every $m \in \text{dom}(\vec{\tau}_n), \vec{\tau}_n(m) \triangleleft \text{even}(x)$; and

4.
$$even(x) \in f^{-1}"X$$
.

Each B_N will therefore be $\exists^{\omega} \prod_n^1 = \prod_n^1$ and so each $G(B_N)$ and $G_{cov}^{1/N}(A \setminus X)$ are determined. By the proof of AD + DC Implies Lebesgue Measurability (28 A • 9), it follows that $A \setminus X$, and hence X, is lebesgue measurable. \dashv

The final result tells us that all sets of reals have the baire property (recall Definition 23 C \cdot 3 that a set has the baire property iff every set is meagre modulo an open set). The proof of this regularity property is perhaps the hardest. The previous proof that all sets of reals are lebesgue measurable under AD + DC required a fair amount of background, but the resulting proof was fairly straightforward. The reverse will be the case for the baire property from AD + DC, and will introduce the first instance of the technique of "copying" a strategy from an "auxiliary" game.

The idea behind copying is that we have two games, G_1 and G_2 , or possibly more. We suppose that one of the players has a strategy in G_2 , or possibly more, and we use this strategy to generate one in G_1 by copying moves.





The game we will use for the baire property will be a *banach-mazur* game, named after the Polish mathematicians Stefan Banach and Stanisław Mazur. The idea behind the game is that I and II play a decreasing sequence of basic open sets. I tries to having the resulting limit (the intersection) intersect a given desired set X while II tries to avoid X.

- 28 A • 12. Definition
Let
$$X \subseteq \mathcal{N}$$
 be given. The *banach-mazur game* for X is the game $G_{bm}(X)$ that takes the form
I: \mathcal{N}_{τ_0} \mathcal{N}_{τ_2} \cdots
II: \mathcal{N}_{τ_1} \mathcal{N}_{τ_3} \cdots
where $\tau_n \triangleleft \tau_{n+1} \in {}^{<\omega}\omega$ for all $n < \omega$. We say I wins iff $X \cap \bigcap_{n < \omega} \mathcal{N}_{\tau_n} \neq \emptyset$.

This is similar to the variant game in the proof of AD + DC Implies PSP (28 A • 2). There, one of the players could play arbitrarily large finite binary sequence where the other could play only one digit at a time. Here, both players can play arbitrarily large finite sequences of integers, so the game is more "balanced" in this way. In this way, the winning condition is similar to number games: if I and II instead just play $\tau_n \in {}^{<\omega}\omega$, the two build up a real $x \in \mathcal{N}$ where then I wins iff $x \in X$.

Recall the following characterization of being nowhere dense from Result 23 C • 6: a set X is nowhere dense iff $X \cap U$ is not dense in (the inherited topology on) U for any open U.

The basic idea now is that II has a winning strategy in the banach–mazur game $G_{bm}(X)$ iff X is meagre, and I has a winning strategy in it iff $\mathcal{N}_{\sigma} \setminus X$ is meagre for some σ . Hence AD allows us to enforce that one or the other happens all the time.

```
28 A • 13. Theorem (AD + DC implies BP)
Assume ZF + DC + AD. Therefore every set X \subseteq \mathcal{N} has the baire property.
```

Proof .:.

Consider the banach–mazur game $G_{bm}(X)$ for $X \subseteq \mathcal{N}$ arbitrary. Firstly, it's not too difficult to see that II wins if X is meagre.
— Claim 1

If X is meagre then II has a winning strategy for $G_{bm}(X)$.

Proof .:.

Suppose X is meagre, and write $X = \bigcup_{n < \omega} X_n$ where each X_n is nowhere dense. We define a winning strategy for **II** using DC and diagonalizing through each X_n . By Result 23 C • 6, inductively, suppose p is a partial play where **I** just played τ and it is now **II**'s nth turn. By Result 23 C • 6, since X_n is nowhere dense, $X_n \cap \mathcal{N}_{\tau}$ is not dense, meaning there is an open set (and hence basic open set by taking a subset) $W \subseteq \mathcal{N}_{\tau}$ where $X_n \cap W = \emptyset$. Such a W can take the form \mathcal{N}_{π} for some π which therefore has $\pi \triangleright \tau_n$. By DC we may define $\sigma(p)$ as \mathcal{N}_{π} . It follows that σ wins for **II** since at **II**'s nth turn in the partial play p, we ensure $X_n \cap \bigcap_{\tau < p} \mathcal{N}_{\tau} = \emptyset$. It follows that in the resulting play $\langle \mathcal{N}_{\tau_n} : n < \omega \rangle$, $X \cap \bigcap_{n < \omega} \mathcal{N}_{\tau_n} = \bigcup_{n < \omega} X_n \cap \bigcap_{n < \omega} \mathcal{N}_{\tau_n} = \emptyset$ and so **II** wins.

Actually this is an equivalence.

— Claim 2

If **II** has a winning strategy for $G_{bm}(X)$ then X is meagre.

Proof .:.

Suppose II has a winning strategy σ . Let p a partial according to σ , where II just played. We can say $x \in X$ is *rejected* at stage p iff $x \in \bigcap_{n < \ln(p)} p(n)$ but for all plays \mathcal{N}_{τ} by $\mathbf{I}, x \notin \bigcap_{n < \ln(p)} p(n) \cap \sigma(p^{\frown}(\mathcal{N}_{\tau}))$, i.e. II playing according to the strategy ensures x isn't in the resulting set.

Since σ is winning, every $x \in X$ is rejected by some partial play *p*—since otherwise using DC to choose I's moves at each stage—there's a play of the game that results in x being in the intersection, contradicting that σ won. We can define

$R_p = \{x \in X : x \text{ is rejected at } p\}.$

Thus $X = \bigcup_{p \in {}^{<\omega_{\omega}}} R_p$. Previously in Claim 2 of AD+DC Implies PSP (28 A • 2), $|R_p| = 1$. Here, however, R_p will be nowhere dense which tells us that $X = \bigcup_{p \in {}^{<\omega_{\omega}}} R_p$ is meagre: the idea is that if x is rejected at p, x can't be rejected at a later stage too. More explicitly, suppose $x \in R_p$ and \mathcal{N}_{τ} is the last move in p, and it's I's turn. Suppose I plays $x \upharpoonright n$ for an arbitrary $n > \ln(\tau)$. Responding with $\sigma(p^{\frown}(\mathcal{N}_{x \upharpoonright n})) = \mathcal{N}_{x \upharpoonright n \frown \tau'}$, it follows that $x \upharpoonright n \frown \langle \tau'(0) + 1 \rangle$ disagrees with σ and with x: the disagreement isn't due to I's turn since it disagrees with x, and the disagreement isn't due to II since it disagrees with σ . Hence any resulting play from that point on wasn't rejected at stage $p: \mathcal{N}_{x \upharpoonright n \frown \langle \tau'(0) + 1 \rangle} \cap R_p = \emptyset$. Thus if $U \subseteq \mathcal{N}$ is (basic) open and $x \in R_p \cap U$, then we have found a (basic) open set $W \subseteq U$ such that $R_p \cap W = \emptyset$ and so R_p is nowhere dense by Result 23 C • 6.

This tells us what happens if **II** wins $G_{bm}(X)$, and tells us X (trivially) has the baire property in such cases. What happens when **I** wins things are slightly less direct because it only gives partial information about X.

— Claim 3

I has a winning strategy for $G_{bm}(X)$ iff $\mathcal{N}_{\sigma} \setminus X$ is meagre for some $\sigma \in {}^{<\omega}$.

DETERMINACY AND POINTCLASS PROPERTIES

Proof .:.

Suppose σ is a winning strategy for I in $G_1 = G_{bm}(X)$. Let $\mathcal{N}_{\sigma(\emptyset)}$ be the first move by I. Now consider the auxiliary game $G_2 = G_{bm}(\mathcal{N}_{\sigma(\emptyset)} \setminus X)$. It follows that II has a winning strategy for this game by copying moves. More precisely, suppose I plays \mathcal{N}_{τ_0} in G_2 . Without loss of generality (since otherwise I immediately loses in that game) suppose τ_0 extends σ . Then we have II copy this move in G_1 and see how I responds there with \mathcal{N}_{τ_1} using σ . We copy that move as player II in G_2 and so on: I in G_1 always responds using σ and everything else is generated using arbitrary plays by I in G_2 .



Following the arrows above gives a winning strategy for **II** in G_2 since both games have the same resulting intersection: $X \cap \bigcap_{n < \omega} \mathcal{N}_{\tau_n} \neq \emptyset$ in G_1 means $\bigcup_{n < \omega} \tau_n = x \in X$ so that $(\mathcal{N}_{\sigma(\emptyset)} \setminus X) \cap \bigcap_{n < \omega} \mathcal{N}_{\tau_n} = \{x\} \setminus X = \emptyset$ and hence **II** has won in G_2 . By Claim 2, $\mathcal{N}_{\sigma(\emptyset)} \setminus X$ is meagre.

The same idea applies in the reverse direction to tell us that if $\mathcal{N}_{\sigma} \setminus X$ is meagre for some σ then I has a winning strategy for $G_{bm}(X)$: let τ be winning for II in $G_{bm}(\mathcal{N}_{\sigma} \setminus X)$ and copy moves as below.



So I in $G_{bm}(X)$ take's II move, plays it as I in $G_{bm}(\mathcal{N}_{\sigma} \setminus X)$ and then uses the strateegy τ to determine II's move there, and then copies it in $G_{bm}(X)$.

So how does this tell us that X has the baire property? The unfortunate fact is that the determinacy of $G_{bm}(X)$ doesn't in general tell us that X has the baire property. We instead need to assume the determinacy of $G_{bm}(X \setminus S)$ for a certain set S. In particular, define

$$S = \bigcup \{ \mathcal{N}_{\sigma} : \sigma \in {}^{<\omega} \omega \land \mathcal{N}_{\sigma} \setminus X \text{ is meagre} \}.$$

In essence, S is the best approximation of $\mathcal{N} \setminus X$ modulo meagre sets. It's not hard to see that S is open and $S \setminus X$ is meagre, given that it's the countable union of meagre sets (cf. Lemma 23 C • 4 (4)).

Thus we have the final proposition to prove the theorem: if $G_{bm}(X \setminus S)$ is determined then X has the baire property. To see this, if I has a winning strategy, then $\mathcal{N}_{\sigma} \setminus (X \setminus S)$ is meagre for some $\sigma \in {}^{<\omega}\omega$ by Claim 3. But then as a subset, $\mathcal{N}_{\sigma} \setminus X \subseteq \mathcal{N}_{\sigma} \setminus (X \setminus S)$ is meagre. Hence $\mathcal{N}_{\sigma} \subseteq S$. Yet this implies $\mathcal{N}_{\sigma} \setminus (X \setminus S) = \mathcal{N}_{\sigma}$, which isn't meagre, a contradiction. So I cannot have a winning strategy, and by determinacy, II wins. By Claim 2, $X \setminus S$ is meagre. Thus $X \bigtriangleup S = (S \setminus X) \cup (X \setminus S)$ is the union of two meagre sets and is hence meagre. Since S is open, it follows that X has the baire property.

As before, we get a ZFC-compatible version of this theorem.

- 28 A • 14. Corollary Assume ZF + DC. Therefore, for $n \in \omega$, $Det(\sum_{n=1}^{n})$ implies $BP(\sum_{n=1}^{n})$, and similarly for $\prod_{n=1}^{n}$. DETERMINACY AND POINTCLASS PROPERTIES

Сн V §28 А

Proof .:.

Let $X \in \sum_{n=1}^{1}$. Let S be as in the proof of AD + DC implies BP (28 A • 13):

$$S = \bigcup \{ \mathcal{N}_{\sigma} : \sigma \in {}^{<\omega}\omega \land \mathcal{N}_{\sigma} \setminus X \text{ is meagre} \}.$$

As an open set, $S \in \sum_{1}^{0}$ and hence $X \setminus S$ is $\sum_{n=1}^{1}$. The determinacy of the banach-mazur game $G_{bm}(A)$ for any $A \subsetneq \mathcal{N}$ is equivalent to the determinacy of the number game G(B) where B consists of all $x \in \mathcal{N}$ such that

1. For each $n < \omega$, $x(n) = \text{code}(\tau_n)$ for some $\tau \in {}^{<\omega}\omega$; and

2. $\tau_0 \tau_1 \tau_2 \cdots \in A$.

Using a computable coding, consider the map taking $x \in \mathcal{N}$ to $\operatorname{code}^{-1}(x(0)) \cap \operatorname{code}^{-1}(x(1)) \cap \cdots$ if each $\operatorname{code}^{-1}(x(n))$ exists, and otherwise x is mapped to some element of $\mathcal{N} \setminus A$. This map $f : \mathcal{N} \to \mathcal{N}$ is clearly continuous and $f^{-1} = B$ therefore has the same projective complexity as A whenever A is projective. Hence the determinacy of $G_{\operatorname{bm}}(X \setminus S)$ is equivalent to the determinacy of G(B) for some $B \in \prod_{n=1}^{n}$. Since $\operatorname{Det}(\prod_{n=1}^{n})$ holds, we get that X has the baire property by the proof of $\operatorname{AD} + \operatorname{DC}$ implies BP (28 A • 13).

We can therefore collect together Corollary 28 A • 4, Corollary 28 A • 10, and Corollary 28 A • 14 to get the following.

– 28 A • 15. Corollary

Assume ZF+DC. Therefore, for $n < \omega$, $Det(\sum_{n=1}^{n})$ implies every $\sum_{n=1}^{n}$ -set has the perfect set property, the baire property, and is lebesgue measurable.

Such a statement is weaker than pure AD, but more interestingly, compatible with ZFC despite being strictly stronger.ⁱⁱⁱ

- 28 A • 16. Definition *Projective determinacy* (PD) is the statement that $\text{Det}(\Sigma_n^1)$ holds for every $n < \omega$.

Relative to the existence of sufficiently large cardinals, PD is consistent with ZFC [?], allowing us to make use of the techniques of determinacy ideas even outside the context where AC fails.

Corollary 28 A • 15 is not the end of the story as far as the implications of $\text{Det}(\sum_{n=1}^{1})$ on regularity properties are concerned. In particular, we can go one step further into $\sum_{n=1}^{1}$.

- 28 A • 17. Theorem Assume $ZF + DC + Det(\sum_{n=1}^{n})$. Therefore every $\sum_{n=1}^{n}$ set has the perfect set property, the baire property, and is lebesgue measurable.

This gives an alternative proof of Corollary 23 A • 21, Corollary 23 C • 10, and Corollary 23 B • 20 using Closed Determinacy (27 A • 3) or Open Determinacy (27 A • 5).^{iv} The general idea behind this is a technique called *unfolding*. In what sense we are "unfolding" something is unclear to me. Nevertheless, the idea behind the technique is to represent $X \in \sum_{n+1}^{1}$ as the projection $\exists^{\mathcal{N}} Y$. In most of the games we play, we would attempt to find a certain element of X. With unfolding, we instead play modified games that have one of the players attempt to find not only $x \in X$ but also guess the y paired with x in the simpler set $Y: \langle x, y \rangle \in Y$. The determinacy of the simpler games about Y allows us to make similar conclusions as before just by ignoring the chosen y.

- Proof of Theorem 28 A • 17 ...

Note that $\text{Det}(\sum_{n=1}^{1})$ implies $\text{Det}(\prod_{n=1}^{1})$ by Result 27 A • 4.

For the perfect set property, argue as in AD + DC Implies PSP (28 A • 2) and Corollary 28 A • 4 to work with an arbitrary $X \in \sum_{n+1}^{1} X \subseteq C$. Let $Y \in \prod_{n=1}^{1} have X = pY, Y \subseteq \mathcal{N} \times C$. Consider the game $\hat{G}'(X)$ which takes the form

 $\hat{G}'(X) \qquad \begin{array}{ccc} \mathbf{I}: & s_0 \in {}^{<\omega}2, y_0 \in \omega & s_1, y_1 & \cdots \\ \mathbf{II}: & x_0 \in 2 & x_1 & \cdots \end{array}$

ⁱⁱⁱIn particular, ZFC + Det(\sum_{1}^{1}) implies PSP(\prod_{1}^{1}) and thus $\omega_{1}^{L[x]} < \omega_{1}$ by Theorem 25 B • 16 and so ω_{1}^{V} is (weakly) inaccessible in L[x] for every $x \in \mathcal{N}$ by Theorem 25 B • 15. Hence the consistency of ZFC + Det(\sum_{1}^{1}) implies the consistency of the existence of a weakly inaccessible cardinal.

^{iv}To be fair, we used Corollary 23 B \cdot 20 in the proof of AD + DC Implies Lebesgue Measurability (28 A \cdot 9).

writing $r = s_0^{(x_0)} (x_1)^{(x_1)} \cdots$ and $y = (y_n : n < \omega)$, we say I wins if $(r, y) \in Y$. As with Corollary 28 A • 4, we can recast this as a number game G(B) for some $B \in \prod_{n=1}^{1}$. The same proof as Claim 1 of AD + DC Implies PSP (28 A • 2) implies that I has a winning strategy σ implies pY has a perfect subset: the map taking $x \in C$ to r from the resulting play (r, y) via σ is continuous, injective, and witnesses that there is a perfect subset of X by Lemma 23 A • 4. Similarly, if II wins, just as before with Claim 2 of AD + DC Implies PSP (28 A • 2), pY will be countable since everything will be rejected at some partial play p. For any fixed y(n) in p, there can be only one possible $x \in C$ rejected at the nth stage p. Since there are only countably many y(n), the set of $x \in C$ rejected at p is countable. Hence the countable union of all $x \in C$ rejected somewhere is countable, and so all of pY = X is countable. By determinacy, one of the players has a winning strategy and thus PSP(X) holds.

For lebesgue measurability, we argue again as in Corollary 28 A • 10 to work with $X \subseteq [0, 1], X \in \sum_{n+1}^{1}$, such that every measurable subset of X is null. Let $X = \mathfrak{p}Y$ for $Y \subseteq [0, 1] \times \mathcal{N}, Y \in \prod_{n=1}^{1}$. Now consider the variant of the covering game $\hat{G}_{cov}^{1/N}(X)$ for $N < \omega$ which takes the form

$$\hat{G}_{cov}^{1/N}(X)$$
 I: $x_0 \in 2, y_0 \in \omega$ x_1, y_1 ...
II: $B_{n_0}^0(1/N)$ $B_{n_1}^1(1/N)$.

For $x = \sum_{n < \omega} x_n / 2^{n+1}$ and $y = \langle y_n : n < \omega \rangle$, we say that I wins iff $\langle x, y \rangle \in Y \setminus (\bigcup_{n < \omega} B_{n_0}^0(1/N) \times \mathcal{N})$. As with Corollary 28 A • 10, this is equivalent to a number game G(B) for some $B \in \sum_n^1$ and is therefore determined. Just as before in Claim 1 from AD + DC Implies Lebesgue Measurability (28 A • 9), I cannot have a winning strategy here: f mapping (indices of) plays by II into x where $\langle x, y \rangle$ is the resulting play by I is continuous and $f''\mathcal{N} \subseteq X$ is $\sum_{i=1}^{1}$, measurable, and so null. Thus $f''\mathcal{N}$ is covered by a play by II, meaning I couldn't win with the strategy. Thus II has a winning strategy, and without loss of generality, can ignore the sequence $\langle y_n : n < \omega \rangle$. The same idea in Claim 2 of AD + DC Implies Lebesgue Measurability (28 A • 9) tells us that the outer measure of X is at most 1/N. To give a sketch, if σ is the strategy for II, let B(p) be the set II plays in response using σ . Without loss of generality, for any $\langle x, y \rangle \in Y$ played by I, x will be in B(p) for some initial segment of the resulting play p. By a counting argument, we can ensure that the outer-measure of what II could play is at most 1/N. Since this holds for each $N \in \omega$, X is null and so measurable.

For the baire property, let $X \in \sum_{n+1}^{1}$, $X \subseteq \mathcal{N}$ be arbitrary. Let S be the open set as in AD + DC implies BP (28 A • 13):

$$S = \bigcup \{ \mathcal{N}_{\sigma} : \sigma \in {}^{<\omega}\omega \land \mathcal{N}_{\sigma} \setminus X \text{ is meagre} \}.$$

Note that $X \setminus S$ is therefore still \sum_{n+1}^{1} , and $S \setminus X$ is meagre. so let $Y \in \prod_{n=1}^{1} X$ witness $X \setminus S = \mathfrak{p}Y$. Consider the variant of the banach-mazur game $\hat{G}_{bm}(Y)$ which takes the form

$$\hat{G}_{\rm bm}(Y) \qquad \begin{array}{cccc} {\bf I}: & y_n \in \omega, \, \mathcal{N}_{\tau_0} & y_1, \, \mathcal{N}_{\tau_2} & \cdots \\ {\bf II}: & \mathcal{N}_{\tau_1} & \mathcal{N}_{t_3} & \cdots \end{array}$$

where $\tau_n \triangleleft \tau_{n+1}$ for all $n < \omega$ and where I wins iff $\langle x, y \rangle \in Y$ where $x = \bigcup_{n < \omega} \tau_n$ and $y = \langle y_n : n < \omega \rangle$. Again as with Corollary 28 A • 14, this can be recast as a number game G(B) for some $B \in \prod_n^1$ and so the game is determined. Similarly to before with Claim 2 of AD + DC implies BP (28 A • 13), if II has a winning strategy, then X is meagre, sense we can in effect ignore y: the points in $X \setminus S$ rejected at each stage is meagre and so $X \setminus S$ will be meagre, meaning $X \bigtriangleup S = (S \setminus X) \cup (X \setminus S)$ is meagre and X has the baire property. If I has a winning strategy σ , then as before, we can just copy this winning strategy in the complement game where we ignore y: $G_{bm}(\sigma(\emptyset) \setminus (X \setminus S))$ to tell us by Claim 2 of AD + DC implies BP (28 A • 13) that $\sigma(\emptyset) \setminus (X \setminus S)$ is meagre. But then $\sigma(\emptyset) \setminus X \subseteq \sigma(\emptyset) \setminus (X \setminus S)$ is meagre so that $\sigma(\emptyset) \subseteq S$, implying $\sigma(\emptyset) \setminus (X \setminus S) = \sigma(\emptyset)$ which isn't meagre, a contradiction. Hence II wins and X has the baire property.

It's important to realize that $\text{Det}(\sum_{n=1}^{1})$ merely implies these regularity properties for $\sum_{n=1}^{1}$, it doesn't imply $\text{Det}(\sum_{n=1}^{1})$ (otherwise Open Determinacy (27 A • 5) would imply PD).

§ 28 B. Basic consequences for measure

Obviously AD, as a statement about real numbers, has a lot to say about sets of real numbers as we've seen. But beyond

DETERMINACY AND POINTCLASS PROPERTIES

this, there are also consequences for the rest of set theory, especially in relation to large cardinal hypotheses and choice principles.

For example, we've thus far been working with the theory ZF + DC + AD, but as stated, it's unknown whether AD already implies DC. Indeed, we already have that AD implies a very simple form of choice: countable choice for sets of reals.

- **28 B · 1.** Result Assume ZF. Therefore AD implies that for every countable family $\{X_n \subseteq \mathcal{N} : n < \omega\}$ such that each $0 < |X_n| \le \aleph_0$, there is a choice function $f \in \prod_{n < \omega} X_n$.

Proof .:.

Consider the game that takes the following form:

I: $n \in \omega$ II: $x(0) \in \omega$ x(1) x(2) ...

where I wins iff $x = \langle x(m) : m < \omega \rangle \notin X_n$. In other words, I chooses which X_n to play, and II tries to get into X_n with their moves. This is clearly equivalent to a number game and is hence determined by AD. Clearly I doesn't have a winning strategy since for any particular $n \in \omega$ that I plays, II can play any particular $x \in X_n \neq \emptyset$. Thus II wins with some strategy σ . If $n \in \omega$ is a play by I, we can let $\zeta(n) \in \mathcal{N}$ be the resulting real played by II's moves using σ . Since σ wins, $\zeta(n) \in X_n$ for each $n < \omega$ and hence ζ is a choice function in $\prod_{n < \omega} X_n$. \dashv

As a result, ω_1 is regular in ZF + AD, which is not as trivial as is the case with ZFC. A decent exercise is to identify precisely where the assumption of AC comes into play in the standard proof that κ^+ is regular for any cardinal κ , Result 5 D • 20.

Note that the baire property holding for all sets of reals implies that there are no ω_1 -length sequences of (distinct) reals, another way in which AC fails in ZF + DC + AD, just as with Corollary 28 A • 3. One can also say that in ZF + DC + AD that \mathbb{R} is *amorphous* in the sense that its cardinality is not directly comparable with cardinals.

- **28 B · 2.** Result Assume ZF + DC + AD. Therefore there is no injection $f : \omega_1 \to \mathcal{N}$.

Proof .:.

Without loss of generality, consider $f : \omega_1 \to \mathbb{R}$ an injection. Consider a construction of the vitali set. In particular, for each $a, b \in f''\omega_1$, write $a \approx b$ iff $a - b \in \mathbb{Q}$. Let $X \subseteq f''A$ be a set of representatives of \approx -equivalence classes. Note that X must be uncountable since otherwise $f''A \subseteq \{x + q : x \in X \land q \in \mathbb{Q}\}$ is contained in a countable set. But X—as a *Vitali* set—is not lebesgue measurable, nor has the baire property by Result 23 B • 16 and Result 23 C • 11. This contradicts AD + DC Implies Lebesgue Measurability (28 A • 9) and AD + DC implies BP (28 A • 13).

– 28 B•3. Corollary -

Assume ZF + DC + AD. Therefore there is no ordinal α such that $\mathcal{N} =_{\text{size}} \alpha$.

Proof .:.

By Result 28 B • 2, the only bijection $f : \alpha \to \mathcal{N}$ that can exist has $\alpha < \omega_1$. Thus the minimal such α is ω . But $\mathcal{N} \neq_{\text{size}} \omega$ by Cantor's Theorem (5 B • 13). \dashv

This has quite a lot of consequences in relation to large cardinals. In particular, ω_1 is measurable in ZF + DC + AD. One might think this is impossible since measurable cardinals need to be limit cardinals and indeed strongly inaccessible. But in the world without choice, such oddities are possible, partially because any non-principal ultrafilter over ω , recast as a subset of \mathcal{N} , doesn't have the baire property and isn't lebesgue measurable [?needcitation].

Proof .:.

Suppose U is a non-principal ultrafilter over ω . Consider the game G that takes the form

where each $s_i \in [\omega]^{<\omega}$ is disjoint from all the previous $s_k, k < i$. We say I wins iff $\bigcup_{i < \omega} s_{2i} \in U$. Hence II wins iff the complement is in U. This game is straight-forwardly recast as a number game G(A) for some $A \subseteq \mathcal{N}$, and hence is determined.

Suppose I has a winning strategy σ . Consider playing an auxiliary game as II and copying moves to generate a winning strategy for II. Write s_0 for $\sigma(\emptyset)$ and for any $s_1 \in [\omega]^{<\omega}$, play as follows.



The resulting play by I in the original is the set $\bigcup_{n < \omega} s_{2n}$ which is disjoint from $\bigcup_{n < \omega} s_{2n+1}$ and in the auxiliary game, I's play is $s_0 \cup \bigcup_{0 < n < \omega} s_{2n+1}$. Because I uses σ in both games, both sets are in U but have intersection $s_0 \in [\omega]^{<\omega}$ in U. But U, being non-principal, contains no finite sets, a contradiction. Hence U could not be non-principal.

- 28 B • 5. Theorem -

Assume ZF + AD. Therefore every non-principal ultrafilter is ω_1 -complete.

Proof .:.

Let U be a non-principal ultrafilter over some set X. Let $X_n \in U$ for $n < \omega$. To see that $\bigcap_{n < \omega} X_n$ is in U, suppose not: $\bigcap_{n < \omega} X_n \notin U$. Without loss of generality,

1. $X_0 = X;$

2. $X_{n+1} \subseteq X_n$ for each $n < \omega$; and

3. $\bigcap_{n < \omega} X_n = \emptyset$ (consider $\bigcap_{i \le n} X_i \setminus \bigcap_{i < \omega} X_i \in U$ instead of X_n for $0 < n < \omega$).

Now let $f: X \to \omega$ be defined by f(x) = n iff *n* is the least such that $x \notin A_n$ so that *f* is surjective. Define an ultrafilter μ over ω by

$$\mu = \{ A \subseteq \omega : f^{-1} " A \in U \}.$$

It's straightforward to see that μ is an filter. To see that μ is an ultrafilter, if $A \notin \mu$ then $f^{-1} A \notin U$ and hence $X \setminus f^{-1} A \in U$. Since $X = f^{-1} \omega, X \setminus f^{-1} A = f^{-1} (\omega \setminus A) \in U$ and hence $\omega \setminus A \in \mu$.

 μ is non-principal by the assumption that $\bigcap_{n < \omega} X_n \notin U$: suppose $\mu = \{A \subseteq \omega : N \in A\}$ for some fixed $N < \omega$. Since, for every $n < \omega$, $X_n \in U$, $f''X_n \in \mu$ and thus $N \in f''X_n$. In particular, $N \in f''X_N$, which is impossible: any x with f(x) = N has $x \notin X_N$ by definition of f. Thus μ is principal, contradicting Lemma 28 B•4. Hence U must be ω_1 -complete.

This alone doesn't tell us that ω_1 is measurable since we have not guaranteed the existence of a non-principal ultrafilter over ω_1 . Nevertheless, there will be, and in fact, there will be many such measures. In fact, in ZF + DC + AD, ω_1 and ω_2 are in fact measurable. A theorem of Steel actually tells us that for "small" uncountable cardinals in L(\mathcal{N}), assuming L(\mathcal{N}) \models AD, regularity is equivalent to measurability in L(\mathcal{N}). This idea in L(\mathcal{N}) is something which an assumption AD⁺ outright implies with a proof due to Woodin.

§ 28 C. Two periodicity theorems

Recall the situation of the prewellordering property, the scale property, uniformization, the reduction property, and the

separation property, all of which we have for Π_1^1 or (Σ_1^1 in the case of separation): cf. PWO(Π_1^1) (25 C • 10), Π_1^1 Scale Property (25 E • 7), Π_1^1 -Uniformization (25 E • 8), Result 25 D • 2, and Result 25 D • 4). Through results like Theorem 25 C • 11, we also have gotten PWO(Σ_2^1) and Scale(Σ_2^1) in the context of ZFC. In the context of AD or PD, we actually get much more as in Figure 25 C • 15 reproduced below.



28 C • 1. Figure: Analytical $\Gamma \neq \Delta_n^1$ such that $ZF + DC + PD \models PWO(\Gamma)$

We also get separation, reduction, uniformization and so forth throughout the analytical and projective hierarchies. We can also propogate the scale property when making use of "very good" scales. These represent two *periodicity* theorems that allow us to propogate the prewellordering and scale properties throughout the analytic and projective hierarchies in a technique due to Yiannis Moschovakis ($\Gamma i \alpha v v \eta \varsigma Mo \sigma \chi o \beta \alpha \kappa \eta \varsigma$), which allows us to go from Σ -like pointclasses to Π -like pointclasses. The original technique from Theorem 25 C • 11 (which required no choice axiom) allows us to go from Π -like pointclasses to Σ -like pointclasses and thus in the context of determinacy, we can go back and forth through the entire analytical hierarchy.

28 C · 2. Theorem (The First Periodicity Theorem) Assume $ZF + DC + Det(\Gamma \cap \neg \Gamma)$ and $PWO(\Gamma)$ where $\Gamma \subseteq \mathcal{P}(\mathcal{N})$ is an adequate pointclass. Therefore $PWO(\forall^{\mathcal{N}}\Gamma)$. *Proof* .:. Let $A \in \forall^{\mathcal{N}}\Gamma$ and $B \in \Gamma$ be such that $A = \forall^{\mathcal{N}}B$. We want to define a $\forall^{\mathcal{N}}\Gamma$ -norm on A. So let φ be a Γ -norm on B and consider the following *sup game*: given $x, y \in A$, let $G_{x,y}$ be the game

where **II** wins iff $\varphi(x, a) \leq \varphi(y, b)$ where $a = \langle a_n : n < \omega \rangle$ and $b = \langle b_n : n < \omega \rangle$. Basically, **II** wants to play the bigger ordinal. Note that this is well defined since $x, y \in A = \forall^{\mathcal{N}} B$ and hence $\langle x, a \rangle, \langle y, b \rangle \in B$ for all $a, b \in \mathcal{N}$. The idea here is that $G_{x,y}$ takes the form of a number game G(X) for some $X \in \Gamma \cap \neg \Gamma$ by Result $25 C \cdot 8: \vec{x} \leq_{\varphi} \vec{y}$ is $\Gamma \cap \neg \Gamma$ whenever $\vec{x}, \vec{y} \in B$ and so

$$X = \{a * b : \varphi(x, a) \le \varphi(y, b)\} \in \Gamma \cap \neg \Gamma.$$

Thus $G_{x,y}$ is determined for each $x, y \in A$. This will give us a relation on A defined by $x \leq^* y$ iff II has a winning strategy in $G_{x,y}$. The result is that \leq^* will be a $\forall^{\mathscr{N}} \Gamma$ -prewellorder, and we now aim to show this.

Note that \leq^* is reflexive since **II** can merely copy **I**'s moves:

The resulting play of each player is $a = \langle a_n : n < \omega \rangle$ where then clearly in $G_{x,x}$, II wins since $\varphi(x, a) \le \varphi(x, a)$.

- Claim 1 - \leq^* is transitive.

DETERMINACY AND POINTCLASS PROPERTIES

Proof ...

Suppose $x \leq^* y \leq^* z$. We want to show II wins $G_{x,z}$ and we do this by playing an intermediary game since \leq_{φ} is already transitive: for some $a, b, c \in \mathcal{N}, \varphi(x, a) \leq \varphi(y, b) \leq \varphi(z, c)$ and so we need to play in a way such that $G_{x,z}$ has the resulting play as a, c. So let $a = \langle a_n : n < \omega \rangle \in \mathcal{N}$ be an arbitrary play by I in $G_{x,z}$. We simultaneously play $G_{x,y}$ and $G_{y,z}$ in a way that generates a winning strategy for II in $G_{x,z}$. Let σ_y be a winning strategy for II in $G_{x,y}$ and let σ_z win for II in $G_{y,z}$. In essence, the strategy for II in $G_{x,y}$ is just " $\sigma_z \circ \sigma_y$ " in a refined, technical sense.



There are thus three reals played: $a, b = \langle b_n : n < \omega \rangle$, and $c = \langle c_n : n < \omega \rangle$. The result is that **II** wins $G_{x,y}$ by using σ_y and thus $\varphi(x, a) \leq \varphi(y, b)$. Since **II** wins $G_{y,z}$ by using σ_z , we also have $\varphi(y, b) \leq \varphi(z, c)$. But these are the reals played in $G_{x,z}$, meaning II has won there— $\varphi(x,a) \leq \varphi(y,b) \leq \varphi(z,c)$ —and the strategy detailed above thus gives a winning strategy in $G_{x,z}$.

Claim 2 -

 \leq^* is total over A: $x \leq^* y$ or $y \leq^* x$ for all $x, y \in A$.

Proof .:.

Suppose $x \not\leq^* y$ so that II doesn't have a winning strategy for $G_{x,y}$. By $\mathsf{Det}(\Gamma \cap \neg \Gamma)$, I wins $G_{x,y}$ with some strategy σ . Using σ , we show II wins $G_{y,x}$ merely by copying I's moves as follows, writing $\sigma(\emptyset) = a_0$, and I's moves $b = \langle b_n : n < \omega \rangle \in \mathcal{N}$ in $G_{x,y}$ as arbitrary:



Hence there are only two reals played: $a = \langle a_n : n < \omega \rangle$ and b. Since I wins $G_{y,x}, \varphi(y,a) \nleq \varphi(x,b)$ which means $\varphi(x, b) \leq \varphi(y, a)$ and thus **II** has won $G_{x,y}$.

So that that remains to show \leq^* is a prewellorder is that it is well-founded.

Claim 3

 \leq^* is well-founded

DETERMINACY AND POINTCLASS PROPERTIES

Proof .:.

Suppose $\langle x_n : n < \omega \rangle$ is <*-decreasing, meaning II wins G_{x_{n+1},x_n} but loses $G_{x_n,x_{n+1}}$. Since I therefore wins $G_{x_n,x_{n+1}}$, using DC, we choose a strategy σ_n for I in $G_{x_n,x_{n+1}}$ for each $n < \omega$. In each $G_{x_n,x_{n+1}}$, the first move $\sigma_n(\emptyset)$, which we call $a_n(0)$, is always determined. But then we just play in $G_{x_n,x_{n+1}}$ the previous move by I in $G_{x_{n+1},x_{n+2}}$:



The resulting plays define $a_i = \langle a_i(n) : n < \omega \rangle$. In all of the boards, I plays with a winning strategy and thus always wins: $\varphi(x_n, a_n) \nleq \varphi(x_{n+1}, a_{n+1})$, meaning $\varphi(x_n, a_n) > \varphi(x_{n+1}, a_{n+1})$. But since this happens for all $n < \omega$, we get a strictly decreasing sequence of ordinals, a contradiction. \dashv

Thus \leq^* is a prewellorder, and all that remains is that \leq^* defines a $\forall^{\mathcal{N}}\Gamma$ -norm. So it suffices to find relations $\leq_1 \in \forall^{\mathcal{N}}\Gamma$ and $\leq_0 \in \neg \forall^{\mathcal{N}}\Gamma = \exists^{\mathcal{N}}\neg \forall^{\mathcal{N}}\Gamma$ as in Result 25 C • 8: for $y \in A$,

 $x \in A \land x \leq^* y$ iff $x \leq_0 y$ iff $x \leq_1 y$.

To do this, notice that we can phrase this in terms of one of the players having or not having a winning strategy:

$$x \in A \land x \leq^{*} y \quad \text{iff} \quad \mathbf{II} \text{ wins } G_{x,y} \text{ with some } \sigma$$

$$iff \quad \exists \sigma \underbrace{\forall a \ (\varphi(x,a) \leq \varphi(y,a*\sigma))}_{\neg \forall^{\mathcal{N}} \Gamma}$$

$$x \in A \land x \leq^{*} y \quad \text{iff} \quad \mathbf{I} \text{ does not win } G_{x,y} \text{ with any } \sigma$$

$$iff \quad \forall \sigma \underbrace{\exists b \ (\varphi(x,\sigma*b) \leq \varphi(y,b))}_{\forall^{\mathcal{N}} \Gamma}$$

It follows by Result 25 C • 8 that \leq^* is a $\forall^{\mathcal{N}} \Gamma$ -norm on A, and since $A \in \forall^{\mathcal{N}} \Gamma$ was arbitrary, $\mathsf{PWO}(\forall^{\mathcal{N}} \Gamma)$ holds. \dashv

- 28 C • 3. Corollary –

Assume $\mathsf{ZF} + \mathsf{DC} + \mathsf{Det}(\Delta^1_n(X))$ for $n < \omega$ and $X \subseteq \mathcal{N}$. Therefore

1. $\mathsf{PWO}(\Pi_m^1(X))$ for odd $m \leq n$, and

2. $\mathsf{PWO}(\Sigma_m^1(X))$ for even $m \le n+1$.

In particular, ZF + DC + PD implies $PWO(\prod_{i=1}^{1}(X))$ and $PWO(\sum_{i=1}^{1}(X))$ for all odd $i < \omega$ and even $j < \omega$.

Proof .:.

Proceed by induction on $n < \omega$. For n = 0, we already have $\mathsf{PWO}(\Sigma_0^1(X))$ by $\mathsf{PWO}(\Sigma_0^1)$ (25 C • 9). Indutively, $\mathsf{Det}(\Delta_{n+1}^1(X))$ implies $\mathsf{Det}(\Delta_n^1(X))$ and therefore $\mathsf{PWO}(\Pi_m^1(X))$ holds for odd $m \le n$ and $\mathsf{PWO}(\Sigma_m^1(X))$ holds for even $m \le n$.

- Suppose *n* is odd. Hence we only need to show $PWO(\Sigma_{n+1}^1(X))$. But this holds by Theorem 25 C 11 and $PWO(\prod_n^1(X))$: $\exists^{\mathcal{N}} \forall^{\mathcal{N}} \prod_n^1(X) = \Sigma_{n+1}^1(X)$.
- If n is even, we need to show PWO(Π¹_{n+1}(X)) and PWO(Σ¹_{n+2}(X)). The first holds by The First Periodicity Theorem (28 C 2) since inductively PWO(Σ¹_n(X)) and Det(Δ¹_n(X)) hold. PWO(Σ¹_{n+2}(X)) therefore follows by PWO(Π¹_{n+1}(X)) and Theorem 25 C 11. ⊢

This also implies separation and reduction in a similar pattern to Figure $28 \text{ C} \cdot 1$ as a result of Result $25 \text{ D} \cdot 2$ (for reduction) and Result $25 \text{ D} \cdot 4$ (for separation in the dual pointclasses).

The second periodicity theorem tells us about how the scale property propogates through the analytic and projective hierarchy, which is exactly the same as with the prewellordering property in Figure 28 C • 1. Recall from Definition 25 E • 1 that a scale is just a sequence of prewellorders that work nicely together and that have a nice limit property. A Γ -scale for Γ a pointclass, from Definition 25 E • 5, is just a scale where the relations on $\langle x, y, n \rangle$ defined by $x \leq_n y$ and $x <_n y$ is in Γ where \leq_n is the *n*th prewellorder of the scale.

For many purposes, arbitrary scales are insufficient, and we need to refine ourselves to work with very good scales.

– 28 C • 4. Definition —

- We say $\vec{\varphi} = \langle \varphi_n : n < \omega \rangle$ is a very good scale on a set $X \subseteq \mathcal{N}$ iff
 - 1. each φ_n is a norm on *X*;
 - 2. for all $x \in {}^{\omega}X$ such that each $\varphi_n \circ x$ is eventually constant, we have
 - a. $\lim x$ exists and $\lim x \in X$, and
 - b. $\forall n < \omega \ (\varphi_n(\lim x) \le \lim(\varphi_n \circ x))$
 - 3. For all $n < \omega, x, y \in X$, if $\varphi_n(x) \le \varphi_n(y)$ then $\forall m < n \ \varphi_m(x) \le \varphi_m(y)$.
- We say $\vec{\varphi}$ is a very good Γ -scale on X iff $\vec{\varphi}$ is a very good scale and a Γ -scale.

Note that (2) this is stronger than the condition in a standard scale from Definition 25 E • 1. For a scale, we require every *convergent* sequence $x \in {}^{\omega}X$ to have (2a) and (2b) hold. But here, the mere fact that $\varphi_n \circ x$ is eventually constant implies that x is convergent. Luckily for us, the property of having a Γ -scale is equivalent to having a very good Γ -scale whenever Γ is adequate. This is basically mimics some of the nice properties of the Π_1^1 -scale on WO.

– 28 C • 5. Result

Assume ZF. For Γ adequate, there is a Γ -scale on $X \subseteq \mathcal{N}$ iff there is a very good Γ -scale on X.

Proof ...

One direction is obvious. So suppose $\vec{\varphi}$ is a Γ -scale on X. Consider $\vec{\psi} = \langle \psi_n : n < \omega \rangle$ defined by, for $x \in X$ and $n < \omega$,

 $\psi_n(x) = \operatorname{code}(\varphi_0(x), x(0), \varphi_1(x), x(1), \cdots, \varphi_n(x), x(n)) \in \operatorname{Ord}$

through a simple coding, which is really the rank of an order we describe, sometimes called *short* lexicographic. We order based on length (2(n + 1) for a fixed n) and then lexicographically, and the rank function for this defines $\vec{\psi}$. It follows that $\psi_n(x) < \psi_n(y)$ iff there is some $m \le n$ such that $x \upharpoonright m = y \upharpoonright m$, $\varphi_i(x) = \varphi_i(y)$ for all i < m, and

• $\varphi_0(x) < \varphi_n(y)$; or

• $\varphi_0(x) = \varphi_n(y)$ and x(n) < y(n).

Similarly, $\psi_n(x) \le \psi_n(y)$ iff the above occurs but with the last inequality replaced with " $\varphi_n(x) \le \varphi_n(y)$ ".

– Claim 1

Assume $\vec{\psi}$ is a scale. Therefore $\vec{\psi}$ is a Γ -scale.

Proof .:.

We need to show that the following relations on triples (x, y, n) are in Γ :

$$x \leq_{\psi_n} y$$
 iff $x \in X \land (y \in X \to \psi_n(x) \leq \psi_n(y))$

$$x <_{\psi_n} y$$
 iff $x \in X \land (y \in X \to \psi_n(x) < \psi_n(y)).$

But this follows from the characterization above: $\psi_n(x) < \psi_n(y)$ iff $x \in X \in \Gamma$ and $\exists m \leq n$ such that

- 1. $x \upharpoonright m = y \upharpoonright m$ (a computable relation); and
- 2. $\forall i < m \ (x \leq_{\varphi_i} y \land y \leq_{\varphi_i} x)$ (a $\forall^{<\omega} \Gamma \land \Gamma = \Gamma$ relation); and
- 3. $x <_{\varphi_m} y \text{ or } x \leq_{\varphi_m} y \land y \leq_{\varphi_m} x \land x(m) < y(m) \text{ (a } \Gamma \lor (\Gamma \land \Gamma) = \Gamma \text{ relation)}.$

Since Γ is adequate, this is in Γ and defines $x <_{\psi_n} y$ and similarly for $\psi_n(x) \le \psi_n(y)$: it asserts $x \in X$, and if $y \notin X$ then m = 0 witnesses the result; whereas if $y \in X$, then clearly the other conditions assert $\psi_n(x) < \psi_n(y)$.

So it suffices to show that $\vec{\psi}$ is a very good scale. Clearly each ψ_n is a norm on X so (1) of Definition 28 C • 4 holds. (3) is straightforward since we have coded previous information into later information: $\psi_m(x)$ is coded in $\psi_n(x)$ for m < n and $\psi_n(x) \le \psi_n(y)$ iff there's a disagreement coded into some previous stage, and that disagreement gives the same inequality. The stages before the least disagreement all have equality, and the stages after the least disagreement all realize that there's an earlier disagreement just like $\psi_n(x)$, $\psi_n(y)$.

For (2) of Definition 28 C • 4, suppose $\vec{x} = \langle x_n : n < \omega \rangle \in {}^{\omega}X$ is such that $\langle \psi_n(x_k) : k < \omega \rangle$ is eventually constant for each $n < \omega$. It follows that $\langle \varphi_n(x_k) : k < \omega \rangle$ is eventually constant and $\langle x_k : k < \omega \rangle$ is convergent since $\varphi_n(x_k)$ and initial segments of x_k are coded in $\psi_n(x_k)$. Thus since $\lim \vec{x}$ exists and $\vec{\varphi}$ is a scale, $\lim \vec{x} \in X$ and (2a) holds.

For (2b), we use the lower semi-continuity of $\vec{\varphi}$. In particular, write $x = \lim \vec{x}$. We thus have for any $n < \omega$ and $m \le n$, $\varphi_m(x) \le \varphi_m(x_k)$ for sufficiently large $k < \omega$. Hence we can choose k such that $x \upharpoonright n = x_k \upharpoonright n$ and so

 $\psi_n(x) = \operatorname{code}(\varphi_0(x), x(0), \cdots, \varphi_n(x), x(n))$ $\leq \operatorname{code}(\varphi_0(x_k), x(0), \cdots, \varphi_n(x_k), x(n)) = \psi_n(x_k).$

Since this holds for sufficiently large $k < \omega$, it follows that $\psi_n(x) \le \lim_{k \to \omega} \psi_n(x_k)$. Hence $\vec{\psi}$ is a very good scale, and a Γ -scale by Claim 1.

This allows us to now prove the second periodicity theorem and again, argue in a back and forth style that the scale property holds for various analytical pointclasses. Periodicity allows us to go from Σ_n^1 to Π_{n+1}^1 . Through more traditional means we can go from Π_n^1 to Σ_{n+1}^1 , in particular through the use of very good scales.

- 28 C • 6. Theorem (The Second Periodicity Theorem) Assume $ZF + DC + Det(\Gamma \cap \neg \Gamma)$ where Γ is a pointclass closed under $\exists^{\mathscr{N}}$. Therefore $Scale(\Gamma)$ implies $Scale(\forall^{\mathscr{N}}\Gamma)$.

Proof .:.

We start out similarly as with The First Periodicity Theorem (28 C • 2). Let $A \in \forall^{\mathcal{N}} \Gamma$ where $B \in \Gamma$ is such that $A = \forall^{\mathcal{N}} B$. Let $\vec{\varphi}$ be a Γ -scale on B where without loss of generality by Result 28 C • 5, we can assume $\vec{\varphi}$ is very good. We want to show A has a $\forall^{\mathcal{N}} \Gamma$ -scale. For $x, y \in A$ and $n < \omega$, consider the game $G_{x,y}^n$ similar to The First Periodicity Theorem (28 C • 2) that takes the following form:

$$G_{x,y}^n$$
 I: $a(0)$ $a(1)$...
II: $b(0)$ $b(1)$...

Write $s_n \in \omega^{<\omega}$ for the finite sequence coded by $n \in \omega$. We say **II** wins iff $\varphi_n(x, s_n \cap a) \leq \varphi_n(y, s_n \cap b)$ so that again, **II** wants to play the larger ordinal with y. Again from The First Periodicity Theorem (28 C • 2), we define $x \leq_n^n y$ iff **II** has a winning strategy in $G_{x,y}^n$. By the proof of The First Periodicity Theorem (28 C • 2), we know that each $G_{x,y}^n$ is determined and each \leq_n^* is a $\forall^{\mathscr{N}} \Gamma$ -prewellorder of A. So the trouble is mostly going to be with the limit property of scales since we intend to show the rank functions $\vec{\psi} = \langle \psi_n : n < \omega \rangle$ from $\langle \leq_n^* : n < \omega \rangle$

forms a $\forall^{\mathcal{N}} \Gamma$ -scale. Note that by definition of $\vec{\psi}$ being a rank function,

$$\psi_n(x) \le \psi_n(y) \to \mathbf{I}$$
 has a winning strategy in $G_{x,y}^n$. (*)

In particular, $\psi_n(x) = \psi_n(y)$ implies II wins $G_{x,y}^n$ and $G_{y,x}^n$. This is particularly useful if $\psi_n(x_k)$ is eventually constant for large $k < \omega$.

Claim 1 Suppose $\vec{x} = \langle x_i : i < \omega \rangle \in A^{\omega}$ converges to x such that $\psi_n \circ \vec{x}$ is eventually constant. Therefore $x \in A$.

Proof ::.

Let α_n be the eventually constant value of $\langle \psi_n(x_k) : k < \omega \rangle$. By removing entries from \vec{x} if necessary, assume $\psi_n(x_k) = \alpha_n$ for $k \ge n$. We need to show $\langle x, b \rangle \in B$ for each $b \in \mathcal{N}$. The idea is for any given b, we want a sequence $\langle b_n : n < \omega \rangle$ such that $\langle \langle x_n, b_n \rangle : n < \omega \rangle$ converges to $\langle x, b \rangle$ and witnesses via $\vec{\varphi}$ and its very good scale limit properties that $\langle x, b \rangle \in B$. The strategy first will be to define $\langle b_n : n < \omega \rangle$ in a way such that $\varphi(x_{n+1}, b_{n+1}) \le \varphi(x_n, b_n)$ for each n so that the sequence is eventually constant. (Really we use a subsequence of \vec{x} though.)

Let $b \in \mathcal{N}$ be arbitrary. We will play many games but to make sure they all cohere together (that we're using the same initial segments each time), for each $k < \omega$ let $n(k) \in \omega$ be such that $s_{n(k)} = b \upharpoonright k$. Since each $\psi_n(x_k) = \alpha_n$ for $k \ge n$, we have $\psi_{n(k)}(x_{n(k)}) = \psi_{n(k)}(x_{n(k+1)}) = \alpha_{n(k)}$. And so **II** wins $G_{x_n(k+1),x_n(k)}^{n(k)}$ with some strategy σ_k by (*) and DC to choose the strategies. Now let's consider playing infinitely many games at once to define b_n for $n < \omega$. The first move by **I** in $G_{x_n(k+1),x_n(k)}^{n(k)}$ will be b(k). We then just translate downward and copy moves as below.



This defines $b_n(m)$ for $m \ge n$ explicitly and we can now define $b_n \upharpoonright n = b \upharpoonright n$ so we have $b_n \in \mathcal{N}$. More importantly, **II**'s play in $G_{x_n(k+1),x_n(k)}^{n(k)}$ is $b_k \upharpoonright [k,\omega)$ whereas **I**'s play is $b(k)^{\frown}b_{k+1} \upharpoonright [k+1,\omega)$. Thus after adjoining $s_{n(k)}$ at the beginning, we arrive at b_k and b_{k+1} . So since **II** wins in each board, we get that $\varphi_{n(k)}(x_{n(k+1)}, b_{k+1}) \le \varphi_{n(k)}(x_{n(k)}, b_k)$. As a very good scale, since $n(j) \le n(k)$ for $j \le \kappa$, we get that for all $j \le k$, $\varphi_{n(j)}(x_{n(k+1)}, b_{k+1}) \le \varphi_{n(j)}(x_{n(k)}, b_k)$. In fact, for all $j \le n(k)$, and hence all j < k,

$$\varphi_j(x_{n(k+1)}, b_{k+1}) \le \varphi_j(x_{n(k)}, b_k).$$

Thus $\langle \varphi_j(x_{n(k)}, b_k) : j < k < \omega \rangle$ is a decreasing sequence of ordinals and is therefore eventually constant. Since $\vec{\varphi}$ is very good on *B*, the limit of $\langle \langle x_{n(k)}, b_k \rangle : k < \omega \rangle$, which is $\langle x, b \rangle$, is in *B*. Since *b* was arbitrary, $x \in A$.

Now we need to show lower semi-continuity: if $\vec{x} = \langle x_k : k < \omega \rangle \in {}^{\omega}A$ converges to $x \in A$ with each $\psi_n \circ \vec{x}$ eventually constant, then that the limit x has norms $\psi_n(x) \leq$ than the constant value $\alpha_n = \lim_{k \to \omega} \psi_n(x_k)$, meaning that we want to show that **II** has a winning strategy for G_{x,x_n}^n for each $n < \omega$.

Again, without loss of generality by deleting entries if necessary, assume $\psi_n(x_k) = \alpha_n$ for $k \ge n$ so that II has a winning strategy in G_{x_k,x_n}^n by (*).

Now let $b \in \mathcal{N}$ be arbitrary. For $k \in \omega$, consider $n_k \in \omega$ such that $s_{n_k} = s_n \frown b \upharpoonright k$ (so in particular, $n_k > n$ and $n_0 = n$). Since $n_{k+1} \ge n_k$, it follows (again from (*) and ψ_{n_k} being constant after x_{n_k}) that **II** wins $G_{x_{n_k+1},x_{n_k}}^{n_k}$ with some strategy σ_k . Now consider playing infinitely many games to tell us how to play in G_{x_n,x_n}^n .



We want to show **II** wins G_{x,x_n}^n . The first move in $G_{x_{n_{k+1}},x_{n_k}}^{n_k}$ by **I** is b(k) for each $k < \omega$. This is good because implicitly, in that game we are adjoining s_n by $b \upharpoonright k$, so the next game which considers $b \upharpoonright k + 1$ will match. The result is that we define b_k as $s_n \frown b \upharpoonright k = s_{n_k}$ followed by **II**'s moves in $G_{x_{n_{k+1}},n_k}^{n_k}$. Note that therefore $\langle \langle x_{n_k}, b_k \rangle : k < \omega \rangle$ converges to $\langle x, b \rangle$ because $b_k \upharpoonright k = b \upharpoonright k$. Since **I** copies the next move's board, we also have that b_{k+1} is $s_n \frown b \upharpoonright k = s_{n_k}$ followed by **I**'s moves in $G_{x_{n_{k+1}},x_{n_k}}^{n_k}$. Since **II** wins in each board above G_{x,x_n}^n , it follows that for each $k < \omega$,

$$\varphi_{n_k}(x_{n_{k+1}}, b_{k+1}) \le \varphi_{n_k}(x_{n_{k+1}}, b_k).$$

As a very good scale, the same follows for previous indices of φ , and hence for each $j \in \omega$, the sequence $\langle \varphi_j(x_{n_k}, b_k) : k \in [j, \omega) \rangle$ is a decreasing sequence of ordinals. Therefore this sequence is eventually some constant value β_j . So as $\vec{\varphi}$ is a scale, lower semi-continuity implies that for all $n < \omega$,

$$\varphi_n(x,b) \leq \beta_n = \varphi_{n_0}(x_{n_0},b_0) = \varphi_n(x_n,b_0).$$

Hence in the above board, **II** has won G_{x,x_n}^n . Hence the strategy described wins for **II** and $x \leq_n^* x_n$ for each n and $\psi_n(x) \leq \lim_{k \to \infty} \psi_n(x_k)$ for each n. This establishes lower continuity for $\vec{\psi}$ and so $\vec{\psi}$ is a scale.

 $\vec{\psi}$ forms a $\forall^{\mathcal{N}} \Gamma$ scale since

$$x \leq_{n} y \quad \text{iff} \quad x \in A \land (y \in A \to \mathbf{I} \text{ does not win } G_{x,y}^{n})$$

$$\text{iff} \quad x \in A \land \forall \sigma \exists z ((x, \sigma * z) \leq_{\varphi_{n}} (y, z)))$$

$$\Gamma$$

$$\forall^{\mathcal{N}} \exists^{\mathcal{N}} \Gamma = \forall^{\mathcal{N}} \Gamma$$

$$x <_{\psi_{n}} y \quad \text{iff} \quad x \in A \land (y \in A \to \mathbf{I} \text{ does not win } G_{x,y}^{n} \land \mathbf{II} \text{ does not win } G_{y,x}^{n})$$

$$\text{iff} \quad x \in A \land \forall \sigma \exists z ((x, \sigma * z) \leq_{\varphi_{n}} (y, z)) \land \forall \sigma \exists z ((x, z) \leq_{\varphi_{n}} (y, z * \sigma)))$$

$$\Gamma$$

$$\forall^{\mathcal{N}} \Gamma$$

If $x, y \in A$ then this works: I doesn't have a winning strategy in $G_{x,y}^n$ iff II does iff $x \leq_n^* y$ iff $\psi_n(x) \leq \psi_n(y)$, and similarly $x <_n^* y$ iff $x \leq_n^* y \land y \not\leq_n^* x$. If $x \in A$ and $y \notin A$, then there is some z with $\langle y, z \rangle \notin A$ and thus $\langle x, a \rangle \leq_{\varphi_n} \langle y, z \rangle$ is true for any $a \in \mathcal{N}$. This similarly holds for \langle_{ψ_n} . Thus $\vec{\psi}$ is a $\forall^{\mathcal{N}} \Gamma$ -scale. \dashv

Similar to the ideas around the prewellordering property, we get the following.

- 28 C • 7. Corollary Assume ZF + DC + PD. Therefore Scale($\Sigma_n^1(X)$) implies Scale($\Pi_{n+1}^1(X)$) for any $X \subseteq \mathcal{N}$.

Indeed, we get the same pattern as with Figure 28 C • 1: under ZF + DC + PD, Scale($\sum_{n=1}^{1}$) holds for even $n < \omega$ while Scale($\prod_{n=1}^{1}$) holds for odd $n < \omega$.



28 C • 8. Figure: Analytical $\Gamma \neq \Delta_n^1$ such that $ZF + DC + PD \models Scale(\Gamma)$

There are actually three periodicity theorems, although the third is mostly unnecessary for our purposes. It concerns the complexity of the strategies that win games, and basically says that if I has a winning strategy for G(X) for $X \in \Gamma$ —where Γ is analytical with Scale(Γ) and Det($\Gamma \cap \neg \Gamma$)—then I has a winning strategy for G(X) in $\Im \Gamma$. The proof of the theorem is quite difficult and technical, and can be found in [23] (Sections 6D and 6E).

One explanation why we get the back and forth propogation of these pointclass properties is the following involving the so-called "game" quantifier.

- 28C•9. Definition For $X \subseteq \mathcal{N}^2$, $n < \omega$, we write $\partial X \subseteq \mathcal{N}$ for the set of all $x \in \mathcal{N}$ such that I wins $G(\{y \in \mathcal{N} : \langle x, y \rangle \in X\})$. We also write $\partial \Gamma$ for $\{\partial X : X \in \Gamma\}$ whenever Γ is a pointclass.

Another way to view this game quantifier is realizing that, under AD,

 $\begin{aligned} x \in \Im X & \text{iff} \quad \mathbf{I} \text{ has a winning strategy for } G(\{y \in \mathcal{N} : \langle x, y \rangle \in X\}) \\ & \text{iff} \quad \exists \sigma \forall y \ \langle x, \sigma * y \rangle \in X \\ & \text{iff} \quad \mathbf{II} \text{ does not have a winning strategy for } G(\{y \in \mathcal{N} : \langle x, y \rangle \in X\}) \\ & \text{iff} \quad \neg \exists \sigma \forall y \ \langle x, y * \sigma \rangle \notin X. \end{aligned}$

It follows that if $X \in \Gamma$ for some adequate Γ , then $\partial X \in \exists^{\mathcal{N}} \forall^{\mathcal{N}} \Gamma$ and $\partial X \in \forall^{\mathcal{N}} \exists^{\mathcal{N}} \Gamma$. So for example, under AD, if $X \in \Pi_1^1$ then $\partial X \in \Sigma_2^1$. In particular, $\partial \Pi_1^1 = \Sigma_2^1$ and likewise $\partial \Sigma_2^1 = \forall^{\mathcal{N}} \exists^{\mathcal{N}} \Sigma_2^1 = \Pi_3^1$. Generalizing this in AD, we get that if Γ is closed under $\forall^{\mathcal{N}}$, $\partial \Gamma \subseteq \exists^{\mathcal{N}} \Gamma$. Moreover, we also get that if $\exists^{\mathcal{N}} \Gamma \subseteq \Gamma$, then $\partial \Gamma \subseteq \forall^{\mathcal{N}} \Gamma$. And so this gives some explanation of the periodicity pattern we've seen with Figure 28 C • 1 and Figure 28 C • 8: $\partial \Sigma_n^1 = \Pi_{n+1}^1$, and $\partial \Pi_n^1 = \Sigma_{n+1}^1$. The other periodicity theorems are then about moving properties of Γ to $\partial \Gamma$ under mild determinacy assumptions.

Exercises

Section 29. Measure, Degrees, and Coding

One of the most unfortunate aspects of studying AD is that some of the most fundamental results require a very large amount of effort to understand and work though. The First Periodicity Theorem $(28 \text{ C} \cdot 2)$ and The Second Periodicity Theorem $(28 \text{ C} \cdot 6)$, for example, take quite a bit of work to prove, and the situation is even worse for the third periodicity theorem [23]. Nevertheless, such theorems are very useful for their consequences and for their uses. An example of this is from the "coding" lemma. We shall use this theorem in our proofs before actually proving it, partly to motivate why it is important. There are three versions we shall use, one being easier to prove, but also much weaker as a statement.

```
29.1. Theorem (Cheapo Coding Lemma) 
Assume ZF + DC + AD. If there is a surjection f : \mathcal{N} \to \kappa then there is a surjection g : \mathcal{N} \to \mathcal{P}(\kappa).
```

The others are harder to prove, but much stronger, more technical statements.

§ 29 A. Consequences of the coding lemma

Section 30. Exercises

Exercises

Chapter VI. An Overview of Forcing

The goal of this chapter is to present the basics of forcing in a more or less self contained way. This is especially so, because the content is adapted from the notes I wrote for a talk introducing forcing for the 2020 GOST (Graduate Organized Set Theory) Seminar at Rutgers.

Section 31. The Purpose, Motivation, and Terminology of Forcing

The main idea behind forcing is to expand a model of set theory by a new set. Moreover, we should do this in a *minimal* way, and we should hope to preserve the membership relation, meaning that the new model should be transitive.

31 • 1. Theorem
Let V ⊨ ZFC be a transitive model. Let P ∈ V be a poset.
A generic extension of V via P by a G ∉ V, written V[G], has the following properties:
1. G ⊆ P;
2. V[G] ⊨ ZFC is transitive;
3. V[G] is the ⊆-least transitive model M of ZFC with V ⊆ M and G ∈ M.

(1) is a really where \mathbb{P} comes into play: we attempt to find a set G not in V, but which still has some intelligible structure to it. (2) is just a nice result of \mathbb{P} being a set. (3) is the most important and motivating idea for us. The idea is that, despite G not being in V, we carry out a bunch of potential constructions of V[G] inside V (so-called \mathbb{P} -names). It is only through using G as a kind of oracle that allows us to form V[G] by interpreting these constructions in V.

The general picture can be understood through the diagram below: V is an inner model of V[G], equivalently, V[G] is an "outer-model" of V.



31.2. Figure: The ground model V and its generic extension V[G]

To figure out which $G \subseteq \mathbb{P}$ are appropriate, we have the following theorem relating truth in V[G] with \mathbb{P} in V. Here $p \Vdash \varphi$ is a notion definable in V which we will introduce later: it's the *forcing* relation.

31.3. Theorem Let $\mathbb{P} \in V$ be a preorder and φ a FOL-sentence. Let $G \subseteq \mathbb{P}$ be "generic". Therefore, $V[G] \models \varphi$ iff there is some $p \in G$ with $p \Vdash \varphi$.

Interpretting the forcing relation requires a lot of work, and there are many perspectives to take on it. Regardless, one can always take the formal approach, using the definition of it in V from Appendix C.

If we're working with the actual universe of sets, the existence of these generic sets $G \notin V$ is called into question. This worry can be alleviated in several ways. Firstly, we can regard ourselves as working in a relatively small inner model $V \subseteq V$, and assume that for each $\mathbb{P} \in V$ and $p \in \mathbb{P}$, there are generic Gs with $p \in G$ (and we will see that for most preorders, $G \notin V$). L, for example, consistently has relatively few sets, so it should be a little more understandable for there to be these $G \notin L$, although this still depends on what sets exist in the ambient universe V.

Many authors choose to think about *countable* transitive models of ZFC as sort of toy models to play with. With such models, the existence of these Gs for every \mathbb{P} and $p \in \mathbb{P}$ is provable from ZFC. Unfortunately, the existence of these models does not follow from ZFC by Gödel incompleteness. This sub-worry can be alleviated when we acknowledge what the purpose of forcing is: consistency results. In particular, the existence of such models is consistent iff ZFC is consistent. So if we're trying to show the consistency of some theory, Con(T), assuming Con(ZFC), it suffices to show Con(T) from Con(ZFC + "there is a countable transitive model of ZFC"). Equivalently, we just take countable, transitive models of sufficiently large finite fragments of ZFC + T for some theory T, demonstrating that ZFC + T is consistent by the compactness theorem (applied in the real world rather than in ZFC).

There is a third interpretation of the generic extension by way of Boolean valued models, where truth value is not taken to be either 0 or 1, but instead an element of a Boolean algebra. Under this interpretation, a suitable ultrafilter G tells us how to interpret the non-strictly-true and non-strictly-false statements as either true or false. Alternatively, if we forgo the existence of such a G, the resulting Boolean algebra can still allow us to see whether a certain theory T is consistent relative to ZFC: in the Boolean valued model, each formula of ZFC has truth value 1, and perhaps so do all the elements of T. If so, one can show using logic that this shows the consistency of ZFC + T.

We will mostly just consider $\mathbf{V} = \langle V, \in \rangle$ to be a transitive inner model in the ambient universe $\mathbf{V} = \langle V, \in \rangle$, and we just assume that every poset or preorder $\mathbb{P} \in V$ and $p \in \mathbb{P}$ that we consider will have a generic G with $p \in G$.ⁱⁱ This is contained in the following, non-standard definition. Again, we are delaying what exactly a "generic" is for later, because the definition needs significant motivation.

31.4. Definition Let M be a transitive model with $\mathbb{P} \in M$ a poset or preorder. We say M can be forced over with \mathbb{P} iff for every $p \in \mathbb{P}$, there is a \mathbb{P} -generic G with $p \in G$. We say M can be forced over iff M can be forced over with \mathbb{P} for every poset or preorder $\mathbb{P} \in M$.

First we will consider what V[G] will actually be, and then we will consider truth in V[G]. The idea behind what V[G] will be is a bunch of conditional constructions that, once we have access to G is, can be thinned out to yield what we were trying to construct dependent on G. This yields a kind of *forcing* to be true in that knowing just a bit about G can already determine the outcome of some constructions.

§ 31 A. Names

For $G \notin L$, there is already the notation of L[G]. Recall that L is defined recursively:

$$L_0 = \emptyset \qquad L_{\alpha+1} = \{ x \subseteq L_\alpha : x \text{ is definable over } \langle L_\alpha, \in \rangle \} \qquad L_\gamma = \bigcup_{\alpha < \gamma} L_\alpha, \text{ for } \gamma \text{ a limit.}$$

So at each stage, we're taking definable subsets. What is the natural model of L that includes G (given that $G \subseteq \mathbb{P} \in L$)? Well, we just make G a definable subset of \mathbb{P} : rather than consider definable subsets of $\langle L_{\alpha}, \in \rangle$, we allow membership in G as a predicate:

$$L_0[G] = \emptyset$$

ⁱThe existence of such models follows from Corollary 7 D • 8.

ⁱⁱThe notation here of V versus V is deliberately similar. Often in the literature, V is the starting point and V[G] is the extension with $V \subsetneq V[G]$ and $G \in V[G] \setminus V$. This is impossible with our framework with how we've defined V: as the collection of all sets, G couldn't exist. Nevertheless, we could have considered ourselves as working in an inner model V of the "true" universe V[G] = V. So in essence, we might as well assume all of these are inner models of the actual universe of sets. And rather than work over V—dealing with the odd philosophical ideas about expanding to V[G]—we just assume we're working over some inner model V and expanding to another inner model $V[G] \subseteq V$. Again, we also could consider the other interpretations of forcing.

$$L_{\alpha+1}[G] = \{ x \subseteq L_{\alpha}[G] : x \text{ is definable over } \langle L_{\alpha}[G], \in, G \rangle \}$$
$$L_{\gamma}[G] = \bigcup_{\alpha < \gamma} L_{\alpha}[G], \text{ for } \gamma \text{ a limit.}$$

As $G \subseteq \mathbb{P}$, G is a definable subset of \mathbb{P} over $\langle L_{\alpha}[G], \in, G \rangle$ whenever $\mathbb{P} \in L_{\alpha}$. Clearly if G were in L already, this wouldn't make a difference: L[G] = L. Moreover, any inner model M with $G \in M$ can do this construction: $L[G] \subseteq M$ if $G \in M$. Hence L[G] is the least inner model M of ZFC with $G \in M$ (assuming $G \subseteq \mathbb{P} \in L$). This is kind of the gold standard we want to emulate when forming the generic extension, and it motivates the idea of a *name*.

Suppose V has access to G. What sets can V form from G? V doesn't know what G is, but it can at consider constructions from \mathbb{P} and then we can thin these out with access to G. The idea is to tag elements at each stage of the construction with elements of \mathbb{P} : look at things of the form $\langle x, p \rangle$ for $p \in \mathbb{P}$. Once we have access to G, if $p \notin G$, we throw out x, and if $p \in G$, we include it. For example, $\{\langle p, p \rangle : p \in \mathbb{P}\}$ will be thinned out to G, as we only include the first coordinate of $\langle p, p \rangle$ where $p \in G$. This is the motivation, but it doesn't precisely work because we also need to think about potential constructions that depend on other potential constructions. This gives a cumulative hierarchy of names.

The idea is that each set is tagged with an element of \mathbb{P} , and we just consider the elements tagged with an element of *G*. We can also iterate this concept:

$$\{\left(\left\{\left\langle\emptyset, p_{0}\right\rangle\right\}, p_{1}\right), \left\langle\emptyset, p_{0}\right\rangle\right\} \text{ will be thinned out to } \begin{cases}\{\left\{\emptyset\right\}, \emptyset\right\} & \text{if } p_{0}, p_{1} \in G, \\ \left\{\emptyset\right\} & \text{if } p_{1} \in G \text{ but } p_{0} \notin G, \\ \left\{\emptyset\right\} & \text{if } p_{0} \in G \text{ but } p_{1} \notin G, \\ \left\{\emptyset\right\} & \text{otherwise.} \end{cases}$$

Note that we aren't going to assume any set-theoretic axioms of V in these definitions. If V doesn't think there are any such sets, the resulting concept is just left as undefined. We also don't assume \mathbb{P} is a poset or anything with intelligible structure: it's just a set we're using to tag elements.

31 A • 1. Definition Let V be a transitive class. Let $\mathbb{P} \in V$. A \mathbb{P} -name is defined by recursion on rank: define • $V_0^{\mathbb{P}} = \emptyset$; • $V_{\gamma}^{\mathbb{P}} = \bigcup_{\alpha < \gamma} V_{\alpha}^{\mathbb{P}}$ for γ a limit; • $V_{\alpha+1}^{\mathbb{P}} = (\mathcal{P}(V_{\alpha}^{\mathbb{P}} \times \mathbb{P}))^{\mathbb{V}}$. Say that τ is a \mathbb{P} -name (of V) iff τ is in $V^{\mathbb{P}} = \bigcup_{\alpha \in \text{Ord}} V^{\mathbb{P}} \subseteq V$.

As a result, any P-name is a collection of pairs $\langle \tau, p \rangle$ where $p \in \mathbb{P}$ and τ is another P-name.

Note that this creates a \mathbb{P} -name rank similar to the rank of a set in the cumulative hierarchy: at each stage we don't just take collections of previous stages, we take collections that have been marked with elements of \mathbb{P} . The result is that we can perform induction on \mathbb{P} -name rank.

— 31 A • 2. Result -

Let V be a transitive class. Let $\mathbb{P} \in V$ and τ a \mathbb{P} -name. Therefore there is some least α with $\tau \in V_{\alpha+1}^{\mathbb{P}} \setminus V_{\alpha}^{\mathbb{P}}$ called the \mathbb{P} -rank of τ .

Each $\tau \in V^{\mathbb{P}}$ can be "thinned out" once we're given access to G. To make this more precise, we iteratively only consider the names tagged with elements of G. This process will be well-defined because at some point we reach $\emptyset \in V_1^{\mathbb{P}}$, which is just to say that \mathbb{P} -names are well-founded as a result of their \mathbb{P} -rank.

- 31 A • 3. Definition

Let V be a transitive class. Let $\mathbb{P} \in V$ be a set, and $G \subseteq \mathbb{P}$, possibly not in V. For a \mathbb{P} -name τ , define by recursion on \mathbb{P} -names

 $\tau_G = \{ \sigma_G : \exists p \in G \ (\langle \sigma, p \rangle \in \tau) \}.$

So, for example, \emptyset is a \mathbb{P} -name, as is $\tau = \{\langle 0, p \rangle : p \in \mathbb{P}\}$. For $G \neq \emptyset$, this is a \mathbb{P} -name for $\tau_G = \{0\} = 1$. Note that there can be multiple \mathbb{P} -names for a single set. For example, for $p, q \in G \{\langle x, p \rangle\}_G = \{\langle x, p \rangle, \langle x, q \rangle\}_G = \{x_G\}$. The

first component of a pair $\langle x, p \rangle \in \tau \in V^{\mathbb{P}}$ tells us *what* the pair is transformed to while the second component tells us *whether* we include the transformation x_G in τ_G .

Note that since G might not be in V, there's no reason to expect $\tau_G \in V$. In the end, our expanded model V[G] will be the collection of these interpretations: $V[G] = \{\tau_G : \tau \in V^{\mathbb{P}}\}$. So it's important to realize that when working with G, we are almost never working inside V.

We can try to get a \mathbb{P} -name for G itself just with $\{\langle p, p \rangle : p \in \mathbb{P}\}$, but this is transformed into $\{p_G : p \in G\}$ rather than $\{p : p \in G\}$. So we need an inverse of $\tau \mapsto \tau_G$ whenever we can. Such a thing isn't possible in general, but we can get the next best thing: an easily definable name for any given $x \in V$.

- 31 A • 4. Definition Let V be a transitive class. Let $\mathbb{P} \in V$. For $x \in V$ define the *canonical name* or *check-name* $\check{x} \in V^{\mathbb{P}}$ recursively to be $\{\check{y} : y \in x\} \times \mathbb{P}$.

For example,

• $\check{\emptyset} = \emptyset \times \mathbb{P} = \emptyset$ is a \mathbb{P} -name for \emptyset ;

- $\check{1} = \{\check{0}\} \times \mathbb{P}$ is a \mathbb{P} -name for 1;
- $\check{2} = \{\check{0}, \check{1}\} \times \mathbb{P}$ is a \mathbb{P} -name for 2;
- $(x \cup y) = \check{x} \cup \check{y}$ is a \mathbb{P} -name for $x \cup y$;
- $\{x, y\} = \{\langle \check{x}, p \rangle, \langle \check{y}, p \rangle : p \in \mathbb{P}\}$ is a \mathbb{P} -name for $\{x, y\}$; etc.

One can easily see that \check{x} is well-defined (by induction on \mathbb{P} -name rank), is inductively a \mathbb{P} -name, and that its interpretation is just x.

— 31 A • 5. Result -

Let V be a transitive class. Let \mathbb{P} be a set and $\emptyset \neq G \subseteq \mathbb{P}$. Let $x \in V$. Therefore $(\check{x})_G = x$.

Proof .:.

Proceed by induction on the rank of x. Inductively, $(\check{y})_G = y$ for each y of lower rank, in particular, for $y \in x$. Hence we can calculate,

 $(\check{x})_G = \{ \langle \check{y}, p \rangle : y \in x \land p \in \mathbb{P} \}_G = \{ (\check{y})_G : \exists p \in G \ (\langle \check{y}, p \rangle \in \check{x}) \}$

But the existence of a $p \in G$ with $\langle \check{y}, p \rangle \in \check{x}$ is just always true if $\check{y} \in \text{dom}(\check{x})$ since $G \cap \mathbb{P} \neq \emptyset$ and $\langle \check{y}, p \rangle \in \check{x}$ for every $p \in \mathbb{P}$. Thus $(\check{x})_G = \{(\check{y})_G : y \in x\} = \{y : y \in x\} = x$. \dashv

This allows us to give a \mathbb{P} -name for $G: \{\langle \check{p}, p \rangle : p \in \mathbb{P}\}$. Hence, by the following definition, $V = \{(\check{x})_G : x \in V\} \subseteq V[G]$ and $G \in V[G]$.

— 31 A•6. Theorem —

Let V be a transitive class. Let $\mathbb{P} \in V$ be a set and $\emptyset \neq G \subseteq \mathbb{P}$. Define $V[G] = \{\tau_G : \tau \in V^{\mathbb{P}}\}$. Therefore,

- 1. $V \subseteq V[G]$ and $G \in V[G]$;
- 2. V[G] is transitive; and
- 3. Any transitive $M \models ZF$ with $V \subseteq M$ and $G \in M$ has $V[G] \subseteq M$.

Proof .:.

- 1. \check{x} has $(\check{x})_G = x \in V[G]$ so that $V \subseteq V[G]$. We have the name $\dot{G} = \{\langle \check{p}, p \rangle : p \in \mathbb{P}\}$ for G so that $G = (\dot{G})_G \in V[G]$.
- 2. To see that V[G] is transitive, let $x \in \tau_G \in V[G]$ where $\tau \in V^{\mathbb{P}}$. Therefore, $x = \sigma_G$ for some $\sigma \in V^{\mathbb{P}}$ so that $x = \sigma_G \in V[G]$.
- 3. Any such M with $V \subseteq M$ can construct each \mathbb{P} -name: $V^{\mathbb{P}} \subseteq M$. Note that V need not be a class for this to hold; M doesn't need to consider the class $V^{\mathbb{P}}$, just each individual $\tau \in V^{\mathbb{P}} \subseteq V \subseteq M$. Indeed, usually $M^{\mathbb{P}}$ will be significantly bigger than $V^{\mathbb{P}}$. Now since $G \in M$, for each \mathbb{P} -name τ of V, we can construct τ_G in M (M knows enough set theory to carry out these constructions). Hence $V[G] \subseteq M$.

So already we have a kind of "minimal" model by expanding V to V[G]. But it's not obvious how we can know

whether $V[G] \models ZFC$ or not. Indeed, it's not at all obvious how to calculate truth in V[G]. Just from it being transitive, finding the right P-names allows us to argue by absoluteness that V[G] models pairing, union, and some of the other simple axioms. But going further than this isn't easy, especially because V[G]'s FOL-theory *can* differ significantly from V.

§31 B. Posets, information, and forcing

Recall the definition of a preorder and poset (a partially ordered set).

31 B•1. **Definition** A *preorder* is a structure $\mathbb{P} = \langle \mathbb{P}, \leq \rangle$ such that

• (reflexive) for all $p \in \mathbb{P}$, $p \leq p$;

• (transitive) for all $p, q, r \in \mathbb{P}$, $p \leq q \leq r$ implies $p \leq r$.

The more common notion is a *poset*: \mathbb{P} is a poset iff it's a preorder that is also ani-symmetric (for all $p, q \in \mathbb{P}$, $p \leq q \leq p$ implies p = q).

For $p, q \in \mathbb{P}$, we say q extends p, is an extension of p, is stronger than p, or is below p iff $q \leq p$. We frequently call elements of \mathbb{P} conditions.

In some sense, we don't really need reflexivity, as any non-reflexive $\langle \mathbb{P}, < \rangle$ has a corresponding reflexive version $\langle \mathbb{P}, \leq \rangle$ where we just consider $\leq = \langle \cup \{ \langle p, p \rangle : p \in \mathbb{P} \}$, and similarly from a reflexive $\langle \mathbb{P}, \leq \rangle$ we can consider the *irreflexive* $\langle = \leq \rangle$ id. The concept of a poset should be fairly familiar, as it corresponds to directed graphs (that are transitive) with no loops. Similarly, we don't actually care about anti-symmetry since we can just take equivalence relation $x \approx y \leftrightarrow x \leq y \land y \leq x$ and look consider \mathbb{P}/\approx as a poset (as we will see at the end of the next section).

All of this is just to say that we really only care about preorders. Elsewhere in the literature, the word "poset" is often used to denote what are actually preorders, it's just that "poset" is a much more familiar concept to mathematicians than preorders. Again, the distinction won't actually matter in the end because all posets are preorders, and we may easily consider the poset "version" of a preorder instead of the preorder itself.

We will be viewing posets and preorders as coding information. Consider the following analogy with knowledge and discovery. Currently, we have a fair amount of knowledge p. At a later point in time, we could make discoveries such that we know q or r. In this way, we have an ordering on our knowledge. Given that more precise, specific information is less likely to be true in general, we say that $p^* \leq p$ to represent that p^* has more information than p. This gives a kind of poset, and motivates the terminology of $p \in \mathbb{P}$ as a "condition". Moreover, this analogy allows us to introduce the forcing relation already: p forces something to be true if p has enough information to determine it.ⁱⁱⁱ

- 31 B • 2. Definition

Let V be a transitive class we can force over. Let $\mathbb{P} \in V$ be a preorder with $p \in \mathbb{P}$. Let φ be a formula and let $\vec{\tau}$ be \mathbb{P} -names. We say p forces " $\varphi(\vec{\tau})$ ", written $p \Vdash "\varphi(\vec{\tau})$ ", iff $p \in G$ implies $V[G] \vDash \varphi(\vec{\tau}_G)$ for all $G \subseteq \mathbb{P}$ "generic" over V (to be defined later).

- 31 B•3. Motivation

Let V be a transitive class we can force over. Let $\mathbb{P} \in V$ be a preorder. For $p \in \mathbb{P}$, write $p^* \in \mathbb{P}$ for an arbitrary $p^* \leq^{\mathbb{P}} p$ (an arbitrary point in time after p). Therefore, for all formulas φ with \mathbb{P} -name parameters,

1. if $p \Vdash \varphi$ then every $p^* \Vdash \varphi$;

- 2. $p \Vdash \neg \varphi$ " iff every $p^* \not\models \varphi$, i.e. you can conclude it's false iff you will never discover that it's true;
- 3. $p \Vdash "\varphi \land \psi"$ iff $p \Vdash \varphi$ and $p \Vdash \psi$;
- 4. if $p \Vdash \exists x \varphi(x)$ then there is some $p^* \leq^{\mathbb{P}} p$ and τ where $p^* \Vdash \varphi(\tau)$; and
- 5. if $p \Vdash \varphi$, and φ is logically equivalent to ψ , then $p \Vdash \psi$;

ⁱⁱⁱThere is a competing notation in some relatively small circles where conditions with more information are considered "larger": $p^* \ge p$. We do not adopt this notation as it is less widespread and also counter-intuitive with the boolean algebra interpretation of forcing. In either notation, p^* is said to be *stronger* than p, or p^* is an *extension* of p. This can help disambiguate when other sources use a different notation.

Note that this is very intuitionistic. There is actually a fairly close connection between forcing and intuitionistic logic [9]. Note, for example, $p \not\models \varphi$ is not equivalent to $p \not\models \neg \varphi$ ". The motivation behind (4) is that if we know something is true, we should be able to discover an example.

− 31 B • 4. Corollary Assume Motivation 31 B • 3. Let *V* be a transitive class we can force over. Let $\mathbb{P} \in V$ be a preorder, $p \in \mathbb{P}$, and φ a formula. Therefore, $p \Vdash \varphi$ iff every extension $p^* \Vdash \varphi$.

Proof .:.

The " \rightarrow " direction is clear. For the " \leftarrow " direction, suppose $p \not\models \varphi$, i.e. $p \not\models$ " $\neg \neg \varphi$ ". By (2) of Motivation 31 B • 3, there is then an extension $p^* \models$ " $\neg \varphi$ ", contradicting that every $p^* \models \varphi$.

The idea given by the proof also shows the following: if it's currently unclear whether something is true, it will be decided later.

− 31 B • 5. Corollary Assume Motivation 31 B • 3. Let *V* be a transitive class we can force over. Let $\mathbb{P} \in V$ be a preorder, $p \in \mathbb{P}$, and φ a formula. Therefore, $p \not\models \varphi$ and $p \not\models$ "¬ φ " implies there are $q, r \leq^{\mathbb{P}} p$ where $q \models \varphi$ and $r \models$ "¬ φ ".

In the end, our goal will be the following.

- 31 B • 6. Theorem Assume Motivation 31 B • 3. Let V be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder. Let $G \subseteq \mathbb{P}$ be "generic" over V. Therefore, $V[G] \models \varphi$ iff there is some $p \in G$ with $p \Vdash \varphi$.

Proof .:.

Clearly if $p \in G$ has $p \Vdash \varphi$ then $V[G] \vDash \varphi$ by Definition 31 B•2. So suppose $V[G] \vDash \varphi$, but every $p \in G$ has $p \not\Vdash \varphi$. In particular, for every $p \in G$, every $p^* \not\Vdash \varphi$ and thus $p \Vdash \neg \varphi$ " by the unproven Motivation 31 B•3. Therefore $V[G] \vDash \neg \varphi$, a contradiction.

So really we need a notion that ensure the forcing relation as defined in Definition 31 B • 2 obeys Motivation 31 B • 3. So this partially motivates what G should look like: we obviously need all $p \in G$ to be compatible with each other in a precise sense. We also need G to interact nicely with extensions: we need to be able to extend elements of G with certain properties as needed. We will see later that this amounts to being a filter and intersecting dense sets.

To state all of this, we need to think about the topology of \mathbb{P} . Many of the topological concepts have been introduced in Section 21, but we only care about characterizations relevant and useful for preorders.

§ 31 C. Poset topology

Properties of the generic G will be induced by the topology of the corresponding preorder \mathbb{P} . The topology on preorders is given by their ordering relation. In particular, basic open sets are just sets that are closed downward.

— 31 C • 1. Definition -

Let \mathbb{P} be a preorder. For $p \in \mathbb{P}$, the *basic open neighborhood* around p is $\mathbb{P}_{\leq \mathbb{P}_p} = \{q \in \mathbb{P} : q \leq^{\mathbb{P}} p\}$, i.e. everything below p. The *preorder topology* on \mathbb{P} is the topology generated by these: the topology with basis $\{\mathbb{P}_{\leq \mathbb{P}_p} : p \in \mathbb{P}\}$.

- 31 C • 2. Corollary

Let \mathbb{P} be a preorder. Therefore $U \subseteq \mathbb{P}$ is open iff U is closed downward: $p \in U$ and $q \leq^{\mathbb{P}} p$ implies $q \in U$.

Proof .:.

If U is closed downward, then for every $p \in U$, $\mathbb{P}_{\leq \mathbb{P}_p} \subseteq U$ and hence U is open. If U is open, note that U has been generated by the basic open neighborhoods, which are closed downward. To be generated by these basic

open sets, U must be an arbitrary union of finite intersections of these $\mathbb{P}_{\leq \mathbb{P}_p}$ s. These finite intersections are easily seen to be closed downward, and unions of sets closed downward are also closed downward. Hence U is closed downward.

Without appealling to topology, one can make the following result a definition.

— 31 C·3. Corollary

Let \mathbb{P} be a preorder. A set $D \subseteq \mathbb{P}$ is *dense* iff D intersects every open set. Thus $D \subseteq \mathbb{P}$ is dense iff for every $p \in \mathbb{P}$, there is some $p^* \leq \mathbb{P}$ where $p^* \in D$. In other words, D is dense iff we can always extend a p to a $p^* \in D$.

Proof .:.

Suppose D is dense with $p \in \mathbb{P}$. Since $D \cap \mathbb{P}_{\leq \mathbb{P}_p} \neq \emptyset$, there is some $p^* \leq^{\mathbb{P}} p$ with $p^* \in D$.

For the most part, the above ideas are not used: we do not care about topological definitions in general. We really only care about sets closed downward, and sets where we can always go downward into the set. In the analogy with knowledge and discovery, these correspond to things that always remain true (closed downward), and things that always have the potential to be true (density) in that for any point of time p, it's always possible to discover at a later time p^* that it's true.

The notion of being able to extend an element is incredibly important for us. We thus have two additional notions for preorders.

— 31 C•4. Definition -

Let \mathbb{P} be a preorder. Let $p, q \in \mathbb{P}$. p and q are *compatible* iff there is a common extension $r \leq^{\mathbb{P}} p, q$. p and q are *incompatible*, written $p \perp q$, otherwise: there is no common extension.

Easy examples of compatible elements include any two comparable elements: $p^* \leq p$ implies p and p^* are compatible. The basic pictures of compatibility and incompatibility are below.



31 C \cdot 5. Figure: Compatibility of *p* and *q* in example preorders

There need not be a common predecessor to p and q if $p \perp q$, as the figure above suggests. But we will only consider preorders where this occurs, sometimes artificially adding a maximal element to \mathbb{P} to ensure that this happens.

In the context of forcing, if p and q are compatible, then there are no conflicts with what they force: there is no φ with $p \Vdash \varphi$ and $q \Vdash \neg \varphi^n$. This follows from (1) of Motivation 31 B • 3: a common extension $r \leq p, q$ would need to have $r \Vdash \varphi \land \neg \varphi^n$, which would mean any G with $r \in G$ has $V[G] \models \varphi \land \neg \varphi^n$, which is impossible. This still, of course, depends on some knowledge about what G can be, but it provides some motivation on *what* we want G to be.

So we now have the fundamental concepts with preorders: extending individual elements, and extending perhaps incomparable (but still compatible) elements. The notion of density is closely connected with the idea that "most" elements have the property or at least are compatible with the property in a loose sense.

```
— 31 C • 6. Lemma
```

Let \mathbb{P} be a preorder. Let D_0 and D_1 be open, dense sets in \mathbb{P} . Therefore $D_0 \cap D_1$ is open, dense (and in particular, non-empty).

Proof .:.

Since both D_0 and D_1 are closed downwards, so is $D_0 \cap D_1$, meaning it's open. To show density, suppose $p \in \mathbb{P}$. We can extend this to some $p^* \in D_0$ and then to some $p^{**} \in D_1$. Since D_0 is closed downward, $p^{**} \in D_0$. By

transitivity, $p^{**} \leq \mathbb{P} p$ so that $D_0 \cap D_1$ is dense.

§ 31 D. Generic filters and solidifying the motivations

We are now in a position to say what properties G needs to have. Firstly, consider the following theorem of ZFC.

31 D • 1. Theorem Let \mathbb{P} be a preorder with $p \in \mathbb{P}$. Let \mathcal{D} be a countable collection of dense sets. Therefore, there is a filter $G \subseteq \mathbb{P}$ where $p \in G$ and $G \cap D \neq \emptyset$ for every $D \in \mathcal{D}$.

Proof .:.

Enumerate $\mathcal{D} = \{D_n : n < \omega\}$. As D_0 is dense, let $p_0 \leq^{\mathbb{P}} p$ be in D_0 . As in Lemma 31 C • 6, just by continually expanding, we get a sequence $\langle p_n \in \mathbb{P} : n \in \omega \rangle$, where $p_{n+1} \leq^{\mathbb{P}} p_n \leq^{\mathbb{P}} p$ and $p_n \in D_n$. Taking the upward closure of this chain $G = \{q \in \mathbb{P} : \exists n \in \omega \ (p_n \leq^{\mathbb{P}} q)\}$ yields a filter (a set closed upward where all elements are compatible) where $p \in G$, and $p_n \in G \cap D_n$ for each $n < \omega$.

The question then becomes: how many dense sets can we intersect? The generic G is one that intersects all dense sets of V. Of course, \mathbb{P} itself also has this property, but we require in addition that all the elements of G are compatible to ensure Motivation 31 B • 3 holds.

- **31 D • 2.** Definition Let \mathbb{P} be a preorder. Let \mathcal{D} be a collection of dense sents. A set $G \subseteq \mathbb{P}$ is said to be \mathbb{P} -generic over \mathcal{D} iff

- $G \cap D \neq \emptyset$ for every $D \in \mathcal{D}$; and
- G is a filter over \mathbb{P} in that
 - $-p^* \in G$ and $p^* \leq^{\mathbb{P}} p$ implies $p \in G$; and
 - $-p, q \in G$ implies there is some $r \leq^{\mathbb{P}} p, q$ with $r \in G$ (so all elements of G are compatible.)

We say that G is generic over V iff G is \mathbb{P} -generic over $\{D \in V : D \text{ is dense}\}$.

Theorem $31 D \cdot 1$ then tells us we can force over countable, transitive models whose existence is independent of ZFC, but nevertheless consistent (relative to the consistency of ZFC).

– 31 D•3. Corollary

Let $V \vDash$ ZFC be a countable, transitive model. Therefore we can force over V.

Proof .:.

For any preorder $\mathbb{P} \in V$ and $p \in \mathbb{P}$, the set of dense subsets in V is countable (since V is) and therefore Theorem 31 D • 1 tells us there is a G that is \mathbb{P} -generic over V with $p \in G$. \dashv

It will turn out that $G \notin V$ if \mathbb{P} satisfies some weak requirements. So what preorders are appropriate to use in forcing? The following terminology is non-standard, and is really just short-hand for the concept.

31 D • 4. Definition
A preorder P is appropriate for forcing iff
there is a ≤^P-maximal element 1^P; and
for every p ∈ P, there are q, r ≤^P p where q ⊥ r.
Hence we usually refer to preorders as FOL({≤, 1})-structures ⟨P, ≤^P, 1^P⟩ where 1^P is a maximal element.

If \mathbb{P} has no maximal element, we can artificially consider $\mathbb{P}' = \mathbb{P} \cup \{1\}$ and take $1 \ge p$ for each $p \in \mathbb{P}$. \mathbb{P}' then has a maximal element. The reason we want these properties is the following.

 \neg

— 31 D • 5. Theorem -

Let V be a transitive class. Let $\mathbb{P} \in V$ be appropriate for forcing. Therefore, there is no $G \in V$ that is \mathbb{P} -generic over V.

Proof .:.

Suppose $G \cap D \neq \emptyset$ for every dense $D \in V$. Clearly $G \neq \mathbb{P}$, as there are incompatible elements in \mathbb{P} , but all elements of G must be compatible. So consider $\mathbb{P} \setminus G$. This set will be dense.

To see that $\mathbb{P} \setminus G$ is dense, let $p \in \mathbb{P}$ be arbitrary. There are then two incompatible conditions $q \perp r$ below p. Since any two elements of G are compatible, we cannot have both $q, r \in G$. So one of these is in $\mathbb{P} \setminus G$, meaning we have an extension of p in $\mathbb{P} \setminus G$. Hence $\mathbb{P} \setminus G$ is dense, and thus $G \notin V$ as otherwise genericity over V implies $G \cap (\mathbb{P} \setminus G) \neq \emptyset$, a contradiction. \dashv

This is in contrast to Theorem 31 D \cdot 1. In particular, if V is countable in our universe (and thus isn't an inner model, but just some countable transitive model of ZFC), then there are only countably many dense sets in V and hence a generic exists in the real world. It's just that V just doesn't see this subset of \mathbb{P} .

We will never actually confirm that a preorder is appropriate for forcing, because we rarely care whether there is a generic already in V: sometimes V = V[G]. This is the case with trivial preorders, for example: a preorder $\mathbb{P} = \{1\}$ has $G = \mathbb{P} \in V$ as generic over V. So whether $G \in V$ or not is practically irrelevant: we care more what properties V[G] has.

We've already defined the generic extension in Subsection 31 A, but we repeat it here for ease of reference. Recall that a \mathbb{P} -name is just a potential construction hinging on *G* in the sense that it is (inductively) some potential constructions marked with elements of \mathbb{P} . When interpretting a \mathbb{P} -name τ , we just take those elements tagged with an element in *G*: $\tau_G = \{\sigma_G : \exists p \in G \ (\langle \sigma, p \rangle \in \tau)\}$. Note that the collection of \mathbb{P} -names depends on the sets *V* can construct, hence the notation " $V^{\mathbb{P}}$ " for the collection of \mathbb{P} -names in *V*.

- **31 D**•6. Definition Let V be a transitive class we can force over. Let $\mathbb{P} \in V$ be a preorder and G P-generic over V. The generic extension $V[G] = \{\tau_G : \tau \text{ is a } \mathbb{P}\text{-name}\}$. We also call V the ground model.

We can then state the reasons why we only want to consider preorders appropriate for forcing as in Definition 31 D • 4:

When forcing without a maximal element in the preorder P, we can just add a new element 1 as a maximum without any harm: we get the same generic extensions and the generics are easily changed as well: going from P to P ∪ {1}, G → G ∪ {1} identifies the generics of the two preorders.

• When forcing with a preorder \mathbb{P} such that everything below $p \in \mathbb{P}$ is compatible, there is a \mathbb{P} -generic $G \in V$. The second of these is then a kind of converse to Theorem 31 D • 5, showing that such forcings are irrelevent in our goal.

- 31 D • 7. Result

Let $V \models ZFC$ be a transitive model. Let $\mathbb{P} \in V$ and $p \in \mathbb{P}$. Suppose that any two q, r below p are compatible. Therefore G being the upward closure of $\mathbb{P}_{\leq p}$ is \mathbb{P} -generic over V with $G \in V$ and so V[G] = V.

Proof .:.

As $\mathbb{P}_{\leq p}$ is open, any dense set must intersect it. $G = \{q \in \mathbb{P} : \exists r \leq p \ (r \leq q)\}$ is clearly closed upward and any two elements are compatible with p. Hence G is \mathbb{P} -generic over V. But G is in V. As a result, each $\tau \in V^{\mathbb{P}}$ can be decoded as τ_G inside V already. Hence $V[G] \subseteq V$. Since $\check{x} \in V^{\mathbb{P}}$ for each $x \in V$, $\check{x}_G = x \in V[G]$ so that $V \subseteq V[G]$ and hence we have equality. \dashv

Of course, we maybe have other generics that are not in the ground model, but the point is that we aren't guaranteed a new set. And we can rectify this just by continually removing such points (and everything below them) from the preorder until we're left with one appropriate for forcing (or else the empty set). So these are the reasons for considering preorders appropriate for forcing: to work with a maximal element for simplicity, and to guarantee that we're adding new sets. That being said, it might be useful to consider forcing with a trivial preorder later where we force over and over, and at some stages we don't want to do anything. The trivial preorder we have in mind is just $\{1\}$ ordered by $\{\langle 1, 1 \rangle\}$.

Back on topic of formalizing our previous motivations, we now can formally define the forcing relation as in Definition $31 B \cdot 2$.

- 31 D • 8. Definition Let $V \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder and $p \in \mathbb{P}$. Write $p \Vdash \varphi$, iff $V[G] \models \varphi$ for every G that is \mathbb{P} -generic over V with $p \in G$.

This allows us to confirm the results of Motivation $31 B \cdot 3$. The proof of this fact is quite long, however, and is only given at the end of Appendix C. A very thorough treatment of the forcing relation in general can be found in Appendix C as well as in Chapter VII of [20] (which the appendix is based on).

— 31 D•9. Lemma

Let $\mathbf{V} \models \mathsf{ZF}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder and φ a formula. Therefore the relation $\{\langle p, \vec{\tau} \rangle : p \Vdash ``\varphi(\vec{\tau})"\}$ is FOL-definable over \mathbf{V} .

- 31 D • 10. Corollary

Let V be a transitive class we can force over. Let $\mathbb{P} \in V$ be appropriate for forcing and $p \in \mathbb{P}$. Write $p^* \in \mathbb{P}$ for an arbitrary $p^* <^{\mathbb{P}} p$. Therefore,

- 1. $p \Vdash \varphi$ iff every $p^* \Vdash \varphi$;
- 2. $p \Vdash \neg \varphi$ iff every $p^* \not\models \varphi$;
- 3. $p \Vdash ``\varphi \land \psi$ " iff $p \Vdash \varphi$ and $p \Vdash \psi$;
- 4. $p \Vdash \exists x \ \varphi(x)$ iff there is some \mathbb{P} -name τ and extension $p^* \leq \mathbb{P}$ where $p^* \Vdash \varphi(\tau)$; and
- 5. For φ and ψ logically equivalent, $p \Vdash \varphi$ iff $p \Vdash \psi$.

As before with Theorem 31 B \cdot 6, this allows us to characterize truth in V[G].

- 31 D • 11. Corollary Let V be a transitive class we can force over. Let $\mathbb{P} \in V$ be a preorder. Let G be \mathbb{P} -generic over V. Therefore $V[G] \models \varphi$ iff $\exists p \in G \ (p \Vdash \varphi)$.

Moreover, through tedious checking, we can confirm each individual axiom of ZFC in V[G].

— 31 D•12. Theorem –

Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder. Let G be \mathbb{P} -generic over V. Therefore $V[G] \vDash \mathsf{ZFC}$. In fact, more generally

- $V \vDash ZF^-$ (i.e. $V \vDash ZF P + Col$) implies $V[G] \vDash ZF P$;
- $V \vDash \mathsf{ZF}$ implies $V[G] \vDash \mathsf{ZF}$;
- $\mathbf{V} \models \mathsf{ZFC}$ implies $\mathbf{V}[G] \models \mathsf{ZFC}$.

The proof of the model theoretic implications are quite tedious, but can also be found later in Appendix C and in Chapter VII of [20].

Overall, we have the following properties of the generic extension.

— 31 D • 13. Theorem

Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder. Let G be \mathbb{P} -generic over V. Therefore

1. $V[G] = \{\tau_G : \tau \in V^{\mathbb{P}}\}$ is transitive with $V \cup \{G\} \subseteq V[G]$.

2. $V[G] \models ZFC$.

- 3. V[G] is the least transitive model of ZF where $V \cup \{G\} \subseteq V[G]$.
- 4. For any FOL-formula φ and \mathbb{P} -names $\vec{\tau}$, $V[G] \models "\varphi(\vec{\tau}_G)$ " iff $p \Vdash "\varphi(\vec{\tau})$ " for some $p \in G$.

Proof .:.

- 1. This follows from Definition $31 D \cdot 6$ and Theorem $31 A \cdot 6$ (1) and (2).
- 2. This follows from Theorem 31 D 12.

 \dashv

- 3. This follows from Theorem 31 A 6 (3) and Theorem 31 D 12.
- 4. This follows from Corollary 31 D 11.

And these are the main properties we actually care about. The properties of the forcing relation we care about are covered in Definition $31 D \cdot 8$, Lemma $31 D \cdot 9$, and Corollary $31 D \cdot 10$.

Section 32. Examples of Forcing

If forcing is all about adding new objects into the universe, we should think about what sorts of objects we want to add. There are all sorts of preorders that generically add in all sorts of objects. There are, of course, limits to what we can add with forcing,^{iv} but commonly we add in functions and subsets. So this will be the purpose of the first few preorders: add in subsets and functions.

One immediate question that pops up is how do we choose what preorder to force with? Commonly, the idea begins with the goal in mind: we want to add in some $G \subseteq V$. Our preorder will often consist of V's approximations to G where $p \leq q$ iff p approximates more than q. In the context of functions and subsets, this ordering is usually containment: $p \leq q$ iff $p \supseteq q$.

§ 32 A. Collapsing cardinals

Recall that cardinals are really just special ordinals. They are determined by what functions^v the model V has. For example, to calculate ω_1 , the general idea is that V looks at each ordinal α , determines whether there's a bijection with ω . Then, the first place it has no bijection, it stops and says "this is ω_1 ".

But because this is all based on what functions V has, if we add in a bijection with ω and, say, $\alpha = \omega_1^V$, we can show this ordinal α is countable in V[G], meaning $\omega_1^V = \alpha < \omega_1^{V[G]}$. We can also generalize this, but let's stick with "collapsing" a cardinal to ω for now.

32A·1. Definition
Let κ be an infinite ordinal. The preorder Col(ℵ₀, κ) = ⟨Col(ℵ₀, κ), ≤⟩ consists of functions f where
|f| < ℵ₀; and
dom(f) ⊆ ω, im(f) ⊆ κ.
We write f ≤ g iff f ⊇ g.

We say f is a *partial function* from ω to κ , written elsewhere in this document as $f : \omega \rightarrow \kappa$, in the sense that $f : A \rightarrow \kappa$ for some subset $A \subseteq \omega$. This shorthand is quite useful as each f is an approximation to a full-fledged function from ω to κ .

The first thing to confirm is that this gives us what we want: G codes a surjection from \aleph_0 onto κ . We will show this slowly with all the detail. The main idea is that $\bigcup G = g$ inherets properties from the approximations in $Col(\aleph_0, \kappa)$. So since we can always add in an $n < \omega$ into the domain of these approximations, and always add in an $\alpha < \kappa$ into the range, it follows that these form dense sets. Hence each $n < \omega$ is in the domain of g, and each $\alpha < \kappa$ is in the range of g.

- 32 A · 2. Theorem

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let κ be an infinite ordinal. Let G be $\mathsf{Col}(\aleph_0, \kappa)^{\mathbf{V}} = \mathbb{P}$ -generic over V. Therefore $\bigcup G = g$ is a surjection from ω to κ . In particular, $\mathbf{V}[G] \models "|\kappa| = \aleph_0$ ".

Proof .:.

We need to confirm several things: that g is in fact a function, that dom $(g) = \omega$, and that im $(g) = \kappa$ so that g is a surjection.

^{iv}For example, we can't add ordinals with forcing.

^vIn particular, what *bijections*

EXAMPLES OF FORCING

 $\int_{g} f$ Claim 1 $\int_{g} f$ Claim 1

Proof .:.

This is a simple application of compatibility of *G*. In particular, if $\langle n, \alpha \rangle$, $\langle n, \beta \rangle \in g$ for some $n < \omega$ and $\alpha, \beta < \kappa$, there are some $f, g \in G$ with $f(n) = \alpha$ and $g(n) = \beta$. By compatibility, there is some $h \in G$ with $h \leq f, g$, meaning a finite, partial function $h \supseteq f, g$. But then $\langle n, \alpha \rangle$, $\langle n, \beta \rangle \in f \cup g \subseteq h$ requires that $\alpha = \beta$ for *h* to be a function at all. Hence $\langle n, \alpha \rangle$, $\langle n, \beta \rangle \in g$ implies $\alpha = \beta$ and thus *g* is a function. \dashv

Just by definition of $g = \bigcup G \subseteq \omega \times \kappa$, we have that g has domain dom $(g) \subseteq \omega$ and im $(g) \subseteq \kappa$. The issue, however, is whether we have equality. This is where dense sets come into play.

Proof .:.

We need to show that for each $n < \omega$, there is some α with $\langle n, \alpha \rangle \in \bigcup G$. The only real way we have to ensure something is in G is to find an appropriate dense set. Then G intersects it, and we have a witness. So for our case, we need an $f \in G$ where $\langle n, \alpha \rangle \in f$ for some α . For each $n < \omega$, consider the set of such f:

$$D_n = \{ f \in \mathbb{P} : n \in \operatorname{dom}(f) \}$$

Note that this is dense in \mathbb{P} , since for any $p \in \mathbb{P}$, if $n \in \text{dom}(p)$, we're done. If $n \notin \text{dom}(p)$, then we just choose some $\alpha < \kappa$ not already in im(p) (p is finite while κ is infinite, so this is possible), and then consider $q = p \cup \{\langle n, \alpha \rangle\} \leq p$. This $q \in D_n$ and extends our arbitrary $p \in \mathbb{P}$, so each D_n is dense.

In particular, $G \cap D_n \neq \emptyset$ for each *n*, and thus $n \in \text{dom}(f) \subseteq \text{dom}(g)$ for some $f \in G$, implying each $n \in \text{dom}(g)$. Therefore $\omega \subseteq \text{dom}(g)$. Since clearly $\text{dom}(g) \subseteq \omega$, we have equality. \dashv

So all that remains to be shown is that g is surjective. To see this, we proceed exactly like in Claim 2 for the range. Let $\alpha < \kappa$ be arbitrary. Consider the set

$$E_{\alpha} = \{ f \in \mathbb{P} : \alpha \in \operatorname{im}(f) \}.$$

This set is dense by the same reason as above: since $p \in \mathbb{P}$ is finite, take $n \in \omega \setminus \text{dom}(p)$ and add in $\langle n, \alpha \rangle$: $q = p \cup \{\langle n, \alpha \rangle\} \leq p$ has $q \in E_{\alpha}$ and thus E_{α} is dense. Therefore there is some $f \in G \cap E_{\alpha}$ and so $\alpha \in \text{im}(f) \subseteq \text{im}(g)$. As $\alpha < \kappa$ was arbitrary, $\kappa \subseteq \text{im}(g)$. We obviously have $\text{im}(g) \subseteq \kappa$, and thus equality. This means $g : \omega \to \kappa$ is a surjection. It follows that $\mathbf{V}[G] \models ``|\kappa| = \aleph_0$ ''.

Where exactly did the forcing relation come into play here? The idea is that $f \in \mathbb{P}$ has $f \Vdash ``f \subseteq \dot{g}`$, where \dot{g} is a name for g^{vi} and it's this sense that our $f \in \mathbb{P}$ is an approximation to g.

The above forcing notion gives us the idea of "collapsing" a cardinal in the following sense. This allows us to consider other forcing notions that *do not* collapse cardinals. Generally such notions are said to *preserve* cardinals, though perhaps it would be better to say that they preserve the $\aleph_{\alpha}s$, as other cardinalities can change.

- 32 A • 3. Definition

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder. We say that \mathbb{P} preserves cardinals iff every $\alpha \in V$ such that $\mathbf{V} \models ``\alpha = |\alpha|$ '' has $\mathbb{I}^{\mathbb{P}} \Vdash ``\check{\alpha} = |\check{\alpha}|$ ''. Equivalently, \mathbb{P} preserves cardinals iff for every \mathbb{P} -generic G over $V, \mathbf{V} \models ``\alpha = |\alpha|$ '' iff $\mathbf{V}[G] \models ``\alpha = |\alpha|$ ''.

The above definition also makes sense for other properties. For example, one can say that \mathbb{P} preserves cofinalities whenever \mathbf{V} and $\mathbf{V}[G]$ agree on the function $\alpha \mapsto \operatorname{cof}(\alpha)$. Similarly, \mathbb{P} preserves stationary sets whenever $S \in V$ is stationary implies $\mathbf{V}[G] \models$ "S is stationary". So it should be obvious that $\operatorname{Col}(\omega, \kappa)$ does not preserve cardinality when $\kappa > \aleph_0$. But $\operatorname{Col}(\omega, \kappa)$ does preserve cardinals $\leq \aleph_0$.^{vii}

^{vi} for example, $\{\langle \langle n, \alpha \rangle, p \rangle \in (\omega \times \kappa) \times \mathbb{P} : p(n) = \alpha\}$ (where "(x)" is just a way of writing "x" when "x" is too large written out).

^{vii}All forcings do this as ω and $n < \omega$ are absolute between transitive models of set theory.

32A·4. Definition Let $\lambda < \kappa$ be infinite ordinals. The preorder $Col(\lambda, \kappa) = \langle Col(\lambda, \kappa), \leqslant \rangle$ consists of functions f where: • $|f| < \lambda$; and • $dom(f) \subseteq \lambda$, $im(f) \subseteq \kappa$. We write $f \leqslant g$ iff $f \supseteq g$.

We of course have a similar property as before:

— 32 A • 5. Theorem —

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\lambda < \kappa$ be infinite cardinals of \mathbf{V} . Let G be $\mathsf{Col}(\lambda, \kappa)^{\mathbf{V}} = \mathbb{P}$ generic over V. Therefore $\bigcup G = g$ is a surjection from λ to κ . In particular, $\mathbf{V}[G] \models ``|\kappa| = |\lambda|$ ''.

Proof .:.

By the same reasoning as before, we already know g is a function with domain dom(g) $\subseteq \lambda$ and image im(g) $\subseteq \kappa$. So we want to show equality of each of these. For each $\alpha < \lambda$ and each $\beta < \kappa$, consider the sets

$$D_{\alpha} = \{ f \in \mathbb{P} : \alpha \in \operatorname{dom}(f) \}$$
$$E_{\beta} = \{ f \in \mathbb{P} : \beta \in \operatorname{im}(f) \}.$$

Since $\kappa > \lambda$, any $f \in \mathbb{P}$ (which then has size $|f| < \lambda$) is not a bijection: dom $(f) \subsetneq \lambda$ and im $(f) \subsetneq \kappa$. If $\alpha \notin \text{dom}(f)$ and $\beta \notin \text{im}(f)$, $f^* = f \cup \{\langle \alpha, \beta \rangle\} \leqslant f$ has $f^* \in D_{\alpha} \cap E_{\beta}$ and thus each is dense. Hence $G \cap D_{\alpha} \neq \emptyset$ and $G \cap E_{\beta} \neq \emptyset$ for each $\alpha < \lambda$ and $\beta < \kappa$. In particular, this yields that $\alpha \in \text{dom}(g)$ and $\beta \in \text{im}(g)$ for each $\alpha < \lambda$ and $\beta < \kappa$, meaning $g : \lambda \to \kappa$ is a surjection.

This shows $Col(\lambda, \kappa)$ collapses (the cardinality of) κ to λ . $Col(\lambda, \kappa)$ still preserves some cardinals by the following fact.

— 32 A · 6. Definition ——

Let α be an ordinal. A preorder \mathbb{P} is $< \alpha$ -closed iff for every $\leq^{\mathbb{P}}$ -decreasing sequence $\langle p_{\xi} \in \mathbb{P} : \xi < \gamma \rangle$ of length $\gamma < \alpha$, there is some condition $p \in \mathbb{P}$ below each entry: $p \leq^{\mathbb{P}} p_{\xi}$ for each $\xi < \gamma$.

- 32 A • 7. Lemma - Col(λ, κ) is $< cof(\lambda)$ -closed.

Proof .:.

Let $\langle p_{\alpha} : \alpha < \gamma \rangle$ be as in the statement. Thus $p = \bigcup_{\alpha < \gamma} p_{\alpha}$ is a partial function from λ to κ . Moreover, as $\gamma < \operatorname{cof}(\lambda)$ and each p_{α} has size $|p_{\alpha}| < \lambda$, it follows that this union has size $|p| \le |\gamma| \cdot \sup_{\alpha < \gamma} |p_{\alpha}| < \lambda$. Hence p is a condition in \mathbb{P} , and clearly lies below each p_{α} .

This gives the following corollary. It's also a nice exercise to see how the result should change if λ is not regular.

— 32 A • 8. Lemma

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let λ be a regular, infinite cardinal of \mathbf{V} . Let $\mathbb{P} \in V$ be a preorder that is $< \lambda$ -closed in V. Therefore \mathbb{P} preserves cardinals $\leq \lambda$. In other words, for G \mathbb{P} -generic over V and $\theta < \lambda$, $\mathbf{V} \models "\theta = |\theta|$ " iff $\mathbf{V}[G] \models "\theta = |\theta|$ ".

Proof .:.

Just by downward absoluteness, if $V[G] \models "\theta = |\theta|"$, then clearly $V \models "\theta = |\theta|"$, because if V[G] has no bijections from smaller ordinals to θ then neither does V. So suppose $V \models "\theta = |\theta|"$, but $V[G] \models "|\theta| = \delta < \theta"$ as witnessed by a bijection $f : \delta \to \theta$ in V[G]. $f \in V[G]$ has a name $\dot{f} \in V^{\mathbb{P}}$ and by Corollary 31 D • 11, there is some $p_0 \in G$ where $p_0 \Vdash "\dot{f}$ is a function from $\check{\delta}$ to $\check{\theta}"$.

We'd like to do the following that doesn't actually work. It *does* provide motivation, however, of continually deciding more and more of f. The $\langle \lambda$ -closure of \mathbb{P} allows us to decide all of f with a single condition. Construct a $\leq^{\mathbb{P}}$ -decreasing sequence $\langle p_{\alpha} \in G : \alpha < \delta \rangle$ where $\exists \beta < \theta \ (p_{\alpha} \Vdash ``\dot{f}(\check{\alpha}) = \check{\beta}`')$. We would do this recursively: starting with p_0 as above and for $f(\alpha) = \beta$, let $q_{\alpha} \in G$ be such that $q_{\alpha} \Vdash ``\dot{f}(\check{\alpha}) = \check{\beta}''$. Then we find a common extension $p_{\alpha+1} \leq^{\mathbb{P}} q_{\alpha}, p_{\alpha}$. At limit stages $\gamma < \delta$, we would appeal to $\langle \lambda$ -closure to find a $p_{\gamma} \leq^{\mathbb{P}} p_{\alpha}$ for all $\alpha < \gamma$. Then a $p^* \leq^{\mathbb{P}} p_{\alpha}$ for all $\alpha < \delta$ (which is supposed to exist by $\langle \lambda$ -closure) has

$$f = \{ \langle \alpha, \beta \rangle : p^* \Vdash ``\dot{f}(\check{\alpha}) = \check{\beta}" \} \in V.$$

This implies $f \in V$ is a bijection from δ to θ , contradicting that $\mathbf{V} \models "\theta = |\theta|"$.

The issue with this approach is that this construction of $\langle p_{\alpha} : \alpha < \delta \rangle$ takes place outside of V. In other words, because V doesn't have access to G, it cannot form this sequence. This is especially obvious when V is some countable transitive model where it's clear V doesn't contain all countable sequences (V thinks P is < λ -closed just in case all sequences in V have lower bounds, but many sequences outside V may not). So we must use a slightly different approach with the same motivating idea just translated into terms of dense sets. But if the reader understands the idea above, this is enough, as the actual approach below doesn't add much understanding.^{viii}

For each $\alpha < \delta$, consider

$$D_{\alpha} = \{ q \leq^{\mathbb{P}} p_0 : \exists \beta < \theta \ (q \Vdash ``\dot{f}(\check{\alpha}) = \check{\beta}") \}$$

This will be open by Corollary 31 D • 10 (1), and also dense below p_0 . To see density, let $q \leq^{\mathbb{P}} p_0$ be arbitrary. Let H be \mathbb{P} -generic over V with $q \in H$. Thus $V[H] \models "f(\alpha) = \beta$ " for some β . There is then some $r \in H$ forcing this: $r \Vdash "\dot{f}(\check{\alpha}) = \check{\beta}$ ". A common extension $q^* \leq^{\mathbb{P}} r, q$ has $q^* \in D_{\alpha}$. So D_{α} is dense below p_0 .

– Claim 1 –

For each $\gamma \leq \delta$, $\bigcap_{\alpha < \gamma} D_{\alpha}$ is open (by Corollary 31 D • 10 (1)) and dense below p_0 .

Proof ::.

Suppose the result holds for all ordinals below γ . Let $q \leq^{\mathbb{P}} p_0$ be arbitrary. Choose $p'_{\beta} \in \bigcap_{\alpha < \beta} D_{\alpha}$ for $\beta < \gamma$. Without loss of generality, choose the p'_{β} s so that they are $\leq^{\mathbb{P}}$ -decreasing (density allows this) with $p'_0 \leq^{\mathbb{P}} q$. As $\gamma < \delta < \lambda$, by $< \lambda$ -closure, there is some $q^* \leq^{\mathbb{P}} p'_{\beta}$ for every $\beta < \gamma$. By density of D_{γ} , there is some $p'_{\gamma} \leq^{\mathbb{P}} q^* \leq^{\mathbb{P}} q$ in D_{γ} and in fact in $\bigcap_{\alpha < \gamma} D_{\alpha}$. Thus $\bigcap_{\alpha < \gamma} D_{\alpha}$ is dense below p_0 . \dashv

Note that each $p_0^* \in D_\delta$ decides all of \dot{f} . In particular, for $p_0^* \in G \cap D_\delta$,

$$f = \{ \langle \alpha, \beta \rangle : p_0^* \Vdash ``\dot{f}(\check{\alpha}) = \check{\beta}'' \} \in V,$$

tells us that $f \in V$ and we get the contradiction as in the faulty proof.

 \dashv

- 32 A • 9. Corollary

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let λ be a regular, infinite cardinal of \mathbf{V} with $\lambda < \kappa \in V$. Therefore $\mathbb{P} = \mathsf{Col}(\lambda, \kappa)^{\mathsf{V}}$ preserves cardinals $\leq \lambda$.

§32B. Forcing –CH

The above forcing shows that cardinals are not absolute between transitive models of set theory. Now we will show both that CH is independent of ZFC,^{ix} and that we can both preserve cardinals and change *cardinality*. To argue this, we will need a little more technology. Lemma 32 A • 8 tells us that $< \lambda$ -closure of a preorder implies preserving cardinals $\leq \lambda$. To argue whether a preorder preserves cardinals $> \lambda$, we need to talk more about names in the next subsection.

But first, let's introduce the preorder for forcing $\neg CH$, and then show this gives what we want. Recall that $\mathcal{P}(\omega)$ can be identified with characterisitic functions $f : \omega \to 2$: $X \subseteq \omega$ is just $\{n < \omega : \chi_X(n) = 1\}$ where $\chi_X(n) = 1$ if $n \in X$ and 0 otherwise. If we want to add a subset of ω , we could then consider the preorder which looks at finite

viii it does, however, motivate the definition of the *distibutivity* of a preorder.

^{ix}Since L \models ZFC + CH, this shows CH is relatively consistent with ZFC, so we will merely show ZFC + \neg CH is relatively consistent with ZFC.

approximations of characteristic functions: finite partial functions from ω to 2. If we want to add a *lot* of partial functions to bump up the size of $\mathcal{O}(\omega)$, we can instead index them: adding in $\{f_{\alpha} \in {}^{\omega}2 : \alpha < \kappa\}$ for κ some cardinal. This is equivalent to adding in a single function $f : \kappa \times \omega \to 2$ where each f_{α} is just the slice $n \mapsto f(\alpha, n)$. So this is what we are approximating.

 $-32B \cdot 1. \text{ Definition}$ Let κ, λ be a cardinals. Define $\text{Add}(\lambda, \kappa) = \langle \text{Add}(\lambda, \kappa), \leq \rangle$ by $\text{Add}(\lambda, \kappa) = \{p : \kappa \times \lambda \rightarrow 2 : |p| < \lambda\} \text{ where } p \leq q \text{ iff } p \supseteq q.$

We will mostly consider $Add(\aleph_0, \kappa)^{\mathsf{x}}$ since we're focusing on CH.

If we consider the forcing relation on $\operatorname{Add}(\aleph_0, \kappa)$, for $g = \bigcup G$, $p \Vdash \check{p} \subseteq \dot{g}$. As with $\operatorname{Col}(\lambda, \kappa)$, we can show this does what we want. Note that we freely identify $g : \kappa \times \omega \to 2$ as a function $g : \kappa \to {}^{\omega}2$ just by taking α to the map g_{α} defined by $g_{\alpha}(n) = g(\alpha, n)$.

— 32B•2. Theorem

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\kappa > |\mathcal{P}(\omega)|^{\mathsf{V}}$ be an infinite, regular cardinal of \mathbf{V} . Let G be $\mathsf{Add}(\aleph_0, \kappa)^{\mathsf{V}} = \mathbb{P}$ generic over V. Therefore $g = \bigcup G \in V[G]$ yields an injection from κ to ^{ω}2. In particular, $\mathbf{V}[G] \models "|\mathcal{P}(\omega)| \ge |\kappa|$ ".

Proof .:.

It should be clear that g is a function by compatibility of G. By considering for each $\alpha < \kappa$ and $n < \omega$

$$D_{\alpha,n} = \{ p \in \mathbb{P} : \langle \alpha, n \rangle \in \operatorname{dom}(p) \},\$$

which is clearly dense (recall each $p \in \mathbb{P}$ is finite, so we can just add $\langle \langle \alpha, n \rangle, 1 \rangle$ to p if $p \notin D_{\alpha,n}$ and get an extension in $D_{\alpha,n}$), it should be clear that $g : \kappa \times \omega \to 2$: $f \in D_{\alpha,n} \cap G$ has $\langle \alpha, n \rangle \in f \subseteq g$.

It then suffices to show that g is injective, or rather the map $\alpha \mapsto g_{\alpha}$ is injective, where $g_{\alpha}(n) = g(n, \alpha)$. To do this, for distinct $\alpha, \beta < \kappa$, consider the set

$$E_{\alpha,\beta} = \{ p \in \mathbb{P} : \exists n < \omega(\langle \alpha, n \rangle, \langle \beta, n \rangle \in \operatorname{dom}(p) \land p(\alpha, n) \neq p(\beta, n)) \}$$

So if $p \in E_{\alpha,\beta} \cap G$, then then g_{α} and g_{β} disagree somewhere. Note that each $E_{\alpha,\beta}$ is dense, since each $p \in \mathbb{P}$ is finite: there are only finitely many n where $\langle \alpha, n \rangle, \langle \beta, n \rangle \in \text{dom}(p)$. Hence some n beyond all of these yields an extension $p^* = p \cup \{\langle \langle \alpha, n \rangle, 1 \rangle, \langle \langle \beta, n \rangle, 0 \rangle\} \leq p$ with $p^* \in E_{\alpha,\beta}$. Therefore, $G \cap E_{\alpha,\beta} \neq \emptyset$ for each α, β , implying $g_{\alpha} \neq g_{\beta}$ for each $\alpha \neq \beta < \kappa$. So V[G] has an injection from κ to $({}^{\omega}2)^{V[G]}$, and therefore to $\mathcal{P}(\omega)^{V[G]}$. So $V[G] \models "|\mathcal{P}(\omega)| \geq |\kappa|$ ".

This doesn't tell us, however, that κ is preserved. A priori, we could have $V[G] \models "|\kappa| = \aleph_1$ " so that the above theorem says $V[G] \models "|\mathcal{O}(\omega)| \ge \aleph_1$ ", which we already know is true since $V[G] \models \mathsf{ZFC}$. To show that κ is preserved, we basically need to show that there aren't any bijections from smaller cardinals. This, in essence, amounts to showing that there aren't too many "choices" our preorder can allow, and this is related to the concept of antichains.

§32C. Antichains

Density has a close connection to *antichains*. Again, this is a general topological concept that one can prove is equivalent to the definition below in the context of preorders. But we have no need for the general definition.

— 32C•1. Definition –	
Let P be a preorder. A se	t <i>A</i> ⊆ \mathbb{P} is an <i>antichain</i> iff any two distinct <i>p</i> , <i>q</i> ∈ <i>A</i> have <i>p</i> ⊥ <i>q</i> .

Maximal antichains also have some nice properties with forcing. One thing that will show this is the following result.

^x frequently read as "add to the powerset of $\aleph_0 \kappa$ -many subsets"

— 32C•2. Lemma -

Let \mathbb{P} be a preorder. Let $A \subseteq \mathbb{P}$ be an antichain. Therefore A is \subseteq -maximal iff every $p \in \mathbb{P}$ has some $q \in A$ where p, q are compatible.

Proof .:.

Suppose A is maximal. Suppose $p \in \mathbb{P}$ is incompatible with every element of A. Thus $A \cup \{p\}$ is an antichain extending A, contradicting maximality.

So suppose A is not maximal. Hence there is some antichain $A \subseteq \mathbb{P}$ with $A \subsetneq A$. Any $p \in A \setminus A$ has, since A is an antichain, $p \perp q$ for each $q \in A$. Thus there is an element of \mathbb{P} with no $q \in A$ where p and q are compatible. This is the contrapositive of the desired direction. \dashv

Note that this has a similar flavor to density: a dense set D allows you to always extend to enter D. Similarly, a maximal antichain A allows you to always find an incompatible element to enter A.

— 32 C • 3. Result

Let \mathbb{P} be a preorder. Suppose every $p \in \mathbb{P}$ has an extension $p^* \in \mathbb{P}$ (meaning there are no bottom nodes).

• Let $A \subseteq \mathbb{P}$ be an antichain. Therefore $\mathbb{P} \setminus A$ is dense.

• In fact, every dense set contains a \subseteq -maximal antichain.

Proof .:.

To show that $\mathbb{P} \setminus A$ is dense, let $p \in \mathbb{P}$ be arbitrary. If $p \in A$, then an extension $p^* \leq^{\mathbb{P}} p$ cannot be in A (as p^* and p are compatible with the obvious common extension p^*). If $p \notin A$, then $p \leq^{\mathbb{P}} p$ has $p \in \mathbb{P} \setminus A$. Hence $\mathbb{P} \setminus A$ is dense.

Now suppose $D \subseteq \mathbb{P}$ is open and dense. We will show that there is a \subseteq -maximal antichain $A \subseteq D$. We know by Zorn's lemma that there is a \subseteq -maximal element in the set of antichains $\{A \subseteq D : A \text{ is an antichain of } \mathbb{P}\}$. So it suffices to show that this maximal element A is a maximal antichain in the context of the rest of \mathbb{P} . So let $p \in \mathbb{P}$ be arbitrary. As D is dense, there is some $p^* \in D$ extending p. Now working in D the same reasoning in Lemma 32 C • 2 tells us that there is some element $q \in A$ compatible with p^* . But then q is compatible with p: there is an $r \leq^{\mathbb{P}} q$ and $r \leq^{\mathbb{P}} p^* \leq^{\mathbb{P}} p$.

So there is a nice interplay between dense sets, and maximal antichains. How does this help us? Well, antichains represent choices: if G is P-generic over V with $A \in V$ a maximal antichain, then $G \cap A$ is a singleton. Moreover, the above result tells us that for G to be generic, G must intersect all maximal antichains.

- 32 C • 4. Corollary

Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force over and $\mathbb{P} \in V$ be a preorder. Let A be a maximal antichain of V. Let G be \mathbb{P} -generic over V. Therefore, $|G \cap A| = 1$.

Proof .:.

Clearly as any two elements of *G* are compatible while any two elements of $A \in V$ are incompatible, $|G \cap A| \le 1$. Consider the downward closure of *A*, $A \downarrow$, which is dense by Lemma 32 C • 2: any $p \in \mathbb{P}$ has some $q \in A$ where p and q are compatible, and therefore there is some common extension $q^* \in A \downarrow$. By density, $G \cap A \downarrow \neq \emptyset$ and so there is some $q^* \in G \cap A \downarrow$. Since *G* is closed upward, $q \in G$, showing $|G \cap A| \ge 1$ and so we have equality.

In fact, this idea gives an alternative characterization of what it means to be generic.

— 32 C • 5. Theorem —

Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force over and $\mathbb{P} \in V$ a preorder. Therefore G is \mathbb{P} -generic over V iff G is a filter over \mathbb{P} such that $|G \cap A| = 1$ for every maximal antichain $A \in V$.

EXAMPLES OF FORCING

Proof .:.

We've shown the (\rightarrow) direction with Corollary 32 C • 4. So let *G* be a filter that intersects each maximal antichain of *V* at exactly one point. Let *D* be dense so that *D* contains a maximal antichain $A \subseteq D$ by Result 32 C • 3. It follows that $\emptyset \neq G \cap A \subseteq G \cap D$, as desired.

This is useful, because we can now talk about what kinds of antichains \mathbb{P} has.

— 32 C • 6. Definition -

Let \mathbb{P} be a preorder and κ be a cardinal. \mathbb{P} is κ -cc (has the κ -chain condition) iff every antichain A of \mathbb{P} has size $|A| < \kappa$ (in the ground model). We say \mathbb{P} is ccc iff it is \aleph_1 -cc.

We introduce this, because κ -cc preorders preserve cardinals and cofinalities $\geq \kappa$ for κ regular. We show this only for ccc preorders, but the proof generalizes. Again, it's a good exercise to see how this changes for κ -cc preorders when κ is singular.

- 32C•7. Theorem

Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be ccc in V. Therefore, \mathbb{P} preserves all cardinals and cofinalities, meaning if $V \vDash ``\alpha = \operatorname{cof}(\beta)"$, then $V[G] \vDash ``\alpha = \operatorname{cof}(\beta)"$, for any G that is \mathbb{P} -generic over V.

Proof .:.

Suppose not. We have two possibilities.

- A cofinality $cof^{V}(\beta)$ is not preserved. Thus $cof^{V}(\beta)$ —a regular cardinal in V—is not regular in V[G].
- A cardinal κ is not preserved. If κ was a limit cardinal in V, then for $\mathbf{V}[G] \models "|\kappa| = \lambda < \kappa$ " has $(\lambda^+)^{\mathbf{V}}$ a regular cardinal of V no longer regular in $\mathbf{V}[G] \models "|(\lambda^+)^{\mathbf{V}}| \le |\kappa| = \lambda$ ". Similarly, if κ is a successor cardinal in V, then it's no longer regular in V[G].

So it suffices to show that every regular $\kappa \in V$ is regular in V[G]. Let $x \in V[G]$ be a λ -length sequence in κ with $\lambda < \kappa$. We will show this is bounded in V[G]. Let \dot{x} be a \mathbb{P} -name for x, and let $p \in \mathbb{P}$ force that \dot{x} is a function from $\dot{\lambda}$ to $\check{\kappa}$. In V, we can consider the possible values of $\dot{x}(\check{\xi})$ for each $\xi < \lambda$:

 $A_{\xi} = \{\beta < \kappa : \exists p \in \mathbb{P} \ (p \Vdash ``\dot{x} \text{ is a function from } \check{\lambda} \text{ to } \check{\kappa} \text{ and } \dot{x}(\check{\xi}) = \check{\beta}``)\}.$

Note that this is the result of an antichain of $p \in \mathbb{P}$: no two compatible p, q can force different values of $\dot{x}(\xi)$. In other words, for each $\beta \in A_{\xi}$, let $p_{\beta} \leq^{\mathbb{P}} p$ have $p_{\beta} \Vdash "\dot{x}(\xi) = \check{\beta}"$. Therefore, $\mathcal{A}_{\xi} = \{p_{\beta} \in \mathbb{P} : \beta < \kappa\}$ is an antichain. Since \mathbb{P} is ccc, \mathcal{A}_{ξ} is countable, and thus so is A_{ξ} for each $\xi < \lambda$. But then $\sup_{\xi < \lambda} A_{\xi}$ is bounded in κ since κ is regular, $\lambda < \kappa$, and each $|A_{\xi}| \leq \aleph_0$. Therefore for $\kappa > \beta > \sup_{\xi < \lambda} A_{\xi}$, each $q \in \mathbb{P}$ forces " $\dot{x}(\xi) < \check{\beta}$ ", meaning $\dot{x}_G = x$ can't be unbounded in κ .

The proof above actually gives a stronger result: if we force something to be bounded in an uncountable, regular cardinal, then we can actually find a *uniform* bound in the ground model. We don't just find some *name* τ for an ordinal (which might be forced to be different ordinals depending on the generic), we get one of the form $\check{\alpha}$ for some $\alpha < \kappa$.

- 32 C • 8. Corollary

Let $\mathbf{V} \vDash \mathsf{ZFC}$ be a transitive model we can force over. Let \mathbb{P} be ccc in \mathbf{V} . Suppose $p \in \mathbb{P}$ has $p \Vdash ``\tau \subseteq \check{\kappa} \land |\tau| < \check{\kappa}$ '' for some \mathbb{P} -name τ and regular $\kappa > \omega$. Therefore there is some $\alpha < \kappa$ where $p \Vdash ``\tau \subseteq \check{\alpha}$ ''.

Note that being ccc or κ -cc is a statement about the ground model: V may have fewer antichains than V[G] does and so a preorder $\mathbb{P} \in V$ may be ccc in the ground model, but not ccc in the generic extension. We may generalize Theorem 32 C • 7 in that κ -cc preorders preserves cardinals $\geq \kappa$. Cofinality might still be changed, however, if the cofinality of a cardinal $\lambda > \kappa$ is below κ : $\lambda > \kappa > \operatorname{cof}(\kappa)$.

§ 32 D. Showing we actually did force ¬CH

Let's return to $Add(\aleph_0, \kappa)$ where κ is regular. We can pretty easily show that $Add(\aleph_0, \kappa)$ is ccc.

- 32 D • 1. Lemma For every ordinal κ , $\mathbb{P} = \mathsf{Add}(\aleph_0, \kappa)$ is ccc.

Proof .:.

Clearly if \mathbb{P} is countable (e.g. if κ is countable), then every antichain is countable. So let $A \subseteq \mathbb{P}$ be an uncountable subset of \mathbb{P} . Consider the set of domains of $p \in A$: $D = \{ dom(p) : p \in A \}$. D must also be uncountable, since each $d \in D$ has only countably many (in fact, finitely many) functions from d to 2, so if D were countable, then would A be too.

- Claim 1 (The Δ -System Lemma) -

There is an uncountable $D' \subseteq D$ and $r \in \mathbb{P}$ such that any two distinct $p, q \in B$ have $p \cap q = r$. In other words, D' forms a Δ -system.

Now we can lift this into A: $A' = \{p \in A : dom(p) \in D'\}$, another uncountable set, but this time, for any two distinct $p, q \in A$, dom $(p) \cap$ dom(q) is some fixed, finite $X \subseteq \kappa \times \omega$. Since there are only countably many functions from X to 2, there must be some $p: X \to 2$ with uncountably many $q, q' \in A'$ with $q \upharpoonright X = q' \upharpoonright$ x = p.

Note that this implies any two $p, q \in A'$ are compatible: $p \cup q$ is a function since p and q never disagree. $p \cup q$ is still a finite partial function from $\kappa \times \omega$ to 2, and thus A cannot be an antichain. So the only antichains of P are \neg countable.

Proving the Δ -system lemma isn't particularly interesting nor difficult, but it's very useful when dealing with forcing. It's a common idea that can be found in most any standard, introductory reference for set theory, e.g. [20]. We include its proof here merely for the sake of being thorough: combinatorics is mostly off-topic for this document.^{xi}

- 32 D \cdot 2. Lemma (The Δ -System Lemma)

Let A be an uncountable family of finite sets. Therefore, there is an uncountable $B \subseteq A$ (in particular, we can take |B| to be any regular cardinal $\leq |A|$) where $\exists r \forall p, q \in B \ (p \neq q \rightarrow p \cap q = r)$ (i.e. B forms a Δ -system). Proof .:.

Without loss of generality, let $|A| = \theta$ be regular (taking a subset of the original A if necessary). Enumerate $A = \{A_{\alpha} : \alpha < \theta\}$. Since $\bigcup A \leq \aleph_0 \cdot \theta = \theta$, through a bijection we will assume without loss of generality that $A \subseteq \mathcal{P}(\theta)$. So consider $S = \{\alpha < \theta : cof(\alpha) = \omega\}$, a stationary subset of θ . (To see this, any ω -length, increasing sequence \vec{x} of elements of a club $C \subseteq \theta$ has $\sup \vec{x} < \theta$ with cofinality ω and is in C because C is closed, and therefore $\sup \vec{x} \in S \cap C \neq \emptyset$.) We can then define $f: S \to \theta$ by $f(\alpha) = \sup(\alpha \cap A_{\alpha})$ for $\alpha \in S$. This f is is then regressive since $|A_{\alpha}| < cof(\alpha)$. So by Fodor's Lemma (11 B • 5), there is some stationary $S_0 \subseteq S$ where $f''S_0 = \{\delta\}$ for some $\delta < \theta$. This means that $\sup(\alpha \cap A_\alpha) = \delta$ for stationarily many α .

Note also that $C = \{ \alpha < \theta : \forall \xi < \alpha \text{ (sup } A_{\xi} < \alpha) \}$ is club in θ : closure is immediate; and for $\alpha_0 < \theta$ arbitrary, the ω -length sequence given by $\alpha_{n+1} = \sup(A_{\alpha_n})$ yields an element of C that is $\geq \alpha_0$, meaning C is unbounded. Note that for $\alpha < \beta \in C$, $A_{\alpha} \subseteq \beta$.

As a result, $S_1 = C \cap S_0$ is also stationary (the intersection of any club D with C is also a club and hence $\emptyset \neq (D \cap C) \cap S_0 = D \cap (C \cap S_0)$). Note also that for $\alpha \in S_1$ and $\beta < \alpha$, $A_\alpha \cap A_\beta \subseteq \delta$. Now for $\alpha \in S_1$, consider $A_{\alpha} \cap \delta$. There are only $|[\delta]^{<\omega}| = |\delta| < \theta$ many finite subsets of δ . So since θ is regular, there must be some $D \subseteq S_1$ of size $|D| = \theta$ such that $A_{\alpha} \cap \delta$ is the same for all $\alpha \in D$.

We will show that the set $B = \{A_{\alpha} : \alpha \in D\}$ a Δ -system. For $\alpha, \beta \in D$, let $\alpha < \beta$ for the sake of definiteness and note that $A_{\beta} \cap A_{\beta} \subseteq \delta$. To see this, $\alpha < \beta$ has $A_{\alpha} \subseteq \beta$ and any $\gamma \in (A_{\beta} \cap A_{\alpha}) \setminus \delta$ then has $\delta \leq \gamma < \beta$ with

xⁱThe proof of The Δ -System Lemma (32 D • 2) also generalizes to the so-called *Generalized A-System Lemma* where if we have uncountable cardinals $\kappa < \theta$ with θ regular such that $|\{x \subseteq \alpha : |x| < \kappa\}| < \theta$ for all $\alpha \in \theta$, then any θ -sized family A of $< \kappa$ -sized sets has a $B \subseteq A$ of size θ that forms a Δ -system. This combinatorial restriction is a bit odd, so we just work with $\kappa = \aleph_0$ below since there there $\langle \kappa$ -sized subsets are just finite and so $|\{x \subseteq \alpha : |x| < \kappa\}| = |\alpha| < \theta$ for all $\alpha \in \theta$. Under GCH, we can easily find examples of cardinals for which this restriction holds: for any infinite cardinal $\lambda, \kappa = \lambda^+$ and $\theta = \lambda^{++}$ work.

EXAMPLES OF FORCING

 \neg

 $\gamma \in A_{\beta}$ meaning $\sup(\beta \cap A_{\beta}) \neq \delta$, a contradiction. As a result, for any distinct $\alpha, \beta, \gamma \in D$, $A_{\alpha} \cap A_{\beta} \subseteq \delta \cap A_{\alpha} = \delta \cap A_{\gamma} \subseteq A_{\gamma}$, meaning $A_{\alpha} \cap A_{\beta} \subseteq A_{\gamma} \cap A_{\eta}$ for all $\alpha, \beta, \gamma, \eta \in D$ and therefore equality.

A result of all of this is the ability to confirm that we have forced $\neg CH$. In particular, we have also forced "V \neq L".

- 32 D • 3. Corollary Let κ be an uncountable cardinal in V. Let G be Add $(\aleph_0, \kappa) = \mathbb{P}$ -generic over V. Therefore κ is not collapsed. In particular, $\kappa = \omega_2$ has $V[G] \models "|\mathcal{P}(\omega)| \ge \aleph_2$ ".

Proof .:.

By Lemma 32 D • 1 Add(\aleph_0, κ) is ccc. So by Theorem 32 C • 7, κ is still a cardinal in V[G]. For $\kappa = \omega_2^V$, we then have $\kappa > \omega_1^V = \omega_1^{V[G]}$ and thus $V[G] \models "|\mathcal{P}(\omega)| \ge \kappa = \omega_2$ " by Theorem 32 B • 2. So $V[G] \models \mathsf{ZFC} + \neg \mathsf{CH}$. \dashv

This just yields a lower bound on 2^{\aleph_0} in V[G]. But how do we actually calculate $(2^{\aleph_0})^{V[G]}$? The answer lies in counting names for subsets of ω . Note that for any particular $x \subseteq \omega$ in V[G], there are a proper class of names for x just by adding junk information that is thrown out by our particular G: e.g. for $\dot{x} \in V^{\mathbb{P}}$ a \mathbb{P} -name for x, we can consider $\dot{x} \cup \{\langle \check{\alpha}, p \rangle\}$ for each $\alpha \in \text{Ord}$ and $p \notin G$ as another name for x.

§ 32 E. Nice names

A "nice name" is just a name that has the sort of properties you would want it to have as a subset of another name. There are a variety of different kinds of names one can consider. Firstly, consider the following.

- **32E**•1. **Definition** Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder. Let *G* be \mathbb{P} -generic over *V*. Let $x \in V[G]$ be arbitrary with name \dot{x} . Let $y \subseteq x$ be in V[G]. A *kinda nice name for* y *as a subset of* x is a $\tau \in V^{\mathbb{P}}$ such that $\tau_G = y$ and dom $(\tau) \subseteq \operatorname{dom}(\dot{x})$.

— 32 E • 2. Result

Let $V \vDash ZFC$ be a transitive model we can force over and $\mathbb{P} \in V$ be a preorder. Let G be \mathbb{P} -generic over V. Let $y \subseteq x \in V[G]$ be arbitrary with $y \in V[G]$. Therefore there is a kinda nice name for y as a subset of x.

Proof .:.

We know y has some name $\dot{y} \in V^{\mathbb{P}}$. Consider

 $\pi = \{ \langle \sigma, p \rangle \in \operatorname{dom}(\dot{x}) \times \mathbb{P} : p \Vdash ``\sigma \in \dot{y} \land \sigma \in \dot{x}" \}.$

Clearly $\pi_G \subseteq y$, since any $p \in G$ with $\langle \sigma, p \rangle \in \pi$ has $\sigma_G \in \dot{y}_G = y$. Similarly, any $\sigma_G \in y$ has some $\sigma' \in \operatorname{dom}(\dot{x})$ with $V[G] \models "\sigma_G = \sigma'_G$ ". This is forced by some $p \in \mathbb{P}$ where then $\langle \sigma', p \rangle \in \pi$ and so $\sigma'_G \in \pi$. Thus $y \subseteq \pi_G$, and so we have equality. \dashv

The benefit of (kinda) nice names is that they allow us to consider just names of a certain form rather than all names, which again form a proper class. In particular, we have the following result.

- 32 E • 3. Result Let $V \models$ ZFC be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder. For any \mathbb{P} -name \dot{x} , there are at most $(2^{|\operatorname{dom}(\dot{x}) \times \mathbb{P}|})^V$ kinda nice names for subsets of \dot{x} .

Proof .:.

Every kinda nice name for a subset of \dot{x}_G (where G is generic) is in $\mathcal{P}(\operatorname{dom}(\dot{x}) \times \mathbb{P})^V$.

 \neg

In particular, if \mathbb{P} is countable, then $2^{\aleph_0} = (2^{\aleph_0})^{V}$ in V[G]. To see this, in V, there are at most $2^{|\operatorname{dom}(\check{\omega}) \times \mathbb{P}|} = 2^{\aleph_0}$
kinda nice names for subsets of ω . In particular, since \mathbb{P} is ccc and so preserves cardinals, if $\mathbf{V} \models CH$, then $\mathbf{V}[G] \models$ " $2^{\aleph_0} = (2^{\aleph_0})^{\mathbf{V}} = \aleph_1^{\mathbf{V}} = \aleph_1$ " and so $\mathbf{V}[G] \models CH$.

Another kind of nice name uses antichains in conjunction with chain conditions to do a better job at counting.

- 32E•4. Definition

Let $V \models ZFC$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder. Let G be \mathbb{P} -generic over V. Let $x \in V[G]$ be arbitrary with name \dot{x} . Let $y \subseteq x$ be in V[G]. A nice name for y as a subset of x is a $\dot{y} \in V^{\mathbb{P}}$ such that

- dom $(\dot{y}) \subseteq$ dom (\dot{x}) , i.e. \dot{y} is a kinda nice name; and
- for each $\tau \in \text{dom}(\dot{y}), \{p \in \mathbb{P} : \langle \tau, p \rangle \in \dot{y}\}$ is an antichain.

Equivalently, \dot{y} is of the form $\bigcup_{\sigma \in \text{dom}(\dot{x})} \{\sigma\} \times A_{\sigma}$ where each A_{σ} is an antichain of \mathbb{P} or else \emptyset .

- 32 E • 5. Result

Let $V \vDash ZFC$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder. Let G be \mathbb{P} -generic over V. Let $y \subseteq x \in V[G]$ be arbitrary with $y \in V[G]$. Therefore there is a nice name for y.

Proof .:.

We know there is a kinda nice name $\dot{y} \in V^{\mathbb{P}}$ for y. For each $\sigma \in \text{dom}(\dot{y}) \subseteq \text{dom}(\dot{x})$, let $A_{\sigma} = \{p \in \mathbb{P} : p \Vdash$ " $\sigma \in \dot{y}$ "}. A_{σ} is non-empty, of course, since $\langle \sigma, p \rangle \in y$ implies $p \in A_{\sigma}$. Of the antichains contained in this set, let A_{σ} be maximal among the subsets of A_{σ} . Therefore, every $p \in \mathbb{P}$ that forces " $\sigma \in \dot{y}$ " is compatible with an element of A_{σ} . So consider the name

$$\pi = \bigcup_{\sigma \in \operatorname{dom}(\dot{v})} \{\sigma\} \times \mathcal{A}_{\sigma}.$$

This is clearly a nice name, so it suffices to show $\pi_G = y$.

To show $y \subseteq \pi_G$, let $\sigma_G \in y$ have a $p \in G$ forcing $\sigma \in \dot{y}$, meaning $p \in A_{\sigma}$. There must then be some $q \in A_{\sigma} \cap G$ compatible with p and thus $\langle \sigma, q \rangle \in \pi$, meaning $q \Vdash "\sigma \in \pi"$ so $\sigma_G \in \pi_G$.

Similarly, for $\sigma_G \in \pi_G$, we have $\langle \sigma, p \rangle \in \pi$ for some $p \in G$, meaning $p \in A_{\sigma}$ and thus $p \Vdash "\sigma \in \dot{y}$ " so that $\sigma_G \in y$. Hence $\pi_G \subseteq y$, and so we have equality: $\pi_G = y$. This means π is a nice name for y.

The above has been stated in a somewhat concrete way, but alternatively, we can say that for any two names $\dot{x}, \dot{y} \in V^{\mathbb{P}}$, there is a nice name $\tau \in V^{\mathbb{P}}$ for a subset of \dot{x} such that $p \Vdash "\dot{y} \subseteq \dot{x} \to \tau = \dot{y}$ " for every $p \in \mathbb{P}$.

— 32 E • 6. Corollary -

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a κ^+ -cc preorder of \mathbf{V} , and let \dot{x} be a \mathbb{P} -name. Therefore, there are at most $(|\mathbb{P}|^{\kappa \cdot |\operatorname{dom}(\dot{x})|})^{\mathsf{V}}$ nice names for subsets of \dot{x} .

Proof .:.

Work in V. There are at most $|\mathbb{P}|^{\kappa} \leq \kappa$ -sized subsets of \mathbb{P} . Hence there are at most that many antichains. Since each nice name is given by a function from dom (\dot{x}) to antichains of \mathbb{P} , there are at most $|\mathbb{P}|^{\kappa \cdot |\operatorname{dom}(\dot{x})|}$ many nice names for subsets of \dot{x} .

In particular, in V, there is a bijection between this ordinal $(|\mathbb{P}|^{\kappa \cdot |\operatorname{dom}(\dot{x})|})^{V}$ and the nice names for subsets of \dot{x} . So in V[G], there is still this bijection that—with the help of G to interpret the \mathbb{P} -names—yields a surjection from this ordinal to $\mathcal{P}(x)$. Hence we can say $\mathbf{V}[G] \models "|\mathcal{P}(x)| \le \left| \left(|\mathbb{P}|^{\kappa \cdot |\operatorname{dom}(\dot{x})|} \right)^{V} \right|^{n}$. If \mathbb{P} is ccc and thus preserves cardinals, this simplies to $\mathbf{V}[G] \models "|\mathcal{P}(x)| \le \left(|\mathbb{P}|^{|\operatorname{dom}(\dot{x})|} \right)^{V}$.

- 32 E•7. Corollary Let $V \models ZFC$ be a transitive model we can force over. Let κ be a regular, uncountable cardinal of V such that $V \models "\kappa^{\aleph_0} = \kappa"$. Let G be Add (\aleph_0, κ) -generic over V. Therefore $V[G] \models "2^{\aleph_0} = \kappa"$.

EXAMPLES OF FORCING

Proof .:.

By Lemma 32 D•1, Add(\aleph_0, κ) is ccc. We know from Theorem 32 B•2 that $V[G] \models "2^{\aleph_0} \ge |\kappa|$ ", and $V[G] \models$ " $|\kappa| = \kappa$ " by preservation of cardinals: Theorem 32 C•7. By counting nice names for subsets of $\check{\omega}$ —which has $|\operatorname{dom}(\check{\omega})| = \aleph_0$ —it follows by Corollary 32 E•6 that V[G] has at most $(\kappa^{\aleph_0})^V = \kappa$ subsets of ω , meaning $V[G] \models "2^{\aleph_0} \le \kappa$ ", and thus we have equality.

In particular, if we can force over V = L or some other transitive model of GCH, then $V[G] \models "2^{\aleph_0} = \kappa$ " whenever G is Add $(\aleph_0, \kappa)^V$ -generic over V.

§ 32 F. Forcing CH

We've seen that we can force \neg CH pretty easily, but it took some work to *confirm* that CH fails in the generic extension. Similarly, we can pretty easily force that CH holds in the generic extension, but it will take some work to show this. We will take the expected approach: add in some surjection from \aleph_1 to $\mathcal{P}(\omega)$ of the ground model. A worry one might have is that both ω_1 and $\mathcal{P}(\omega)$ might change in the generic extension: perhaps one of the following holds:

1.
$$\omega_1^{\mathsf{V}[G]} \neq \omega_1^{\mathsf{V}}$$
; or
2. $\mathfrak{P}(\omega)^{\mathsf{V}[G]} \neq \mathfrak{P}(\omega)^{\mathsf{V}}$.

We will need to confirm that this doesn't happen: ω_1 isn't collapsed, and we don't add too many subsets of ω . Note that $\mathcal{P}(\omega)^{\mathbf{V}} = \mathcal{P}(\omega)^{\mathbf{V}[G]}$ implies $\omega_1^{\mathbf{V}} = \omega_1^{\mathbf{V}[G]} \cdot \mathbf{x}_{ii}$

We have seen that $\leq \omega$ -closed preorders preserve \aleph_1 , but they also preserve $\mathcal{P}(\omega)$.

- 32 F • 1. Lemma Let $V \models ZFC$ be a transitive model we can force over. Let κ be a cardinal of V. Suppose $\mathbb{P} \in V$ is a $\leq \kappa$ -closed preorder in V. Suppose G is \mathbb{P} -generic over V. Therefore, $\mathcal{P}(\kappa)^{V[G]} = \mathcal{P}(\kappa)^{V}$.

Proof .:.

The basic idea is that \mathbb{P} being $\leq \kappa$ -closed means that we can collect together κ -much information in V already. The motivating idea is as follows, although the real argument is in the next paragraph. In particular, for $y \subseteq \kappa$ with $y \in V[G]$, we have a kinda nice name $\dot{y} \in V^{\mathbb{P}}$ for y. For each $\alpha < \kappa$, we either have $\mathbf{V}[G] \models "\alpha \in y"$ or $\mathbf{V}[G] \models "\alpha \notin y"$ and thus we have some element of the preorder p_{α} that either forces " $\check{\alpha} \in \dot{y}$ " for " $\check{\alpha} \notin \dot{y}$ ". By continually expanding, we get a \leq -decreasing sequence of elements in the preorder which continually decide more and more of \dot{y} . Hence there is some $p \in \mathbb{P}$ with $p \leq p_{\alpha}$ for each $\alpha < \kappa$. This p then decides whether any α is in $y: y = \{\alpha : p \Vdash "\check{\alpha} \in \dot{y}"\} \in V$, implying $\mathcal{P}(\kappa)^{\mathbf{V}[G]} \subseteq \mathcal{P}(\kappa)^{\mathbf{V}}$. The other containment is obvious since $V[G] \subseteq V$.

As with Corollary 32 A • 9, the above argument actually needs to be translated in terms of dense sets. A terse argument in that style is given below: let \dot{y} be a kinda nice name for a subset of κ . For each α , consider $D_{\alpha} = \{p \in \mathbb{P} : p \Vdash ``\check{\alpha} \in \dot{y}"$ or $p \Vdash ``\check{\alpha} \notin \dot{y}"\}$. Each $D_{\alpha} \in V$ is dense and open. By $\leq \kappa$ -closure, $\bigcap_{\alpha < \kappa} D_{\alpha} = D_{\kappa} \in V$ is also dense and open. Any $p \in D_{\kappa} \cap G$ yields that $\dot{y}_G = y = \{\alpha : p \Vdash ``\check{\alpha} \in \dot{y}"\} \in V$, meaning $\mathcal{O}(\kappa)^{\mathbf{V}[G]} \subseteq \mathcal{O}(\kappa)^{\mathbf{V}}$.

- 32 F • 2. Corollary

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a $\leq \omega$ -closed preorder of \mathbf{V} . Suppose G is \mathbb{P} -generic over V. Therefore $\omega_1^{\mathbf{V}} = \omega_1^{\mathbf{V}[G]}$ and $\mathcal{P}(\omega)^{\mathbf{V}} = \mathcal{P}(\omega)^{\mathbf{V}[G]}$.

So we will consider the following preorder, sometimes written $Col(\aleph_1, 2^{\aleph_0})$, adding a bijection between \aleph_1 and $\mathcal{P}(\omega)$ of the ground model. To make this countably closed, we can't work with finite functions as we have been doing before:

^{xii}Recall that we can code $\alpha < \omega_1$ by relations on ω which can be coded by subsets of ω , and hence agreeing on $\mathcal{P}(\omega)$ entails agreeing on ω_1 as the least ordinal that can't be coded in this way. Indeed, this holds more generally: agreeing on $\mathcal{P}(\kappa)$ entails agreeing on what κ^+ is.

the countable union of finitely many functions isn't necessarily finite. So we use the next best idea: *countable* partial functions. Since these will still have relatively small domain compared to \aleph_1 , we have enough flexibility when using them as approximations.

- 32 F • 3. Definition Define $\operatorname{Fn}_{<\aleph_1}(\omega_1, \mathcal{P}(\omega)) = \langle \operatorname{Fn}_{<\aleph_1}(\omega_1, \mathcal{P}(\omega)), \leqslant \rangle$ by $\operatorname{Fn}_{<\aleph_1}(\omega_1, \mathcal{P}(\omega)) = \{p : \omega_1 \rightharpoonup \mathcal{P}(\omega) : |p| < \aleph_1\}$ where $p \leqslant q$ iff $p \supseteq q$.

It should be clear that $\operatorname{Fn}_{<\aleph_1}(\omega_1, \mathfrak{P}(\omega))$ is countably closed, since the union of any countable chain is still countable, and is obviously still a partial function from ω_1 to $\mathfrak{P}(\omega)$.

- 32 F•4. Theorem Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force over. Let G be $\mathbb{P} = \mathsf{Fn}_{<\aleph_1}(\omega_1, \mathcal{P}(\omega))^V$ -generic over V. Therefore $V[G] \vDash \mathsf{CH}$.

Proof .:.

Really this just amounts to showing that $\bigcup G = g$ is a surjection from ω_1^{V} to $\mathcal{P}(\omega)^{\mathsf{V}}$. By countable closure, Lemma 32 F • 1 tells us that $\mathcal{P}(\omega)^{\mathsf{V}[G]} = \mathcal{P}(\omega)^{\mathsf{V}}$ and $\omega_1^{\mathsf{V}} = \omega_1^{\mathsf{V}[G]}$, meaning g would be a surjection from $\omega_1^{\mathsf{V}[G]}$ to $\mathcal{P}(\omega)^{\mathsf{V}[G]}$ and so $\mathsf{V}[G] \models \mathsf{CH}$. Since the two interpretations are equal, we just write " $\mathcal{P}(\omega)$ " and " \aleph_1 ".

But that $g: \omega_1 \to \mathcal{P}(\omega)$ is a surjection is clear: for each $\alpha < \omega_1$ and each $x \in \mathcal{P}(\omega)$, the following are dense $D_{\alpha} = \{p \in \mathbb{P} : \alpha \in \operatorname{dom}(p)\}$ and $E_x = \{p \in \mathbb{P} : x \in \operatorname{im}(p)\}.$

To see that D_{α} is dense, just extend any $p \in \mathbb{P}$ with $\langle \alpha, \emptyset \rangle$. To see that E_x is dense, just note that $p \in \mathbb{P}$ being countable implies dom $(p) \neq \omega_1$ and thus we can choose some $\alpha \in \omega_1 \setminus \text{dom}(p)$ and extend p with $\langle \alpha, x \rangle$. This new (partial) function remains countable and so is in E_x .

But this means each $x \in \mathcal{P}(\omega)$ has a $p \in G \cap E_x$ where then $x \in \operatorname{im}(p) \subseteq \operatorname{im}(g)$ so that $\mathcal{P}(\omega) \subseteq \operatorname{im}(g)$ and g is surjective. Given that dom $(g) = \omega_1$ (by the density of the D_{α} s), we get the result: $V[G] \models "|\mathcal{P}(\omega)| = \aleph_1$ ". \dashv

There are actually a great number of preorders that force CH. For example, $Add(\aleph_0, 1)$ does this. In fact, $Add(\aleph_0, 1)$ forces a princple known as \Diamond . It's not a bad exercise (although moderately difficult) to show that this holds, assuming the reader knows the definition of \Diamond .^{xiii}

In general, we have many different options when it comes to adding a generic G with certain properties. This is in part due to the vagueness of "approximation" when using a preorder of sets supposed to approximate G. Many of these preorders turn out to be equivalent in the sense that a generic $G \subseteq \mathbb{P}$ yields a generic $H \subseteq \mathbb{Q}$ where V[G] = V[H]. For example, the forcing we used with $Add(\aleph_0, \kappa)$ —Cohen forcing—is equivalent to the subpreorder where all conditions have domains that not only are finite subsets of ω , but are actual natural numbers: dom(p) = n for some $n < \omega$.

There are also many preorders that are not equivalent, but that *can* give similar generics. For example, forcing with $Col(\lambda, \kappa)$ collapses $|\kappa|$ to λ , but leaves $|\lambda| = \lambda$ in the generic extension. If instead of conditions of size $< \lambda$ we consider *finite* conditions, we still end up with a generic G with $\bigcup G = g$ as a surjection from λ to κ , but we also end up with a surjection from ω to κ . This means κ is (and so subsequently all cardinals $\leq \kappa$, including λ , are) collapsed down to ω .^{xiv}

This is all just to say that it's generally not difficult to come up with a preorder that adds some object serving whatever purpose you want *in the ground model*. But it's far more difficult to show it doesn't muck things up in the generic extension. This is the purpose of the discussion of antichains, nice names, and further ideas.

^{xiii}Because I'm certainly not going to talk about it here.

^{xiv}To see that we collapse κ to ω if we consider the preorder \mathbb{P} of *finite* partial functions from λ to κ (ordered by inclusion), just consider $g \upharpoonright \omega = \bigcup G \upharpoonright \omega$. Each $p \in \mathbb{P}$ is finite, and so dom $(p) \subsetneq \omega$. For each $\alpha < \kappa$, there is some $n \in \omega \setminus \text{dom}(p)$ where then $p \cup \{(n, \alpha)\} \in D_{\alpha} = \{p \in \mathbb{P} : \alpha \in \text{im}(p \upharpoonright \omega)\}$. This implies each D_{α} is dense and so $p \in G \cap D_{\alpha}$ implies $\alpha \in \text{im}(p \upharpoonright \omega) \subseteq \text{im}(g \upharpoonright \omega)$ implying $\kappa \subseteq \text{im}(g \upharpoonright \omega)$ and therefore $g \upharpoonright \omega$ is a surjection onto κ , meaning $V[G] \models ``|\kappa| = \aleph_0$ ''. This argument doesn't work with $\text{Col}(\lambda, \kappa)$, since conditions there can be infinite and so contain all of ω in their domains, preventing us from extending into D_{α} .

Section 33. More General Topics and Theory

The above discussion has primarily focused on showing the independence of CH from ZFC using the posets $Add(\aleph_0, \kappa)$ and $Fn_{<\aleph_1}(\omega_1, \mathcal{P}(\omega))$. In doing so, we've shown a lot of results in these restricted circumstances. Questions we have not answered include:

- Can we generalize the above to change 2^{κ} just as we changed 2^{\aleph_0} ?
- Did we need to use the posets above or would other, similar posets work?
- In what sense can two posets be the same for forcing, and (more importantly) how can we know in the ground model?
- It's possible to force twice to go from V to V[G] and then to V[G][H], but is it possible to go from V[G] to some intermediate submodel V ⊊ V[H] ⊊ V[G]?

All of these questions and more will be investigated here.

§ 33 A. Changing 2^{κ} for regular κ

The above discussion has primarily focused on showing the independence of CH from ZFC using the posets $Add(\aleph_0, \kappa)$ and $Fn_{<\aleph_1}(\omega_1, \mathcal{P}(\omega))$. In doing so, we've shown a lot of results in these restricted circumstances. For the sake of a better, more general understanding, we state the generalizations of these. ^{xv}

33A•1. Theorem Let $\kappa, \lambda > 0$ be cardinals of $\mathbf{V} \models \mathsf{ZFC}$ a transitive model; $I, J \in V$; and \mathbb{P} a preorder. Define $\mathsf{Fn}_{<\kappa}(I, J) = \langle \{p : I \rightarrow J : |p| < \kappa \}, \supset, \emptyset \rangle$. Therefore, all interpreted in \mathbf{V} ;

- 1. a κ -cc preorder preserves cardinals and cofinalities $\geq \kappa$;
- 2. a < κ -closed preorder preserves cardinals and cofinalities $\leq \kappa$;
- 3. a κ^+ -cc preorder \mathbb{P} and \mathbb{P} -name τ gives $\mathbb{1}^{\mathbb{P}} \Vdash "|\mathcal{O}(\tau)| \leq \check{\delta}$ " where $\delta = |\mathbb{P}|^{\kappa \cdot |\operatorname{dom}(\tau)|}$ is the number (in the ground model) of nice names for subsets of τ .
- 4. $\operatorname{Fn}_{<\kappa}(I, J)$ is $< \operatorname{cof}(\kappa)$ -closed for any I, J;
- 5. $\operatorname{Fn}_{<\kappa}(I, J)$ is κ^+ -cc whenever κ is regular, $|J| \le \kappa$, and $2^{<\kappa} = \kappa$;
- 6. $\operatorname{Col}(\kappa, \lambda) = \operatorname{Fn}_{<\kappa}(\kappa, \lambda)$ is $< \operatorname{cof}(\kappa)$ -closed where $\kappa < \lambda$;
- 7. Add $(\kappa, \lambda) = \operatorname{Fn}_{<\kappa}(\lambda \times \kappa, 2)$ is $< \kappa$ -closed and κ^+ -cc whenever $\kappa = \operatorname{cof}(\kappa) = 2^{<\kappa}$;

We also get a number of variants of The Δ -System Lemma (32 D • 2).

— 33 A • 2. Theorem (The Generalized Δ-System Lemma) —

We say a family of sets B is or forms a Δ -system iff $\exists r \forall p, q \in B \ (p \cap q = r)$.

- (The Δ -System Lemma) if A is an uncountable family of finite sets, then there is an uncountable Δ -system $B \subseteq A$.
- (The Not-As-Generalized Δ-System Lemma) if κ = cof(κ) = 2^{<κ} and if A is a > κ-sized family of sets of size < κ, then there is a Δ-system B ⊆ A with |B| > κ.
- (The Generalized Δ -System Lemma) if $\aleph_0 \le \kappa < \theta = cof(\theta)$ has $|\{x \le \alpha : |x| < \kappa\}| < \theta$ for all $\alpha < \theta$ and

^{xv}Another easy consequence from these notions is that if we have a transitive model M we can force over, then V isn't a generic extension by any preorder $\mathbb{P} \in M$, and in fact $|M| \le \aleph_1$. Ultimately these ideas are not relevant to our discussion, because they are a quirky result of being lazy with very strong propositions. All these results don't need that we can force with *every* preorder of the ground model, but just the preorders used in the statements. We will always be able to force with countable, transitive models, and indeed, if we can force over M, then every set in M is countable, just because we can force with Col(ω, κ) for each $\kappa \in \text{Ord} \cap M$. The converse of this—that if every set of M is countable then we can force over M—is easy by Corollary 31 D • 3. Note that M need not be countable for this to happen, e.g. if $\omega_1 = \omega_1^{V}$ is inaccessible in L, then $L_{\omega_1} \models \text{ZFC}$ and every set in L_{ω_1} is countable, but $|L_{\omega_1}| = \aleph_1$.

More General Topics and Theory

if A is a $|A| = \theta$ -sized family of sets of size $\langle \kappa$, then there is a Δ -system $B \subseteq A$ with $|B| = |A| = \theta$.

The proofs of all the results of Theorem 33 A • 1 and The Generalized Δ -System Lemma (33 A • 2) are almost identical to the special cases given before in Section 32. One easy consequence of the above results is that we may easily force almost any change to the continuum function $\kappa \mapsto 2^{\kappa}$ at a regular, infinite κ .

33A·3. Corollary
Let V ⊨ ZFC be a transitive model we can force over. Let κ, λ be infinite cardinals of V. Let P = Add(κ, λ)^V ∈ V with G P-generic over V. Suppose further that
i. V ⊨ "κ is regular";
ii. V ⊨ "2^{<κ} = κ";
iii. V ⊨ "λ^κ = λ".

Therefore $V[G] \models "2^{\kappa} = \lambda$ ", and all cardinals of V are cardinals of V[G].

Proof .:.

Argue in V. By Theorem 33 A • 1 (7), (4), (2), and (i), $\mathbb{P} = \operatorname{Fn}_{<\kappa}(\kappa \times \lambda, 2)$ is $< \operatorname{cof}(\kappa) = \kappa$ -closed and so all cardinals $\leq \kappa$ are preserved. Moreover by (7), (5), (ii), and (1), $\mathbb{P} = \operatorname{Fn}_{<\kappa}(\kappa \times \lambda, 2)$ is κ^+ -cc and so cardinals $\geq \kappa^+$ are preserved. \mathbb{P} yields a generic coding an injection from λ to κ^2 and hence $V[G] \models "2^{\kappa} \leq \lambda$ ". By Theorem 33 A • 1 (3) and (iii), since \mathbb{P} is κ^+ -cc, there are only

$$|\mathbb{P}|^{\kappa \cdot |\operatorname{dom}(\kappa)|} \le |\{x \subseteq \kappa \times \lambda \times 2 : |x| < \kappa\}|^{\kappa} = |\{x \subseteq \lambda : |x| < \kappa\}|^{\kappa} \le (\lambda^{\kappa})^{\kappa} = \lambda^{\kappa} = \lambda$$

nice names for subsets of $\check{\kappa}$ and therefore $V[G] \vDash "2^{\kappa} \leq \lambda$ " so that equality holds.

Н

– 33 A•4. Corollary —

Let $\mathbf{V} \models \mathsf{ZFC} + \mathsf{GCH}$ be a transitive model we can force over. Let $\operatorname{cof}(\kappa)^{\mathbf{V}} = \kappa < \operatorname{cof}(\lambda)^{\mathbf{V}} \le \lambda$ be infinite cardinals of \mathbf{V} . Therefore $\mathbb{1}^{\operatorname{Add}(\kappa,\lambda)^{\mathbf{V}}} \Vdash "2^{\check{\kappa}} = \check{\lambda}"$.

Proof .:.

Argue in V. We have κ is regular by hypothesis. As a model of GCH, any cardinal $\mu < \kappa$ has $\mu^+ = 2^{\mu} \le \kappa$ and hence the supremum over these μ yields $2^{<\kappa} \le \kappa$ and so obviously $\kappa \le 2^{<\kappa}$ yields equality. Thus (i) and (ii) of Corollary 33 A • 3 are satisfied. To see that $\lambda^{\kappa} = \lambda$, we again use GCH with Theorem 5 E • 6. Thus all the hypothesis of Corollary 33 A • 3 apply and so in any generic extension of V, $2^{\kappa} = \lambda$ holds.

Forcing changes to 2^{κ} for *singular* κ is much more difficult in general. This is partly due to theorems like Silver's Theorem, which says that the first failure to GCH can't be a singular cardinal of uncountable cofinality. More precisely, if $\{\lambda < \kappa : 2^{\lambda} = \lambda^+\}$ is stationary in $\kappa > cof(\kappa) > \omega$, then $2^{\kappa} = \kappa^+$. The proof of this combinatorial result isn't covered here, and is included mostly to show that the theory of forcing can be quite difficult and subtle, if that wasn't already obvious.

§ 33 B. Small topics

Returning to the more general theory of forcing, closure properties of preorders are nice for preserving sufficiently small cardinals and cofinalities, but more generally they don't add sufficiently short sequences. This concept is equivalent to the following we quickly explore.

— 33 B•1. Definition -

Let κ be an infinite cardinal. A preorder \mathbb{P} is $< \kappa$ -distributive^{xvi} iff for every collection \mathcal{D} of $< \kappa$ -many open dense sets of \mathbb{P} , $\bigcap \mathcal{D}$ is open dense.

^{xvi}This is also called " κ -distributive" and " (κ, ∞) -distributive" elsewhere.

— 33 B • 2. Corollary

Let κ be an infinite cardinal. Let \mathbb{P} be $< \kappa$ -closed. Therefore \mathbb{P} is $< \kappa$ -distributive.

Proof .:.

Proceed by induction on κ . Let $\mathcal{D} = \{D_{\alpha} : \alpha < \lambda\}$ be a collection of $\lambda < \kappa$ -many open dense sets of \mathbb{P} . It's obvious $\bigcap \mathcal{D}$ is open, so we must show it's dense. Let $p \in \mathbb{P}$ be arbitrary. Choose inductively, for $\alpha, \gamma < \lambda$,

$$p_{0} \in D_{0} \cap \mathbb{P}_{\leq p},$$

$$p_{\alpha+1} \in D_{\alpha+1} \cap \mathbb{P}_{\leq p_{\alpha}}, \text{ and}$$

$$p_{\gamma} \in \bigcap_{\xi < \gamma} \mathbb{P}_{\leq p_{\xi}} \text{ for } \gamma \text{ a limit ordinal.}$$

So $\langle p_{\alpha} : \alpha < \lambda \rangle$ is decreasing. At limit stages, such p_{γ} exist by $< \kappa$ -closure and are in $\bigcap_{\xi < \gamma} D_{\xi}$ since the D_{ξ} s are open. Also by $< \kappa$ -closure, there is some $p_{\lambda} \leq^{\mathbb{P}} p_{\alpha}$ for each $\alpha < \lambda$, and the same idea as before gives $p_{\lambda} \in \bigcap_{\alpha < \lambda} D_{\alpha}$ with $p_{\lambda} \leq p$. Hence $\bigcap \mathcal{D}$ is dense. \dashv

Distributivity is just what we need to ensure we don't add small sequences by the same ideas as with Lemma 32 A • 8.

— 33 B • 3. Result ·

- Let $V \vDash$ ZFC be a transitive model we can force over.
- Suppose $\mathbb{P} \in V$ is $< \kappa$ -distributive for some cardinal κ of \mathbf{V} .
- Let G be \mathbb{P} -generic over V.
- Let $f \in V[G]$ be such that $f : \lambda \to V$ for $\lambda < \kappa$.

Therefore $f \in V$. In other words, \mathbb{P} adds no new $< \kappa$ -length sequences of elements of V.

Proof .:.

We proceed similarly to Lemma 32 A • 8. Note that im $f \subseteq X = V_{\rho}^{\vee} \in V$ for some sufficiently large ρ . Let \dot{f} be a name for f and let $p \in G$ be such that $p \Vdash "\dot{f}$ is a function from $\check{\lambda}$ to \check{X} ". For each $\alpha < \lambda$, consider $D_{\alpha} = \{q \leq^{\mathbb{P}} p_0 : \exists x \in X \ (q \Vdash "\dot{f}(\check{\alpha}) = \check{x}")\}$ Since $\bigcap_{\alpha < \lambda} D_{\alpha}$ is dense by $< \kappa$ -distributivity, let $p^* \in G \cap \bigcap_{\alpha < \lambda} D_{\alpha}$. It follows that $f = \{\langle \alpha, x \rangle : p^* \Vdash "\dot{f}(\check{\alpha}) = \check{x}"\} \in V.$

We have a similar notion of "covering" with chain conditions. This is weaker than containing every relevant sequence, but still quite useful.

− 33 B • 4. Definition Let $V \subseteq W \models$ ZFC be two transitive models with κ a cardinal. We say V, W satisfy $< \kappa$ -covering iff every $< \kappa$ -sized subset $X \subseteq V$ in W can be covered by a $< \kappa$ -sized set $X \subseteq Y \in V$ in V. We similarly define $\leq \kappa$ -covering as $< \kappa^+$ -covering.

Generic extensions by $\leq \kappa$ -closed preorders then satisfy $\leq \kappa$ -covering (with the ground model) just by distributivity. But chain conditions actually give much more covering as the following shows. Ultimately, covering is useful in arguments as it shows that the extension isn't too far off from the ground model in a precise way, and chain conditions then tell us this closeness continues into the higher parts of the universes.

- 33 B • 5. Result (Chain Condition Covering)

- Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over.
- Let $\mathbb{P} \in V$ be κ -cc for some cardinal κ of \mathbf{V} .
- Let G be \mathbb{P} -generic over V.

Therefore V, V[G] satisfy $\leq \lambda$ -covering for every $\lambda \geq \kappa$.

More General Topics and Theory

Proof .:.

Let $\lambda \ge \kappa$ be arbitrary. Let $X = \{x_{\alpha} : \alpha < \delta\} \in V[G]$ have $X \subseteq V$ with size $\delta \le \lambda$. Let \dot{X} be a nice name for X meaning $\dot{X} = \bigcup_{\alpha < \delta} \{\dot{x}_{\alpha}\} \times A_{\alpha}$ where each $A_{\alpha} \subseteq \mathbb{P}$ is a maximal antichain and each \dot{x}_{α} is a name for x_{α} . Since \mathbb{P} is κ -cc, each A_{α} has size $< \kappa$. Since there are $< \lambda$ -many such A_{α} s, it follows that

$$Y = \bigcup_{\alpha < \$} \{ x \in V : \exists p \in A_{\alpha} \ (p \Vdash ``\check{x} = \dot{x}_{\alpha}") \} \in V$$

is the $\delta \leq \lambda$ -sized union of $< \kappa$ -sized sets and thus has size $\leq \delta \cdot \kappa \leq \lambda$. Clearly $X \subseteq Y$, so this works. \dashv

Chain conditions also preserve closure properties of inner models like ultrapowers.

— 33 B•6. Lemma

An inner model N of ZFC is closed under κ -sequences iff it contains all κ -sequences of ordinals: ^{κ}Ord $\subseteq N$.

Proof .:.

The (\rightarrow) direction is clear. So suppose N conatins all κ -sequences of ordinals. Let $f : \kappa \to N$ be a κ -length sequence. Note that $f : \kappa \to V_{\lambda}^{\mathsf{N}}$ for some λ . Enumerate in $\mathsf{N} V_{\lambda}^{\mathsf{N}} = \{g(\alpha) : \alpha < |V_{\lambda}|^{\mathsf{N}}\}$ by some bijection $g : |V_{\lambda}|^{\mathsf{N}} \to V_{\lambda}^{\mathsf{N}}$ in N. It follows that $g^{-1} \circ f : \kappa \to Ord$ must be in N by hypothesis. Since $g \in N$, it follows that $g \circ (g^{-1} \circ f) = f \in N$.

— 33 B • 7. Result —

- Let $V \models$ ZFC be a transitive model we can force over, and let κ be a regular cardinal of V.
- Let $\mathbf{M} \subseteq \mathbf{V}$ be an inner model closed under κ -sequences of $V: \{f \in V: f : \kappa \to M\} \subseteq M$.
- Let $\mathbb{P} \in V \cap M$ be κ -cc in V.
- Let G be \mathbb{P} -generic over V.

Therefore G is P-generic over M, and M[G] is closed under κ -sequences of V[G].

Proof .:.

That G is P-generic over M is immediate: M contains fewer dense sets than V and G intersects all of them. By Lemma 33 B • 6, it suffices to show $V[G] \models "^{\kappa} \text{Ord} \subseteq M[G]$ ". Let $f : \kappa \to \lambda$ be a sequence in V[G] with name \hat{f} as forced by some $p \in G$: $p \Vdash "\hat{f}$ is a $\check{\kappa}$ -length sequence of ordinals". We must find a name for f in M. For each $\alpha < \kappa$, define

$$F_{\alpha} = \{ \beta \in \text{Ord} : \exists p^* \leq p \ (p^* \Vdash ``\dot{f}(\check{\alpha}) = \check{\beta}``) \}.$$

For each $\beta \in F_{\alpha}$, let $p_{\beta} \leq p$ witness this: $p_{\beta} \Vdash "\dot{f}(\check{\alpha}) = \check{\beta}"$. It follows that $A_{\alpha} = \{p_{\beta} \in \mathbb{P} : \beta \in F_{\alpha}\}$ is an antichain and so has size $< \kappa$ and thus so does F_{α} . It follows that each $A_{\alpha}, F_{\alpha} \in M$ and the maps $\alpha \mapsto A_{\alpha}, \alpha \mapsto F_{\alpha}$ are in M. More importantly, the map $\beta \in F_{\alpha} \mapsto p_{\beta}$ is also in M as F_{α} is, through coding, just an ordinal $< \kappa$ and each p_{β} is in M. But then we can form a name \ddot{f} for f in M without direct reference to \dot{f} :

$$\ddot{f} = \bigcup_{\alpha < \kappa} \left\{ \left\langle \langle \langle \check{\alpha}, \check{\beta} \rangle \rangle, p_{\beta} \right\rangle \colon p \in A_{\alpha} \land \beta \in F_{\alpha} \right\}.$$

It's not hard to see that $\ddot{f}_G = f$ and so $f \in M[G]$.

Moving on to a new topic, a very useful principle is the Maximum Principle. This principle strengthens Corollary 31 D • 10 (4) in that we don't need to extend from p to a $p^* \leq p$ to get a witness to existential statements. Note that this relies highly on AC in the ground model. In fact, it holding for every preorder in the ground model is *equivalent* to AC in the ground model.

— 33 B • 8. Result (Maximum Principle) —

Let \mathbb{P} be a preorder with $p \in \mathbb{P}$. Let φ be a FOLp-formula with \mathbb{P} -name parameters. Suppose $p \Vdash \exists x \varphi(x)$. Therefore there is some \mathbb{P} -name σ where $p \Vdash \varphi(\sigma)$.

 \neg

More General Topics and Theory

Proof .:.

Suppose $p \Vdash \exists x \varphi(x)$. This means $D = \{p^* \leq p : \exists \sigma \in V^{\mathbb{P}} (p^* \Vdash \varphi(\sigma))\}$ is dense below p. So for each $q \in D$, let σ_q be such a name. There is an antichain $\mathcal{A} \subseteq D$ that is maximal by Result $32 C \cdot 3$. Consider the name

 $\sigma = \{ \langle \tau, q^* \rangle : \exists r, q \ (q^* \leq q \in \mathcal{A} \land q^* \leq r \land \langle \tau, r \rangle \in \sigma_q) \}.$ It follows that $p \Vdash ``\varphi(\sigma)``$. To see this, let G be P-generic over V with $p \in G$.

 $|G \cap A| = 1$ by Theorem 32 C • 5 so there is some $a \in G \cap A$. Thus

 $\sigma_G = \{\langle \tau, a^* \rangle : a^* \leq a \land \exists r \ (a^* \leq r \land \langle \tau, r \rangle \in \sigma_a)\}_G = \{\langle \tau, r \rangle \in \sigma_a : r \text{ is compatible with } a\}_G.$

Since all $r \in G$ are already compatible with a, this is just $\{\tau_G : \exists r \in G \ (\langle \tau, r \rangle \in \sigma_a)\} = (\sigma_a)_G$. Since $a \Vdash ``\varphi(\sigma_a)$, we thus have $V[G] \vDash ``\varphi(\sigma_G)$, and so $p \Vdash ``\varphi(\sigma)$.

It's not too difficult to show the equivalence with AC in the ground model as follows.

- **33 B • 9. Result** Under ZF, AC is equivalent to MP where MP is the statement that for every preorder \mathbb{P} , $p \in \mathbb{P}$, and FOLp-formula φ with \mathbb{P} -name parameters, if $p \Vdash \exists x \varphi(x)$ then there is some \mathbb{P} -name σ where $p \Vdash \varphi(\sigma)$.

Proof .:.

Maximum Principle (33 B • 8) shows AC \rightarrow MP. So let F be a non-empty family of non-empty sets. Consider the preorder $\mathbb{F} = \langle F \cup \{\emptyset\}, \{\langle p, p \rangle, \langle p, \emptyset \rangle : p \in F\}, \emptyset \rangle$ meaning $p \leq q$ iff $q = \emptyset$ or p = q. The only two dense sets are $F \cup \{\emptyset\}$ and F (which is an antichain). So any G \mathbb{P} -generic must have $|G \cap F| = 1$ and therefore $G = \{p, \emptyset\}$ for some $p \in F$ implying G is in the ground model and so the generic extension is just the ground model.

To construct a choice function for F, note that since $p \in F$ is non-empty, $\mathbb{1}^{\mathbb{F}} \Vdash \exists x \ (x \in \check{p})$ ". By MP, there is some \mathbb{P} -name τ where $\mathbb{1}^{\mathbb{F}} \Vdash \tau \in \check{p}$ ". So any $p \in F$ forces the same. But forcing with p determines the generic: $G = \{\emptyset, p\}$ and therefore we can evaluate $\tau_G \in p$. This means $p \Vdash \tau = \check{x} \in \check{p}$ " for exactly one $x \in p$. As a result, $C = \{\langle p, x \rangle : p \in F \land p \Vdash \tau = \check{x}"\}$ is a choice function for F.

Part of why the properties we've been looking at so far have been nice is their ability to decide information. Although Maximum Principle (33 B • 8) says we can find names witnessing existential statements, it's not often we can force those names τ to be of the form \check{x} for some x in the ground model even if $\mathbb{1} \Vdash \tau \in \check{y}$ '' for some y in the ground model. This is mostly used for when we want a name for an ordinal: just because $\mathbb{1} \Vdash \tau < \check{\kappa}$ '' doesn't mean there's any $\beta < \kappa$ where $\mathbb{1} \Vdash \check{\beta} < \check{\kappa}$ ''. In the case of ccc preorders, we have Corollary 32 C • 8 to give us this. As a side note, it suffices in many cases to just consider nice names when counting names, even if we use some other kind of name (and we will later).

— 33 B•10. Lemma

Let \mathbb{P} be a preorder. Let τ be a \mathbb{P} -name. Therefore there is a nice-name σ such that $\mathbb{1}^{\mathbb{P}} \Vdash ``\tau = \sigma"$.

Proof .:.

Use Result 32 E • 5. In particular, for each $\pi \in \text{dom}(\tau)$, set \mathcal{A}_{π} to be a maximal antichain contained in $\{p \in \mathbb{P} : p \Vdash ``\pi \in \tau"\}$. Define $\sigma = \bigcup_{\pi \in \text{dom}(\tau)} \{\pi\} \times \mathcal{A}_{\pi}$ which is a nice name for for τ , the reasoning of Result 32 E • 5 tells us $\mathbb{1}^{\mathbb{P}} \Vdash ``\tau = \sigma"$.

We will not use this result for now, as it just restates Result 32 E • 5 in a less concrete but more understandable way.

Another thing to consider is minimality in a different sense than the minimality of the generic extension compared to the ground model. In particular, although there is no model W with $V \subseteq W \subseteq V[G]$ such that $G \in W$, it may still be possible that $V \subseteq V[H] \subseteq V[G]$ for some other generic H over a poset in the ground model V (potentially the *same* poset as with G). Alternatively, it may be that there is *no* such generic extension and so V[G] is truly minimal as a generic extension. We give a negative example and a positive example, although we only prove the negative later in talking about product forcing.

— 33 B•11. Result —

Cohen forcing generic extensions are never minimal. Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} = \mathsf{Add}(\aleph_0, 1)^V$ with $G \mathbb{P}$ -generic over V. Therefore there is a $H \mathbb{P}$ -generic over V with $V \subsetneq V[H] \subsetneq V[G]$.

In particular, after forcing with Cohen forcing, we get an infinite descending chain of generic extensions $V \subsetneq \cdots \subsetneq V[G_2] \subseteq V[G_1] \subseteq V[G_0]$. The basic idea behind the proof is that when we add a real g, the even digits of g don't determine the odd digits and thus $V \subsetneq V[even(g)] \subsetneq V[g]$.

A positive example for minimality is Sacks forcing where we force with perfect trees (cf. Theorem 15.34 of [17]).

- 33 B • 12. Theorem — Sacks forcing extensions are minimal over the ground model in that for any generic extension V[G] over the ground model V, any set of ordinals $X \in V[G]$ has $X \in V$ or $G \in V[X]$.

This shows that there is no intermediary model $V \subseteq W \subseteq V[G]$ because inner models are determined by the sets of ordinals they contain (just by appropriately coding things). So if W contained a set of ordinals $X \in V[G] \setminus V$ then $V[G] \subseteq W$ implying V[G] = W. Such examples are rare, however, and more commonly (although still rare) preorders give generic extensions minimal with respect to some smaller class of objects like reals, saying that if $x \in V[G] \setminus V$ is a real then $G \in V[x]$. Cohen forcing doesn't have this property in the slightest because Result 33 B • 11 shows the even entries of the new real are unable to define the odd entries.

§ 33 C. Homomorphisms and forcing equivalence

Another important idea is how flexible some forcing notions are. There are a great number of posets that do essentially the same thing when forcing with them. For example, we could have defined Add(\aleph_0 , 1) = ${}^{<\omega}\omega$, ensuring that our partial functions have their domain a natural number rather than merely any finite subset of ω . It seems pretty clear that forcing with this new preorder this would do the same thing as Add(\aleph_0 , 1) defined before. Certainly they both add a new function $g : \omega \to \omega$. But how do we show that these two preorders do everything else the same as well? Indeed, there are many complicated forcings that add subsets of ω , but also do other things. For example, Col(\aleph_0 , \aleph_1) adds a subset of ω coding a countable well-order of order-type (the ground model's) ω_1 .

— 33 C • 1. Definition -

Let $V \models ZFC$ be a transitive model we can force over. Let $\mathbb{P}, \mathbb{Q} \in V$ be preorders appropriate for forcing. We say \mathbb{P} and \mathbb{Q} are *forcing equivalent* iff their generic extensions are the same: for every $G \mathbb{P}$ -generic over V, there is an $H \mathbb{Q}$ -generic over V such that V[G] = V[H], and vice versa.

Note that this can be reformulated in a first-order way: \mathbb{P} is forcing equivalent to \mathbb{Q} iff their boolean algebras of regular open sets are isomorphic (which we will not investigate here).

Often the goal is to figure our when two preorders are forcing equivalent. One way to do this is with identifying one as a dense set in the other. Some standard terminology is the following.^{xvii}

— 33 C • 2. Definition -

Let \mathbb{P} , \mathbb{Q} be preorders. A function $f : \mathbb{P} \to \mathbb{Q}$ is an *incompatibility homomorphism* iff it's a homomorphism preserving incompatibility, i.e.

• $f(\mathbb{1}^{\mathbb{P}}) = \mathbb{1}^{\mathbb{Q}};$

- for all $p, p' \in \mathbb{P}, p \leq^{\mathbb{P}} p'$ implies $f(p) \leq^{\mathbb{Q}} f(p')$;
- for all $p, p' \in \mathbb{P}$, p, p' are incompatible in \mathbb{P} implies f(p), f(p') are incompatible in \mathbb{Q} ;

We also say that f is an *incompatibility embedding* iff f is injective and the above hold with "implies" replaced by "iff". In addition,

• f is a dense homomorphism iff f is an incompatibility homomorphism and $f^{"}\mathbb{P}$ is dense in \mathbb{Q} ;

^{xvii}Most sources just say "homomorphism" instead of "incompatibility homomorphism" used here, in defiance to preorders appropriate for forcing merely being FOL($\{\leq, 1\}$)-models and not FOL($\{\leq, 1, \bot\}$)-models.

 \neg

- f is a *complete homomorphism* iff f is an incompatibility homomorphism and for all $\mathcal{A} \subseteq \mathbb{P}$, \mathcal{A} is a maximal antichain in \mathbb{P} implies $f \ \mathcal{A}$ is a maximal antichain in \mathbb{Q} ;
- \mathbb{P} is a *complete suborder* of \mathbb{Q} iff $\mathbb{P} \subseteq \mathbb{Q}$ and id $\upharpoonright \mathbb{P} : \mathbb{P} \to \mathbb{Q}$ is a complete homomorphism.

Whenever there is a dense embedding from one preorder to another, they are forcing equivalent because the map allows us to translate between the generics. A complete homomorphism allows us, however, to only go one way with this: all generic extensions by \mathbb{Q} yield generic extensions by \mathbb{P} , but the reverse need not hold.

Proof .:.

Let \mathbb{P} and \mathbb{Q} be preorders and $f : \mathbb{P} \to \mathbb{Q}$ a dense homomorphism. Since f is an incompatibility homomorphism, we need to show preservation of maximal antichains. Clearly by preservation of incompatibility, any antichain \mathcal{A} of \mathbb{P} has $f''\mathcal{A}$ as an antichain of \mathbb{Q} . Since $f''\mathbb{P}$ is dense, it contains a maximal antichain \mathcal{A}' with $f''\mathcal{A} \subseteq \mathcal{A}' \subseteq f''\mathbb{P}$ by Result 32 C • 3. Moreover, if $\mathcal{A}' \setminus f''\mathcal{A} \neq \emptyset$, there is a $q \in f''\mathbb{P}$ below an element of \mathcal{A}' . Let $p \in \mathbb{P}$ witness f(p) = q, and consider $\mathcal{A} \cup \{p\}$. This will be an antichain in \mathbb{P} since if $r \leq^{\mathbb{P}} p$, p' for some $p' \in \mathcal{A}$, $r \in \mathbb{P}$, then $f(r) \leq^{\mathbb{Q}} f(p), f(p') \in \mathcal{A}'$, contradicting that \mathcal{A}' is an antichain. But this contradicts that \mathcal{A} is maximal in \mathbb{P} . Hence $\mathcal{A}' \setminus f''\mathcal{A} = \emptyset$ so that $f''\mathcal{A} = \mathcal{A}'$ is maximal.

Complete homomorphisms (and complete suborders) are mostly used in building up a notion of forcing and showing that we don't muck up what we were trying to accomplish with previous stages. This not-mucking-things-up property ensures that we still are able to get generics for the "smaller" preorders, i.e. the domains of complete homomorphisms.

— 33 C • 4. Result

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P}, \mathbb{Q} \in V$ be preorders with $f : \mathbb{P} \to \mathbb{Q}$ a complete homomorphism in V. Let G be \mathbb{Q} -generic over V. Therefore f^{-1} "G is \mathbb{P} -generic over V and $V[f^{-1}] \subseteq V[G]$.

Proof .:.

It's easy to see that $f^{-1}"G$ is a filter: for upward closure, $p^* \leq \mathbb{P}$ p with $p^* \in f^{-1}"G$ yields $f(p^*) \leq \mathbb{Q}$ $f(p) \in G$ and therefore $p \in f^{-1}"G$. For compatibility in $f^{-1}"G$, if $p, p' \in f^{-1}"G$ then p and p' are at least compatible (otherwise $f(p) \perp f(p') \in G$). From here it suffices to show $f^{-1}"G$ intersects every dense set of \mathbb{P} since in particular, it needs to intersect $(\mathbb{P}_{\leq p} \cap \mathbb{P}_{\leq p'}) \cup \{r \in \mathbb{P} : r \perp p \lor r \perp p'\}$ and by compatibility of its members, must intersect $\mathbb{P}_{\leq p} \cap \mathbb{P}_{\leq p'}$.

Let $D \subseteq \mathbb{P}$ be dense so that there is a maximal antichain $\mathcal{A} \subseteq D$ by Result 32 C • 3. It follows that $f''\mathcal{A}$ is maximal in \mathbb{Q} and so $G \cap (f''\mathcal{A}) \neq \emptyset$ by Theorem 32 C • 5. Thus $\emptyset \neq (f^{-1}''G) \cap \mathcal{A} \subseteq (f^{-1}''G) \cap D$, as desired.

From here, $f \in V$ so that $f, G \in V[G]$ implies $f^{-1}"G \in V[G]$ and hence $V[f^{-1}"G] \subseteq V[G]$.

We also can prove the stronger statement that dense homomorphisms yield forcing equivalent preorders.

33C • 5. Theorem (Dense Forcing Equivalence)
Let V ⊨ ZFC be a transitive model we can force over. Let P, Q ∈ V be preorders with f : P → Q a dense homomorphism in V. Therefore P and Q are forcing equivalent.
More precisely, for G P-generic over V, and H Q-generic over V,
1. f"G↑ = {q ∈ Q : ∃p ∈ f"G (p ≤^Q q)} is Q-generic over V with V[G] = V[f"G↑]; and
2. f⁻¹"H is P-generic over V with V[f⁻¹"H] = V[H].

Proof .:.

For the sake of notation, write $X \uparrow$ for the upward closure and $X \downarrow$ for the downward closure (in the relevant

More General Topics and Theory

preorder) of a set $X: X \downarrow = \{p : \exists x \in X \ (p \leq x)\}$, and similarly for $X \uparrow$.

1. $f "G \uparrow$ is clearly closed upward. For compatibility, if $q, q' \in f "G \uparrow$ then there are $p, p' \in G$ with $f(p) \leq^{\mathbb{Q}} q$ and $f(p') \leq^{\mathbb{Q}} q'$. Since p, p' are compatible, there is a common extension which then maps to a common extension of q, q'.

For genericity, let $D \subseteq \mathbb{Q}$ be dense and open. Consider $f^{-1}"D$. This is dense in \mathbb{P} , because if $p \in \mathbb{P}$ then f(p) can be extended to an element $q \in D$ and then, since $f"\mathbb{P}$ is dense, we can extend to an element $f(p^*) \leq \mathbb{Q}$ q with therefore $f(p^*) \in D$ is compatible with f(p). As an incompatibility homomorphism, p^* is therefore compatible with p. But any $r \leq^{\mathbb{P}} p^*$, p yields $f(r) \in D$ since D is open and thus $r \in f^{-1}"D$ is below p, showing $f^{-1}"D$ is dense. As a result, $G \cap f^{-1}"D \neq \emptyset$ and this yields an element of $(f"G\uparrow) \cap D$.

To see that $V[G] = V[f"G\uparrow]$, we have $V[f"G\uparrow] \subseteq V[G]$ just because we can construct $f"G\uparrow$ from $G \in V[G]$ and $f \in V \subseteq V[G]$. To show $V[G] \subseteq V[f"G\uparrow]$, note that $f^{-1}"(f"G\uparrow)$ is \mathbb{P} -generic over V by Result 33 C • 4. It's not too difficult to show $f^{-1}"(f"G\uparrow) = G$ because $G \subseteq f^{-1}"(f"G\uparrow)$ and any element $p \in f^{-1}"(f"G\uparrow)$ has $f(p) \geq^{\mathbb{Q}} f(p')$ for some $p' \in G$ and therefore p is compatible with p'. If we consider $\mathbb{P}_{\leq p} \cup \{q \in \mathbb{P} : q \perp p\}$, this is dense in \mathbb{P} and so G intersects it at some p^* . We have $f(p^*) \leq^{\mathbb{Q}} f(p') \leq^{\mathbb{Q}} f(p)$ so that $p^* \in G \cap (\mathbb{P}_{\leq p} \cup \{q \in \mathbb{P} : q \perp p\})$ is compatible with p and hence below p, meaning $p \in G$. As a result, $G \in V[f"G\uparrow]$ and so $V[G] = V[f"G\uparrow]$.

2. Lemma 33 C • 3 and Result 33 C • 4 implies f^{-1} "*H* is \mathbb{P} -generic over *V* with $V[f^{-1}"H] \subseteq V[H]$. To see that $V[H] \subseteq V[f^{-1}"H]$, (1) implies $f"(f^{-1}"H)\uparrow$ is \mathbb{Q} -generic over *V* with $V[f^{-1}"H] = V[f"(f^{-1}"H)\uparrow]$ and $f"(f^{-1}"H)\uparrow = H$, because clearly $f"(f^{-1}"H)\uparrow \subseteq H$ and any element $q \in H$ has $\{f(p) \leq \mathbb{Q} q : p \in \mathbb{P}\}$ as dense so that it intersects *H* giving an element $p \in f^{-1}"H$ with $f(p) \leq \mathbb{Q} q$ and therefore $q \in f"(f^{-1}"H)\uparrow$. This shows $f"(f^{-1}"H)\uparrow = H$ and therefore *H*, being constructed from $f, \mathbb{Q} \in V$ and $f^{-1}"H \in V[f^{-1}"H]$ has $H \in V[f^{-1}"H]$. It follows that $V[H] \subseteq V[f^{-1}"H]$ and hence equality.

It's also not hard to see that if $\mathbb{P} \subseteq \mathbb{Q}$ has \mathbb{P} dense in \mathbb{Q} then the identity map id $\upharpoonright \mathbb{P} : \mathbb{P} \to \mathbb{Q}$ is a dense homomorphism. This then yields a great number of preorders as forcing equivalent, just by showing one is dense in the other. A clear example is two alternative definitions of $\mathrm{Add}(\aleph_0, \kappa)$.

- 33 C • 6. Example

By Dense Forcing Equivalence (33 C • 5), the following preorders are forcing equivalent: 1. Add($\aleph_0, 1$) = $\langle \{p : \omega \rightarrow 2 : |p| < \aleph_0\}, \supseteq, \emptyset \rangle$; 2. $\langle {}^{\langle \omega} 2, \supseteq, \emptyset \rangle$; 3. $\langle {}^{\langle \omega} \omega, \supseteq, \emptyset \rangle$; 4. $\langle \{p : \omega \rightarrow \omega : |p| < \aleph_0\}, \supseteq, \emptyset \rangle$; and 5. $\langle {}^{\langle \omega}[\omega], \triangleright, \emptyset \rangle$ (where \triangleright is end extension: $p \triangleright q$ iff $p \supseteq q$ and $x \in p \setminus q$ has $x > \max q$). *Proof* .:. (1,2) It should be clear that the universe of (2) is a dense subset of the universe of (1) and the identity map is then easily check to be a dense homomorphism. Hence the two are forcing equivalent. (2,3) Consider the incompatibility homomorphism $f : {}^{\langle \omega} \omega \rightarrow {}^{\langle \omega} 2$ defined by $f(\emptyset) = \emptyset$ and otherwise $f(p) = \operatorname{code}(p(0))^{\frown} \operatorname{code}(p(1))^{\frown} \operatorname{code}(p(2))^{\frown} \cdots^{\frown} \operatorname{code}(p(n))$, where $n = \operatorname{dom}(n) - 1$ and where $\operatorname{code}(x)$ is just the string of 0s and 1s that is $x \in \omega$ written in binary.

where $n = \operatorname{dom}(p) - 1$ and where $\operatorname{code}(x)$ is just the string of 0s and 1s that is $x \in \omega$ written in binary. That f is preserves \supseteq should be clear. That f preserves incompatibility should also be clear since code is injective. That $f'' = \omega \omega$ is dense follows from the fact that f is actually surjective. To see this, any $p \in {}^{<\omega} 2$ without leading zeros (i.e. $p(0) \neq 0$) can be seen as the binary representation of a single natural number $n = \sum_{m \in \operatorname{dom}(p)} p(m) \cdot 2^m$ with then $f(\langle 0, n \rangle) = p$. If p has m leading zeroes, we just take the previous n and consider $f((\operatorname{const}_0 \upharpoonright m) \cap \{\langle m, n \rangle\}) = p$. It follows from f being a dense homomorphism that (2) and (3) are forcing equivalent.

- (3,4) It should be clear that (3) is dense subset of (4) and the identity map is easily a dense homomorphism. Hence the two are forcing equivalent.
- (2,5) We identify $p \in {}^{<\omega}[\omega]$ with its characteristic function: for $n \leq \max(p)$,

$$\chi(p)(n) = \begin{cases} 1 & \text{if } n \in p \\ 0 & \text{if } n \notin p \end{cases}$$

This map $p \mapsto \chi(p) \upharpoonright (\max(p) + 1)$ is a surjective embedding, i.e. an isomorphism and thus a dense homomorphism. \dashv

Another example of this is our ability to merely use posets instead of preorders.

- **33C**•7. Theorem Let \mathbb{P} be a preorder. Define $p \approx q$ iff $p \leq^{\mathbb{P}} q \wedge q \leq^{\mathbb{P}} p$ for $p, q \in \mathbb{P}$. Therefore \approx is an equivalence relation and $\mathbb{P}/\approx = \langle \mathbb{P}/\approx, \leq^{\mathbb{P}/\approx}, \mathbb{1}^{\mathbb{P}/\approx} \rangle = \langle \{[p]_{\approx} : p \in \mathbb{P}\}, \{\langle [p]_{\approx}, [q]_{\approx} \rangle : p \leq^{\mathbb{P}} q\}, [\mathbb{1}^{\mathbb{P}}]_{\approx} \rangle$ is a preorder and in fact a poset. Moreover, \mathbb{P} and \mathbb{P}/\approx are forcing equivalent.

Proof .:.

That \approx is transitive and reflexive follows from $\leq^{\mathbb{P}}$ being transitive and reflexive. That \approx is symmetric is trivial. Thus \approx is an equivalence relation. Furthermore, $\leq^{\mathbb{P}/\approx}$ is well defined—i.e. its definition doesn't depend on the choice of representatives of the equivalence class—since if $p \approx p'$ and $q \approx q'$ with $p \leq q$ then by transitivity, $p' \leq p \leq q \leq q'$ implies $p' \leq q'$.

As a result, reflexivity and transitivity of $\leq^{\mathbb{P}/\approx}$ is immediate by the reflexivity and transitivity of $\leq^{\mathbb{P}}$. Similarly, $[\mathbb{1}^{\mathbb{P}}]_{\approx}$ being the maximal element of \mathbb{P}/\approx follows easily from $\mathbb{1}^{\mathbb{P}}$ being maximal in \mathbb{P} . For anti-symmetry, suppose $[p]_{\approx} \leq^{\mathbb{P}/\approx} [q]_{\approx} \leq^{\mathbb{P}/\approx} [p]_{\approx}$. It follows that $p \leq^{\mathbb{P}} q \leq^{\mathbb{P}} p$ and therefore $[p]_{\approx} = [q]_{\approx}$, as desired. This shows \mathbb{P}/\approx is a poset.

To see that \mathbb{P} and \mathbb{P}/\approx are forcing equivalent we use Dense Forcing Equivalence (33 C • 5) with a choice function for the equivalence classes: $f: \mathbb{P}/\approx \to \mathbb{P}$ has $f([p]_{\approx}) \approx p$. This will be a dense homomorphism. To see this, it's clearly an embedding by definition of $\leq^{\mathbb{P}/\approx}$ and $\mathbb{1}^{\mathbb{P}/\approx}$. To see that it preserves incompatibility, suppose $f([p]_{\approx}), f([q]_{\approx})$ are compatible in \mathbb{P} with $r \leq^{\mathbb{P}} f([p]_{\approx}), f([q]_{\approx})$. Since $r \approx f([r]_{\approx})$, we get

 $f([r]_{\approx}) \leqslant^{\mathbb{P}} f([p]_{\approx}), f([q]_{\approx}) \quad \text{iff} \quad [r]_{\approx} \leqslant^{\mathbb{P}/\approx} [p]_{\approx}, [q]_{\approx},$

meaning $[p]_{\approx}$ and $[q]_{\approx}$ are compatible. Trivially, the image of $p \mapsto [p]_{\approx}$ is dense in \mathbb{P}/\approx because it's surjective by definition. Thus f is a dense homomorphism and so \mathbb{P} is forcing equivalent to \mathbb{P}/\approx .

Another useful result in talking about forcing equivalent preorders is not only can we find names for the same object, we can translate names in an effective way.

— 33 C • 8. Theorem (Name Translation Theorem) –

Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force over. Let $f : \mathbb{P} \to \mathbb{Q}$ be an incompatibility homomorphism in V. Therefore, there is a function $T : V^{\mathbb{P}} \to V^{\mathbb{Q}}$ such that

- 1. for all H Q-generic over V and all $\tau \in V^{\mathbb{P}}$, $T(\tau)_H = \tau_{f^{-1}H}$;
- 2. If f is an embedding or dense homomorphism, then for all G P-generic over V and all $\tau \in V^{\mathbb{P}}$, $\tau_G = T(\tau)_{f^*G\uparrow}$;
- 3. If f is a dense homomorphism, then for all H Q-generic over V, $V[H] = \{T(\tau)_H : \tau \in V^{\mathbb{P}}\}.$

In particular, if f is a dense homomorphism, for every formula φ and $p \in \mathbb{P}$, $p \Vdash "\varphi(\tau)"$ iff $f(p) \Vdash "\varphi(T(\tau))"$.

Proof .:.

We define $T(\tau)$ by induction on \mathbb{P} -name rank. Set $T(\emptyset) = \emptyset$. Now suppose $T(\sigma)$ has been defined for all $\sigma \in \operatorname{dom}(\tau) \subseteq V^{\mathbb{P}}$. Define

$$T(\tau) = \{ \langle T(\sigma), f(p) \rangle : \langle \sigma, p \rangle \in \tau \}.$$

It's clear that then $T(\sigma)$ is a Q-name. Now let G be P-generic over V and H Q-generic over V so that if f is a dense homomorphism,

- $f"G\uparrow = \{q \in \mathbb{Q} : \exists p \in G \ (f(p) \leq^{\mathbb{Q}} q)\}$ is \mathbb{Q} -generic over V by Dense Forcing Equivalence (33 C 5); and
- f^{-1} "*H* is P-generic over *V* by Result 33 C 4.

Showing this T works isn't too difficult. Clearly $T(\emptyset) = \emptyset$ has (2) and (1) hold, so we may assume the result holds for elements of lower \mathbb{P} or \mathbb{Q} -name rank. Note that in showing (1), we don't need f^{-1} "H to be generic over \mathbb{P} .

- 1. To see $T(\tau)_H = \tau_{f^{-1}"H}$, let $\sigma_{f^{-1}"H} \in \tau_{f^{-1}"H}$ for some $\langle \sigma, p \rangle \in \tau$ and $p \in f^{-1}"H$. It follows that $f(p) \in H$ so that $\sigma_{f^{-1}"H} = T(\sigma)_H \in T(\tau)_H$: $\tau_{f^{-1}"H} \subseteq T(\tau)_H$. Similarly, if $T(\sigma)_H \in T(\tau)_H$ for some $\langle T(\sigma), f(p) \rangle \in T(\tau)$ for $f(p) \in H$, then $p \in f^{-1}"H$ with $\langle \sigma, p \rangle \in \tau$ implying $\sigma_{f^{-1}"H} = T(\sigma)_H \in \tau_H$: $T(\tau)_H \subseteq \tau_{f^{-1}"H}$.
- 2. Suppose f is dense. To see $\tau_G = T(\tau)_{f^*G\uparrow}$, we first show (\subseteq). Let $\sigma_G \in \tau_G$ so that $\langle \sigma, p \rangle \in \tau$ for some $p \in G$. It follows that $\langle T(\sigma), f(p) \rangle \in T(\tau)$ with $f(p) \in f^*G\uparrow$ and therefore $\sigma_G = T(\sigma)_{f^*G\uparrow} \in T(\tau)_{f^*G\uparrow}$ and hence $\tau_G \subseteq T(\tau)_{f^*G\uparrow}$.

Now suppose $\zeta_{f^{"}G\uparrow} \in T(\tau)_{f^{"}G\uparrow}$ so that $\langle \zeta, f(p) \rangle = \langle T(\sigma), f(p) \rangle \in T(\tau)$ for some $\langle \sigma, p \rangle \in \tau$. Our goal is now to show $p \in G$ since then $\zeta_{f^{"}G\uparrow} = T(\sigma)_{f^{"}G\uparrow} = \sigma_G \in \tau_G$ shows $T(\tau)_{f^{"}G\uparrow} \subseteq \tau_G$, and hence equality. Because $f(p) \in f^{"}G\uparrow, f(p) \geq^{\mathbb{Q}} f(q)$ for some $q \in G$.

- If f is an embedding, this implies p ≥^ℙ q ∈ G so that p ∈ G. As noted, this completes the proof of T(τ)_{f"G↑} ⊆ τ_G and hence T(τ)_{f"G↑} = τ_G.
- If f is a dense homomorphism, $f(p) \geq^{\mathbb{Q}} f(q)$ merely implies that p and q are compatible. So it's not immediate that $p \in G$. But if we consider $D = \mathbb{P}_{\leq p} \cup \{r \in \mathbb{P} : r \perp p\}$, this will be open and dense in \mathbb{P} and so we have a $p^* \in G \cap D$. Without loss of generality, we may take $p^* \leq^{\mathbb{P}} q$ by compatibility of G. Thus $f(p^*) \leq^{\mathbb{Q}} f(p)$ so p^* and p are compatible, and as $p^* \in D$, $p^* \leq^{\mathbb{P}} p$. Hence $p^* \in G$ gives $p \in G$.
- 3. Suppose f is a dense homomorphism and let $\sigma \in V^{\mathbb{Q}}$ be arbitrary. Consider the name

 $\tau = \{ \langle \varsigma, p \rangle : f(p) \Vdash_{\mathbb{Q}} ``T(\varsigma) \in \sigma" \}$

It's not difficult to see that $T(\tau)_H = \sigma_H$ since $x \in T(\tau)_H$ iff $x = T(\varsigma)_H$ for some $\langle \varsigma, p \rangle \in \tau$ and $f(p) \in H$ which therefore forces $T(\varsigma)_H \in \sigma_H$, meaning $T(\tau)_H \subseteq \sigma_H$.

Similarly, $x \in \sigma_H$ iff $x = \varsigma_H$ for some $\langle \varsigma, q \rangle \in \sigma$ with $q \in H$. Inductively, there is some $\tau' \in V^{\mathbb{P}}$ where $\varsigma_H = T(\tau')_H$ and therefore a $q^* \leq q$ in H forcing their equivalence and therefore $q^* \Vdash "T(\tau') \in \sigma"$. Since f is dense, $f'' \mathbb{P} \cap H \neq \emptyset$ and therefore there is some $q^{**} \leq \mathbb{Q}$ q^* with $q^{**} \in f'' \mathbb{P} \cap G$. This gives $q^{**} = f(p)$ for some $p \in \mathbb{P}$ and therefore $f(p) \Vdash "T(\tau') \in \sigma"$ meaning $\langle \tau', p \rangle \in \tau$ with $x = \varsigma_H = T(\tau')_H \in T(\tau)$. This means $\sigma_H \subseteq T(\tau)_H$ and therefore equality holds.

We also can show that what's forced gets translated too. Suppose f is a dense homomorphism.

- Suppose $p \Vdash "\varphi(\tau)$ ". Let H be Q-generic over V with $f(p) \in H$. We then have $p \in f^{-1}$ "H and by (1) and Dense Forcing Equivalence (33 C 5), $V[H] = V[f^{-1}"H] \vDash "T(\tau)_H = \tau_{f^{-1}"H} \land \varphi(\tau_{f^{-1}"H})$ ". It follows that $V[H] \vDash "\varphi(T(\tau)_H)$ ", and since H was arbitrary with $f(p) \in H$, $f(p) \Vdash "\varphi(T(\tau))$ ".
- Suppose $f(p) \Vdash "\varphi(T(\tau))$ ". Let G be P-generic over V with $p \in G$. Therefore have by (2) and Dense Forcing Equivalence (33 C • 5) $V[G] = V[f"G\uparrow] \vDash "T(\tau)_{f"G\uparrow} = \tau_G \land \varphi(T(\tau)_{f"G\uparrow})$ ", meaning $V[G] \vDash$ " $\varphi(\tau_G)$ ". As G was arbitrary with $p \in G$, $p \Vdash "\varphi(\tau)$ ".

The idea is that (1) and (2) show T indeed translates names, and (3) shows that this T is "surjective" in the sense that it covers all the names used by a generic extension by \mathbb{Q} .

§ 33 D. A word on class forcing

Sometimes we want to force with a preorder that is not a set but is still easily definable. In doing so, we must be careful, because the various theorems about forcing with sets need no longer apply. An easy example to show this would be the preorder $Col(\aleph_0, Ord)$ which collapses the class of ordinals to be countable:

 $\operatorname{Col}(\aleph_0, \operatorname{Ord}) = \{p : \omega \rightarrow \operatorname{Ord} : |p| < \aleph_0\} = {}^{<\omega}\operatorname{Ord}.$

For any **Col**(\aleph_0 , Ord)-generic *G* over the ground model *V*, $\bigcup G$ will be a surjection from \aleph_0 to Ord^{*V*} and therefore *V*[*G*] has a bijection from ω to Ord^{*V*} = Ord^{*V*[*G*]}. It follows by replacement that Ord^{*V*[*G*]} $\in V[G]$, meaning *V*[*G*] $\not\models$ ZFC.

So we need to be careful about what preorders we use if we want to preserve ZFC. Additionally, we also need to be careful about what exactly we mean by "generic" since no set will be dense in $Col(\aleph_0, Ord)$, for example, only classes. To modify some definitions given earlier, a generic filter will need to intersect every dense *class* of the class preorder.^{xviii} Similarly we can define class antichains and so forth.

The way to show a generic extension by a class preorder still satisfies ZFC is to require that the preorder is "set-like" in a certain sense. In particular, we may use the following, adopted from [11].

— 33 D•1. Definition

Let \mathbb{P} be a (class) preorder. We call $D \subseteq \mathbb{P}$ predense (below $p \in \mathbb{P}$) iff every $q \in \mathbb{P}$ (below p) is compatible with an element of D.

We call \mathbb{P} pretame iff for all $p \in \mathbb{P}$, and every set X, and class sequence^{xix} $\langle D_x : x \in X \rangle$ of dense classes D_x for $x \in X$, there is a set $\{d_x \subseteq D_x : x \in X\}$ of sets predense below some $q \leq^{\mathbb{P}} p$. We call \mathbb{P} tame iff \mathbb{P} is pretame and $\mathbb{1}^{\mathbb{P}} \Vdash \mathbb{P}$, the powerset axiom.

We can phrase tame-ness without the use of the forcing relation, but showing this requires a great amount of effort and additional technology for what is basically a long footnote. That said, we are still justified in using the forcing relation because for class preorders, it's definable when the preorder is pretame. In general with class forcing, there isn't a guarantee the forcing relation $p \Vdash \varphi$ for any given formula φ is definable like with set preorders. Showing that pretame class preorders do admit such a relation again requires too much effort to include here.

— 33 D•2. Theorem

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P}, \leq, \{1\} \subseteq V$ yield a class preorder \mathbb{P} such that \mathbb{P} is tame over V. Therefore for any G \mathbb{P} -generic over V has $\mathbf{V}[G] \models \mathsf{ZFC}$.

It turns out that basically all the preorders we would want to use have this property, partly because they are built up in a natural way from smaller preorders, and to confirm pretame-ness, we just restrict our attention to some initial stages of how the dense classes are built up. Confirming powerset is harder but can be done using complicated technology similar to the definition of pretameness. We, of course, will not confirm any of this here and will instead work with just with set preorders. For a more thorough treatment and discussion of class forcing, see [11], [1], or [25].

We can see for example that $Col(\aleph_0, Ord)$ isn't tame nor even pretame since each

$$D_n = \{ f : \omega \to \operatorname{Ord} : n \in \operatorname{dom}(f) \}$$

is a class dense in $\text{Col}(\aleph_0, \text{Ord})$ for $n \in \omega$ (as slices of the definable class $D = \{\langle f, n \rangle : n \in \omega \land n \in \text{dom}(f) \land f : \omega \rightarrow \text{Ord}\}$). But for any $p = \emptyset$, there is not a $q \leq p$ and collection of sets $\{d_n : n \in \omega\}$ such that each $d_n \subseteq D_n$ is predense below q. To see this, let $q : \omega \rightarrow \text{Ord}$ with $\text{dom}(q) \subseteq n$ and suppose $\{d_m : m \in \omega\}$ has each $d_m \subseteq D_m$ predense below q. As sets, we have a bound $\beta_n = \sup \bigcup_{r \in d_n} r(n)$ on what the predense set d_n allows. In particular, $q \cup \{\langle n, \beta_n + 1 \rangle\} \leq q$ is incompatible with every element of d_n so that d_n isn't predense.

xviii This is equivalent for set preorders since any such class will need to be a subset of the preorder and therefore be a set by comprehension.

^{xix}Note that since each D_x is a class, $\langle D_x : x \in X \rangle$ doesn't make much sense since it's a collection of classes (and really a collection of ordered pairs of classes). So by $\langle D_x : x \in X \rangle$ existing as a class, we really mean that a certain class $D \subseteq V \times X$ exists where each slice $D_x = \{a : \langle a, x \rangle \in D\}$ is dense for $x \in X$.

Section 34. Iterated Forcing

We now move away from the more concrete examples and turn back to the general theory of forcing. The idea of forcing allows us to expand a transitive model of set theory V to a generic extension V[G] by some $G \notin V$. The concept of iterated forcing is just doing this multiple times in a single step. More precisely, in the generic extension V[G] we might force with a preorder in V[G] to another generic extension V[G][H] and again with a preorder now in V[G][H] to V[G][H][K]. The idea behind iterated forcing is to find a single preorder *in* V with generic G * H * K so that V[G * H * K] = V[G][H][K]. And from here, we can continue to all sorts of longer iterations of arbitrary length $\alpha \in$ Ord.

First we work with two-step (and subsequently finite-step) iterations. Then we generalize this to iterations of arbitrary length. While the idea behind two-step iterations is somewhat canonical, we have many choices of what to do at limit stages in more general iterations. This yields a great number of topics in the subject of forcing trying to figure out how to force a lot of things at once for the sake of consistency results.

§34A. Two-step iterations

We first note that any preorder we're dealing with has a name that is *always* forced to be a preorder. Usually in the generic extension, we only get truth forced by *some* element of the preorder. But actually we can often get a potentially different name that is forced by $\mathbb{1}^{\mathbb{P}}$ to have the property.

— 34 A • 1. Lemma (Conditional Name Lemma) –

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be appropriate for forcing and $G \mathbb{P}$ -generic over V. Suppose $\mathbf{V}[G] \models \varphi(\tau_G)$ for some $\tau \in V^{\mathbb{P}}$ and formula φ . Suppose $\mathsf{ZFC} \vdash ``\exists x \varphi(x)``$. Therefore there is a $\sigma \in V^{\mathbb{P}}$ where $\tau_G = \sigma_G$ and in fact

$$\mathbb{1}^{\mathbb{P}} \Vdash ``\varphi(\sigma) \land (\varphi(\tau) \to \tau = \sigma)".$$

Proof .:.

Let H be \mathbb{P} -generic over V. We have that

$$\mathbf{V}[H] \models ``\exists x \ (\varphi(x) \land (\varphi(\tau_H) \to \tau_H = x))''.$$
(*)

The reason is just that if $V[H] \models "\varphi(\tau_H)$ ", then $x = \tau_H$ witnesses (*). And otherwise, we already know $V[H] \models ZFC$ so by hypothesis $V[H] \models "\exists x \ \varphi(x)$ ". Any such x also witnesses (*). Since H was arbitrary, $\mathbb{I}^{\mathbb{P}} \Vdash "\exists x \ (\varphi(x) \land (\varphi(\tau) \to \tau = x))$ ".

By Maximum Principle (33 B • 8), there is some \mathbb{P} -name σ where $\mathbb{1}^{\mathbb{P}} \Vdash "\varphi(\sigma) \land (\varphi(\tau) \to \tau = \sigma)"$. \dashv

So by considering e.g. the infinite binary tree $\langle {}^{<\omega}2, \supseteq \rangle$ if our desired P-name isn't actually appropriate for forcing, we get the following.

This simplifies the situation because we don't need to worry about whether the \mathbb{P} -name for our preorder is only a preorder when forcing below some $p \in \mathbb{P}$ and otherwise is some other object entirely. The bottom line is that it suffices to consider only certain kinds of \mathbb{P} -names that allow us to simply assume basic properties of the preorder in the forcing

relation.

The idea behind the definition of $\mathbb{P} * \dot{\mathbb{Q}}$ is that as we go further down in \mathbb{P} , we decide more and more of $\dot{\mathbb{Q}}$ and this allows us to also go down in \mathbb{Q} . The result is that dense sets can be separated in a natural way into a dense set in \mathbb{P} and a dense set of \mathbb{Q} . Recall the notation $\langle \langle x, y \rangle \rangle$ for a \mathbb{P} -name for $\langle x_G, y_G \rangle$.^{xx}

- **34A**•**3**. **Definition** Let $\mathbb{P} = \langle \mathbb{P}, \leq^{\mathbb{P}} \rangle$ be appropriate for forcing. $\dot{\mathbb{Q}}$ is a \mathbb{P} -name for a preorder appropriate for forcing iff $\mathbb{1}^{\mathbb{P}} \Vdash ``\dot{\mathbb{Q}}$ is appropriate for forcing". For a \mathbb{P} -name $\dot{\mathbb{Q}} = \langle \langle \dot{\mathbb{Q}}, \leq^{\dot{\mathbb{Q}}} \rangle \rangle$, define $\mathbb{P} * \dot{\mathbb{Q}} = \langle \mathbb{P} * \mathbb{Q}, \leq \rangle$ by $\mathbb{P} * \dot{\mathbb{Q}} = \{ \langle p, \dot{q} \rangle : p \in \mathbb{P} \land \dot{q} \in \operatorname{dom}(\dot{\mathbb{Q}}) \land p \Vdash ``\dot{q} \in \dot{\mathbb{Q}}" \}$, with $\langle p^*, \dot{q}^* \rangle \leq \langle p, \dot{q} \rangle$ iff $p^* \leq^{\mathbb{P}} p$ and $p^* \Vdash ``\dot{q}^* \leq^{\dot{\mathbb{Q}}} \dot{q}"$.

Given the definability of forcing in the ground model, the above yields that $\mathbb{P} * \dot{\mathbb{Q}}$ is a preorder *in the ground model* whereas \mathbb{Q} is usually only in the generic extension. Of course, it's not difficult to confirm that this is indeed a preorder.

Let $\mathbb{P} = \langle \mathbb{P}, \leq^{\mathbb{P}} \rangle$ be appropriate for forcing. Let $\dot{\mathbb{Q}} = \langle \! \langle \dot{\mathbb{Q}}, \leq^{\dot{\mathbb{Q}}} \rangle \! \rangle$ be a \mathbb{P} -name for a preorder appropriate for forcing. Therefore $\mathbb{P} * \dot{\mathbb{Q}}$ is appropriate for forcing.

Proof .:.

Assuming $\mathbb{P} * \dot{\mathbb{Q}} = \langle \mathbb{P} * \mathbb{Q}, \leq \rangle$ is a preorder, it's trivial to confirm that it's appropriate for forcing: the maximal element being $\langle \mathbb{1}^{\mathbb{P}}, \dot{\mathbb{1}}^{\dot{\mathbb{Q}}} \rangle$ is immediate and the incompatibility follows from the incompatibility of \mathbb{P} . So it suffices to show that $\mathbb{P} * \dot{\mathbb{Q}}$ is a preorder.

Reflexivity is immediate since it holds of \mathbb{P} and $\mathbb{1}^{\mathbb{P}}$ forces the reflexivity of $\leq^{\dot{\mathbb{Q}}}$. For transitivity, suppose $\langle p^{**}, \dot{q}^{**} \rangle \leq \langle p, \dot{q} \rangle$, aiming to show $\langle p^{**}, \dot{q}^{**} \rangle \leq \langle p, \dot{q} \rangle$. By the transitivity of $\leq^{\mathbb{P}}$, $p^{**} \leq^{\mathbb{P}} p$. We also have, being below p^{**} and $\mathbb{1}^{\mathbb{P}}$,

$$p^{**} \Vdash ``\dot{q}^{**} \leqslant^{\dot{\mathbb{Q}}} \dot{q}^{*} \leqslant^{\dot{\mathbb{Q}}} \dot{q} \text{ and } \leqslant^{\dot{\mathbb{Q}}} \text{ is transitive}''$$

and therefore $p^{*} \Vdash ``\dot{q}^{**} \leqslant^{\dot{\mathbb{Q}}} \dot{q}^{''}$. Hence $\langle p^{**}, \dot{q}^{**} \rangle \leqslant \langle p, q \rangle$. \dashv

We unfortunately don't get generally that $\mathbb{P} * \dot{\mathbb{Q}}$ is a poset even if \mathbb{P} is and $\dot{\mathbb{Q}}$ is forced to be a poset. The issue is that we might have $p \Vdash \dot{q}_0 = \dot{q}_1$ so that $\langle p, \dot{q}_0 \rangle \leq \langle p, \dot{q}_1 \rangle \leq \langle p, \dot{q}_0 \rangle$, but $\langle p, \dot{q}_0 \rangle \neq \langle p, \dot{q}_1 \rangle$. This is why we have adopted the convention here of using *preorders* rather than posets. Posets are nicer to work with, but this added nicety usually doesn't come up. Preorders work just as well.

Regardless, we are more interested in what the generic extensions of $\mathbb{P} * \hat{\mathbb{Q}}$ look like. To do this, we will often be translating dense sets into different contexts and sometimes reforming them. We also do the same for generics, and so it's useful to have a notation to express how generics are broken down and reformed.

- **34A**•5. Definition Let \mathbb{P} be a preorder and $\dot{\mathbb{Q}}$ a \mathbb{P} -name for a preorder. For $A \subseteq \mathbb{P}$ and $B \subseteq \dot{\mathbb{Q}}_A$, define $A * B = \{ \langle p, \dot{q} \rangle \in \mathbb{P} * \dot{\mathbb{Q}} : p \in A \land \dot{q}_A \in B \}.$

This doesn't exactly mesh with the same use of "*" when defining $\mathbb{P} * \dot{\mathbb{Q}}$, but it's quite useful and intuitive. Generally, we will only define A * B when A is \mathbb{P} -generic over the ground model. On the topic of dense sets, another corollary of Conditional Name Lemma (34 A • 1) is that if we want to consider a \mathbb{P} -name for some particular dense set D of a preorder \mathbb{Q} in the generic extension, we can choose a name \dot{D} which is always forced to be dense. This allows us to easily think about the dense sets of the two-step iteration.

^{xx}Since $\langle x_G, y_G \rangle = \{\{x_G\}, \{x_G, y_G\}\}$, an easy canonical example of such a name is

$$\langle\!\langle x,y\rangle\!\rangle = \left\{\!\left(\left\{\langle x,\mathbb{1}^{\mathbb{P}}\rangle\right\},\mathbb{1}^{\mathbb{P}}\right),\left(\left\{\langle x,\mathbb{1}^{\mathbb{P}}\rangle,\langle y,\mathbb{1}^{\mathbb{P}}\rangle\right\},\mathbb{1}^{\mathbb{P}}\right)\!\right\}.$$

— 34 A • 6. Theorem (Two-Step Iterated Forcing) —

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be appropriate for forcing and $\dot{\mathbb{Q}} \in V^{\mathbb{P}}$ a \mathbb{P} -name for a preorder appropriate for forcing. Let $G \subseteq \mathbb{P} * \dot{\mathbb{Q}}$. Therefore,

- 1. G is $\mathbb{P} * \dot{\mathbb{Q}}$ -generic over V iff $G = G_{\mathbb{P}} * G_{\mathbb{Q}}$ for some $G_{\mathbb{P}}$ that is \mathbb{P} -generic over V and some $G_{\mathbb{Q}}$ that is $\dot{\mathbb{Q}}_{G_{\mathbb{P}}}$ -generic over $V[G_{\mathbb{P}}]$.
- 2. In fact, we can take
 - $G_{\mathbb{P}} = \mathfrak{p}_{\mathbb{P}}G = \{ p \in \mathbb{P} : \exists \dot{q} \ (\langle p, \dot{q} \rangle \in G) \}, \text{ and } G_{\mathbb{Q}} = \{ \dot{q}_{G_{\mathbb{P}}} : \exists p \in \mathbb{P} \ (\langle p, \dot{q} \rangle \in G) \}.$
- 3. Moreover, $V[G] = V[G_{\mathbb{P}}][G_{\mathbb{Q}}]$ for $G_{\mathbb{P}}$ and $G_{\mathbb{Q}}$ as in (2).

Proof .:.

1. The direction (\rightarrow) will be shown in (2). For the (\leftarrow) direction, let $G_{\mathbb{P}}$ be \mathbb{P} -generic over V and $G_{\mathbb{Q}} = \dot{\mathbb{Q}}_{G_{\mathbb{P}}}$ -generic over $V[G_{\mathbb{P}}]$. Write

$$G' = G_{\mathbb{P}} * G_{\mathbb{Q}} = \left\{ \langle p, \dot{q} \rangle \in \mathbb{P} * \dot{\mathbb{Q}} : p \in \mathbb{P} \land \dot{q}_{G_{\mathbb{P}}} \in G_{\mathbb{Q}} \right\}.$$

Let $D \in V$ be dense in $\mathbb{P} * \dot{\mathbb{Q}}$. For each $\dot{q} \in \text{dom}(\dot{\mathbb{Q}})$, consider the conditions forcing an extension in D:

$$D_{\dot{q}} = \left\{ p \in \mathbb{P} : \exists \dot{q}^* \left(p \Vdash ``\dot{q}^* \leq^{\mathbb{Q}} \dot{q} \land \langle p, \dot{q}^* \rangle \in \check{D}'' \right) \right\} \subseteq \mathbb{P}.$$

This is dense in \mathbb{P} because any $p \in \mathbb{P}$ has $\langle p, \dot{q} \rangle \in \mathbb{P} * \dot{\mathbb{Q}}$ with an extension $\langle p^*, \dot{q}^* \rangle \in D$ so that $p^* \Vdash ``\dot{q}^* \leq \dot{\mathbb{Q}}$ and $\langle p^*, \dot{q}^* \rangle \in D$ means that p^* clearly forces this. In particular, $p^* \in D_{\dot{q}}$ and $D_{\dot{q}} \in V$ is dense in \mathbb{P} . Now consider in the generic extension the corresponding qs from $G_{\mathbb{P}}$ and D:

$$D_{\mathbb{Q}} = \left\{ \dot{q}_{G_{\mathbb{P}}} : \exists p \in G_{\mathbb{P}} \left(\langle p, \dot{q} \rangle \in D \right) \right\} \subseteq \mathbb{Q}.$$

This is dense in \mathbb{Q} since any $\dot{q}_{G_{\mathbb{P}}} \in \mathbb{Q}$ has a p forcing this and therefore an extension $p^* \in D_{\dot{q}} \cap G_{\mathbb{P}}$ which then forces an extension of $\dot{q}_{G_{\mathbb{P}}}$ into $D_{\mathbb{Q}}$. As a result, there is some element $\dot{q}_{G_{\mathbb{P}}} \in D_{\mathbb{Q}} \cap G_{\mathbb{Q}}$ which then yields a $p \in G_{\mathbb{P}}$ with $\langle p, \dot{q} \rangle \in D \cap (G_{\mathbb{P}} * G_{\mathbb{Q}}) \neq \emptyset$. Hence $G_{\mathbb{P}} * G_{\mathbb{Q}}$ is $\mathbb{P} * \dot{\mathbb{Q}}$ -generic over V.

2. It should be clear as defined that $G = G_{\mathbb{P}} * G_{\mathbb{Q}}$. Let D be dense in \mathbb{P} . We can identify this with a subset of the iteration $D' = \{\langle p, \dot{q} \rangle \in \mathbb{P} * \dot{\mathbb{Q}} : p \in D\}$ which is dense in $\mathbb{P} * \dot{\mathbb{Q}}$. In particular, $G \cap D' \neq \emptyset$ which yields an element of $\mathfrak{p}_{\mathbb{P}}(G \cap D') \subseteq \mathfrak{p}_{\mathbb{P}}G \cap \mathfrak{p}_{\mathbb{P}}D' = G_{\mathbb{P}} \cap D$. So $G_{\mathbb{P}}$ is \mathbb{P} -generic over V.

Similarly, if D is dense in $\mathbb{Q} = \dot{\mathbb{Q}}_{G_{\mathbb{P}}}$ then we can choose a \mathbb{P} -name \dot{D} with $\mathbb{1}_{\mathbb{P}} \Vdash ``\dot{D}$ is dense in $\dot{\mathbb{Q}}$ " by Conditional Name Lemma (34 A • 1). In particular, $D' = \{\langle p, \dot{q} \rangle \in \mathbb{P} * \dot{\mathbb{Q}} : p \Vdash ``\dot{q} \in \dot{D}"\}$ is dense in $\mathbb{P} * \dot{\mathbb{Q}}$ where then $G \cap D' \neq \emptyset$. An element $\langle p, \dot{q} \rangle \in G \cap D'$ has $p \in G_{\mathbb{P}}$ with $p \Vdash ``\dot{q} \in \dot{D}"$ and therefore $\dot{q}_{G_{\mathbb{P}}} \in \dot{D}_{G_{\mathbb{P}}} = D$. And by definition $\dot{q}_{G_{\mathbb{P}}} \in G_{\mathbb{Q}} : G_{\mathbb{Q}} \cap D \neq \emptyset$. So $G_{\mathbb{Q}}$ is \mathbb{Q} -generic over $V[G_{\mathbb{P}}]$.

3. Since both generic extensions are transitive models containing V and satisfying ZFC, the containments follow from Theorem 31 A • 6 (3) by showing that each generic extension contains G, $G_{\mathbb{P}}$ and $G_{\mathbb{Q}}$. In V[G], from G, we can construct $G_{\mathbb{P}}$ and $G_{\mathbb{Q}}$ as in (2) so clearly $V[G_{\mathbb{P}}][G_{\mathbb{Q}}] \subseteq V[G]$. But given $G_{\mathbb{P}}$ and $G_{\mathbb{Q}}$, we can construct $G = G_{\mathbb{P}} * G_{\mathbb{Q}}$. Hence $V[G] \subseteq V[G_{\mathbb{P}}][G_{\mathbb{Q}}]$ and so equality holds.

This then allows us to start from a model of CH, force to a model of \neg CH, and then force CH again! The applications are endless! But really the point is that this makes some constructions and arguments easier because we might know what each preorder does individually. For instance, we might want to force $2^{\aleph_0} = \aleph_3$ and $2^{\aleph_1} = \aleph_3$, and this allows us to do it in one step by first forcing the first and then forcing the other. Of course, there is some careful checking to make sure we don't screw up what we forced in the previous step. And this question becomes more complicated and harder to answer with longer iterations. For now, we focus on the fact that iterating two ccc preorders yields a ccc preorder. Recall the use of ccc preorders in preserving cardinals and cofinalities from Theorem 32 C • 7 to motivate why this might be useful.

— 34 A • 7. Lemma

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a ccc preorder appropriate for forcing in \mathbf{V} . Let $\hat{\mathbb{Q}}$ be a \mathbb{P} -name for a preorder appropriate for forcing such that $\mathbb{1}^{\mathbb{P}} \Vdash "\dot{\mathbb{Q}}$ is ccc". Therefore $\mathbb{P} * \dot{\mathbb{Q}}$ is ccc.

Proof .:.

Note that forcing with \mathbb{P} yields that \aleph_1 is still regular in the generic extension. In particular, for $\sigma \in V^{\mathbb{P}}$, if σ is forced to be a small subset of \aleph_1^{V} , i.e. $\mathbb{1}^{\mathbb{P}} \Vdash "\sigma \subseteq \check{\aleph}_1 \land |\sigma| < \check{\aleph}_1$ ", then there is some $\alpha < \omega_1^{\mathsf{V}}$ such that $\mathbb{1}^{\mathbb{P}} \Vdash "\sigma \subseteq \check{\alpha}$ ".

So suppose $\mathbb{P} * \dot{\mathbb{Q}}$ is not ccc in **V**: let $A = \{\langle p_{\alpha}, \dot{q}_{\alpha} \rangle : \alpha < \kappa\}$ be an antichain witnessing this with $\kappa \geq \aleph_{1}^{\mathsf{V}}$. Consider in $V^{\mathbb{P}}$ the name $\sigma = \{\langle \check{\alpha}, p_{\alpha} \rangle : \alpha < \kappa\}$ so that $\sigma_{G} = \{\xi < \kappa : p_{\xi} \in G\} \subseteq \kappa$ whenever $G \subseteq \mathbb{P}$. So let G be \mathbb{P} -generic over V, writing $\dot{\mathbb{Q}}_{G} = \mathbb{Q}$ and $(\dot{q}_{\alpha})_{G} = q_{\alpha}$ for $\alpha < \kappa$.

The corresponding set $\{q_{\alpha} : \alpha \in \sigma_G\}$ is an antichain of \mathbb{Q} . To see this, for distinct $\alpha, \beta \in \sigma_G$, if q_{α} and q_{β} were compatible, then this would be forced by some element of $G: p \Vdash ``\dot{r} \leq \overset{\hat{\mathbb{Q}}}{=} \dot{q}_{\alpha}, \dot{q}_{\beta}$ '' for some $p \in G$ and $\dot{r} \in \operatorname{dom}(\dot{\mathbb{Q}})$. As G is a filter, we may freely extend p and assume without loss of generality that $p \leq^{\mathbb{P}} p_{\alpha}, p_{\beta}$. But then $\langle p, \dot{r} \rangle$ extends both $\langle p_{\alpha}, \dot{q}_{\alpha} \rangle, \langle p_{\beta}, \dot{q}_{\beta} \rangle \in A$ in $\mathbb{P} * \dot{\mathbb{Q}}$, contradicting that A is an antichain.

But $\mathbb{1}^{\mathbb{P}} \Vdash ``\dot{\mathbb{Q}}$ is ccc", and so in particular, $V[G] \models ``|\{q_{\alpha} : \alpha \in \sigma_G\}| < \kappa$ ". As *G* was arbitrary, $\mathbb{1}^{\mathbb{P}} \Vdash ``\sigma \subseteq \kappa \land |\sigma| < \kappa$ " and thus σ is forced to be bounded by some $\beta < \kappa$, meaning there are at most $|\beta| < \kappa$ elements of σ in the ground model, contradicting construction. \dashv

By the same proof above, Lemma 34 A • 7 actually generalizes to κ -cc preorders in that if \mathbb{P} is κ -cc and $\mathbb{1}^{\mathbb{P}} \Vdash$ " $\dot{\mathbb{Q}}$ is $\check{\kappa}$ -cc", then $\mathbb{P} * \dot{\mathbb{Q}}$ is κ -cc. We don't care as much about κ -cc preorders in general because ccc preorders are often much more useful as we'll see in the next subsection.

Iterated forcing also can be called *product forcing* in the case that $\dot{\mathbb{Q}} = \check{\mathbb{Q}}$ for some \mathbb{Q} in the ground model. In this case, we actually get commutativity: $\mathbb{P} * \check{\mathbb{Q}} \cong \mathbb{Q} * \check{\mathbb{P}}$ —which we call $\mathbb{P} \times \mathbb{Q}$ —and so the same generic extensions. Of course, one needs to be careful with using the above results: \mathbb{Q} may be, for example, ccc in the ground model but not in the generic extension by \mathbb{P} .^{xxi} As a result, the product of ccc preorders might not be ccc without additional information about \mathbb{P} . So really, one needs to understand the product forcing version of Lemma 34 A • 7 as saying a stronger statement: more than both \mathbb{P} and \mathbb{Q} being ccc in the ground model, if in addition \mathbb{Q} is *still ccc* in the generic extension by \mathbb{P} , then $\mathbb{P} \times \mathbb{Q}$ is ccc.

§ 34 B. Longer iterations

The ideas of the previous subsection allow us to pretty easily define iterations of finitely many preorders, and more generally the *successor* stage in an iteration of potentially infinitely many preorders. The limit stage is where things aren't uniquely defined: do we take the ω th stage to be the direct limit of all previous preorders? If successor stages act like the product, do we just take the "product" of all of the previous preorders? This results in the "inverse limit" and is basically the largest the limit stage could be whereas the direct limit is the smallest the limit stage could be, and there are many different posets that fall somewhere in between. Ultimately, we have a large degree of freedom of coming up with a limit preorder \mathbb{P}_{ω} whose generic extensions contain generic extensions of previous preorders. In other words, if the goal of iterated forcing is to get a chain of models $V \subseteq V[G_1] \subseteq V[G_2] \subseteq \cdots$, what exactly $V[G_{\omega}]$ should be isn't determined by the previous preorders and generics. Answering these questions and nailing down the final generic extension fundamentally requires us to think about what our conditions look like and what their *support* is.

Up until this point, adding a maximal element 1 to our preorders has been practically uncessary: $\mathbb{1}^{\mathbb{P}} \Vdash \varphi$ can be restated as $\forall p \in \mathbb{P} \ (p \Vdash \varphi)$ and instead of $\langle x, \mathbb{1}^{\mathbb{P}} \rangle$ in the name τ we just need to have $\{\langle x, p \rangle : p \in \mathbb{P}\} \subseteq \tau$. So most of the above results do not depend on 1 in an essential way. This changes now, since we require the ability to "do nothing" with our conditions. How often we do *something* is considered the "support".

To give some motivation for the very technical and lengthy definitions to follow, our current technology allows us

^{xxi}For example, a suslin tree (which hasn't been introduced in this document) \mathbb{P} is ccc but $\mathbb{P} \times \mathbb{P}$ isn't [24]. In other words, \mathbb{P} is ccc in the ground model, but not in the generic extension, precisely because \mathbb{P} adds an uncountable branch through \mathbb{P} that can be used in the generic extension to form an uncountable antichain.

ITERATED FORCING

to start with \mathbb{P}_0 and iterate with a \mathbb{P}_0 -name $\dot{\mathbb{Q}}_0$ to get $\mathbb{P}_1 = \mathbb{P}_0 * \dot{\mathbb{Q}}_0$ and then given a \mathbb{P}_1 -name $\dot{\mathbb{Q}}_1$, we can define $\mathbb{P}_2 = \mathbb{P}_1 * \dot{\mathbb{Q}}_1$. The general perspective is that $\langle \dot{\mathbb{Q}}_{\alpha} : \alpha < \kappa \rangle$ is the sequence of (names of) preorders we want to force with. When forcing with them we get initial segments of our iteration: $\langle \mathbb{P}_{\alpha} : \alpha < \kappa \rangle$ has \mathbb{P}_{α} as the iteration of the first α -many $\dot{\mathbb{Q}}_s$ meaning $\mathbb{P}_{\alpha+1} = \mathbb{P}_{\alpha} * \dot{\mathbb{Q}}_{\alpha}$.

For example, the series of iterations $((1 * \dot{Q}_0) * \dot{Q}_1) * \dot{Q}_2$ —where 1 is the trivial preorder with just one element—can be thought of as the result of the sequence $\langle P_i : i < 4 \rangle$ given by the sequence $\langle \dot{Q}_i : i < 3 \rangle$ where

$$\begin{split} \mathbb{P}_0 &= \mathbb{1} \\ \mathbb{P}_1 &= \mathbb{P}_0 * \dot{\mathbb{Q}}_0 \end{split} \qquad \begin{split} \mathbb{P}_2 &= \mathbb{P}_1 * \dot{\mathbb{Q}}_1 \\ \mathbb{P}_3 &= \mathbb{P}_2 * \dot{\mathbb{Q}}_2 = ((\mathbb{P}_0 * \dot{\mathbb{Q}}_0) * \dot{\mathbb{Q}}_1) * \dot{\mathbb{Q}}_2 \end{split}$$

Formally, the elements of these will be pairs of pairs of pairs and so on, but we can instead view them as sequences of elements or choice function into the names of conditions—meaning into the domains of the $\dot{\mathbb{Q}}_n$ s. In this way, elements of \mathbb{P}_3 will be choices function for $\{\operatorname{dom}(\dot{\mathbb{Q}}_0), \operatorname{dom}(\dot{\mathbb{Q}}_1), \operatorname{dom}(\dot{\mathbb{Q}}_2)\}$, meaning elements of $\prod_{n<3} \operatorname{dom}(\dot{\mathbb{Q}}_n)$. Sequences in larger \mathbb{P}_{α} s extend each other when the corresponding values in the $\dot{\mathbb{Q}}_{\alpha}$ s (are forced to) extend each other.

Notationally, the conditions of $\mathbb{P}_{n+1} = \dot{\mathbb{Q}}_0 * \dot{\mathbb{Q}}_1 * \ldots * \dot{\mathbb{Q}}_n$ before would take the form $\langle \langle \langle \langle \emptyset, \dot{q}_0 \rangle, \dot{q}_1 \rangle \cdots \rangle, \dot{q}_n \rangle$. The ordering would be defined by

$$\langle \langle \langle \langle \emptyset, \dot{q}_0^* \rangle, \dot{q}_1^* \rangle \cdots \rangle, \dot{q}_n^* \rangle \leq_{n+1} \langle \langle \langle \langle \emptyset, \dot{q}_0 \rangle, \dot{q}_1 \rangle \cdots \rangle, \dot{q}_n \rangle$$

 $\text{iff} \quad \langle \langle \langle \emptyset, \dot{q}_0^* \rangle, \dot{q}_1^* \rangle \cdots \rangle \leqslant_n \langle \langle \langle \emptyset, \dot{q}_0 \rangle, \dot{q}_1 \rangle \cdots \rangle \text{ and } \langle \langle \langle \emptyset, \dot{q}_0^* \rangle, \dot{q}_1^* \rangle \cdots \rangle \Vdash ``\dot{q}_n^* \leqslant^{\dot{\mathbb{Q}}_n} \dot{q}_n ``.$

Viewed a sequences, we identify $\langle \langle \langle \langle \emptyset, \dot{q}_0 \rangle, \dot{q}_1 \rangle \cdots \rangle, \dot{q}_n \rangle$ with the map $p = n \mapsto \dot{q}_n$. In this way we can rephrase the ordering as

 $p^* \leq_{n+1} p$ iff $p^* \upharpoonright n \leq_n p \upharpoonright n$ and $p^* \upharpoonright n \Vdash p^*(n) \leq^{\hat{\mathbb{Q}}_n} p(n)^n$,

which basically says that we go down in each component in the only way that makes sense. One may note that this is partly why we start with 1: we need $p \upharpoonright 0 = \emptyset \in 1$ to force statements about elements of $\dot{\mathbb{Q}}_0$.

Under this sequence view, we have the following results or perhaps definition or characterization depending on how the reader wants to think of it.

34B•1. Result
Let ⟨Q̂_n : n < N⟩, N < ω, be a sequence of names where
1. Q̂_n is a P_n-name for a preorder for each n < N;
2. we define P₀ = 1, P_{n+1} = P_n * Q̂_n for each n < N, and we view elements as sequences as above;
Therefore, for all n < N,
3. Each p ∈ P_n is a function p ∈ ∏_{n<N} dom(Q̂_n);

4. If $p \in \mathbb{P}_n$ and m < n, then $p \upharpoonright m \in \mathbb{P}_m$;

- 5. $p \in \mathbb{P}_{n+1}$ iff $p \upharpoonright n \in \mathbb{P}_n$ and $p(n) \in \operatorname{dom}(\dot{\mathbb{Q}}_n)$ such that $p \upharpoonright n \Vdash "p(n) \in \dot{\mathbb{Q}}_n$ ";
- 6. $p^* \leq \mathbb{P}_{n+1} p$ iff $p^* \upharpoonright n \leq_n p \upharpoonright n$ and $p^* \upharpoonright n \Vdash p^*(n) \leq \hat{\mathbb{Q}}_n p(n)^n$.

This generalizes to larger iterations supposing we know what to do at limit stages. In the end, the limit stage α will be determined not by the order—which trivially has $p^* \leq_{\alpha} p \leftrightarrow \forall \xi < \alpha \ (p^* \upharpoonright \xi \leq_{\xi} p \upharpoonright \xi)$ —but instead by what sequences $p \in \prod_{\xi < \alpha} \text{dom}(\dot{\mathbb{Q}}_{\xi})$ we allow. Support is the way we restrict such conditions.

- **34B**•2. Definition Let κ be an ordinal. For $\alpha < \kappa$, let \mathbb{P}_{α} be a preorder and let $\dot{\mathbb{Q}}_{\alpha} = \langle \langle \dot{\mathbb{Q}}_{\alpha}, \leq_{\alpha}', \dot{\mathbb{I}}_{\alpha}' \rangle \rangle$ be a \mathbb{P}_{α} -name for a preorder with $\mathbb{1}^{\mathbb{P}_{\alpha}} \Vdash :\dot{\mathbb{I}}_{\alpha}'$ is maximal in $\dot{\mathbb{Q}}_{\alpha}''$. Then the *support* of any $p \in \prod_{\alpha < \kappa} \operatorname{dom}(\dot{\mathbb{Q}}_{\alpha})$ is $\operatorname{sprt}(p) = \{\alpha < \kappa : \mathbb{1}^{\mathbb{P}_{\alpha}} \not\models :p(\alpha) = \dot{\mathbb{I}}_{\alpha}'''\}$

And from here we can give the very technical and long definition of longer iterations. Note that really we're defining the iteration \mathbb{P}_{α} s based on our choice of $\dot{\mathbb{Q}}_{\alpha}$ s. So there are no real restrictions on the $\dot{\mathbb{Q}}_{\alpha}$ s except that they are \mathbb{P}_{α} -names for preorders.

— 34 B • 3. Definition

Let κ be an ordinal. Let $I \subseteq \mathcal{O}(\kappa)$ be some collection we will think of as allowed supports with $\emptyset \in I$. A κ -stage iterated forcing with supports in I is a pair of sequences $\langle \mathbb{P}_{\alpha} : \alpha \leq \kappa \rangle$ and $\langle \dot{\mathbb{Q}}_{\alpha} : \alpha < \kappa \rangle$ such that for all $\xi \leq \alpha < \kappa$,

- 1. Each $\mathbb{P}_{\alpha} = \langle \mathbb{P}_{\alpha}, \leq_{\alpha}, \mathbb{1}_{\alpha} \rangle$ is appropriate for forcing (or else trivial) with trivial $\mathbb{P}_{0} = \mathbb{1} = \langle \{\emptyset\}, =, \emptyset \rangle$.
- 2. Each $\dot{\mathbb{Q}}_{\alpha} = \langle \langle \dot{\mathbb{Q}}_{\alpha}, \leq_{\alpha}', \dot{\mathbb{I}}_{\alpha}' \rangle \rangle$ is a \mathbb{P}_{α} -name for preorder appropriate for forcing (or else trivial in that $|\operatorname{dom}(\mathbb{Q}_{\alpha})| = 1$).
- 3. Each $\mathbb{1}_{\alpha} = \langle \hat{\mathbb{1}}'_{\xi} : \xi < \alpha \rangle$.
- 4. Each element of \mathbb{P}_{α} is a function $p \in \prod_{\xi < \alpha} \operatorname{dom}(\dot{\mathbb{Q}}_{\xi})$.
- 5. If $p \in \mathbb{P}_{\alpha}$, then $p \upharpoonright \xi \in \mathbb{P}_{\xi}$.
- 6. For $\alpha = \xi + 1$ a successor, we essentially set $\mathbb{P}_{\xi+1} = \mathbb{P}_{\xi} * \dot{\mathbb{Q}}_{\xi}$:
 - $\mathbb{P}_{\xi+1}$ is defined by $p \in \mathbb{P}_{\xi+1}$ iff $p \upharpoonright \xi \in \mathbb{P}_{\xi}$, $p(\xi) \in \operatorname{dom}(\dot{\mathbb{Q}}_{\xi})$, and $p \upharpoonright \xi \Vdash_{\mathbb{P}_{\xi}} "p(\xi) \in \dot{\mathbb{Q}}_{\xi}$ "; and
 - $\leq_{\xi+1}$ is defined by $p^* \leq_{\xi+1} p$ iff $p^* \upharpoonright \xi \leq_{\xi} p \upharpoonright \xi$ and $p^* \upharpoonright \xi \Vdash "p^*(\xi) \leq'_{\xi} p(\xi)"$.
- 7. For α a limit, we essentially require support in I with stronger conditions having stronger initial segments:
 - \mathbb{P}_{α} is defined by $p \in \mathbb{P}_{\alpha}$ iff $p \upharpoonright \xi \in \mathbb{P}_{\xi}$ for every $\xi < \alpha$ and sprt $(p) \in I$;
 - \leq_{α} is defined by $p^* \leq_{\alpha} p$ iff $p^* \upharpoonright \xi \leq_{\xi} p \upharpoonright \xi$ for every $\xi < \alpha$.

Above, $\Vdash_{\mathbb{P}_{\alpha}}$ is just there to help disambiguate which preorder we're forcing with, although the only preorder $p \upharpoonright \xi$ is in \mathbb{P}_{ξ} (and hence the only preorder the forcing relation makes sense for there is with \mathbb{P}_{ξ}). The above notation is fairly standard in the literature, but we may suggestively write $\mathbf{X}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ instead of \mathbb{P}_{κ} . Of course, this notation doesn't say what happens at limit stages, so this must be described in the surrounding context.

Really all this means is that elements of $\mathbb{P}_{\kappa} = \mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ are choice functions and the ordering is just forced pointwise extension such that we don't enlarge the support too much: for q a condition, $p \leq_{\kappa} q$ iff $\operatorname{sprt}(p) \in I$ and $p \upharpoonright \alpha \Vdash$ " $p(\alpha) \leq_{\alpha}' q(\alpha)$ " for each $\alpha < \kappa$.

It should also be clear that for a κ -stage iteration $\langle \mathbb{P}_{\alpha} : \alpha \leq \kappa \rangle$, $\langle \dot{\mathbb{Q}}_{\alpha} : \alpha < \kappa \rangle$, if we cut off the iteration at some stage α , we get an α -stage iteration $\langle \mathbb{P}_{\xi} : \xi \leq \alpha \rangle$, $\langle \dot{\mathbb{Q}}_{\xi} : \xi < \alpha \rangle$.

For $\kappa = 2$, this pretty clearly meshes with the previous definition from Definition 34 A • 3: there $\mathbb{P} * \mathbb{Q}$ would have to be written as follows:

	$\mathbb{P}_0 = \mathbb{1}$
$\dot{\mathbb{Q}}_0 = \check{\mathbb{P}}$	$\mathbb{P}_1 = \mathbb{1} \ast \dot{\mathbb{Q}}_0$
$\dot{\mathbb{Q}}_1 = \dot{\mathbb{Q}}$	$\mathbb{P}_2 = \mathbb{1} * \dot{\mathbb{Q}}_0 * \dot{\mathbb{Q}}_1 = \mathbb{1} * \check{\mathbb{P}} * \dot{\mathbb{Q}} \cong \mathbb{P} * \dot{\mathbb{Q}}.$

For longer iterations, in the case that $I = \{\emptyset\}$, then \mathbb{P}_{α} will simply be trivial for limit α , so the requirement that I isn't *just* this is necessary below. We also need $\emptyset \in I$ to ensure that $\mathbb{1}_{\alpha} = \langle \dot{\mathbb{1}}'_{\xi} : \xi < \alpha \rangle$ —whose support is empty—is in fact in \mathbb{P}_{α} for limit α .^{xxii}

— 34 B•4. Corollary

Let κ be an ordinal and $\{\emptyset\} \subsetneq I \subseteq \mathbb{P}(\kappa)$. Therefore a κ -stage iterated forcing with supports in I, $\langle \mathbb{P}_{\alpha} : \alpha \leq \kappa \rangle$ with $\langle \dot{\mathbb{Q}}_{\alpha} : \alpha < \kappa \rangle$ as in Definition 34 B • 3, results in \mathbb{P}_{κ} which is a trivial preorder (if all the $\dot{\mathbb{Q}}_{\alpha}$ s are trivial for $\alpha < \kappa$) or a preorder appropriate for forcing. In fact, each \mathbb{P}_{α} has this property for $\alpha < \kappa$.

Proof .:.

If all of the $\dot{\mathbb{Q}}_{\alpha}$ s are \mathbb{P}_{α} -names for trivial preorders—i.e. $|\operatorname{dom}(\dot{\mathbb{Q}}_{\alpha})| = 1$ —then it's not difficult to show each \mathbb{P}_{α} is trivial by induction (because there is only ever one function in $\prod_{\alpha < \kappa} \operatorname{dom}(\dot{\mathbb{Q}}_{\alpha})$).

Otherwise \mathbb{P}_{ξ} is trivial up to some $\xi \leq \alpha < \kappa$ but $\dot{\mathbb{Q}}_{\alpha}$ is a \mathbb{P}_{α} -name for a preorder appropriate for forcing (and

^{xxii}Alternatively, one may say that according to our definitions, there simply are no iterations with support in I if $\emptyset \notin I$ since $\mathbb{1}_{\alpha}$ as defined wouldn't be in \mathbb{P}_{α} .

ITERATED FORCING

thus nontrivial) in which case there is some preorder in the ground model \mathbb{Q}_{α} where $\mathbb{1}_{\alpha} \Vdash ``\check{\mathbb{Q}}_{\alpha} = \dot{\mathbb{Q}}_{\alpha}$ " and hence $\mathbb{P}_{\alpha+1} \cong \mathbb{P}_{\alpha} * \check{\mathbb{Q}}_{\alpha} \cong \mathbb{Q}_{\alpha}$ is appropriate for forcing by Corollary 34 A • 4. To see that $\langle \dot{\mathbb{1}}'_{\xi} : \xi < \alpha + 1 \rangle$ is indeed maximal in \mathbb{P}_{α} , inductively, $a = \langle \dot{\mathbb{1}}'_{\xi} : \xi < \alpha \rangle$ is maximal in \mathbb{P}_{α} , and it's not hard to see that $a^{\frown} \langle \dot{\mathbb{1}}'_{\alpha} \rangle$ is then maximal in $\mathbb{P}_{\alpha+1}$.

For the limit step \mathbb{P}_{α} , we first must show it's a preorder and that it's appropriate for forcing. Transitivity follows by transitivity of the previous \mathbb{P}_{ξ} s inductively. Reflexivity follows from the fact that $p \leq_{\alpha} p^* \leq_{\alpha} p$ implies $p \upharpoonright \xi \leq_{\xi} p^* \upharpoonright \xi \leq_{\xi} p \upharpoonright \xi$ for every $\xi < \alpha$. This implies $p = p^*$ by the reflexivity of each \leq_{ξ} : $p \upharpoonright \xi = p^* \upharpoonright \xi$ for every $\xi < \alpha$ implies $p = \bigcup_{\xi < \alpha} p \upharpoonright \xi = \bigcup_{\xi < \alpha} p^* \upharpoonright \xi = p^*$.

 \mathbb{P}_{α} has a maximal element $\langle \hat{\mathbb{I}}'_{\xi} : \xi < \alpha \rangle$ since inductively this restricted to ξ is the maximal element of \mathbb{P}_{ξ} . For $p \in \mathbb{P}_{\alpha}$, to show there are always incompatible extensions of p, we must break into cases.

- If $\xi \in \operatorname{sprt}(p)$, then as $\dot{\mathbb{Q}}_{\xi}$ is forced to be appropriate for forcing, choose two extensions $p_0(\xi)$ and $p_1(\xi)$ forced to be incompatible: $\mathbb{1}_{\xi} \Vdash "p_0(\xi), p_1(\xi) \leq_{\xi}' p(\xi) \wedge p_0(\xi) \perp p_1(\xi)$ ". Therefore $p_0 = p \upharpoonright \xi^{\frown}(p_0(\xi))^{\frown} p \upharpoonright (\alpha \setminus \xi)$ and $p_1 \upharpoonright \xi^{\frown}(p_1(\xi))^{\frown} p \upharpoonright (\alpha \setminus \xi)$ have $\operatorname{sprt}(p_0) = \operatorname{sprt}(p_1) = \operatorname{sprt}(p) \in I$ and therefore $p_0, p_1 \in \mathbb{P}_{\alpha}$ are two incompatible extensions of p.
- If sprt(p) = Ø, then p = l_α. For a non-empty X ∈ I, take ξ < α arbitrarily and note that any two incompatible extensions p₀, p₁ ≤_ξ p ↾ ξ in P_ξ yields p₀ ∩ p ↾ (α \ ξ) and p₁ ∩ p ↾ (α \ ξ) as two incompatible extensions of p.

Note that for (7) and (5) in Definition $34 \mathbb{B} \cdot 3$ to make sense for κ with lots of limit ordinals below it, we will want *I* to be closed under subsets: $\alpha < \kappa$ with $\operatorname{sprt}(p \upharpoonright \alpha) \subseteq \operatorname{sprt}(p) \in I$ with α a limit requires $\operatorname{sprt}(p \upharpoonright \alpha) \in I$. In general, we want *I* to be an *ideal*, the dual concept of a filter in that $F \subseteq \mathcal{P}(X)$ is a filter iff $\{X \setminus A : A \in F\}$ is an ideal. The idea is that we usually want our supports to be "small" in some sense, and we use an ideal to make this concept coherent.

- **34B**•5. Definition Let X be a set and $I \subseteq \mathcal{P}(X)$. $\emptyset \subsetneq I \subsetneq \mathcal{P}(X)$ is an *ideal* iff • $X \subseteq Y \in I$ implies $X \in I$; and • $X, Y \in I$ implies $X \cup Y \in I$.

 $\mathcal{P}(\kappa)$ is sometimes called an *improper ideal*. We will not adopt this convention here to preserve the fact that every ideal corresponds to a filter. Some natural examples of ideals are given from size:

— 34 B • 6. Example — _____

- For κ a cardinal and $\lambda \leq \kappa$ another cardinal,
 - 1. $\{X \subseteq \kappa : |X| \le \lambda\}$ is an ideal.
 - 2. $\{X \subseteq \kappa : |X| < \lambda\}$ is an ideal.
 - 3. $\{X \subseteq \kappa : X \text{ is bounded in } \kappa\}$ is an ideal.

Some more trivial ways of getting ideals are as follows:

- $\{\emptyset\}$ is an ideal.
- For any filter $F \subseteq \mathcal{O}(\kappa)$, $\{\kappa \setminus X : X \in F\}$ is an ideal.
- For any $Y \subseteq X \neq \emptyset$, $\mathcal{P}(Y) \subseteq \mathcal{P}(X)$ is an ideal.
- For any $X \neq \emptyset$ and $x \in X$, $\{Y \subseteq X : x \notin Y\} \subseteq \mathcal{P}(X)$ is an ideal called a *principal* ideal, corresponding to principal filters.

(7) and (5) of Definition $34B \cdot 3$ tells us that *I* should be closed under subsets. Definition $34B \cdot 3$ (6) tells us that *I* should be non-principal in that at successor stages, we can always increase the support by one element and not cause any problems. Of course, the definition doesn't entail these exactly, but they provide motivation for focusing on non-principal ideals.

Occasionally in our iterations we will talk of *full support* where we take support in $\mathcal{O}(\kappa)$, i.e. no restrictions. This is not an ideal as defined in Definition 34B • 5, but this doesn't matter much at all for our purposes. For the most part,

the popular ideals for supports are *bounded support* as in Example 34 B • 6 (3) (or (2) if $\kappa = \lambda$ is regular), *countable support* as in (1) with $\lambda = \aleph_0$, and *finite support* as in (2) with $\lambda = \aleph_0$. Other supports, such as easton support will be composed of these, allowing for different supports depending on the stage of the iteration.

The primary goal of iterated forcing is the following, where we use the notation $\mathbf{x}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ for \mathbb{P}_{α} in iterations. One downside to using this notation is that whereas \mathbb{P}_{α} is more like a parameter—a particular preorder from the ground model— $\mathbf{x}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is more like a defined notion which requires interpretation and therefore might differ when defined in the ground model versus the generic extension in the same way that the preorder $Col(\aleph_0, \aleph_2)$ will have a different interpretation in the ground model versus its generic extension. This is especially a problem for iterations since we might be working with *many* different generic extensions. But the reader just has to know that below we are *always*

interpretting the preorder as defined in the ground model. It's just that writing $\left(\mathbf{x}_{\xi<\alpha}\,\dot{\mathbf{Q}}_{\xi}\right)^{\mathbf{v}}$ is a bit too much while \mathbb{P}_{α}

is a bit less transparent. As such, we say $\mathbf{a}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ is a κ -stage iterated forcing rather than the more formally correct pair of sequences $\langle \mathbf{a}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi} : \alpha \leq \kappa \rangle$, $\langle \dot{\mathbf{Q}}_{\alpha} : \alpha < \kappa \rangle$. The notation $\mathbf{a}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ makes it easy to see what the sequences are anyway. In doing so, we let the support be defined by context rather than integrating it into the notation to further clutter things.

— 34 B • 7. Theorem (Iterated Forcing) -

Assume the following:

- $V \vDash$ ZFC is a transitive model we can force over.
- $\kappa \in \operatorname{Ord} \cap V$
- $I \in V$ with $\{\emptyset\} \subsetneq I \subseteq \mathcal{P}(\kappa)$ be a non-principal ideal or $\mathcal{P}(\kappa)$ itself.
- $\mathbf{k}_{\alpha < \kappa} \, \dot{\mathbb{Q}}_{\alpha} \in V$ is a κ -stage iterated forcing with supports in I of V.
- G is $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ -generic over V.

Therefore for each $\alpha < \kappa$,

1. $G \upharpoonright \alpha = \{p \upharpoonright \alpha : p \in G\}$ is $\mathbf{k}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ -generic over *V*.

2. $G_{\alpha+1} = \{p(\alpha)_{G \upharpoonright \alpha} : p \in G\}$ is $(\dot{\mathbb{Q}}_{\alpha})_{G \upharpoonright \alpha}$ -generic over $V[G \upharpoonright \alpha]$.

Proof .:.

Since G is a filter, and p^{*} ≤_κ p implying p^{*} ↾ α ≤_α p ↾ α, it follows that any two elements of G ↾ α have a common extensions in G ↾ α. G ↾ α is also pretty easily closed upward (since every r ∈ *****_{ξ < α} Q_ξ has an extension in *****_{ξ < κ} Q_ξ and thus in G just by adding a tail of i'_ε's). Hence G ↾ α is always a filter.

To show genericity, let $D \subseteq \bigstar_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ be dense in *V*. We can take $D' = \{q \in \bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi} : q \upharpoonright \alpha \in D\}$. Note that if $D' \cap G \neq \emptyset$, then $D \cap (G \upharpoonright \alpha) \neq \emptyset$ by restricting to elements to be in $\bigstar_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$. So it suffices to show *D'* is dense in $\bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$. To see this, let $p \in \bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ be arbitrary. Consider that $p \upharpoonright \alpha$ has an extension $p^* \in D$. Note that $\operatorname{sprt}(p^* \cap (p \upharpoonright (\kappa \setminus \alpha))) \subseteq \operatorname{sprt}(p) \cup \operatorname{sprt}(p^*)$ and therefore is in *I*. Hence $p^* \cap (p \upharpoonright (\kappa \setminus \alpha))$ is a condition of $\bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ which extends *p* and is in *D'*. Therefore $G \cap D' \neq \emptyset$ and so $(G \upharpoonright \alpha) \cap D \neq \emptyset$.

Write Q_α = ⟨Q_α, ≤'_α, 1'_α⟩ for (Q_α)_{G ↾α} where α < κ. First we show that G_α is a filter and then that it's generic. That any two elements of G_α are can be extended to an element of G_α follows immediately from the fact that all elements of G have this property: a common extension r ≤_κ p, q ∈ G with r ∈ G yields r ↾ (α + 1) ≤_{α+1} p ↾ (α + 1), q ↾ (α + 1) and therefore r ↾ α ⊨ "r(α) ≤'_α p(α), q(α)". In particular, r(α)_{G ↾α} ∈ G_α as a common extension to p(α)_{G ↾α} and q(α)_{G ↾α} in Q_α.

To show the upward closure of G_{α} , let $p^*(\alpha)_{G \upharpoonright \alpha} \leq_{\alpha}' p = \dot{p}_{G \upharpoonright \alpha}$ with $p^* \in G$ and $\dot{p} \in \text{dom}(\mathbb{Q}_{\alpha})$. Change p^* to $q = (p^* \setminus \{ \langle \alpha, p^*(\alpha) \rangle \}) \cup \{ \langle \alpha, \dot{p} \rangle \}$. This only (potentially) decreases the support by one element and is therefore a condition since I is an ideal. Moreover, $p^* \leq q$ and therefore $q \in G$ and so $q(\alpha)_{G \upharpoonright \alpha} = p \in G_{\alpha}$. Hence G_{α} is a filter. To show genericity of G_{α} , let $D \subseteq \mathbb{Q}_{\alpha}$ be dense in $V[G \upharpoonright \alpha]$. There is some condition $p_D \in G$ with $p_D \upharpoonright \alpha \in G \upharpoonright \alpha$ forcing this: $p_D \upharpoonright \alpha \Vdash ``D`$ is dense in $\dot{\mathbb{Q}}_{\alpha}$.". It's not hard to see that

$$D' = \{q \leq_{\kappa} p_D : p_D \upharpoonright \alpha \Vdash ``q(\alpha) \in D"\}$$

is dense below p_D in $\mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$. To see this, any $q \leq p_D$ has $q \upharpoonright \alpha \Vdash \exists x \ (x \leq'_{\alpha} q(\alpha) \land x \in D)$ ". We can then extend $q \upharpoonright \alpha$ to $q^* \upharpoonright \alpha$ to get a name $\tau \in \operatorname{dom}(\dot{\mathbb{Q}}_{\alpha})$ where $q^* \upharpoonright \alpha \Vdash \exists x \ (x \leq'_{\alpha} q(\alpha) \land x \in D)$ ". It follows that $q^* = q^* \upharpoonright \alpha \frown \langle \tau \rangle \frown q \upharpoonright (\kappa \setminus (\alpha + 1))$ has $\operatorname{sprt}(q^*) \subseteq \operatorname{sprt}(q^* \upharpoonright \alpha) \cup \{\alpha\} \cup \operatorname{sprt}(p) \in I$ with $q \in D'$. Hence D' is dense below p_D in $\mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$.

So by genericity of G, there is some element $q \in G \cap D'$ with then $q(\alpha)_{G \upharpoonright \alpha} \in G_{\alpha} \cap D$. Hence G_{α} is \mathbb{Q}_{α} -generic over $V[G \upharpoonright \alpha]$.

Unfortunately, a converse to the above isn't possible in the sense that we may have $\mathbf{*}_{\alpha < \kappa} \hat{\mathbf{Q}}_{\alpha}$ with sequences of generics $\langle G_{\alpha} : \alpha < \kappa \rangle$ and $\langle G \upharpoonright \alpha : \alpha < \kappa \rangle$ where

- $G \upharpoonright \alpha = \prod_{\xi < \alpha} G_{\alpha} \cap \mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\alpha}$ that is $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\alpha}$ -generic over V; and
- $(G_{\alpha+1})_{G \upharpoonright \alpha}$ is $(\dot{\mathbb{Q}}_{\alpha})_{G \upharpoonright \alpha}$ -generic over $V[G \upharpoonright \alpha]$;

but there is no generic G having these sections as above.^{xxiii} Put another way, there may be a chain of generic extensions $V \subseteq V[G_0] \subseteq V[G_1] \subseteq \cdots$ with no V[G] containing $\{G_n : n < \omega\}$ such that G is generic over the iteration in V, because G may code something nasty.^{xxiv} If we want to have a better idea of what these limit generics will look like when they *do* exist, we need to look further into the supports of iteration. These are obviously important because they shape the limit iterations.

§ 34 C. A diversion into model theory and support

Let us now investigate why supports are important by considering what their iterations look like. Then we can look at their generic extensions.

— 34 C • 1. Definition

- Let κ be an ordinal and $I \subseteq \mathcal{P}(\kappa)$. Let $\bigstar_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ be a κ -stage iterated forcing with supports in I. We say this is a • *finite support* iteration iff $I = \{X \subseteq \kappa : |X| < \aleph_0\};$
 - Junite support iteration in $T = \{X \subseteq k : |X| < x_0\},\$
 - *bounded support* iteration iff $I = \{X \subseteq \kappa : X \text{ is bounded in } \kappa\} = \{X \subseteq \kappa : \sup X < \kappa\};$
 - *full support* iteration iff $I = \mathcal{P}(\kappa)$.

These three kinds of support are not the only important ones in the subject of iterated forcing: countable support and revised countable support are both very important for technical arguments regarding preorders that preserve certain kinds of stationary sets, also called "proper forcings", motivating the proper forcing axiom (PFA). But such topics will not be covered here. Instead, we will focus on the above three because the resulting iterations have a simple form to describe in terms of the initial segments of the iteration.

Recall that the direct limit dir $\lim_{\mathcal{F}} \mathcal{A}$ of a directed system of models $\langle \mathcal{A}, \mathcal{F} \rangle$ (\mathcal{A} is the set of models and \mathcal{F} is the set of embeddings) is the "least" model that A embeds into dir $\lim_{\mathcal{F}} \mathcal{A}$ for each $A \in \mathcal{A}$ and if A embeds into M for each $A \in \mathcal{A}$, then dir $\lim_{\mathcal{F}} \mathcal{A}$ embeds into M too as in Figure 6 A • 8, reproduced below:

^{xxiii}To see this, let $\mathbf{V} \models \mathsf{ZFC}$ be a countable transitive model. Hence there is some real $r \in {}^{\omega}2$ coding the countable ordinal $(\operatorname{Ord} \cap V, \in)$. We may consider $\mathbb{Q}_0 = \mathsf{Add}(\aleph_0, 1)$ with generic G_0 over $V, \mathbb{Q}_1 = \mathsf{Add}(\aleph_0, 1)^{\mathsf{V}[G_0]}$ with generic G_1 over $V[G_0], \mathbb{Q}_2 = \mathsf{Add}(\aleph_0, 1)^{\mathsf{V}[G_0*G_1]}$ with generic G_1 over $V[G_0 * G_1]$, and so on, continually adding a subset of ω in the generic extension. Taking names for these preorders, the result is a preorder $\mathbf{*}_{n<\omega} \dot{\mathbb{Q}}_n$. It's not too difficult to show that we may modify the generics G_n to G'_n so that their members' first value is r(n). If there were a resulting generic G' to $\mathbf{*}_{n<\omega} \dot{\mathbb{Q}}_n$ that works with these G'_n s, we could reconstruct r and get $r \in V[G']$, which would show that $V[G'] \models "\operatorname{Ord} \cap V = \operatorname{Ord} \cap V[G']$ is countable", contradicting $V[G'] \models \mathsf{ZFC}$.

^{xxiv}One may trivially show this just by taking a countable, cofinal sequence of ordinals $\langle \alpha_n : n < \omega \rangle$ in a countable, transitive model $V \models ZFC$ and then iteratively collapsing these ordinals down to be countable with $\mathbb{Q}_n = \operatorname{Col}(\aleph_0, \alpha_n)^{V[G_0 * \dots * G_{n-1}]}$. We then get $V \subseteq V[G_0] \subseteq V[G_0 * G_1] \subseteq \cdots$ with no generic extension of the iteration V[G] because such a G codes the countability of V and V[G]'s ordinals. The issue, of course, is that this sequence of ordinals $\langle \alpha_n : n < \omega \rangle$ isn't in the ground model, and hence the iteration $\mathbf{x}_{n<\omega} \dot{\mathbb{Q}}_n$ isn't even definable over the ground model V. So the construction in the previous footnote is a more substantive example.



34 C • 2. Figure: The direct limit embeddings

It turns out that with finite support iterations, $\mathbf{*}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha} = \operatorname{dir} \lim_{\alpha < \kappa} \mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$, that is the iteration is the direct limit of the previous iterations. We always have embeddings from the initial segments into the final iteration. Here we use "embedding" to be the usual meaning between models, and here preorders are FOL(\leq , 1)-structures. Elsewhere in the literature, frequently homomorphisms and embeddings between preorders regard preorders instead as FOL(\leq , \perp , 1)-structures, meaning the homomorphism and embedding must preserve incompatibility. This restricted sense is unnecessary although it can clean-up some theorem statements.

— 34 C • 3. Lemma

Let κ be an ordinal, $I \subseteq \mathcal{P}(\kappa)$, and $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ a κ -stage iteration with support in a non-principal ideal I or $\mathcal{P}(\kappa)$ itself.^{XXV} Therefore for each $\alpha < \beta \leq \kappa$, there is an incompatibility embedding $\iota_{\alpha,\beta}$ from $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ to $\mathbf{*}_{\xi < \beta} \dot{\mathbb{Q}}_{\xi}$ defined by $\iota_{\alpha,\beta}(p) = p^{-} \langle \dot{\mathbb{I}}'_{\xi} : \alpha \leq \xi < \beta \rangle$.

Proof ...

We have $\iota_{\alpha,\beta}(p) \in \mathbf{*}_{\xi < \beta} \dot{\mathbb{Q}}_{\xi}$ since $\operatorname{sprt}(\iota_{\alpha,\beta}(p)) = \operatorname{sprt}(p) \in I$. It's also not difficult to see that if $p^* \leq_{\alpha} p$ then $\iota_{\alpha,\beta}(p^*) \leq_{\beta} \iota_{\alpha,\beta}(p)$, and the converse clearly holds since restricting to α yields $p^* = \iota_{\alpha,\beta}(p^*) \upharpoonright \alpha \leq_{\alpha} \iota_{\alpha,\beta}(p) \upharpoonright \alpha = p$. It should also be clear that $\iota_{\alpha,\beta}$ is injective and hence is an embedding. That the embedding is complete is obvious in that a common extension $r \leq_{\beta} \iota_{\alpha,\beta}(p), \iota_{\alpha,\beta}(q)$ yields $r \upharpoonright \alpha \leq_{\alpha} p, q$ and the converse is already established since $\iota_{\alpha,\beta}$ is an embedding. \dashv

Note that by Name Translation Theorem $(33 \text{ C} \cdot 8)$, we can then "easily" translate names over previous iterations into names over later iterations.

Really this tells us what the direct limit should look like since it should be composed of as little as possible. This makes sense since any non-principal ideal will contain the ideal of finite sets. Put in another way, the iteration is composed *only* of the values of these embeddings: $\mathbf{a}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha} = \bigcup_{\alpha < \kappa} \iota_{\alpha,\kappa} \mathbf{a}_{\beta} \mathbf{a}_{\beta}$. Equivalently, $p \in \mathbf{a}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ iff there is some $\alpha < \kappa$ where $p = \iota_{\alpha,\kappa}(p \upharpoonright \alpha)$. This is the main idea behind why the resulting model is the direct limit. We also then do away with specifying the embeddings, as they will always be these $\iota_{\alpha,\beta}$ s which just add a tail of $\dot{1}$'s which then obviously commute: $\iota_{\alpha,\gamma} = \iota_{\beta,\gamma} \circ \iota_{\alpha,\beta}$ for $\alpha < \beta < \gamma$.

— 34 C • 4. Definition –

Let κ be an ordinal and $\mathbf{*}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ a κ -stage iteration with support in some $I \subseteq \mathcal{O}(\kappa)$. We say $\mathbf{*}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ is the *direct limit of the previous iterations* iff it's the direct limit of the system of embeddings $\{\{\mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi} : \alpha < \kappa\}, \{\iota_{\alpha,\beta} : \alpha \leq \beta < \kappa\}$ where $\iota_{\alpha,\beta} : \mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi} \rightarrow \mathbf{*}_{\xi < \beta} \dot{\mathbf{Q}}_{\xi}$ is defined by $\iota_{\alpha,\beta}(p) = p^{-} \langle \dot{\mathbf{I}}'_{\xi} : \alpha \leq \xi < \beta \rangle$.

With this simplification, we can state when we take direct limits generally. Note that $I \cap \mathcal{P}(\alpha) \subseteq \{x \subseteq \alpha : \sup x < \alpha\}$ is the same as saying $I \cap \mathcal{P}(\alpha) = \bigcup_{\beta < \alpha} I \cap \mathcal{P}(\beta)$, meaning that we don't allow any new supports at stage α .

^{xxv}We want *I* to be a non-principal ideal here because we need $\iota_{\alpha,\beta}(p)$ to be an element of $\mathbf{x}_{\xi<\beta}\dot{\mathbb{Q}}_{\xi}$ even when α is a successor. As we've set things up, successor stages always allow us to increase the support by one element. So when we restrict the support in the limit stage, we need the previous successor supports to be in *I*.

— 34C•5. Theorem (Support of Direct Limits) -

Let κ be an ordinal and $\mathbf{*}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ a non-trivial, κ -stage iteration with support in a non-principal ideal $I \subseteq \mathcal{O}(\kappa)$ or $\mathcal{O}(\kappa)$ itself. Therefore, for limit α , $\mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is the direct limit of previous iterations iff $I \cap \mathcal{O}(\alpha) = \bigcup_{\beta < \alpha} I \cap \mathcal{O}(\beta)$, i.e.

$$I \cap \mathcal{P}(\alpha) \subseteq \{x \subseteq \alpha : \sup x < \alpha\}.$$

Proof .:.

(\leftarrow) Let \mathbb{P} be such that there is an embedding $\iota'_{\beta} : \mathbf{*}_{\xi < \beta} \dot{\mathbb{Q}}_{\xi} \to \mathbb{P}$ for each $\beta < \alpha$ and these commute with the $\iota_{\beta,\gamma}$ s: $\iota'_{\beta} = \iota'_{\gamma} \circ \iota_{\beta,\gamma}$. We must show that there is an embedding from $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ to \mathbb{P} . Since each condition's support in the α th iteration is bounded, define $\beta_p = \sup\{\xi + 1 : \xi \in \operatorname{sprt}(p)\} < \alpha$. It follows that $p \upharpoonright \beta_p$ contains all of the "content" of p in that $\iota_{\beta_p,\alpha}(p \upharpoonright \beta_p) = p$. In particular, we can define $\iota'_{\alpha} : \mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} \to \mathbb{P}$ by $\iota'_{\alpha}(p) = \iota'_{\beta_p}(p \upharpoonright \beta_p)$.

To see that ι'_{α} is an embedding, it's clear that $p^* \leq_{\alpha} p$ implies $\operatorname{sprt}(p^*) \supseteq \operatorname{sprt}(p)$ and therefore $\beta_{p^*} \ge \beta_p$. In particular,

$$p^* \upharpoonright \beta_{p^*} \leq_{\beta_{p^*}} p \upharpoonright \beta_{p^*} = \iota_{\beta_p,\beta_{p^*}}(p \upharpoonright \beta_p)$$

and therefore applying the embedding ι'_{β_n*} , we get

$$\iota_{\alpha'}(p^*) = \iota'_{\beta_{p^*}}(p^* \upharpoonright \beta_{p^*}) \leq^{\mathbb{P}} \iota'_{\beta_{p^*}}(\iota_{\beta_p,\beta_{p^*}}(p \upharpoonright \beta_p)) = \iota'_{\beta_p}(p \upharpoonright \beta_p) = \iota_{\alpha'}(p).$$

We get injectivity similarly: if $p \neq q$, let $\beta_p \geq \beta_q$ for the sake of definiteness so by injectivity of the $\iota'_{\beta}s$,

$$\iota'_{\alpha}(p) = \iota'_{\beta_p}(p \upharpoonright \beta_p) \neq \iota'_{\beta_p}(q \upharpoonright \beta_p) = \iota'_{\beta_p}(\iota_{\beta_q,\beta_p}(q \upharpoonright \beta_q)) = \iota'_{\beta_q}(q \upharpoonright \beta_q) = \iota'_{\alpha}(q).$$

 $(\rightarrow) \text{ Any } p \in \mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} \text{ with unbounded support is not in the image of any } \iota_{\beta,\alpha} \text{ for a } \beta < \alpha \text{ and hence} \\ \mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} \neq \bigcup_{\beta < \alpha} \iota_{\beta,\alpha} " \mathbf{*}_{\xi < \beta} \dot{\mathbb{Q}}_{\xi} \text{ which is necessary for being the direct limit.} \qquad \neg$

This tells us that finite support takes limits at every stage since finite support is always bounded. This idea also allows us to mix things together, taking bounded support and hence direct limits just at certain stages. More generally, finite support iterations take direct limits at every limit stage, and are in fact the only kind of support that does this.

- 34 C • 6. Corollary

Let κ be an ordinal and $\mathbf{*}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ a κ -stage, finite support iteration. Therefore $\mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is the direct limit of the previous iterations for each limit α . Moreover, any non-trivial κ -length iteration that takes direct limits at every limit stage has finite support.

Proof .:.

Clearly finite support is bounded in every limit, so we always take the direct limit by Support of Direct Limits (34 C • 5). To see that finite support iterations are the only ones with this property, we proceed by induction on limit $\alpha < \kappa$ to show that all elements of $\mathbf{x}_{\boldsymbol{\xi} < \alpha} \dot{\mathbb{Q}}_{\boldsymbol{\xi}}$ have finite support.

For $\alpha = \omega$, this is clear: as the direct limit, support is bounded and therefore finite. Inductively, if every limit $\beta < \alpha$ takes finite support and support at the α the-stage is bounded, then support at the α th stage is finite. To see this, any $p \in \mathbf{x}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ has $p = \iota_{\beta,\alpha}(p \upharpoonright \beta)$ for some $\beta < \alpha$. But then the support of both is inductively finite (even if β is a successor, it's only finitely many iterations above a limit and so can only add at most finitely many non- $\dot{\mathbb{I}}$'s to p restricted to the largest limit below β): for I the support of the iteration, $\operatorname{sprt}(p) = \operatorname{sprt}(p \upharpoonright \beta) \in I \cap \mathcal{P}(\beta) \subseteq \{x \subseteq \beta : x \text{ is finite}\}$. Additionally, we must have $I \cap \mathcal{P}(\beta) \supseteq \{x \subseteq \beta : x \text{ is finite}\}$ because at successor stages of the iteration, we allow ourselves to extend the support by one element which eventually yields any finite number of elements.

In contrast to direct limits, we also have *inverse* limits, which is kind of the reverse of the direct limit in that we have projections from $\mathbf{x}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ to previous iterations and this is the "least" such in that any other model with this property

also has a "projection" onto $\mathbf{a}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$. In model theoretic terms, we have the following definition. Recall that a homomorphism is just an embedding without the requirement of injectivity, i.e. $f : A \to B$ has $R^{\mathbf{A}}(\vec{x}) \to R^{\mathbf{B}}(f(\vec{x}))$ and $f(g^{\mathbf{A}}(\vec{x})) = g^{\mathbf{B}}(f(\vec{x}))$ for any relation symbol 'R' and function symbol 'g'.

— 34C•7. Definition

Let \mathcal{A} be a set of FOL(σ)-models for some signature σ . Let \mathcal{F} be a set of homomorphisms between models of \mathcal{A} . $\langle \mathcal{A}, \mathcal{F} \rangle$ is a *projective system of homomorphisms* iff

- for each $A, B \in A$, there is at most one $f : A \to B$ in \mathcal{F} , denoted $f_{A,B}$ with $f_{A,A} = id \upharpoonright A$;
- for each $A, B \in \mathcal{A}$, there is some $C \in \mathcal{A}$ with $f_{C,A}, f_{C,B} \in \mathcal{F}$; and
- if $f_{A,B}$, $f_{B,C} \in \mathcal{F}$, then there is a homomorphism $f_{A,C} \in \mathcal{F}$ with $f_{B,C} \circ f_{A,B} = f_{A,C}$.

For $(\mathcal{A}, \mathcal{F})$ a projective system of homomorphisms, the *inverse limit* is the FOL (σ) -model inv $\lim_{\mathcal{F}} \mathcal{A}$ such that

- 1. there is a homomorphism $f_{\infty,A}$: inv $\lim_{\mathcal{F}} \mathcal{A} \to A$ such that $f_{A,B} \circ f_{\infty,A} = f_{\infty,B}$ whenever $f_{A,B}$ exists; and
- 2. for every M satisfying (1) in place of inv $\lim_{\mathcal{F}} \mathcal{A}$ with homomorphisms $f_{M,A} : M \to A$ for $A \in \mathcal{A}$, there is a unique homomorphism $f_{M,\infty} : M \to inv \lim_{\mathcal{F}} \mathcal{A}$ such that $f_{\infty,A} \circ f_{M,\infty} = f_{M,A}$ for all $A \in \mathcal{A}$;

The idea can be illustrated with a similar figure as with Figure 34 C • 2 just given by reversing the arrows. And similarly to the direct limit, we may view \mathcal{F} as instead the result of a relation where **A** *R* **B** iff $f_{A,B} \in \mathcal{F}$. The result for projective systems is that *R* must be directed from below as in any two **A**, **B** have an *R*-lower bound **C**. The difference from the direct limit is then merely in the direction of the arrows: the inverse limit is what everything going to the models of \mathcal{A} must go through first. In our case, the *R* is linear and in fact is just > on the length of the iteration: $\alpha > \beta$ implies there is a projection (a homomorphism) from $\mathbf{x}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ to $\mathbf{x}_{\xi < \beta} \dot{\mathbf{Q}}_{\xi}$.



34C•8. Figure: The inverse limit homomorphisms

Another thing to keep in mind is that while the direct limit is the result of embeddings, the inverse limit merely has homomorphisms which need not be injective. This allows the inverse limit to be "bigger" while still having these maps into the "smaller" models (and hence why these maps are frequently called "projections"). Like the direct limit, the inverse limit always exists:^{xxvi} Ensuring that the characteristic property holds is similar to forming the completion of a metric space: we take the product of all of the models, restrict to sequences we care about, and can project down by looking at the relevant value.

— 34 C • 9. Theorem -

Let $\langle \mathcal{A}, \mathcal{F} \rangle$ be a projective system of homomorphisms. Therefore the inverse limit inv $\lim_{\mathcal{F}} \mathcal{A}$ exists and both the inverse limit and the homomorphisms $f_{\infty,A}$: inv $\lim_{\mathcal{F}} \mathcal{A} \to A$ for $A \in \mathcal{A}$ are unique up to isomorphism.

Proof ...

Remember that any element of $\prod_{A \in A} A$ is a function from A to $\bigcup_{A \in A} A$ and thus has models as its inputs. Consider the model with universe

 $N = \left\{ x \in \prod_{\mathbf{A} \in \mathcal{A}} A : \forall \mathbf{A}, \mathbf{B} \in \mathcal{A} \ (f_{\mathbf{A}, \mathbf{B}} \in \mathcal{F} \to f_{\mathbf{A}, \mathbf{B}}(x(\mathbf{A})) = x(\mathbf{B})) \right\}.$

Relations, functions, and so forth are evaluated pointwise: for a relation symbol 'R', $R^{N}(\vec{x})$ iff for every $A \in A$,

^{xxvi}Of course, in many situations, the inverse limit won't be like the constituent models which is why in many category theory contexts they must state "if it exists" since the inverse limit may not be in the relevant category.

 $R^{\mathsf{A}}(\vec{x}(\mathsf{A}))$, and similarly $f^{\mathsf{N}}(\vec{x})$ is the function where each input $\mathsf{A} \in \mathcal{A}$ has the value $f^{\mathsf{A}}(\vec{x}(\mathsf{A}))$.

We pretty clearly have projections just as evaluations: $f_{N,A}(x) = x(A)$ for any $x \in N$. This is easily seen to be a homomorphism by definition of the relation and function interpretations of N. Similarly this works with the homomorphisms of \mathcal{F} by definition of N. Thus it suffices to show the "maximality" of N.

Let **M** be arbitrary such that there are homorphisms $f_{M,A} : M \to A$ for each $A \in A$ and these satisfy $f_{M,A} = f_{B,A} \circ f_{M,B}$ whenever $f_{B,A} \in \mathcal{F}$. We may define $f_{M,N} : M \to N$ as follows: $f_{M,N}(x)$ is the map $A \mapsto f_{M,A}(x)$. This is pretty clearly a homomorphism since if $R^{M}(\vec{x})$ then for each $A \in A$, $R^{A}(f_{M,A}(\vec{x}))$ iff $R^{N}(f_{M,N}(\vec{x}))$, and similarly for evaluating functions.

To show the uniqueness of $f_{M,N}$, let $g_{M,N} : M \to N$ also work with the $f_{N,A}s$. For any $x \in M$, $g_{M,N}(x) \in N$ is thus a function from A. But for all $A \in A$, since $f_{N,A} \circ g_{M,N} = f_{M,A}$,

 $g_{\mathsf{M},\mathsf{N}}(x)(\mathsf{A}) = f_{\mathsf{N},\mathsf{A}}(g_{\mathsf{M},\mathsf{N}}(x)) = f_{\mathsf{M},\mathsf{A}}(x) = (\mathsf{A} \mapsto f_{\mathsf{M},\mathsf{A}}(x))(\mathsf{A}) = f_{\mathsf{M},\mathsf{N}}(x)(\mathsf{A}).$

Hence $g_{M,N}(x) = f_{M,N}(x)$ for all $x \in M$ and so $f_{M,N}$ is unique. This establishes N as an inverse limit and thus the existence of inverse limits.

For uniqueness up to isomorphism, if \mathbf{M} , \mathbf{N} both satisfy the definition of being the inverse limit, then there are unique homomorphisms $f_{\mathbf{N},\mathbf{M}}: N \to M$ and $f_{\mathbf{M},\mathbf{N}}: M \to \mathbf{N}$ that both work with the other projections. But then $f_{\mathbf{N},\mathbf{M}} \circ f_{\mathbf{M},\mathbf{N}}: M \to M$ and $f_{\mathbf{M},\mathbf{N}} \circ f_{\mathbf{N},\mathbf{M}}: N \to N$ both work with the other projections:

 $f_{\mathsf{N},\mathsf{A}} \circ (f_{\mathsf{M},\mathsf{N}} \circ f_{\mathsf{N},\mathsf{M}}) = f_{\mathsf{M},\mathsf{A}} \circ f_{\mathsf{N},\mathsf{M}} = f_{\mathsf{N},\mathsf{A}},$

and similarly for **M**. Since the identity id $\upharpoonright N : N \to N$ is also a homomorphism from **N** to itself with this property, by the uniqueness of the homomorphism, it follows that $f_{M,N} \circ f_{N,M} = \text{id} \upharpoonright N$ and similarly $f_{N,M} \circ f_{M,N} = \text{id} \upharpoonright M$ implying that $f_{N,M} = f_{M,N}^{-1}$ are isomorphisms. \dashv

Returning to iterated forcing, this is brought up merely because full support iterations give the inverse limit at every limit stage. More generally, we get the following as a kind of complement to bounded support with Support of Direct Limits ($34C \cdot 5$): if we take all unbounded sets as allowable support at a limit stage α , then we get the inverse limit.

— 34 C • 10. Lemma

Let κ be an ordinal, $I \subseteq \mathcal{P}(\kappa)$, and $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ a κ -stage iteration with support in I. Therefore for each $\alpha < \beta \leq \kappa$, restriction $\pi_{\beta,\alpha} = p \mapsto p \upharpoonright \alpha$ is a homomorphism from $\mathbf{*}_{\xi < \beta} \dot{\mathbb{Q}}_{\xi}$ to $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$.

Proof .:.

 $\pi_{\beta,\alpha}$ is well-defined by Definition 34 B • 3 (5) and is a homomorphism by Definition 34 B • 3 (7) and (3). \dashv

Again, we simplify the terminology since we're only dealing with preorders here.

— 34C•11. Definition

Let κ be an ordinal and $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ a κ -stage iteration of preorders appropriate for forcing. We say $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ is the *inverse limit of the previous iterations* iff it's the inverse limit of the projective system of homomorphisms $\langle \{\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} : \alpha < \kappa \}, \{\pi_{\beta,\alpha} : \alpha \leq \beta < \kappa \} \rangle$ where $\pi_{\beta,\alpha} : \mathbf{*}_{\xi < \beta} \dot{\mathbb{Q}}_{\xi} \rightarrow \mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ is defined by $\pi_{\beta,\alpha}(p) = p \upharpoonright \alpha$.

In the spirit of Theorem 34 C • 9, for iterations we identify the universe of the inverse limit not as (writing \mathbb{P}_{α} for $\mathbf{a}_{\xi<\alpha}\dot{\mathbf{Q}}_{\xi}$ to save space)

$$N = \left\{ x \in \prod_{\alpha < \kappa} \mathbb{P}_{\alpha} : \forall \xi, \alpha < \kappa \ (\xi \le \alpha \to x(\mathbb{P}_{\alpha}) \upharpoonright \xi = x(\mathbb{P}_{\xi})) \right\}$$

but instead as

$$N' = \left\{ x \in \prod_{\alpha < \kappa} \operatorname{dom}(\dot{\mathbb{Q}}_{\alpha}) : \forall \alpha < \kappa \ (x \upharpoonright \alpha \in \mathbb{P}_{\alpha}) \right\}.$$

Here we basically identify $x \in \prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$ with the initial segments of a single function since they're all \subseteq -comparable anyway: $p = \bigcup_{\alpha < \kappa} x(\mathbb{P}_{\alpha})$ has $x = \langle p \upharpoonright \alpha : \alpha < \kappa \rangle$. It's also not hard to see that any $x \in \prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$ is the result of such a $p \in \prod_{\alpha < \kappa} \text{dom}(\dot{\mathbb{Q}}_{\alpha})$. In particular, if we want our iteration to be the inverse limit of previous iterations, we require the support to allow all of these kinds of limits.

— 34 C • 12. Theorem (Support of Inverse Limits)

Let κ be an ordinal, $I \subseteq \mathcal{P}(\kappa)$, and $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ a non-trivial κ -stage iteration with support in I. Let $\alpha < \kappa$ be a limit and I a non-principal ideal or $\mathcal{P}(\kappa)$ itself. Therefore, $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ is the inverse limit of previous iterations iff $\{x \in I \cap \mathcal{P}(\alpha) : \sup x = \alpha\} = \{x \in \mathcal{P}(\alpha) : \sup x = \alpha \land \forall \beta < \alpha \ (x \cap \beta \in I)\}.$

Proof .:.

(\leftarrow) Let \mathbb{P} be such that there is a homomorphism $\pi'_{\beta} : \mathbb{P} \to \bigstar_{\xi < \beta} \dot{\mathbb{Q}}_{\xi}$ for each $\beta < \alpha$ and these commute with the $\pi_{\beta,\gamma}$ s: for $\gamma \leq \beta < \alpha$, $\pi'_{\gamma} = \pi_{\beta,\gamma} \circ \pi_{\beta}$, meaning $\pi'_{\beta}(p) \upharpoonright \gamma = \pi'_{\gamma}(p)$. We must show that there is a homomorphism π'_{α} from \mathbb{P} to $\bigstar_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$. Since the π'_{β} s play nicely with the $\pi_{\beta,\gamma}$ s, for each $p \in \mathbb{P}$, define

$$\pi'_{\alpha} = \bigcup_{\beta < \alpha} \pi'_{\beta}(p).$$

- Claim 1 For each $p \in \mathbb{P}$, $\pi'_{\alpha}(p) \in \mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$. Moreover, $\pi'_{\alpha}(p) \upharpoonright \beta = \pi'_{\beta}(p)$ for $\beta \le \alpha$, meaning $\pi_{\alpha,\beta} \circ \pi'_{\alpha} = \pi'_{\beta}$.

Proof .:.

That $\bigcup_{\beta < \alpha} \pi'_{\beta}(p)$ is a relation with domain α is immediate. Suppose $\langle \beta, y_0 \rangle, \langle \beta, y_1 \rangle \in \pi'_{\alpha}(p)$ for some $\beta < \alpha$, so there are $\beta < \beta_0 \le \beta_1$ with $\langle \beta, y_0 \rangle \in \pi'_{\beta_0}(p)$ and $\langle \beta, y_1 \rangle \in \pi'_{\beta_1}(p)$. Thus

$$\langle \beta, y_0 \rangle, \langle \beta, y_1 \rangle \in \pi'_{\beta_0}(p) \cup \pi'_{\beta_1}(p) \upharpoonright \beta_0 = \pi'_{\beta_0}(p) \cup \pi'_{\beta_0}(p) = \pi'_{\beta_0}(p).$$

Since both $\pi'_{\beta_0}(p)$ is a function, $y_0 = y_1$ and therefore $\pi'_{\alpha}(p) \in \prod_{\xi < \alpha} \operatorname{dom}(\dot{\mathbb{Q}}_{\xi})$. Similar arguments also show that $\pi_{\alpha,\beta}(\pi'_{\alpha}(p)) = \pi'_{\beta}(p)$.

So to see that $\pi'_{\alpha}(p) \in \bigstar_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$, we need $\operatorname{sprt}(\pi'_{\alpha}(p)) \in I$. If $I = \mathfrak{O}(\kappa)$ or $\operatorname{sprt}(\pi'_{\alpha}(p))$ is unbounded in α , this is clear. Otherwise $\operatorname{sprt}(\pi'_{\alpha}(p)) \subseteq \beta$ for some $\beta < \alpha$, so $\beta = \beta^* + n$ for some limit ordinal β^* and $n \in \omega$. In particular, $\operatorname{sprt}(\pi'_{\alpha}(p) \upharpoonright \beta^*) \in I$ and $\operatorname{sprt}(\pi'_{\alpha}(p) \upharpoonright \beta) \setminus \operatorname{sprt}(\pi'_{\alpha}(p) \upharpoonright \beta^*)$ is finite and therefore $\operatorname{sprt}(\pi'_{\alpha}(p)) = \operatorname{sprt}(\pi'_{\alpha}(p) \upharpoonright \beta) \in I$ as a non-principal ideal. \dashv

Because each π'_{β} is a homomorphism, we get that π'_{α} is too:

$$\pi'_{\alpha}(\mathbb{1}^{\mathbb{P}}) = \bigcup_{\beta < \alpha} \pi'_{\beta}(\mathbb{1}^{\mathbb{P}}) = \bigcup_{\beta < \alpha} \langle \dot{\mathbb{1}}'_{\xi} : \xi < \beta \rangle = \langle \dot{\mathbb{1}}'_{\xi} : \xi < \alpha \rangle = \mathbb{1}_{\alpha}$$
$$p \leq^{\mathbb{P}} q \to \forall \beta < \alpha \ \left(\pi'_{\alpha}(p) \upharpoonright \beta = \pi'_{\beta}(p) \leq_{\beta} \pi'_{\beta}(q) = \pi'_{\alpha}(q) \upharpoonright \beta\right) \to \pi'_{\alpha}(p) \leq_{\alpha} \pi'_{\alpha}(q).$$

It's also not difficult to see that π'_{α} is the unique homomorphism in working with these π'_{ξ} for $\xi < \alpha$, precisely because as the restrictions of π'_{α} , it needs to agree with all of them anyway.

- (\rightarrow) Clearly any unbounded subset $x \in I \cap \mathcal{P}(\alpha)$ has $x \cap \beta \in I$ for each $\beta < \alpha$. So suppose $x \cap \beta \in I$ for each $\beta < \alpha$. We construct a $p \in \prod_{\xi < \alpha} \text{dom}(\dot{\mathbb{Q}}_{\xi})$ by transfintie recursion. Set $p_0 = \emptyset$.
 - For $\beta + 1 < \alpha$, if $\beta \in x$ then $p_{\beta+1} = p_{\beta} \cup \{\langle \beta, \dot{q} \rangle\}$ for any $\dot{q} \in \text{dom}(\dot{\mathbb{Q}}_{\beta})$ where $\mathbb{1}_{\beta} \not\Vdash ``\dot{q} = \dot{\mathbb{1}}'_{\beta}$. Otherwise set $p_{\beta+1} = p_{\beta} \cup \{\langle \beta, \dot{\mathbb{1}}'_{\beta} \rangle\}$.
 - At limit stage $\beta < \alpha$, since $x \cap \beta \in I$, we can take $p_{\beta} = \bigcup_{\xi < \beta} p_{\xi}$.

It follows that $p = \bigcup_{\beta < \alpha} p_{\beta}$ has $p \upharpoonright \beta \in \mathbf{*}_{\xi < \beta} \dot{\mathbb{Q}}_{\xi}$ for each $\beta < \alpha$. In particular, p must be in the inverse limit (as in the remark above the theorem statement) and therefore have its support in I. \dashv

As full support obviously has this property, we get inverse limits at every limit stage. And this is more-or-less the only support with this property (certainly the only non-principal ideal with this property).

— 34 C • 13. Corollary —

Let κ be an ordinal and $\mathbf{a}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ a κ -stage, full support iteration. Therefore $\mathbf{a}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is the inverse limit of the previous iterations for each limit α . Moreover, any non-trivial κ -length iteration with support I satisfying

• $\{\alpha\} \in I$ for every $\alpha < \kappa$; and

• $x, y \in I$ implies $x \cup y \in I$;

that takes inverse limits at every limit stage has full support.

Proof .:.

Full support has this property by Support of Inverse Limits (34 C • 12). So suppose we have an iteration as in the statement taking inverse limits at every limit stage. Since *I* is closed under finite unions and $\{\alpha\} \in I$ for every $\alpha < \kappa$, all finite subsets of κ are in *I*. It follows that

 $\{x \in I \cap \mathcal{P}(\omega) : \sup x = \omega\} = \{x \in \mathcal{P}(\omega) : \sup x = \omega \land \forall n < \omega \ (x \cap n \in I)\},\$

meaning that all infinite subsets of ω are in I in addition to all the finite subsets: $I \cap \mathcal{P}(\omega) = \mathcal{P}(\omega)$. Inductively, suppose we have full support at all previous limit stages $< \alpha$, i.e. $I \cap \mathcal{P}(\beta) = \mathcal{P}(\beta)$ for all limit $\beta < \alpha$. It follows that $I \cap \mathcal{P}(\beta) = \mathcal{P}(\beta)$ for all $\beta < \alpha$ including successors, since any successor is at most finitely many ordinals away from a limit: $\beta = \beta^* + n$ for some $n < \omega$ and limit β^* where then, since I is closed under finite unions and contains every finite subset of κ , $\mathcal{P}(\beta) = \{x \cup y : x \subseteq \beta^* \land y \subseteq \beta \subseteq \kappa \text{ is finite}\} \subseteq I$. Since we take the inverse limit at stage $\alpha \le \kappa$ and all bounded subsets of α are in I, it follows as with ω that all bounded and unbounded subsets of α are in $I : I \cap \mathcal{P}(\alpha) = \mathcal{P}(\alpha)$ and therefore $I \cap \mathcal{P}(\kappa) = \mathcal{P}(\kappa)$.

It's still possible to take the inverse limit at every stage without full support if we allow trivial $\hat{Q}_{\alpha}s$ at certain stages $\alpha < \kappa$ and remove elements of *I* that contain α . But these are somewhat *ad hoc*. Natural examples will have *I* be a non-principal ideal or $\mathcal{P}(\kappa)$ itself and such sets are forced to be $\mathcal{P}(\kappa)$ when taking inverse limits at every stage.

The main idea behind this subsection is partly to motivate more complicated supports. Elsewhere in the literature, supports might be defined by where direct limits or inverse limits are taken. Support of Inverse Limits $(34 \text{ C} \cdot 12)$ and Support of Direct Limits $(34 \text{ C} \cdot 5)$ tell us how to translate this in terms of support. This translation is more-orless unnecessary, and it's more intuitive to think of the preorders in terms of their structure and how they relate to previous iterations. The more important translation is the reverse: taking things in terms of support and translating this to understand the iterations.

For example, easton support (sometimes called reverse easton support) is usually defined by

$$I = \{ X \subseteq \kappa : \forall \delta \le \kappa \; (\operatorname{cof}(\delta) = \delta = \aleph_{\delta} \to |X \cap \delta| < \delta) \}.$$

This is hard to unpack on its surface, but if we think about what's happening at each limit stage, we're essentially just taking bounded support at regular, limit cardinal (i.e. weakly inaccessible) stages and full support elsewhere.^{xxvii} In other words, to define the iteration preorder, we take direct limits of previous iterations at weakly inaccessible stages and inverse limits everywhere else. This characterization is much more intelligible when trying to gain an intuition.^{xxviii}

Now we can think about what happens with the generic extensions depending on their support. One might expect that if we take the direct limit of previous iterations, then the generic extension is the direct limit of the previous generic extensions. This is actually always false (if we're iterating non-trivial preorders): if G is $\mathbf{x}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ -generic over V and $\mathbf{x}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ is the direct limit of the previous iterations, then $\{G \upharpoonright \alpha : \alpha < \kappa\}$ is in V[G] but isn't in any of the $V[G \upharpoonright \alpha]$ s and hence isn't in their direct limit. So what good is it to know whether an iteration is the direct or inverse limit of previous iterations? Mostly we will be concerned about what properties are preserved by iterations and not about the generic extension so directly.

For example, finite support iterations of ccc preorders are themselves ccc and hence preserve cardinals and cofinalities, and this is mostly a result of The Δ -System Lemma (32 D • 2), which is often used in iterated forcing to work with

xxvii Typically, one forces with trivial preorders at non-regular cardinal stages, so often in the literature we just consider subsets consisting of regular cardinals.

^{xxviii}Easton support is useful in many ways, usually in showing that the continuum function $\kappa \mapsto 2^{\kappa}$ for regular κ can be *any* function that is \leq -increasing and obeys $\kappa < \operatorname{cof}(2^{\kappa})$, although this is done with product forcing.

supports.

— 34 C • 14. Result

Let κ be an ordinal. Let $\mathbf{a}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ be a κ -stage, finite support iteration such that for all $\xi < \alpha \leq \kappa$, $\mathbb{1}_{\alpha} \Vdash "\dot{\mathbf{Q}}_{\xi}$ is ccc". Therefore $\mathbf{a}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is ccc for every $\alpha \leq \kappa$.

Proof .:.

Proceed by induction on $\alpha \leq \kappa$. Clearly for $\alpha = 0$, $\mathbf{x}_{\xi<0}\dot{\mathbf{Q}}_{\xi} = 1$ is ccc. Inductively, the successor case follows by Lemma 34 A • 7: $\mathbf{x}_{\xi<\alpha+1}\dot{\mathbf{Q}}_{\xi} \cong (\mathbf{x}_{\xi<\alpha}\dot{\mathbf{Q}}_{\xi}) \ast \dot{\mathbf{Q}}_{\alpha}$ where $\mathbf{x}_{\xi<\alpha}\dot{\mathbf{Q}}_{\xi}$ is inductively ccc, a finite support iteration, and $\mathbb{I}_{\alpha} \Vdash \ddot{\mathbf{Q}}_{\alpha}$ is ccc".

So suppose α is a limit but the result fails: $\mathcal{A} \subseteq \bigstar_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ is an uncountable antichain. By The Δ -System Lemma (32 D • 2), since all supports are finite, we may assume all the conditions of \mathcal{A} have the same intersection: for $p, q \in A$, $\operatorname{sprt}(p) \cap \operatorname{sprt}(q) = r$ for some finite $r \subseteq \alpha$. Since any two elements of A are incompatible, they must disagree somewhere in their shared support r. But then restricting to this area yields $\{p \upharpoonright (1 + \max r) : p \in \mathcal{A}\}$ as an uncountable antichain of $\bigstar_{\xi < 1 + \max r} \dot{\mathbb{Q}}_{\xi}$, contradicting the inductive hypothesis.

This proof also generalizes to finite support iterations of (names for) λ -cc preorders. Now although finite support iterations, which take direct limits at every limit stage, have this property, we can actually weaken this hypothesis quite a bit. The only restriction we actually need is when the entire length of the iteration has cofinality λ , in which case we only require the direct limit to be taken stationarily many times.

- 34 C • 15. Theorem (Direct Limit Chain Conditions)

Let κ be a limit ordinal and λ a cardinal. Let $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\kappa}$ be a κ -length iteration with support in some $I \subseteq \mathcal{P}(\kappa)$ an ideal or $\mathcal{P}(\kappa)$ itself. Suppose

- 1. $\mathbf{*}_{\alpha < \kappa} \hat{\mathbb{Q}}_{\alpha}$ is the direct limit of previous iterations;
- 2. $\mathbf{*}_{\boldsymbol{\xi} < \boldsymbol{\alpha}} \dot{\mathbf{Q}}_{\boldsymbol{\xi}}$ is λ -cc for each $\boldsymbol{\alpha} < \kappa$; and

3. $cof(\kappa) = \lambda$ implies { $\alpha < \kappa : \mathbf{k}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is the direct limit of previous iterations} is stationary in κ .

Therefore $\mathbf{a}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ is λ -cc.

Proof .:.

Let $\mathcal{A} \subseteq \bigstar_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ be an antichain and assume without loss of generality for the sake of contradiction that $|\mathcal{A}| = \lambda$. If $cof(\kappa) \neq \lambda$, we may proceed in much the same way: note that

$$\mathcal{A} = \bigcup_{\alpha < \kappa} \{ p \in \mathcal{A} : \operatorname{sprt}(p) \subseteq \alpha \} = \bigcup_{\alpha < \operatorname{cof}(\kappa)} \{ p \in \mathcal{A} : \operatorname{sprt}(p) \subseteq \gamma_{\alpha} \},$$

where $\langle \gamma_{\alpha} : \alpha < \operatorname{cof}(\kappa) \rangle$ is a cofinal sequence in κ .

- If cof(κ) < λ, then one of these sets must have size |A| = λ, meaning for some γ < κ, {p ∈ A : sprt(p) ⊆ γ} has size λ. But then {p ↾ γ : p ∈ A∧sprt(p) ⊆ γ} is a λ-sized antichain of 𝐾_{α<γ} Q_α, contradicting (2). One may check that this is an antichain by the almost-complete embedding ι_{γ,κ} : 𝐾_{ξ<γ} Q_ξ → 𝐾_{α<κ} Q_α as per Lemma 34 C 3.
- Similarly, if cof(κ) > λ, then sup{sup sprt(p) : p ∈ A} < κ. To see this, as the direct limit of previous iterations, all support is bounded in κ. Since there are only λ < cof(κ)-many such supports given from elements of A, sup{sup sprt(p) : p ∈ A} is also bounded above by some γ < κ and hence we again have {p ↾ γ : p ∈ A ∧ sprt(p) ⊆ γ} as a λ-sized antichain of *****_{α<γ} Q

 ^ˆ_α, contradicting (2).

So suppose $\operatorname{cof}(\kappa) = \lambda$ as witnessed by a sequence $\vec{\gamma}$, and enumerate $\mathcal{A} = \{p_{\alpha} : \alpha < \lambda\}$. Assume $\vec{\gamma} = \langle \gamma_{\alpha} : \alpha < \lambda \rangle$ is continuous in the sense that $\sup_{\xi < \alpha} \gamma_{\xi} = \gamma_{\alpha}$ for any limit $\alpha < \lambda$. As a result, $\{\gamma_{\alpha} : \alpha < \lambda\}$ is club in κ . By hypothesis, $S = \{\alpha < \kappa : \mathbf{x}_{\xi < \alpha} \, \dot{\mathbf{Q}}_{\xi} \text{ is the direct limit of previous iterations}\}$ is stationary and hence

 $S \cap \vec{\gamma} = \left\{ \gamma_{\alpha} < \lambda : \bigstar_{\xi < \gamma_{\alpha}} \dot{\mathbb{Q}}_{\xi} \text{ is the direct limit of previous iterations} \right\}$

is stationary. It's not hard to see that then $S_0 = \{ \alpha < \lambda : \gamma_\alpha \in S \cap \vec{\gamma} \}$ is stationary in λ (any club of λ is

ITERATED FORCING

transformed to one of κ via $\alpha \mapsto \gamma_{\alpha}$ and hence intersects $S \cap \vec{\gamma}$ so the pull-back intersects S_0). This is nice because λ is regular and we can then use Fodor's Lemma (11 B • 5). In particular, define $f : S_0 \to \lambda$ by taking

$$f(\alpha)$$
 = the least ξ where sprt $(p_{\alpha} \upharpoonright \gamma_{\alpha}) \subseteq \gamma_{\xi}$.

This is regressive on S_0 since any $\alpha \in S_0$ has $p_\alpha \upharpoonright \gamma_\alpha \in \mathbf{*}_{\xi < \gamma_\alpha} \dot{\mathbb{Q}}_{\gamma_\alpha}$ which is the direct limit, meaning its support is bounded in γ_α . Therefore, there is some $\alpha^* < \lambda$ and stationary $S_1 \subseteq S_0$ where $f''S_1 = \{\alpha^*\}$. Since λ is regular, the γ_α s are cofinal in κ , and all supports are bounded in κ anyway, we take another λ -sized subset S_2 —not caring whether it's stationary or not—where $\operatorname{sprt}(p_\alpha) \subseteq \gamma_\beta$ whenever $\alpha < \beta$ and $\alpha, \beta \in S_2$. Since $\mathbf{*}_{\xi < \gamma_{\alpha^*}} \dot{\mathbb{Q}}_{\xi}$ is λ -cc, $\{p_\beta \upharpoonright \gamma_{\alpha^*} : \in S_2\}$ is not an antichain. And so $p_{\xi_0} \upharpoonright \gamma_{\alpha^*}$ and $p_{\xi_1} \upharpoonright \gamma_{\alpha^*}$ are compatible for some $\xi_0, \xi_1 \in S_2$ as witnessed by some $p^* \leq_{\gamma_{\alpha^*}} p_{\xi_0} \upharpoonright \gamma_{\alpha^*}, p_{\xi_1} \upharpoonright \gamma_{\alpha^*}$. But then we can extend to an extension of p_{ξ_0} and p_{ξ_1} . Explicitly, assume for the sake of definiteness $\xi_0 < \xi_1$ so that $\operatorname{sprt}(p_{\xi_0}) \subseteq \gamma_{\xi_1}$. Now consider

$$p^{**} = p^{*} (p_{\xi_0} \upharpoonright (\gamma_{\xi_1} \setminus \gamma_{\alpha^*})) (p_{\xi_1} \upharpoonright (\kappa \setminus \gamma_{\xi_1})).$$

It follows that $\operatorname{sprt}(p^{**}) \in I$ as I is closed under finite unions. Moreover, $p^{**} \leq_{\kappa} p_{\xi_0}, p_{\xi_1}$, contradicting that $p_{\xi_0}, p_{\xi_1} \in A$ and A is an antichain.

This can be quite useful when used in conjunction with combinatorial properties of certain large cardinals. For example, this gives the following, which is a nice result useful in certain contexts, basically showing that we don't collapse cardinals $\geq \lambda$ (λ -cc preorders have this property) when using long iterations of small forcings.

- 34 C • 16. Corollary

Let κ be a cardinal and $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ a κ -stage iteration taking direct limits at strongly inaccessible stages. Suppose $|\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}| < \kappa$ whenever $\alpha < \kappa$ and $\{\lambda < \kappa : \lambda \text{ is strongly inaccessible}\}$ is stationary (i.e. κ is *mahlo*). Therefore $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ is κ -cc.

Inverse limits do *not* perserve the ccc-ness of previous iterations. The idea why is the same reason why ${}^{\omega}2$ —the inverse limit of $\langle {}^{n}2, \pi_{n,m} : m \leq n < \omega \rangle$ (with $\pi_{n,m}(\tau) = \tau \upharpoonright m$)—is uncountable but ${}^{<\omega}2$ —the direct limit of $\langle {}^{n}2, \iota_{n,m} : n \leq m < \omega \rangle$ (with $\iota_{n,m}(\tau) = (\tau \frown \text{ const } 0) \upharpoonright m$)—is countable. More explicitly, for any infinite support sprt $(p) \in I$ with $\sup(p) = \alpha$ a limit, we can continually extend $\iota_{\xi,\alpha}(p \upharpoonright \xi)$ in sprt $(p) \setminus \xi$ to incompatible elements and so embed the infinite binary tree into the iteration. The branches of these incompatible elements yield an antichain in the inverse limit, and since we can identify these elements as branches of ${}^{<\omega}2$, we get 2^{\aleph_0} -many elements in our antichain.

§ 34 D. Canonical names

Moving on from Direct Limit Chain Conditions $(34 \text{ C} \cdot 15)$ and Result $34 \text{ C} \cdot 14$, we would like to establish similar results for $< \kappa$ -closed preorders: that the iteration of $< \kappa$ -closed preorders is $< \kappa$ -closed. The issue with this occurs even at two-stage iterations, but it was not brought up there in order to avoid overloading the already very technical definitions earlier in Subsection 34 B.

In particular, even if \mathbb{P} is $< \kappa$ -closed and $\mathbb{1}^{\mathbb{P}} \Vdash ``\dot{\mathbb{Q}}$ is $< \check{\kappa}$ -closed", we may not have that $\mathbb{P} * \dot{\mathbb{Q}}$ is $< \kappa$ -closed. To see the issue, if we have a $\gamma < \kappa$ -length sequence of pairs $\langle \langle p_{\alpha}, \dot{q}_{\alpha} \rangle : \alpha < \gamma \rangle$, we have some condition $p^* \leq^{\mathbb{P}} p_{\alpha}$ for every $\alpha < \gamma$, and we know $p^* \Vdash ``\exists q \ \forall \alpha < \check{\gamma} \ (q \leq^{\dot{\mathbb{Q}}} \dot{q}_{\alpha})``$. So ostensibly there is some \mathbb{P} -name \dot{q}^* witnessing this and therefore $\langle p^*, \dot{q}^* \rangle \leq \langle p_{\alpha}, \dot{q}_{\alpha} \rangle$ for each $\alpha < \gamma$. The issue is that \dot{q}^* may not be in dom($\dot{\mathbb{Q}}$) and hence $\langle p^*, \dot{q}^* \rangle$ may not be in $\mathbb{P} * \dot{\mathbb{Q}}$. This issue, we need to ensure that anything forced to be a member of $\dot{\mathbb{Q}}$ is actually forced to be equal to a particular member of $\dot{\mathbb{Q}}$. This amounts to choosing the names in dom($\dot{\mathbb{Q}}$) more carefully and one common way to do this is with so-called *canonical* names.

^{xxix}That said, we can get an extension $p^{**} \leq^{\mathbb{P}} p^*$ and $\tau \in \operatorname{dom}(\dot{\mathbb{Q}})$ where $p^{**} \Vdash ``\tau = \dot{q}^{**}$ and so $\langle p^{**}, \tau \rangle \in \mathbb{P} * \dot{\mathbb{Q}}$. To see this, in any generic extension V[G] with $p \in G$, $V[G] \models ``\exists q \in \dot{\mathbb{Q}}_G$ below every $(\dot{q}_{\alpha})_G$ and therefore some τ_G witnesses this for $\tau \in \operatorname{dom}(\dot{\mathbb{Q}})$. Some $p^{**} \in G$ then forces this and we may take $p^{**} \leq p^*$ by compatibility and arrive at $\langle p^{**}, \tau \rangle$ as desired. This idea certainly works in this case, but the added argumentation may not hold in general if we require $\mathbb{1}^{\mathbb{P}}$ to force something, and problems may occur with longer iterations with potentially infinite supports.

- 34 D • 1. Definition

Let \mathbb{P} be a preorder. Let τ be a \mathbb{P} -name. We say τ is *canonical* iff it has minimal $|\operatorname{trcl}(\tau)|$ in the sense that for every \mathbb{P} -name σ , if $\mathbb{1}^{\mathbb{P}} \Vdash "\sigma = \tau$ " then $|\operatorname{trcl}(\tau)| < |\operatorname{trcl}(\sigma)|$.

If $\dot{\mathbb{Q}} = \langle \langle \dot{\mathbb{Q}}, \leq \dot{\mathbb{Q}}, \dot{\mathbb{I}}^{\dot{\mathbb{Q}}} \rangle \rangle$ is a P-name for a preorder, we *redefine* $\mathbb{P} * \dot{\mathbb{Q}}$ to be the preporder with universe

$$\{\langle p, \dot{q} \rangle : p \in \mathbb{P} \land \dot{q} \text{ is a canonical name } \land \mathbb{1}^{\mathbb{P}} \Vdash ``\dot{q} \in \dot{\mathbb{Q}}"\}$$

with the same ordering and maximal element definitions as before:

 $\langle p^*, \dot{q}^* \rangle \leq \langle p, \dot{q} \rangle$ iff $p^* \leq^{\mathbb{P}} p$ and $p^* \Vdash ``\dot{q}^* = \tau^* \land \dot{q} = \tau \land \tau^* \leq^{\dot{\mathbb{Q}}} \tau$.

Note that this is in line with the original notion of a canonical name for an element of the ground model as per Definition 31 A • 4. There is a slight caveat that $\mathbb{1}^{\mathbb{P}}$ must have small rank, but this poses little difficulty since we really only care about the structure of \mathbb{P} , and we can easily take an isomorphism to reduce the rank of $\mathbb{1}^{\mathbb{P}}$. For iterations this will pose no problem since we can just take the maximal element of each preorder to have minimal rank to get that the maximal element of the iteration has minimal rank.

- 34 D • 2. Corollary

For \mathbb{P} a preorder appropriate for forcing, and assume that $\mathbb{1}^{\mathbb{P}}$ has minimal transitive closure cardinality among the elements of \mathbb{P} . Therefore, for any x, the *check-name* defined iteratively by $\check{x} = \{\langle \check{y}, \mathbb{1}^{\mathbb{P}} \rangle : y \in x\}$ is canonical, and in fact $|\operatorname{trcl}(\check{x})|$ is minimal among all τ where $\exists p \in \mathbb{P} \ (p \Vdash ``\check{x} \subseteq \tau")$.

Proof .:.

For any τ , $|\operatorname{trcl}(\tau)| = \sum_{y \in \tau} |\operatorname{trcl}(\{y\})| = |\tau| \cdot \sup\{|\operatorname{trcl}(y)| + 1 : y \in \tau\}$. It suffices to show

 $|\operatorname{trcl}(\check{x})|$ is minimal among \mathbb{P} -names forced by any element of \mathbb{P} to contain \check{x} . (*)

Proceed by induction on rank. For $x = \emptyset$, $\mathring{\emptyset} = \emptyset$ is clearly satisfies (*). Suppose \check{y} satisfies (*) in place of \check{x} for each $y \in x$. Suppose $p \Vdash ``\check{x} \subseteq \tau$ '' for some \mathbb{P} -name τ and $p \in \mathbb{P}$. Consider

 $\tau' = \{ \langle \zeta, q \rangle \in \tau : \exists p^* \leq^{\mathbb{P}} p \exists y \in x \ (p^* \Vdash ``\zeta = \check{y}") \},\$

which basically consists of everything in τ which might be in x when looking below p. By (*) for each \check{y} and the minimality of $|\operatorname{trcl}(\mathbb{1}^{\mathbb{P}})|$, we have $|\operatorname{trcl}(\langle \check{y}, \mathbb{1}^{\mathbb{P}}\rangle)| \leq |\operatorname{trcl}(\langle \zeta, q \rangle)|$ whenever there's a $p^* \leq \mathbb{P} p$ such that $p^* \Vdash c_{\zeta} = \check{y}^{"}$. In particular, as every $y \in x$ has such a $\zeta \in \text{dom}(\tau')$, it follows that $|\check{x}| = |x| \le |\tau'| \le |\tau|$ and

 $|\operatorname{trcl}(\check{x})| = |\check{x}| \cdot \sup\{|\operatorname{trcl}(\check{y}, \mathbb{1}^{\mathbb{P}})| + 1y \in x\} \le |\tau'| \cdot \sup\{|\operatorname{trcl}(\langle \zeta, q \rangle)| + 1 : \langle \zeta, q \rangle \in \tau'\} \le |\operatorname{trcl}(\tau)|.$ By induction the result holds. -

Again, this was just a motivating result. The important property of canonical names is that they are (in principle) easy to find. In particular, if we have a (name for) a non-empty set, we can find a canonical name forced to be in it. This is useful in iterations $\mathbb{P} * \dot{\mathbb{Q}}$ because properties of elements of $\dot{\mathbb{Q}}$ can be translated to names forced to have those properties.

- 34 D • 3. Lemma (Canonical Name Search)

Let \mathbb{P} be a preorder. Let τ , $\dot{\mathbb{Q}}$ be \mathbb{P} -names such that $p \Vdash ``\tau \in \dot{\mathbb{Q}}$ '' for some $p \in \mathbb{P}$. Therefore there is a canonical name σ such that $p \Vdash "\sigma = \tau$ " and $\mathbb{1}^{\mathbb{P}} \Vdash "\dot{\mathbb{Q}} \neq \emptyset \rightarrow \sigma \in \dot{\mathbb{Q}}$ ".

Proof .:.

Using the technique of Conditional Name Lemma (34 A
$$\cdot$$
 1), we know that there is some ζ such that

$$\mathbb{I}^{\mathbb{P}} \Vdash ``\dot{\mathbb{Q}} \neq \emptyset \to \varsigma \in \dot{\mathbb{Q}}''.$$
(*)

Taking such a name ς , let ς' be a nice name for ς . Explicitly, for each $\rho \in \text{dom}(\varsigma)$, take \mathcal{A}_{ρ} to be a maximal antichain contained in $\{q \perp p : q \Vdash ``\rho \in \varsigma``\}$. Then we set

$$\varsigma' = \{ \langle \rho, q \rangle : q \leq p \land q \Vdash "\rho \in \tau" \} \cup \bigcup_{\rho \in \operatorname{dom}(\varsigma)} \{ \rho \} \times \mathcal{A}_{\rho}.$$

Clearly $p \Vdash \varsigma' = \tau$ and it's not hard to see that (*) holds of ς' in place of ς . But all of this is just to say that

ITERATED FORCING

 \neg

the following set is non-empty:

 $\left\{ |\operatorname{trcl}(\sigma)| : \sigma \in \mathcal{H}_{|\operatorname{trcl}(\tau)|^+} \land p \Vdash ``\tau = \sigma" \land \mathbb{1}^{\mathbb{P}} \Vdash ``\dot{\mathbb{Q}} \neq \emptyset \to \sigma \in \dot{\mathbb{Q}}" \right\}.$

So if we take σ witnessing the minimum value of this, we get σ as canonical.

In particular, since we'll be dealing with non-empty sets, assuming the hypotheses above, we get a canonical name σ which is always forced to be in $\dot{\mathbb{Q}}$ and $p \Vdash \sigma = \tau$ ". This lemma basically says that it suffices to use canonical names when searching for elements in iterated forcing.

- 34 D • 4. Definition

For κ an ordinal, $I \subseteq \mathcal{O}(\kappa)$, we *redefine* a κ -stage iteration with supports in I to be two sequences $\langle \mathbb{P}_{\alpha} : \alpha \leq \kappa \rangle$, $\langle \dot{\mathbb{Q}}_{\alpha} : \alpha < \kappa \rangle$ that satisfy all the requirements as before but where elements of \mathbb{P}_{α} are instead sequences of canonical names: $p \in \mathbb{P}_{\alpha}$ has $p(\xi)$ as a canonical name with $\mathbb{I} \Vdash "p(\xi) \in \dot{\mathbb{Q}}_{\xi}$ ".

Making these changes doesn't actually change any of the above results, since if \dot{Q} is a P-name for a preorder, then we can form another preorder \dot{Q}' such that $\mathbb{P} * \dot{Q}'$ in the old sense is both $\mathbb{P} * \dot{Q}'$ and $\mathbb{P} * \dot{Q}$ in the new sense, and a similar result holds for longer iterations.

— 34 D • 5. Result

Let $\mathbb{P} = \langle \mathbb{P}, \leq, 1 \rangle$ be a preorder and $\dot{\mathbb{Q}}_0 = \langle \! \langle \dot{\mathbb{Q}}_0, \leq_0, \dot{\mathbb{I}}_0 \rangle \! \rangle$ a \mathbb{P} -name for a preorder. Therefore there is a \mathbb{P} -name for a preorder $\dot{\mathbb{Q}}_1$ such that

1. $\mathbb{P} * \dot{\mathbb{Q}}_0 = \mathbb{P} * \dot{\mathbb{Q}}_1$; and

- 2. $\mathbb{P} * \dot{\mathbb{Q}}_1$ is the same as $\mathbb{P} * \dot{\mathbb{Q}}_1$ as in the old sense.
- 3. $\mathbb{P} * \dot{\mathbb{Q}}_0$ in the old sense is forcing equivalent to $\mathbb{P} * \dot{\mathbb{Q}}_1$.

Proof .:.

1. We merely collect the canonical names together, taking

 $\dot{\mathbb{Q}}_1 = \{ \langle \tau, \mathbb{1}^{\mathbb{P}} \rangle : \tau \text{ is canonical} \land \mathbb{1}^{\mathbb{P}} \Vdash ``\tau \in \dot{\mathbb{Q}}_0" \},$

with order $\leq_1 = \leq_0$ To show that the two iterations (in the new sense) are the same, we must show their underlying sets are the same and that their orderings are the same. To show that $\mathbb{P} * \dot{\mathbb{Q}}_1 = \mathbb{P} * \dot{\mathbb{Q}}_0$, we really just need that

 $\operatorname{dom}(\dot{\mathbb{Q}}_1) = \{\tau : \tau \text{ is canonical} \land \mathbb{1}^{\mathbb{P}} \Vdash ``\tau \in \dot{\mathbb{Q}}_1"\}.$

- $(\subseteq) \text{ If } \tau \in \operatorname{dom}(\dot{\mathbb{Q}}_1) \text{ then } \langle \tau, \mathbb{1}^{\mathbb{P}} \rangle \in \dot{\mathbb{Q}}_1 \text{ with } \tau \text{ canonical and hence } \mathbb{1}^{\mathbb{P}} \Vdash ``\tau \in \dot{\mathbb{Q}}_1`'.$
- (2) Suppose τ is canonical with $\mathbb{1}^{\mathbb{P}} \Vdash ``\tau \in \dot{\mathbb{Q}}_1$ " but $\mathbb{1}^{\mathbb{P}} \not\Vdash ``\tau \in \dot{\mathbb{Q}}_0$ ". Therefore there is some $p \in \mathbb{P}$ with $p \Vdash ``\tau \in \dot{\mathbb{Q}}_1 \setminus \dot{\mathbb{Q}}_0$ ". But for some $p^* \leq p$ and $\sigma \in \operatorname{dom}(\dot{\mathbb{Q}}_1)$, we have $p^* \Vdash ``\tau = \sigma$ " and therefore $p^* \Vdash ``\sigma \notin \dot{\mathbb{Q}}_0$ " despite the fact that $\sigma \in \operatorname{dom}(\dot{\mathbb{Q}}_1)$ implies $\mathbb{1}^{\mathbb{P}} \Vdash ``\sigma \in \dot{\mathbb{Q}}_0$ ", a contradiction.

It's immediate that \leq_1 and \leq_0 give the same orderings for $\mathbb{P} * \dot{\mathbb{Q}}_1$ and $\mathbb{P} * \dot{\mathbb{Q}}_0$.

2. This gives that $\mathbb{P} * \dot{\mathbb{Q}}_1$ is the same as in the old sense because the old sense of the iteration has domain, as in Definition 34 A • 3,

 $\{\langle p, \dot{q} \rangle : p \in \mathbb{P} \land \dot{q} \in \operatorname{dom}(\dot{\mathbb{Q}}_1) \land p \Vdash ``\dot{q} \in \dot{\mathbb{Q}}_1"\},\$

and the new sense has domain, by the above results,

 $\{\langle p, \dot{q} \rangle : p \in \mathbb{P} \land \dot{q} \text{ is canonical} \land \mathbb{1}^{\mathbb{P}} \Vdash ``\dot{q} \in \dot{\mathbb{Q}}_1"\}.$

This is because (\subseteq) every element of dom $(\dot{\mathbb{Q}}_1)$ is already canonical and forced by every $p \in \mathbb{P}$ to be an element of $\dot{\mathbb{Q}}_1$; and (\supseteq) every canonical element forced by $\mathbb{1}^{\mathbb{P}}$ is already in dom $(\dot{\mathbb{Q}}_1)$. From this it's easy to see that the ordering is the same as well because we have included all of the canonical names into dom (\leq_1) .

3. Write ★ for iterations in the old sense. To show forcing equivalence, we give a dense homomorphism. Specifically, for each τ ∈ dom(Q̂₀), let f(τ) ∈ dom(Q̂₁) be a canonical name with 1^P ⊢ "τ = f(τ)", and remember that we have 1^P ⊢ "f(τ) ∈ Q̂₀ = Q̂₁". Define F : P ★ Q̂₀ → P * Q̂₁ by F(⟨p,q̀⟩) =

 $\langle p, f(\dot{q}) \rangle$. This is pretty clearly a homomorphism.

F preserves incompatibility since if $\langle p, f(\tau) \rangle$, $\langle p', f(\tau') \rangle \in \mathbb{P} * \dot{\mathbb{Q}}_1$ are compatible, then there is some $\langle p'', \dot{q} \rangle$ where $p'' \leq^{\mathbb{P}} p, p'$ and $p'' \Vdash ``\dot{q} \leq_0 f(\tau), f(\tau)''$. As $\mathbb{I}^{\mathbb{P}} \Vdash ``\dot{q} \in \dot{\mathbb{Q}}_0$ '', there is a name $\tau^* \in \text{dom}(\dot{\mathbb{Q}}_0)$ and extension $p^* \leq p''$ where $p^* \Vdash ``\tau^* = q \leq_0 f(\tau), f(\tau') \land f(\tau) = \tau \land f(\tau') = \tau''$ so that $\langle p^*, \tau^* \rangle$ is a common extension of $\langle p, \tau \rangle, \langle p', \tau' \rangle \in \mathbb{P} \star \dot{\mathbb{Q}}_0$.

The image $F''\mathbb{P} \star \dot{\mathbb{Q}}_0$ is dense in $\mathbb{P} \star \dot{\mathbb{Q}}_1$ since any $\langle p, \dot{q} \rangle \in \mathbb{P} \star \dot{\mathbb{Q}}_1$ has $p \Vdash ``\dot{q} \in \dot{\mathbb{Q}}_0$ '' and we can therefore extend to a $p^* \in \mathbb{P}$ where some $\tau \in \text{dom}(\dot{\mathbb{Q}}_0)$ has $p^* \Vdash ``\dot{q} = \tau$ '' and so $F(\langle p^*, \tau \rangle) = \langle p^*, f(\tau) \rangle \leq \langle p, \dot{q} \rangle$ as desired. By Dense Forcing Equivalence (33 C • 5), the two are forcing equivalent. \dashv

This easily (although even more tediously) generalizes to longer iterations to show we don't lose any of the results related to those either. The reason the analogous result for longer iterations is so technical is that while the general idea is simple—just transforming \dot{Q}_{α} to the \dot{Q}'_{α} as \dot{Q}_0 is transformed to \dot{Q}_1 is above—implementing this idea requires transforming names of the old iteration into names of the new iteration in a way that works at every stage via Name Translation Theorem (33 C • 8). The details of this are then left to any *extremely* dedicated reader.

For us, the main benefit of canonical names is the ability to easily prove things one would expect like that the iteration of $< \kappa$ -closed preorders is $< \kappa$ -closed. Whereas before the issue was we could get a *name* for something below the relevant chain, we had no guarantee that such a name was in the domain of the preorder. With canonical names and Canonical Name Search (34 D • 3), we get actual elements.

– 34 D•6. Lemma

Let $\lambda > 0$ be an ordinal. Let \mathbb{P} be a $< \lambda$ -closed preorder and $\dot{\mathbb{Q}}$ a \mathbb{P} -name for a preorder such that $\mathbb{1}^{\mathbb{P}} \Vdash ``\dot{\mathbb{Q}}$ is $< \dot{\lambda}$ -closed". Therefore $\mathbb{P} * \dot{\mathbb{Q}}$ is $< \lambda$ -closed.

Proof .:.

Write $\mathbb{P} * \dot{\mathbb{Q}} = \langle \mathbb{P} * \dot{\mathbb{Q}}, \leq, 1 \rangle$. Let $\gamma < \lambda$ and $\langle \langle p_{\alpha}, \dot{q}_{\alpha} \rangle : \alpha < \gamma \rangle$ be a γ -length, \leq -decreasing sequence. Since \mathbb{P} is $< \lambda$ -closed, there is some $p^* \leq p_{\alpha}$ for every $\alpha < \gamma$ and this $p^* \Vdash \forall \alpha \leq \beta < \check{\gamma} (\dot{q}_{\beta} \leq^{\dot{\mathbb{Q}}} \dot{q}_{\alpha})^{"}$. Since $\mathbb{I}^{\mathbb{P}} \Vdash "\dot{\mathbb{Q}}$ is $< \check{\lambda}$ -closed",

$$p^* \Vdash ``\exists \dot{q}^* \forall \alpha < \check{\gamma} (q^* \in \dot{\mathbb{Q}} \land q^* \leq^{\mathbb{Q}} \dot{q}_{\alpha})$$
"

and therefore there is some name \dot{q}^* that p^* forces with this property. In fact, by Canonical Name Search (34 D • 3), there is a canonical name τ with $\mathbb{1}^{\mathbb{P}} \Vdash ``\tau \in \dot{\mathbb{Q}}$ '' and $p^* \Vdash ``\dot{q}^* = \tau$ '' and therefore $p^* \Vdash ``\forall \alpha < \check{\gamma} \ (\tau \leq \overset{\hat{\mathbb{Q}}}{q}_{\alpha})$ ''. Hence $\langle p^*, \tau \rangle \in \mathbb{P} * \dot{\mathbb{Q}}$ witnesses the result.

For longer iterations, full support always preserves this property, but we can do a bit better than this analogous to Direct Limit Chain Conditions $(34 \text{ C} \cdot 15)$ with bounded supports.

- 34 D • 7. Theorem (Inverse Limit Closure)

Let κ , λ be ordinals. Let $\mathbf{*}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ be a κ -stage iteration with support in $I \subseteq \mathcal{P}(\kappa)$, an ideal or $\mathcal{P}(\kappa)$ itself, such that

- inverse limits are taken at every limit stage $\alpha \leq \kappa$ with $cof(\alpha) < \lambda$;
- we take either direct or inverse limits at all other limit stages;
- $\mathbb{1}_{\alpha} \Vdash ``\dot{\mathbb{Q}}_{\alpha}$ is $< \dot{\lambda}$ -closed'' for each $\alpha < \kappa$

Therefore $\mathbf{a}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ is $< \lambda$ -closed.

Proof .:.

Proceed by induction on κ . $\kappa = 0$ is trivial as the only sequences are constant 1 sequences. Successors were shown with Lemma 34 D • 6. So let κ be a limit.

Let $\langle p_{\eta} : \eta < \theta \rangle$ be a \leq_{κ} -decreasing sequence of length $\theta < \lambda$, meaning $\langle p_{\eta} \upharpoonright \alpha : \eta < \kappa \rangle$ is \leq_{α} -decreasing for each $\alpha < \theta$. Firstly, define $p \upharpoonright 0 = \emptyset$, as expected. Suppose $\alpha < \kappa$, and thus far we have defined $p \upharpoonright \alpha$ such

ITERATED FORCING

that

$$p \upharpoonright \alpha \in \bigotimes_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} \text{ and } p \upharpoonright \alpha \leq_{\alpha} p_{\eta} \upharpoonright \alpha \text{ for all } \eta < \theta.$$

$$(*)$$

We define $p(\alpha)$ as follows. Since

 $p \upharpoonright \alpha \Vdash `` \forall \eta \leq \mu < \check{\theta} (p_{\mu}(\alpha) \leq'_{\alpha} p_{\eta}(\alpha)) \land \check{\theta} < \check{\lambda} \land \dot{\mathbb{Q}}_{\alpha} \text{ is } < \check{\lambda} \text{-closed}'',$

it follows that $p \upharpoonright \alpha$ forces some name below all of the $p_{\eta}(\alpha)$ s. And by Canonical Name Search (34 D • 3), there is some canonical $\mathbf{x}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ -name τ with $\mathbb{1}_{\alpha} \Vdash ``\tau \in \dot{\mathbb{Q}}_{\alpha}$ '' and $p \upharpoonright \alpha \Vdash ``\tau \leq '_{\alpha} p_{\eta}(\alpha)$ '' for all $\eta < \theta$. Hence setting $p(\alpha) = \tau$ yields (*) for $\alpha + 1$ in place of α . Note that we may take $\tau = \dot{\mathbb{I}}'_{\alpha}$ if we don't need to extend:

$$p(\alpha) = \mathring{1}'_{\alpha} \quad \text{iff} \quad \forall \eta < \theta \ (p_{\eta}(\alpha) = \mathring{1}'_{\alpha}). \tag{\dagger}$$

For limit α , $p \upharpoonright \alpha = \bigcup_{\xi < \alpha} p \upharpoonright \xi$ so that it's clear $p \upharpoonright \alpha \leq_{\alpha} p_{\eta} \upharpoonright \alpha$ for every $\eta < \theta$ assuming $p \upharpoonright \alpha \in \mathbf{X}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$. To show this, we must examine the support of $p \upharpoonright \alpha$. Note that we have both

1. sprt $(p \upharpoonright \alpha) = \bigcup_{\xi < \alpha} \text{sprt}(p \upharpoonright \xi)$; and

2. by (†), sprt($p \upharpoonright \alpha$) = $\bigcup_{n < \theta} sprt(p_{\eta} \upharpoonright \alpha)$.

If we are taking the inverse limit at stage α , we're done: (*) holds of $\xi < \alpha$ inductively with sprt $(p \upharpoonright \xi) \in I$ for $\xi < \alpha$ and so (1) and Support of Inverse Limits (34 C • 12) implies sprt $(p \upharpoonright \alpha) \in I$ hence (*).

So suppose we're taking the direct limit at stage α and therefore $\theta < \lambda \leq cof(\alpha)$. Thus $p_{\eta} \upharpoonright \alpha \in \mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ has bounded support for each $\eta < \theta$. Again because $\theta < cof(\alpha)$, this implies $\bigcup_{\eta < \theta} \sup \operatorname{sprt}(p_{\eta} \upharpoonright \alpha) = \sup \operatorname{sprt}(p \upharpoonright \alpha)$ is bounded and therefore $p \upharpoonright \alpha \in \mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ by Support of Direct Limits (34 C • 5) and hence (*) holds.

It follows that
$$p = p \upharpoonright \kappa \in \bigstar_{\alpha < \kappa} \hat{\mathbb{Q}}_{\alpha}$$
 with $p \leq_{\kappa} p_{\eta}$ for each $\eta < \theta$, witnessing the result. \dashv

Note that this uses canonical names^{XXX} in an essential way compared to the previous way of defining iterations. To explain a little, we don't need canonical names to show the two-step iteration $\mathbb{P} * \dot{\mathbb{Q}}$ is $< \kappa$ -closed whenever \mathbb{P} is and $\mathbb{1}^{\mathbb{P}} \Vdash "\dot{\mathbb{Q}}$ is $< \kappa$ -closed". The reason is that if $p \Vdash "\exists \dot{q} \varphi(\dot{q})$ " for any formula φ , it follows that there is then an extension $p^* \leq p$ where $p^* \Vdash "\varphi(\sigma)$ " for some $\sigma \in \text{dom}(\dot{\mathbb{Q}})$ and therefore we can consider $\langle p^*, \sigma \rangle \in \mathbb{P} * \dot{\mathbb{Q}}$. The issue with longer iterations is that the move from p to p^* is needed to ensure we can find such a σ , but this might screw with the support if we build up p as in Inverse Limit Closure (34 D • 7).

- **34D**•8. Corollary Let κ, λ be ordinals and $\mathbf{*}_{\alpha < \lambda} \dot{\mathbf{Q}}_{\alpha}$ a full support λ -stage iteration. Suppose $\mathbb{1}_{\alpha} \Vdash ``\dot{\mathbf{Q}}_{\alpha}$ is $< \check{\kappa}$ -closed" for each $\alpha < \lambda$. Therefore $\mathbf{*}_{\alpha < \lambda} \dot{\mathbf{Q}}_{\alpha}$ is $< \kappa$ -closed.

The above also generalizes to other properties not described here like being $< \kappa$ -strategically closed, $< \kappa$ -directed closed, and so on.

§ 34 E. Breaking up an iteration

The general intuition behind an iteration $\mathbf{*}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ is that we have a sequence of generic extensions of the ground model V. It would then seem that we could break this process in the middle and consider the rest of the process as its own iteration: go from V to $V[G \upharpoonright \alpha]$ and then from $V[G \upharpoonright \alpha]$ to $V[G \upharpoonright \kappa]$ through its own iteration, effectively thinking of $V[G \upharpoonright \kappa] = V[G \upharpoonright \alpha][G \upharpoonright [\alpha, \kappa)]$. In other words, we would like to say something along the lines of $\mathbf{*}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha} \cong \mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi} * \mathbf{*}_{\alpha \leq \xi < \kappa} \dot{\mathbf{Q}}_{\alpha}$. It turns out that this intuition is more or less correct. The best way to do this is just to restrict the domains of the elements.

— 34 E•1. Definition -

Let κ be an ordinal, $\alpha < \kappa$, and $\mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ a κ -stage iteration. Define the $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ -name for $\mathbf{*}_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ by $\underbrace{\mathbf{*}}_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi} = \left\{ (p \upharpoonright [\alpha, \kappa)) : p \in \mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi} \right\}$

xxx or any sort of collection of names with a result similar to Canonical Name Search (34 D • 3)

ITERATED FORCING

We order these elements with the $\mathbf{k}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ -name

$$\dot{\preccurlyeq}^{\kappa}_{\alpha} = \left\{ \langle \langle\!\langle q, r \rangle\!\rangle, p \rangle : q, r \in \operatorname{dom}\left(\bigstar_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi} \right) \land p \in \bigstar_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} \land p^{\frown}q \leqslant_{\kappa} p^{\frown}r \right\}$$

In essence, for $G \upharpoonright \alpha *_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ -generic over V, $(\bigstar_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi})_{G \upharpoonright \alpha}$ is ordered by $q \preccurlyeq_{\alpha}^{\kappa} r$ iff $p \frown q \preccurlyeq_{\kappa} p \frown r$ for some $p \in G \upharpoonright \alpha$. The maximal element here is clearly $\mathbb{1}_{\kappa} \upharpoonright [\alpha, \kappa)$. It's not difficult to see that this is indeed a name for a preorder.

- **34E**•2. Corollary Let κ be an ordinal, $\alpha < \kappa$, and $\mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ a κ -stage iteration. Therefore $\mathbb{1}_{\alpha} \Vdash \mathbf{*}_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ is a preorder".

More important for us is that the full iteration is forcing equivalent to forcing with $\mathbf{x}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi} * \mathbf{x}_{\alpha < \xi < \kappa} \dot{\mathbf{Q}}_{\xi}$.

- **34E**•3. Result Let κ be an ordinal, $\alpha < \kappa$, and $\mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ a κ -stage iteration with support in $I \subseteq \mathcal{P}(\kappa)$ an ideal or $\mathcal{P}(\kappa)$ itself. Therefore $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} * \mathbf{*}_{\alpha \leq \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ is forcing equivalent to $\mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$.

Proof .:.

Write \leq for the order on $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} * \mathbf{*}_{\alpha \leq \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$. For each p, let $p_{\alpha,\kappa}$ be a canonical name for $p \upharpoonright [\alpha, \kappa)$. Consider the map $f : \mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi} \to \mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} * \mathbf{*}_{\alpha \leq \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ defined by $f(p) = \langle p \upharpoonright \alpha, p_{\alpha,\kappa} \rangle$. This will turn out to be a dense embedding.

This will clearly preserve 1. Moreover, if $p^* \leq_{\kappa} p$ then $p^* \upharpoonright \alpha \leq_{\alpha} p \upharpoonright \alpha$ and as I is an ideal, $p^* = p^* \upharpoonright \alpha^{-} p^* \upharpoonright [\alpha, \kappa) \leq_{\kappa} p^* \upharpoonright \alpha^{-} p \upharpoonright [\alpha, \kappa)$ and therefore $p^* \upharpoonright \alpha \Vdash (p^* \upharpoonright [\alpha, \kappa)) \stackrel{\prec}{\prec} \stackrel{\kappa}{}_{\alpha} (p \upharpoonright [\alpha, \kappa)) \stackrel{\leftrightarrow}{}_{\alpha} (p^*) = \langle p^* \upharpoonright \alpha, p^*_{\alpha,\kappa} \rangle \leq \langle p \upharpoonright \alpha, p_{\alpha,\kappa} \rangle = f(p).$

Thus f is a homomorphism. In fact, if $f(p^*) \leq f(p)$ then $p^* \upharpoonright \alpha \leq_{\alpha} p \upharpoonright \alpha$ with $p^* \upharpoonright \alpha \Vdash$ " $p^*_{\alpha,\kappa} = (p^* \upharpoonright [\alpha,\kappa)) \stackrel{\sim}{\prec} \stackrel{\kappa}{}_{\alpha} (p \upharpoonright [\alpha,\kappa)) \stackrel{\sim}{} = p_{\alpha,\kappa}$ ", meaning

 $p^* = p^* \upharpoonright \alpha^{\frown} p^* \upharpoonright [\alpha, \kappa) \leq_{\kappa} p^* \upharpoonright \alpha^{\frown} p \upharpoonright [\alpha, \kappa) \leq_{\kappa} p \upharpoonright \alpha^{\frown} p \upharpoonright [\alpha, \kappa) = p.$ So we've shown $p^* \leq_{\kappa} p$ iff $f(p^*) \leq f(p)$ telling us that f is an embedding.

We haven't yet shown that f is an incompatibility embedding, but this will follow when we show that $f" \bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ is dense. (Any common extension $r \leq f(p), f(q)$ would have another extension $f(r^*) \leq r$ with then $f(r^*) \leq f(p), f(q)$ iff $r^* \leq_{\kappa} p, q$ by the above argument.) So let $\langle p, \dot{q} \rangle \in \bigstar_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi} \ast \bigstar_{\alpha \leq \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ be arbitrary. There is an extension p^* and $\tau \in \operatorname{dom}(\bigstar_{\alpha \leq \xi < \kappa} \dot{\mathbb{Q}}_{\xi})$ where $p^* \Vdash "\dot{q} = \tau$ ". But any such τ has the form $\tau = (r \upharpoonright [\alpha, \kappa))$ where $r \in \bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$. Since I is an ideal, $s = p^* \cap r \upharpoonright [\alpha, \kappa) \in \bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ with $f(s) = \langle p^*, r_{\alpha,\kappa} \rangle \leq \langle p, \dot{q} \rangle$ as desired. Thus f is a dense embedding and the result follows by Dense Forcing Equivalence (33 C • 5).

This is all just the setup for viewing properties of $\mathbf{k}_{\xi < \kappa} \dot{\mathbf{Q}}_{\xi}$ through forced properties of the tails $\mathbf{k}_{\alpha \leq \xi < \kappa} \dot{\mathbf{Q}}_{\xi}$. Often we can prove results about this by using previous results about iterations, but to do that, we must actually show that $\mathbb{1}_{\alpha} \Vdash \mathbf{k}_{\alpha \leq \xi < \kappa} \dot{\mathbf{Q}}_{\xi}$ is an iteration". Doing this is unfortunately technical, requiring us to translate names since technically these $\dot{\mathbf{Q}}_{\xi}$ s are $\mathbf{k}_{\zeta < \xi} \dot{\mathbf{Q}}_{\xi}$ -names whereas we want to consider what happens when we break up that iteration. In the end, we can translate $\dot{\mathbf{Q}}_{\alpha+\xi}$ to (what is forced by $\mathbb{1}_{\alpha}$ to be) a $\mathbf{k}_{\alpha \leq \zeta < \alpha+\xi} \dot{\mathbf{Q}}_{\zeta}$ -name for a preorder $\dot{\mathbf{Q}}'_{\xi}$. Conceptually this is basically trivial, but formally this is quite involved.

— 34 E • 4. Result

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\kappa \in \operatorname{Ord} \cap V$, $\alpha < \kappa$, and $\mathbf{*}_{\xi < \kappa} \dot{\mathbf{Q}}_{\xi}$ a κ -stage iteration with support in some $I \in V$ being an ideal or $\mathcal{P}(\kappa) \cap V$ itself. Let $G \upharpoonright \alpha$ be $\mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ -generic over V. Therefore

$$V[G \upharpoonright \alpha] \vDash " \bigstar_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi} \text{ is isomorphic to a } (\kappa - \alpha) \text{-stage iteration } \bigstar_{\xi < \kappa - \alpha} \dot{\mathbb{Q}}_{\xi}^{\circ}".$$

Moreover, the support of $\mathbf{k}_{\xi < \kappa - \alpha} \dot{\mathbb{Q}}_{\xi}^{\circ}$ is the shifted support of $I: \{X \subseteq \kappa - \alpha : \{\alpha + \beta : \beta \in X\} \in I\}$.
Proof .:.

We will work in $V[G \upharpoonright \alpha]$ so that $\bigstar_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ is indeed a preorder, ordered by $\preccurlyeq_{\alpha}^{\kappa}$. We will inductively define $\dot{\mathbb{Q}}_{\xi}^{\circ} = \langle \langle \dot{\mathbb{Q}}_{\xi}^{\circ}, \leqslant_{\xi}^{\circ}, \mathbb{I}_{\xi}^{\circ} \rangle \rangle$ that's supposed to be the translation of $\dot{\mathbb{Q}}_{\alpha+\xi}$. In doing so, we will get isomorphisms $\varphi_{\gamma} : \bigstar_{\alpha \le \xi < \alpha+\gamma} \dot{\mathbb{Q}}_{\xi} \to \bigstar_{\xi < \gamma} \dot{\mathbb{Q}}_{\xi}^{\circ}$ for $\gamma \le (\kappa - \alpha)$ that work nicely with projections:

for all
$$\delta < \gamma$$
 and $p \in \bigstar_{\alpha \le \xi < \alpha + \delta} \dot{\mathbb{Q}}^{\circ}_{\xi}$, we have that $\varphi_{\gamma}(p) \upharpoonright \delta = \varphi_{\delta}(p)$. (*)

So to actually get started with the proof, proceed by induction on $\gamma \leq \kappa - \alpha$ to define $\mathbf{x}_{\xi < \gamma} \dot{\mathbf{Q}}_{\xi}^{\circ}$ and show that such an isomorphism φ_{γ} satisfying (*) exists. For $\gamma = 0$, both $\mathbf{x}_{\alpha \leq \xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ and $\mathbf{x}_{\xi < \gamma} \dot{\mathbf{Q}}_{\xi}^{\circ}$ are trivial: 1. So suppose the result holds for all $\delta < \gamma$.

For γ a limit ordinal, we just take support allowed by the normal iteration at stage $\alpha + \gamma$: for all appropriate p, $p \in \bigstar_{\xi < \gamma} \dot{\mathbb{Q}}^{\circ}_{\xi}$ iff there is a $p' \in \bigstar_{\alpha \le \xi < \alpha + \gamma} \dot{\mathbb{Q}}_{\alpha}$ where $\forall \delta < \gamma \ (\varphi_{\delta}(p' \upharpoonright \alpha + \delta) = p \upharpoonright \delta)$. \square Claim 1

For limit
$$\gamma, \varphi_{\gamma} : \mathbf{X}_{\alpha \leq \xi < \alpha + \gamma} \, \dot{\mathbb{Q}}_{\xi} \to \mathbf{X}_{\xi < \gamma} \, \dot{\mathbb{Q}}_{\alpha}^{\circ}$$
 defined by $\varphi_{\gamma}(p) = \bigcup_{\delta < \gamma} \varphi_{\delta}(p \upharpoonright \alpha + \delta)$ is an isomorphism.

Proof .:.

It should be clear that φ_{γ} is injective by the inductive hypothesis. Surjectivity follows by definition of the limit iteration $\mathbf{x}_{\xi < \gamma} \dot{\mathbf{Q}}_{\xi}^{\circ}$ above. That φ_{γ} maps maximal elements to maximal elements is trivial by the inductive hypothesis. That $\varphi_{\gamma}(p^*) \leq_{\gamma} \varphi_{\gamma}(p)$ iff $p^* \preccurlyeq^{\alpha+\gamma}_{\alpha} p$ follows inductively: γ being a limit means $\alpha + \gamma$ is a limit and it should be clear that

$$p^* \preccurlyeq^{\alpha+\gamma}_{\alpha} p \quad \text{iff} \quad \forall \delta < \gamma \ (p^* \upharpoonright \alpha + \delta \preccurlyeq^{\alpha+\delta}_{\alpha} p \upharpoonright \alpha + \delta)$$

$$\text{iff} \quad \forall \delta < \gamma \ (\varphi_{\delta}(p^* \upharpoonright \alpha + \delta) \leqslant_{\delta} \varphi_{\delta}(p \upharpoonright \alpha + \delta))$$

$$\text{iff} \quad \varphi_{\gamma}(p^*) = \bigcup_{\delta < \gamma} \varphi_{\delta}(p^* \upharpoonright \alpha + \delta) \leqslant_{\gamma} \bigcup_{\delta < \gamma} \varphi_{\delta}(p \upharpoonright \alpha + \delta) = \varphi_{\gamma}(p). \quad \dashv$$

With the limit case out of the way, we can start thinking about the more difficult successor case, which basically consists in finding a $\mathbf{x}_{\xi < \gamma} \dot{\mathbf{Q}}_{\xi}^{\circ}$ -name for $\dot{\mathbf{Q}}_{\alpha+\gamma}$. The issue is that $\dot{\mathbf{Q}}_{\alpha+\gamma}$ is a $\mathbf{x}_{\xi < \alpha+\gamma} \dot{\mathbf{Q}}_{\xi}$ -name. Nevertheless, Name Translation Theorem (33 C • 8) and Result 34 E • 3 tell us we can translate the $\mathbf{x}_{\xi < \alpha+\gamma} \dot{\mathbf{Q}}_{\xi}$ -name $\dot{\mathbf{Q}}_{\alpha+\gamma}$ to a $\mathbf{x}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi} * \mathbf{x}_{\alpha \le \xi < \alpha+\gamma} \dot{\mathbf{Q}}_{\xi}$ -name $\ddot{\mathbf{Q}}_{\alpha+\gamma}$ which then gives a $\mathbf{x}_{\alpha \le \xi < \alpha+\gamma} \dot{\mathbf{Q}}_{\xi}$ -name $\ddot{\mathbf{Q}}_{\alpha+\gamma}$ in $V[G \upharpoonright \alpha]$.

Since φ_{γ} is a dense embedding, by Name Translation Theorem (33 C • 8) there is a translation T a function from $V[G \upharpoonright \alpha]^{\star_{\alpha \leq \xi < \alpha + \gamma} \dot{\mathbb{Q}}_{\xi}}$ to $V[G \upharpoonright \alpha]^{\star_{\xi < \gamma} \dot{\mathbb{Q}}_{\gamma}^{\circ}}$ where for any formula ψ , and names $\tau_0, \dots, \tau_n \in V[G \upharpoonright \alpha]^{\star_{\alpha \leq \xi < \alpha + \gamma} \dot{\mathbb{Q}}_{\xi}}$,

$$\mathbb{1}^{\bigstar_{\xi < \gamma} \dot{\mathsf{Q}}_{\xi}^{\circ}} \Vdash ``\psi(T(\tau_{0}), \cdots, T(\tau_{n}))" \quad \text{iff} \quad \mathbb{1}^{\bigstar_{\alpha \le \xi < \alpha + \gamma} \dot{\mathsf{Q}}_{\xi}} \Vdash ``\psi(\tau_{0}, \cdots, \tau_{n})",$$

and also in any generic extension $V[G \upharpoonright \alpha][H]$ by $\mathbf{*}_{\alpha \leq \xi < \alpha + \gamma} \dot{\mathbf{Q}}_{\xi}^{\circ}$, $T(\tau_i)_H = (\tau_i)_{\varphi_{\gamma}^{-1}}$. So let $\dot{\mathbf{Q}}_{\gamma}^{\circ}$ be a canonical name for $T(\ddot{\mathbf{Q}}_{\gamma})$. This then defines $\mathbf{*}_{\xi < \gamma + 1} \dot{\mathbf{Q}}_{\gamma}^{\circ}$. To define $\varphi_{\gamma + 1} : \mathbf{*}_{\alpha \leq \xi < \alpha + \gamma + 1} \dot{\mathbf{Q}}_{\xi} \to \mathbf{*}_{\xi < + 1} \dot{\mathbf{Q}}_{\xi}^{\circ}$, we just set $\varphi_{\gamma + 1}(p) = \varphi_{\gamma}(p \upharpoonright \alpha + \gamma)^{\frown} \langle \tau \rangle$ where τ is a canonical name for $T(p(\gamma))$, and here we take $p(\gamma)$ to be a canonical name for an element of $\ddot{\mathbf{Q}}_{\alpha + \xi}$. It's not difficult to see that $\varphi_{\gamma + 1}$ is an embedding so we merely need to show it's surjective. But this follows by Name Translation Theorem (33 C \cdot 8) (3) since every name is forced to be equivalent to one in the image of T. This establishes the successor case and thus the result. The support being the same but shifted is immediate.

All of this is essentially just to say $\bigstar_{\alpha \leq \xi < \kappa} \dot{Q}_{\xi}$ should be thought of as an iteration and the names for preorders we force with shouldn't be thought of too much since they can be translated into the relevant contexts. Now we can start showing legitimate results of these tail iterations $\bigstar_{\alpha \leq \xi < \kappa} \dot{Q}_{\xi}$.

- 34E.5. Result Let $\mathbf{*}_{\xi < \kappa} \dot{\mathbf{Q}}_{\xi}$ be a κ -stage iteration which is the direct limit of previous iterations. Therefore for each $\alpha < \kappa$, $\mathbb{I}_{\alpha} \Vdash \mathbf{*}_{\mathbf{X} \leq \xi < \kappa} \dot{\mathbf{Q}}_{\xi}$ is the direct limit of previous iterations", by which we mean the direct limit of $\mathbf{*}_{\alpha \leq \xi < \alpha + \gamma} \dot{\mathbf{Q}}_{\xi}$ for each $\gamma < \kappa - \alpha$.

Proof .:.

As the direct limit of previous iterations, the support of each element of $\mathbf{x}_{\xi < \kappa} \hat{\mathbf{Q}}_{\xi}$ is bounded in κ , and cutting off an initial segment to get $\mathbf{x}_{\alpha \leq \xi < \kappa} \hat{\mathbf{Q}}_{\xi}$ clearly still preserves this property.

A more difficult result to establish is $\chi_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ being the inverse limit of the previous iterations whenever $\chi_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ is the inverse limit. Part of the reason it's difficult is because it's false in general. It's possible to get some partial results, but these involve more technology.

- 34E·6. Definition
Let
$$\mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$$
 is a κ -stage iteration with support in some $I \subseteq \mathcal{P}(\kappa)$. For $\alpha < \kappa$, write
 $K_{\alpha} = \{\beta \in (\alpha, \kappa] : \mathbf{*}_{\xi < \beta} \dot{\mathbb{Q}}_{\xi} \text{ is the direct limit of previous iterations} \}.$
For $X \subseteq \kappa$ and $\alpha < \kappa$, we say X is K_{α} -thin iff $\forall \beta \in K_{\alpha} (\sup(X \cap \beta) < \beta).$

Basically, X is K_{α} -thin iff it's close to being an allowed support: the support at direct limit stages requires bounded support there, which X satisfies, but X may not work at the other stages (especially if X isn't in the ground model). In particular, for ideals or full support, if $p \in \mathbf{X}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$, then sprt(p) is K_{α} -thin for every $\alpha < \kappa$.

- 34 E • 7. Result -

Let $\mathbf{x}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ be a κ -stage iteration with support in some $I \subseteq \mathcal{P}(\kappa)$ —a non-principal ideal or $\mathcal{P}(\kappa)$ itself—such that

• inverse limits or direct limits are taken at every limit stage; and

• $\mathbf{x}_{\xi < \kappa} \hat{\mathbf{Q}}_{\xi}$ is the inverse limit of previous iterations.

Suppose further that there is an $\alpha < \kappa$ where

1. If $\mathbb{1}_{\alpha} \Vdash ``\forall X (X \text{ is } \check{K}_{\alpha}\text{-thin} \to \exists Y \in \check{V} (X \subseteq Y \land Y \text{ is } K_{\alpha}\text{-thin}))``.xxi$

Therefore $\mathbb{1}_{\alpha} \Vdash ::_{\alpha < \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ is (forcing equivalent to) the inverse limit of previous iterations".

Proof ...

For the sake of notation, work with a transitive model we can force over V, which is where the result's statement is interpreted. Let G be $\bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ -generic over V. Work in $V[G \upharpoonright \alpha]$. Suppose $p_i \in V[G \upharpoonright \alpha]$ should be in the inverse limit of previous iterations, i.e. dom $(p_i) = [\alpha, \kappa)$ and

$$p_i \upharpoonright \gamma \in \left(\bigstar_{\alpha \le \xi < \gamma} \dot{\mathbb{Q}}_{\xi} \right)_{G \upharpoonright \alpha} = \left\{ p \upharpoonright [\alpha, \gamma) \in V : p \in \bigstar_{\xi < \gamma} \dot{\mathbb{Q}}_{\xi} \right\}$$

for every $\gamma < \kappa$. We'd like to have

$$p_i \in \left(\bigstar_{\alpha \leq \xi < \kappa} \dot{\mathbb{Q}}_{\xi} \right)_{G \upharpoonright \alpha} = \left\{ p \upharpoonright [\alpha, \kappa) \in V : p \in \bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi} \right\}$$

Note that $\operatorname{sprt}(p_i)$ (which might not be in V) is K_{α} -thin in $\mathbf{V}[G \upharpoonright \alpha]$ so there is a K_{α} -thin $Y \in V$ with $\operatorname{sprt}(p_i) \subseteq Y$. We now attempt to define a $q \in \mathbf{k}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ (so in V) where $p_i \approx q \upharpoonright [\alpha, \kappa)$ in the sense that the two are forced to be equal everywhere (and hence in the poset version of the preorders, their equivalence classes are equal).

For each $\beta \in \kappa \setminus Y$, let $q(\beta) = \dot{\mathbb{1}}'_{\beta}$ so that sprt $(q) \subseteq Y$. For $\beta \in Y$, we know $\beta > \alpha$ and $p_i(\beta) = \tau \in V *_{\xi < \beta} \dot{\mathbb{Q}}_{\xi}$ is canonical in V and $V \models "\mathbb{1}_{\beta} \Vdash "\tau \in \dot{\mathbb{Q}}_{\beta}$ ". Of course, because p_i isn't in V, we a priori don't know what $p_i(\beta)$ is. Instead, note that the statement

$$p_i \upharpoonright \gamma \in \left\{ p \upharpoonright [\alpha, \gamma) \in V : p \in \bigstar_{\xi < \gamma} \dot{\mathbb{Q}}_{\xi} \right\}$$

^{xxxi}This implicitly assumes the non-trivial fact that the generic extension can tell whether any given set is actually in the ground model. In other words, whether it makes sense to use " $\tau \in \tilde{V}$ " in forcing. Although this is true, its proof is non-trivial and we do not prove it here since it's used merely for simplicity.

is upward absolute (fixing our interpretation of the iteration from V). Note also that we have a incompatibility homomorphisms from earlier iterations into later ones. In particular, for $\gamma < \eta \leq \kappa$, let $T_{\gamma,\eta}$ translate $\mathbf{x}_{\xi<\gamma} \dot{\mathbb{Q}}_{\xi}$ -names to $\mathbf{x}_{\xi<\eta} \dot{\mathbb{Q}}_{\xi}$ -names as per Name Translation Theorem (33 C • 8) (2). We then have some canonical $\mathbf{x}_{\xi<\beta} \dot{\mathbb{Q}}_{\xi}$ -name τ where

$$\mathbb{1}_{\beta} \Vdash ``T_{\alpha,\beta}(\dot{p}_i)(\check{\beta}) = \tau \land \tau \in \dot{\mathbb{Q}}_{\beta}''.$$

So define $q(\beta) = \tau$ which defines q on all of κ .

To see that $q \in \bigstar_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$, it's clear by definition each $q(\xi)$ is forced to be in $\dot{\mathbb{Q}}_{\xi}$. So we just need to confirm $\operatorname{sprt}(q \upharpoonright \beta) \in I$ for every limit $\beta \le \kappa$. But *Y* is K_{α} -thin: if $\beta < \kappa$ is a limit and we take a direct limit there, $\operatorname{sprt}(q \upharpoonright \beta) \subseteq Y \cap \beta$ (which is empty if $\beta < \alpha$) is bounded in β and hence inductively in *I*. If $\beta < \kappa$ is a limit and we take an inverse limit there, then inductively, $\operatorname{sprt}(q \upharpoonright \beta) = \bigcup_{\xi < \beta} \operatorname{sprt}(q \upharpoonright \xi) \in I$. Since we always take inverse or direct limits, $\operatorname{sprt}(q \upharpoonright \beta) \in I$ for every $\beta \le \kappa$.

Now for $\beta > \alpha$, if $\beta \notin Y$ then $\beta \notin \operatorname{sprt}(q) \cup \operatorname{sprt}(p_i)$ and hence $q(\beta)$ and $p_i(\beta)$ are forced to be equal (although perhaps aren't literally equal). If $\beta \in Y$, then in $V[G \upharpoonright \beta]$, we get $q(\beta) = (T_{\alpha,\beta}(\dot{p}_i(\check{\beta})))_{G\upharpoonright\beta} = (\dot{p}_i)_{G\upharpoonright\alpha}(\beta) = p_i(\beta)$ and hence $p_i \upharpoonright \gamma$ and $q \upharpoonright [\alpha, \gamma)$ extend each other for each $\gamma < \kappa$. In particular, if we instead consider the forcing equivalent *posets* $(\bigstar_{\alpha \leq \xi < \gamma} \dot{\mathbb{Q}}_{\xi})_{G\upharpoonright\alpha} / \approx_{\gamma} (\text{where } p \approx_{\gamma} p' \text{ iff } p \leq_{\gamma} p' \leq_{\gamma} p) \text{ of the iteration preorders, we get <math>[p_i \upharpoonright \gamma]_{\approx_{\gamma}} = [q \upharpoonright [\alpha, \gamma)]_{\approx_{\gamma}}$ and hence in the inverse limit (as caclulated in $V[G \upharpoonright \alpha]$), $[p_i] = [q \upharpoonright [\alpha, \kappa)]$. This means the inverse limit of $V[G \upharpoonright \alpha]$ is isomorphic to the inverse limit in V, which is forcing equivalent to the original $\bigstar_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$.

Mostly the benefit of breaking up iterations is when considering elementary embeddings where an initial segment of our iteration is preserved (being below H_{κ} for κ the critical point) and the tail has some nice properties. Breaking up iterations also allows us to reason as with two-step iterations. For example, if we don't screw up support too badly, $\mathbf{x}_{\xi<\kappa}\dot{\mathbf{Q}}_{\xi}$ is < κ -closed iff for some (equivalently for every) $\alpha < \kappa$, $\mathbf{x}_{\xi<\alpha}\dot{\mathbf{Q}}_{\xi}$ is < κ -closed and $\mathbb{I}_{\alpha} \Vdash \mathbf{x}_{\xi<\kappa}\dot{\mathbf{Q}}_{\xi}$ is < κ -closed.

— 34 E • 8. Result

Let $\mathbf{k}_{\xi < \kappa} \dot{\mathbf{Q}}_{\xi}$ be a κ -stage iteration with support in some $I \subseteq \mathcal{P}(\kappa)$ —a non-principal ideal or $\mathcal{P}(\kappa)$ itself—such that there is some $\alpha < \kappa$ and some $\lambda \in \text{Ord where}$

• inverse limits or direct limits are taken at every limit stage;

• inverse limits are taken at every limit stage $\geq \alpha$ of cofinality $< \lambda$;

- $\mathbb{1}_{\alpha} \Vdash ``\dot{\mathbb{Q}}_{\alpha}$ is $< \lambda$ -closed" for every $\alpha \leq \xi < \kappa$; and
- $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ is λ -cc.

Therefore $\mathbb{1}_{\alpha} \Vdash ``\bigstar_{\alpha \leq \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ is $< \lambda$ -closed".

Proof .:.

Again for the sake of notation, work with a transitive model we can force over V, which is where the result's statement is interpreted. Let G be $\mathbf{*}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ -generic over V. Work in $V[G \upharpoonright \alpha]$. By Result $34 \ge 4$, $\mathbb{P} = (\mathbf{*}_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi})_{G \upharpoonright \alpha}$ is isomorphic to an iteration with shifted support. In particular, we still take inverse limits or direct limits at every limit stage, and inverse limits at now at all limit stages of cofinality $< \lambda$. The last condition ensures that we indeed take inverse limits at those stages by Result $34 \ge 4$, $\mathbb{P} = (\mathbf{*}_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\xi})_{G \upharpoonright \alpha}$. By Inverse Limit Closure $(34 \ge 4)$, the iteration, isomorphic to \mathbb{P} , is $< \lambda$ -closed.

§ 34 F. Martin's Axiom

No overview of iterated forcing is complete without some discussion of forcing axioms like Martin's Axiom^{xxxii} (MA) or the Proper Forcing Axiom (PFA) among many others. These forcing axioms state the closure of V under forcing in

xxxiiNamed after Donald "Tony" Martin

the sense of a strengthening of Theorem 31 $D \cdot 1$: for certain kinds of preorders and for certain collections of dense sets, we can find a generic in V. When moving from the ground model to a generic extension, often establishing properties of the new set we are attempting to add only uses a few dense sets. Essentially, these force axioms say that we already have generics for those dense sets and hence we may use the technique of forcing while staying entirely in the ground model.

- **34F**•1. Definition Let κ be an ordinal. MA(κ) is the statement "for every ccc preorder \mathbb{P} and family \mathcal{D} of open, dense sets of \mathbb{P} , if $|\mathcal{D}| \leq \kappa$ then there is a G \mathbb{P} -generic over \mathcal{D} ". MA is the statement $\forall \kappa < 2^{\aleph_0} MA(\kappa)$.

Theorem 31 D • 1 then says ZFC \vdash MA(\aleph_0). So under CH, MA trivially holds because $\kappa < 2^{\aleph_0}$ implies κ is countable and thus MA(κ) \leftrightarrow MA(\aleph_0). So MA is easily consistent relative to ZFC. But this is really just to say that MA isn't interesting under CH. The true interest in the axiom is with MA + \neg CH.^{xxxiii} Note that MA is also maximal in the sense that MA(2^{\aleph_0}) is provably false.

− 34 F • 2. Result ZFC \vdash "MA(\aleph_0) $\land \neg$ MA(2^{\aleph_0})".

Proof .:.

We already have ZFC \vdash "MA(\aleph_0)" by Theorem 31 D • 1. To see that MA(2^{\aleph_0}) is false, consider Add(\aleph_0 , 1) which is ccc and therefore there are only $|Add(\aleph_0, 1)|^{\aleph_0} = 2^{\aleph_0}$ -many dense sets of Add(\aleph_0 , 1) in V. MA(2^{\aleph_0}) would imply there is a $G \in V$ Add(\aleph_0 , 1)-generic over V. But as a preorder appropriate for forcing, Theorem 31 D • 5 implies $G \notin V$, a contradiction.

The general idea behind showing the consistency of MA is just to continually force with all ccc preorders below a certain regular cardinal κ . When we do so, we end up forcing MA(θ) for every $\theta < \kappa$ and that $2^{\aleph_0} = \kappa$ and hence MA. Assuming $\kappa > \aleph_1$, we would then have MA + \neg CH. The technical details of this require a lot of work. Let's first start with an application of MA which helps establish why we need to go with a regular cardinal κ .

− **34 F • 3**. Theorem ZFC + MA \vdash "2^{\aleph_0} is regular".

To do this, we will basically show $2^{\kappa} = 2^{\aleph_0}$ for every infinite $\kappa < 2^{\aleph_0}$. This would show $cof(2^{\aleph_0}) = cof(2^{\kappa}) > \kappa$ for each $\kappa < 2^{\aleph_0}$ by König's Cofinality Theorem (5 D • 21) and therefore $cof(2^{\aleph_0}) \ge 2^{\aleph_0}$ giving equality. To show this, we will add a function associating subsets of κ to subsets of ω .

```
- 34F•4. Definition
Let \kappa be a cardinal. A subset \mathcal{A} \subseteq \mathcal{P}(\kappa) is an almost disjoint family of subsets of \kappa iff for any two A, B \in \mathcal{A}, |A \cap B| < \kappa.
```

We will be focused on $\kappa = \aleph_0$ because for regular κ , we can always find an almost disjoint family of size 2^{κ} . We show this just in the case of $\kappa = \aleph_0$.

- 34 F • 5. Result

There is an almost disjoint family of subsets of ω of size 2^{\aleph_0} .

Proof .:.

Consider ${}^{<\omega}\omega$ so that $|{}^{<\omega}\omega| = \aleph_0$ and we may instead find an almost disjoint family of subsets of ${}^{<\omega}\omega$ rather than of \aleph_0 . To do this, we just note that for $x \in \mathcal{N}$, $S_x = \{p \in {}^{<\omega}\omega : p \leq x\} \subseteq {}^{<\omega}\omega$. Moreover, for distinct $x, y \in \mathcal{N}$, $|S_x \cap S_y| < \aleph_0$ because any disagreement is carried ever afterword: $x(n) \neq y(n)$ implies $x \upharpoonright m \neq y \upharpoonright m$ for all $m \ge n$. As a result, $\{S_x : x \in \mathcal{N}\} \subseteq \mathcal{O}({}^{<\omega}\omega)$ is an almost disjoint family of size 2^{\aleph_0} . And of course, by passing to a bijection $b : {}^{<\omega}\omega \to \omega$, we can transform this to an almost disjoint family of

^{xxxiii}MA + \neg CH has many applications, one of which is to imply suslin's hypothesis of the non-existence of suslin trees, also known as ccc \aleph_1 aronszajn trees, by building an uncountable branch through them. If we allow MA to talk about non-ccc preorders, then \neg CH would allow us to also kill the non-ccc \aleph_1 -aronszajn trees in the same way, which is a contradiction with the fact that \aleph_1 -aronszajn trees provably exist. In this way, relaxing MA to include all ccc and non-ccc preorders yields a statement *equivalent* to CH.

subsets of ω .

As a result, to show $2^{\kappa} = 2^{\aleph_0}$ for $\kappa < \aleph_0$, we just need to identify subsets a κ -sized almost disjoint family with subsets of ω . We do this one subset X at a time with the preorder Adp(A, X) which says that X can be coded by a subset of ω . In particular, we add a set almost disjoint from all members of X, but *not* almost disjoint from the other members of A: we can characterize $X = \{A \in A : |g \cap A| = \aleph_0\}$ where g is coded by the generic. Hence we need to code both what is in g and what is not in g, and this means we really need to decide the characteristic function in ${}^{\omega}2$. To build up to this characteristic function, we need to make sure we only have finitely many 1s in common with the elements of X. We may do this as follows.

— 34 F • 6. Definition -

34F•7. Lemma

Let $\mathcal{A} \subseteq \mathcal{P}(\omega)$ be an almost disjoint family of sets with $X \subseteq \mathcal{A}$. The preorder $\mathsf{Adp}(\mathcal{A}, X)$ —the almost disjoint preorder on \mathcal{A} given by X—is defined by $p \leq q$ iff $p \supseteq q$, $\mathbb{1} = \emptyset$, and $\operatorname{Adp}(\mathcal{A}, X) = \{p : \omega \rightharpoonup 2 : p^{-1} | \{1\} < \aleph_0 \land \forall A \in X (|A \cap \operatorname{dom}(p)| < \aleph_0)\}.$

Note that the elements of $Adp(\mathcal{A}, X)$ need not be finite like with $Add(\aleph_0, \kappa)$: we just need their domain to be almost disjoint from the elements of X.^{xxxiv} To be able to do anything with this preorder with MA, we need to confirm it's ccc.

Let $\mathcal{A} \subseteq \mathcal{P}(\omega)$ be an almost disjoint family of sets with $X \subseteq \mathcal{A}$. Therefore $\mathsf{Adp}(\mathcal{A}, X)$ is ccc.

Proof .:.

Let $W \subseteq \operatorname{Adp}(\mathcal{A}, X)$ be an antichain. Let $U = \{p^{-1} | \{1\} : p \in W\}$. It follows that |U| = |W| because for $p, q \in \operatorname{Adp}(\mathcal{A}, X)$ to be incompatible, they must disagree on some $n \in \omega$ and therefore $p^{-1} | \{1\} \neq q^{-1} | \{1\}$. By definition of $\operatorname{Adp}(\mathcal{A}, X)$, $|p^{-1} | \{1\}| < \aleph_0$ for each $p \in W$ and therefore $|W| = |U| \le \aleph_0^{<\aleph_0} = \aleph_0$. \dashv

With this, we can continually force with Adp(A, X) and get the regularity of 2^{\aleph_0} from MA, i.e. Theorem $34 \text{ F} \cdot 3$.

- **34F**•8. Theorem For any infinite $\kappa < 2^{\aleph_0}$, MA(κ) implies $2^{\kappa} = 2^{\aleph_0}$ and $cof(2^{\aleph_0}) > \kappa$.

Proof .:.

Let $\kappa < 2^{\aleph_0}$ be arbitrary. Since there is an almost disjoint family of subsets of ω of size 2^{\aleph_0} , we take take a subset of this family of size κ . So assume \mathcal{A} is an almost disjoint family of size κ enumerated as $\mathcal{A} = \{A_\alpha : \alpha < \kappa\}$. Let $X \subseteq \mathcal{A}$ be arbitrary.

- Claim 1 There is a collection \mathcal{D} of $|\mathcal{A}| \cdot \aleph_0$ -many dense sets in V such that any G Adp (\mathcal{A}, X) -generic over \mathcal{D} has $X = \{A \in \mathcal{A} : |g \cap A| = \aleph_0\},$ where $g = (\bigcup G)^{-1}$ " $\{1\} \subseteq \omega$. \neg

^{xxxiv}Note that there might be almost disjoint families that are *maximal* in the sense that there is no subset of ω almost disjoint from any of their members. This means that there may not be infinite elements of Adp(A, X) in the ground model for X = A (which is partly the point in forcing there to be one). In fact, it's consistent for there to be maximal almost disjoint families of size $< 2^{\aleph_0}$! The least size of such a family is frequently denoted α and is one of many so-called *cardinal characteristics of the continuum* in that one can fairly easily show $\aleph_0 < \alpha \le 2^{\aleph_0}$ so that ZFC + CH \vdash " $\alpha = 2^{\aleph_0}$ ". MA and other similar forcing axioms often force these characteristics to actually be 2^{\aleph_0} , and this is no different for α : ZFC + MA \vdash " $\alpha = 2^{\aleph_0}$ ". Nevertheless, it's consistent for $\aleph_0 < \alpha < 2^{\aleph_0}$ to hold.

Proof .:.

Write $Y = \{ \alpha < \kappa : A_{\alpha} \in X \}$ the subset of κ that X codes. For each $n \in \omega$, $\alpha \in Y$, and $\beta \in \kappa \setminus Y$, consider the sets

$$D_{\beta} = \{ p \in \operatorname{Adp}(\mathcal{A}, X) : A_{\beta} \subseteq \operatorname{dom}(p) \},\$$
$$E_{\alpha}^{n} = \{ p \in \operatorname{Adp}(\mathcal{A}, X) : n \leq |A_{\alpha} \cap p^{-1}"\{1\}| \}.$$

Each of these will be dense in Adp(A, X). To see this, each D_{β} is dense since we can freely add 0s while staying a condition: for $p \in Adp(A, X)$, $p^* = p \cup ((A_{\beta} \setminus dom(p)) \times \{0\}) \in D_{\beta}$ is below p. Each E_{α}^n is dense in Adp(A, X) because each $A_{\alpha} \in A$ is infinite, and we are only adding finitely many 1s to bump up the size to at least n. Since we only add finitely many, we stay a condition.

Take $\mathcal{D} = \{D_{\beta}, E_{\alpha}^{n} : n \in \omega \land \alpha \in Y \land \beta \in \kappa \setminus Y\}$ so that $|\mathcal{D}| = \kappa \cdot \aleph_{0}$. For G Adp (\mathcal{A}, X) -generic over \mathcal{D} , it's easy to see $\bigcup G$ will be a partial function from ω to 2. So $g = (\bigcup G)^{-1}$ "{1} makes sense. So now we merely need to check that $A \in X$ iff $|A \cap g| = \aleph_{0}$. For $\alpha, \beta < \kappa$,

- (\rightarrow) Suppose $A_{\alpha} \in X$ so that $\alpha \in Y$ and $G \cap E_{\alpha}^{n} \neq \emptyset$ for every $n < \omega$. This means $|g \cap A_{\alpha}| \ge n$ for every $n < \omega$ and therefore $|g \cap A_{\alpha}| = \aleph_{0}$.
- (←) Suppose $A_{\beta} \notin X$ so that $\beta \in \kappa \setminus Y$ and $G \cap D_{\beta} \neq \emptyset$. Any $p \in G \cap D_{\beta}$ has already decided all values of $\bigcup G$ (the characteristic function of g) on A_{β} and hence has decided the intersection $g \cap A_{\beta}$. Yet as a condition in Adp(A, X), $|p^{-1}|^{+}\{1\}| < \aleph_0$ meaning $g \cap A_{\beta}$ must be finite. \dashv

This shows $\mathsf{MA}(\kappa) \to 2^{\kappa} = 2^{\aleph_0}$ if $\kappa \ge \aleph_0$. This is because each $Y \subseteq \kappa$ yields a ccc preorder $\mathsf{Adp}(\mathcal{A}, \{A_\alpha : \alpha \in Y\})$ by Lemma 34 F • 7. By $\mathsf{MA}(\kappa)$, there is a G_Y generic over the $|\mathcal{A}| \cdot \aleph_0 = \kappa$ -many dense sets of \mathcal{D} described in Claim 1 and therefore a $g_Y \subseteq \omega$ where then $Y = \{\alpha < \kappa : |g_Y \cap A_\alpha| = \aleph_0\}$. Hence we get an injection from $\mathcal{O}(\kappa)$ to $\mathcal{O}(\omega)$ by mapping Y to such a g_Y . Clearly $2^{\aleph_0} \le 2^{\kappa}$ so this gives equality.

But as $\kappa < 2^{\aleph_0}$ is arbitrary and MA(κ) holds for all such κ , we get $2^{\kappa} = 2^{\aleph_0}$ for all $\kappa < 2^{\aleph_0}$. By König's Cofinality Theorem (5 D • 21), $\kappa < cof(2^{\kappa}) = cof(2^{\aleph_0})$ for each $\kappa < 2^{\aleph_0}$. Therefore $cof(2^{\aleph_0}) \ge 2^{\aleph_0}$ and hence equality. \dashv

This is actually necessary for showing the consistency of $MA + \neg CH$ because we need to not only force MA but also $\neg CH$, and in doing so, we need to stop the iteration at some regular cardinal stage. Firstly, we need to show that we can bound the size of the preorders we're working with.

— 34 F • 9. Lemma

Let $\kappa < 2^{\aleph_0}$ be an infinite ordinal. Therefore MA(κ) is equivalent to MA($|\kappa|$) restricted to preorders of size $\leq \kappa$.

Proof .:.

Clearly $\mathsf{MA}(\kappa) \leftrightarrow \mathsf{MA}(|\kappa|)$ so we can assume κ is an infinite cardinal. Clearly full $\mathsf{MA}(\kappa)$ implies $\mathsf{MA}(\kappa)$ restricted to preorders of size $\leq \kappa$. So suppose $\mathsf{MA}(\kappa)$ holds for preorders of size $\leq \kappa$. Let \mathbb{P} be a ccc preorder of arbitrary size with $\mathcal{D} \leq \kappa$ -sized family of dense subsets of \mathbb{P} . Consider \mathcal{D} regarded as a signature for first order logic. Regarding \mathbb{P} as a FOL(\mathcal{D})-model (interpretting $D \in \mathcal{D}$ by $D^{\mathbb{P}} = D$), we can take a skolem hull $\mathsf{Hull}^{\mathbb{P}}(\emptyset)$. By Taking a Skolem Hull ($6 \land 2$), $|\mathsf{Hull}^{\mathbb{P}}(\emptyset)| \leq \kappa$ and $\mathsf{Hull}^{\mathbb{P}}(\emptyset) \leq \mathbb{P}$. In particular,

1. Hull^{\mathbb{P}}(\emptyset) satisfies all the axioms of preorders and is therefore a preorder.

2. $\operatorname{Hull}^{\mathbb{P}}(\emptyset) \models \neg \exists r \ (r \leq p, q)$ iff $\mathbb{P} \models \neg \exists r \ (r \leq p, q)$ for all $p, q \in \operatorname{Hull}^{\mathbb{P}}(\emptyset)$,

so that the identity map from $\operatorname{Hull}^{\mathbb{P}}(\emptyset)$ to \mathbb{P} is an incompatibility homomorphism. Similarly,

3. Hull^P(\emptyset) \models "D(p)" iff $\mathbb{P} \models$ "D(p)" for every $D \in \mathcal{D}$ and $p \in \text{Hull}^{\mathbb{P}}(\emptyset)$.

4. Hull^P(
$$\emptyset$$
) \models " $\forall p \exists x (D(x) \land x \leq p)$ " iff $\mathbb{P} \models$ " $\forall p \exists x (D(x) \land x \leq p)$ " for every $D \in \mathcal{D}$.

This means that $D^{\operatorname{Hull}^{\mathbb{P}}(\emptyset)} = D \cap \operatorname{Hull}^{\mathbb{P}}(\emptyset)$ is dense in $\operatorname{Hull}^{\mathbb{P}}(\emptyset)$ for each $D \in \mathcal{D}$.

5. Hull^P(\emptyset) is ccc since P is and the identity is an incompatibility homomorphism (an antichain $\mathcal{A} \subseteq \operatorname{Hull}^{\mathbb{P}}(\emptyset)$ is also an antichain of P).

It follows by $MA(\kappa)$ that there is a G Hull^P(\emptyset)-generic over \mathcal{D} . Thus $G \uparrow = \{p \in \mathbb{P} : \exists q \in G \ (q \leq p)\}$ is P-generic over \mathcal{D} . Since \mathbb{P} was arbitrary, this tells us $MA(\kappa)$ holds. \dashv Now we may prove the consistency of MA + \neg CH. We will do this by forcing with a finite support iteration, preserving cardinals and cofinalities. As a result, we need to be slightly careful about what cardinal we choose to make 2^{\aleph_0} : Theorem 34 F • 8 tells us that $\kappa = 2^{\aleph_0}$ is regular, and since finite support iterations of (what are forced to be) ccc preorders is itself ccc, this means κ needs to be regular in the ground model. Moverover, Theorem 34 F • 8 also tells us that $2^{<\kappa} = 2^{\aleph_0} = \kappa$ so that a form of CH should hold for κ . To get this, it suffices to just assume GCH for simplicity.

The general idea of the proof is to one-by-one add generics to ccc preorders of size $< \kappa$. Since $2^{<\kappa} = \kappa$, we can ensure there are only κ -many such preorders up to isomorphism (take every preorder of size $\theta < \kappa$ to have domain θ) throughout the procedure. And so we go through them all until we reach the end. The rest of the proof is confirming that this works: any θ -sized preorder for $\theta < \kappa$ and θ -sized family of dense sets in the generic extension is contained in some prior stage, and thus is given a generic at the next stage. And through some combinatorial results about nice names, we get $2^{\aleph_0} = \kappa$, yielding MA.

— 34 F • 10. Theorem (The Consistency of MA without CH)

Let $V \vDash \mathsf{ZFC} + \mathsf{GCH}$ be a transitive model we can force over. Let κ be a regular, uncountable cardinal of V. Therefore there is a preorder $\mathbb{P} \in V$ where $\mathbb{1}^{\mathbb{P}} \Vdash \mathsf{MA} + 2^{\aleph_0} = \kappa$ ". In particular, any generic extension of V by \mathbb{P} models $\mathsf{ZFC} + \mathsf{MA} + \neg \mathsf{CH}$.

Proof .:.

Work in V. We will define a finite support iteration $\mathbb{P} = \mathbf{*}_{\alpha < \kappa} \hat{\mathbb{Q}}_{\alpha}$, giving a ccc preorder in V by Result 34 C • 14 because we will require for each $\xi < \kappa$,

$$_{\xi} \Vdash ``\dot{\mathbb{Q}}_{\xi} \text{ is } \operatorname{ccc} \land |\dot{\mathbb{Q}}_{\xi}| < \check{\kappa}''.$$

$$(*)$$

Without loss of generality, work with preorders \mathbb{Q} that are "standard" in the sense that $\mathbb{Q} = |\mathbb{Q}|$ is an ordinal, and $\mathbb{1}^{\mathbb{Q}} = 0$. First we need the following combinatorial result showing that we don't destroy the fact that $2^{<\kappa} = \kappa$. Doing this will allow us to ensure we don't add too many ccc preorders later.

– Claim 1 ·

Let \mathbb{P} be a ccc preorder in V with $|\mathbb{P}| \leq \kappa$. Therefore $\mathbb{1}^{\mathbb{P}} \Vdash 2^{<\check{\kappa}} \leq \check{\kappa}$ and thus $\mathbb{1}^{\mathbb{P}} \Vdash 2^{\aleph_0} \leq \check{\kappa}$.

Proof .:.

By Corollary 32 E • 6, there are at most $(|\mathbb{P}|^{\aleph_0 \cdot \theta})^V$ -many nice names in V for subsets of $\check{\theta}$ for any cardinal $\theta < \kappa$. Since κ is regular, $\kappa^{<\kappa} = 2^{<\kappa} = \kappa$. Thus there at most κ -many nice names for subsets of θ for each $\theta < \kappa$. Hence $\mathbb{1}^{\mathbb{P}} \Vdash \forall \theta < \check{\kappa} (2^{\theta} \le \kappa)$, giving the result.

This will apply to us given that (*) inductively holds.

- Claim 2

Suppose $\mathbf{x}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is a finite support iteration and (*) holds for every $\xi < \alpha$. Therefore $\mathbf{x}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is ccc of size $\leq \kappa$, and so $\mathbb{1}_{\alpha} \Vdash \mathbb{1}_{2^{<\check{\kappa}}} = \check{\kappa}^{"}$.

Proof .:.

By Result 34 C • 14, we get that $\mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is ccc. Proceed by induction on α to show $\mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ has size $< \kappa$. For $\alpha = 0$, this is obvious. For limit α , as the direct limit, the cardinality is the supremum of previous stages all of which have size $< \kappa$. As κ is regular, this implies the limit iteration also has size $< \kappa$. For $\alpha + 1$, since $\mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is ccc and κ is regular, Corollary 32 C • 8 implies $\mathbf{1}_{\alpha} \Vdash \ddot{\mathbf{Q}}_{\alpha} \mid \leq \check{\theta}$ " for some $\theta < \kappa$. Identifying $\dot{\mathbf{Q}}_{\alpha}$ with $\check{\theta}$, Lemma 33 B • 10 tells us that every name for an element of $\dot{\mathbf{Q}}_{\alpha}$ is forced by $\mathbf{1}_{\alpha}$ to be equivalent to a nice name which is determined by a θ -length sequence of antichains. But since $\mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ is ccc, these antichains are countable and so there are κ^{\aleph_0} -many nice names and therefore at most $\kappa^{\aleph_0} = \kappa$ -many canonical names for elements of $\dot{\mathbf{Q}}_{\alpha}$. Hence $\mathbf{*}_{\xi < \alpha + 1} \dot{\mathbf{Q}}_{\xi}$ has size less than $\kappa \times \kappa = \kappa$.

So suppose $\mathbf{x}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ has been defined for $\alpha < \kappa$. For each $\theta < \kappa$, there are at most $2^{\theta} < \kappa$ non-isomorphic

 \neg

preorders over θ (since each corresponds to a subset of $\theta \times \theta$). Since $\mathbb{1}_{\alpha} \Vdash 2^{<\check{\kappa}} = \check{\kappa}$, it follows that

 $\mathbb{1}_{\alpha} \Vdash$ "there are at most $\check{\kappa}$ -many ccc standard preorders of size $\langle \check{\kappa} \rangle$ ".

In particular, for each $\alpha < \kappa$, let \dot{c}_{α} be a canonical name for a κ -length list of all standard preorders of size $< \kappa$. For $\beta < \kappa$, by $\dot{\mathbb{P}}_{\alpha,\beta}$ we mean a canonical name for the β th entry in this list: $\mathbb{1}_{\alpha} \Vdash ``\dot{\mathbb{P}}_{\alpha,\beta} = \dot{c}_{\alpha}(\check{\beta})``$. We will then force with all of these $\dot{\mathbb{P}}_{\alpha,\beta}$ at some point. To make things precise, we can take a surjection $f : \kappa \to \kappa \times \kappa$ such that $f(\alpha)_0 \le \alpha$ for all $\alpha < \kappa$, which means that at stage α , $\dot{\mathbb{P}}_{f(\alpha)_0, f(\alpha)_1} = \dot{\mathbb{P}}_{f(\alpha)}$ has indeed been defined. (Note that we have used choice quite a lot for this, but in well-ordering H_{λ} for sufficiently large λ , we only need to use choice once, just choosing the least such name each time.)

Now unfortunately, if we want to force with $\dot{\mathbb{P}}_{f(\alpha)}$ at stage α , we need to translate the name. This is because although $f(\alpha)_0 \leq \alpha$ so $\dot{\mathbb{P}}_{f(\alpha)}$ was defined at a previous stage, $\dot{\mathbb{P}}_{f(\alpha)}$ is merely a $\mathbf{*}_{\xi < f(\alpha)_0} \dot{\mathbb{Q}}_{\xi}$ -name, not a $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ -name. So we must be able to freely translate names from earlier stages to later stages. We do this as in Name Translation Theorem (33 C • 8) with the incompatibility homomorphisms $\iota_{\beta,\alpha}$ for $\beta < \alpha$ as in Lemma 34 C • 3: just adding a bunch of 1s to the end. This is just to say we have a map $T_{\beta,\alpha} : V^*_{\xi < \beta} \dot{\mathbb{Q}}_{\xi} \to V^*_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ by setting $T_{\beta,\alpha}(\emptyset) = \emptyset$, and inductively,

$$T_{\beta,\alpha}(\tau) = \{ \langle T_{\beta,\alpha}(\sigma), \iota_{\beta,\alpha}(p) \rangle : \langle \sigma, p \rangle \in \tau \}.$$

In pushing these names up to later iterations, we don't change the interpretations by Name Translation Theorem $(33 \text{ C} \cdot 8)$.

— Claim 3

Let $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ be a finite support α -stage iteration for some ordinal α . Let $G \upharpoonright \alpha$ be $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}$ -generic over V and $\beta < \alpha$. Therefore $T_{\beta,\alpha}(\tau)_{G \upharpoonright \alpha} = \tau_{G \upharpoonright \beta}$.

Again, suppose $\mathbf{k}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ has been defined. To define $\dot{\mathbf{Q}}_{\alpha}$ (and therefore $\mathbf{k}_{\xi < \alpha+1} \dot{\mathbf{Q}}_{\xi}$), consider $T_{f(\alpha)_{0,\alpha}}(\dot{\mathbf{P}}_{f(\alpha)})$. It may be that this is no longer ccc, which is fine, as we just let $\dot{\mathbf{Q}}_{\alpha}$ be a canonical name such that

 $\mathbb{1}_{\alpha} \Vdash ``\dot{\mathbb{Q}}_{\alpha} \text{ is a standard, ccc preorder of size } < \check{\kappa} \land (T_{f(\alpha)_0,\alpha}(\dot{\mathbb{P}}_{f(\alpha)}) \text{ is too} \rightarrow \dot{\mathbb{Q}}_{\alpha} = T_{f(\alpha)_0,\alpha}(\dot{\mathbb{P}}_{f(\alpha)}))''.$

Taking finite supports, this procedure defines $\mathbf{k}_{\xi < \alpha} \dot{\mathbb{Q}}_{\alpha}$ for all $\alpha \leq \kappa$ and we have defined $\mathbb{P} = \mathbf{k}_{\xi < \kappa} \dot{\mathbb{Q}}_{\alpha}$. Note that Claim 2 and Claim 1 imply $\mathbb{1}^{\mathbb{P}} \Vdash 2^{\aleph_0} \leq \check{\kappa}^n$. To show $\mathbb{1}^{\mathbb{P}} \Vdash 2^{\aleph_0} \geq \check{\kappa}^n$, we must show the more difficult result that $\mathbb{1}^{\mathbb{P}} \Vdash \mathsf{MA}(\check{\theta})^n$ for every $\theta < \kappa$.

Let G be \mathbb{P} -generic over V. Let $\mathbb{Q} \in V[G]$ be a standard, ccc preorder of size $< \kappa$ in $\mathbb{V}[G]$ and let $\mathcal{D} \in V[G]$ be a collection of dense subsets of \mathbb{Q} of size $< \kappa$.

— Claim 4

There is some $\alpha < \kappa$ where $\mathbb{Q}, \mathcal{D} \in V[G \upharpoonright \alpha]$.

Proof .:.

Let $\hat{\mathbb{Q}}$ be a $\mathbf{*}_{\xi < \kappa} \hat{\mathbb{Q}}_{\xi}$ -name for \mathbb{Q} . The idea is that as the direct limit of previous iterations, the supports of all elements in ran($\hat{\mathbb{Q}}$) are bounded in κ which is still regular in V[G] since \mathbb{P} is ccc. Moreover, since there are only $\theta < \kappa$ many elements, the supremum of all of these is some $\theta' < \kappa$ which then has $\mathbb{Q} \in V[G \upharpoonright \theta']$.

More explicitly, for each $\langle q^*, q \rangle \in \leq^{\mathbb{Q}}$, let $\alpha_{\langle q^*, q \rangle} = \operatorname{sup} \operatorname{sprt}(p)$ be minimal among the p such that $p \Vdash ``\dot{q}^* \leq^{\dot{\mathbb{Q}}} \dot{q}$ " for canonical names \dot{q}^*, \dot{q} . Let $\alpha = \operatorname{sup}_{q^* \leq q} \alpha_{\langle q^*, q \rangle}$. Since $|\leq^{\dot{\mathbb{Q}}}| = \theta < \kappa$ and κ is regular, $\alpha < \kappa$. But then in $V[G \upharpoonright \alpha]$,

$$\leq^{\mathbb{Q}} = \{ \langle q^*, q \rangle \in \theta \times \theta : \exists p \in G \upharpoonright \alpha \ (\iota_{\alpha,\kappa}(p) \Vdash ``\check{q}^* \leq^{\mathbb{Q}} \check{q}") \} \in V[G \upharpoonright \alpha].$$

The same idea works for \mathcal{D} by considering the graph $\{\langle \alpha, D \rangle : \alpha \in D \in \mathcal{D}\} \subseteq |\mathbb{Q}| \times |\mathbb{Q}|$.

As a result, \mathbb{Q} , \mathcal{D} appears in our long list of standard, ccc preorders. In particular, $\mathbb{Q} = \mathbb{P}_{\alpha,\beta}$ for some $\beta < \kappa$ and α as in Claim 4. But then as $f : \kappa \to \kappa \times \kappa$ was surjective, there is some $\alpha' \ge \alpha$ where

 $\mathbb{1}_{\alpha'} \Vdash ``\dot{\mathbb{Q}}_{\alpha'}$ is a standard, ccc preorder of size $\langle \check{\kappa} \land (T_{\alpha,\alpha'}(\dot{\mathbb{P}}_{\alpha,\beta}) \text{ is too } \rightarrow \dot{\mathbb{Q}}_{\alpha'} = T_{\alpha,\alpha'}(\dot{\mathbb{P}}_{\alpha,\beta}))''$.

By Claim 3, $T_{\alpha,\alpha'}(\dot{\mathbb{P}}_{\alpha,\beta})_{G \upharpoonright \alpha'} = (\dot{\mathbb{P}}_{\alpha,\beta})_{G \upharpoonright \alpha} = \mathbb{Q}$ which is indeed a standard, ccc preorder of size $< \kappa$ (because this is true in V[G] and being ccc is downward absolute between transitive models of ZFC). As a result,

 $V[G \upharpoonright \alpha' + 1] \vDash "G \upharpoonright \alpha' + 1$ is Q-generic over $\mathcal{D} \subseteq \{D \in V[G \upharpoonright \alpha'] : D \text{ is dense in } Q\}$ ".

V[G] then models the same as being generic over \mathcal{D} is upward absolute. As \mathbb{Q} and \mathcal{D} were arbitrary of size $< \kappa$, it follows that $\mathbf{V}[G] \models \text{``MA}(\theta)$ '' for every $\theta < \kappa$ and hence $\mathbf{V}[G] \models \text{``2}^{\theta} = 2^{\aleph_0} \le 2^{<\kappa} = \kappa$ '' and by Claim 1, $\mathbf{V}[G] \models \text{``2}^{\aleph_0} = \kappa$ ''' so that $\mathbf{V}[G] \models \text{MA} + \text{``2}^{\aleph_0} = \kappa$ '''. As G was arbitrary, $\mathbb{1}^{\mathbb{P}}$ forces this. \dashv

There are an enormous amount of applications of MA $(+\neg CH)$ that are interesting to set theorists. There are actually entire books on the subject [12]. We are not so interested in MA here for its ability to show the consistency of things but instead as a way of showing how long iterated forcing can be useful. There are many generalizations of MA stating the existence of generics for different kinds of preorders, and all of them use the same over-simplified procedure to show their consistency: force with all of those preorders and do some proper book keeping. Of course, there is very often more to their proofs than that since frequently these axioms require large cardinals to construct the iteration. But MA serves as a nice introduction to that field.

§ 34 G. Product forcing, homogeneity, and the failure of AC

Another application of iterated forcing is with forcing the failure of AC, which is really done with *product forcing*, i.e. iterated forcing where we use names for preorders already in the ground model. Now, as stated, that isn't quite right: we know from Theorem 31 D • 12 that forcing over a model of ZFC gives a model of ZFC. So we aren't forcing $\neg AC$, but instead we're forcing that an *inner model* of the generic extension satisfies $ZF + \neg AC$. In particular, $HOD^{V[G]} \models ZF + \neg AC$ for some generic extension V[G] of V.

First we should introduce the notion of a product forcing, as mentioned earlier.

- **34G**•1. **Definition** Let \mathcal{P} be a family of preorders. Let $I \subseteq \mathcal{P}(\mathcal{P})$.

- For $x \in \prod_{\mathbb{P} \in \mathcal{P}} \mathbb{P}$, the cartesian product, define $\operatorname{sprt}(x) = \{\mathbb{P} \in \mathcal{P} : x(\mathbb{P}) \neq \mathbb{1}^{\mathbb{P}}\}$.
- Define $(\prod \mathcal{P}, \leq)$, the product of \mathcal{P} with support in I, by $\prod \mathcal{P} = \{x \in \prod_{\mathbb{P} \in \mathcal{P}} \mathbb{P} : \operatorname{sprt}(x) \in I\}$, ordered entry-wise: $x \leq y$ iff for every $\mathbb{P} \in \mathcal{P}, x(\mathbb{P}) \leq^{\mathbb{P}} y(\mathbb{P})$.

We also write $\prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$ for $\prod \{\mathbb{P}_{\alpha} : \alpha < \kappa\}$ and $\mathbb{P} \times \mathbb{Q}$ for $\prod \{\mathbb{P}, \mathbb{Q}\}$. It's not difficult to see that products are isomorphic to iterations, but where the preorders are in the ground model. Note that this tells us the product is indeed a preorder, given that iterations are preorders.

— 34 G • 2. Lemma

Let $\kappa \in \text{Ord}$ and $\mathbb{P}_{\alpha} = \langle \mathbb{P}_{\alpha}, \leq_{\alpha}', \mathbb{1}_{\alpha}' \rangle$ be a preorder for $\alpha < \kappa$. Let $I \subseteq \kappa$ be closed under $X \mapsto X \cup Y$ where Y is any finite subset of κ .

Therefore, $\prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$ with support in some *I* is isomorphic to the iteration $\mathbf{*}_{\alpha < \kappa} \check{\mathbb{P}}_{\alpha}$ with support in *I*, where each $\check{\mathbb{P}}_{\alpha}$ is the canonical $\mathbf{*}_{\xi < \alpha} \check{\mathbb{P}}_{\xi}$ -name for \mathbb{P}_{α} .

Proof .:.

Recall that the $\mathbf{*}_{\xi < \alpha} \dot{\mathbb{P}}_{\xi}$ -name $\check{\mathbb{P}}_{\alpha} = \{\langle \check{p}, \mathbb{1}_{\alpha} \rangle : p \in \mathbb{P}_{\alpha} \}$. Proceed by induction on κ to show the map φ defined by $\langle p_{\alpha} \in \mathbb{P}_{\alpha} : \alpha < \kappa \rangle \mapsto \langle \check{p}_{\alpha} : \alpha < \kappa \rangle$ is an isomorphism from $\prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$ to $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{P}}_{\alpha}$. Note that $\varphi(\vec{p})(\alpha) = \dot{\mathbb{I}}_{\alpha}' = \check{\mathbb{I}}^{\mathbb{P}_{\alpha}}$ iff $\vec{p}(\alpha) = \mathbb{I}^{\mathbb{P}_{\alpha}}$. In particular, $\operatorname{sprt}(\varphi(\vec{p})) = \operatorname{sprt}(\vec{p})$ and therefore $\varphi(\vec{p})$ is a condition of the iteration iff \vec{p} is a condition of the product.

By Corollary 34 D • 2, elements of $\mathbf{*}_{\alpha < \kappa} \dot{\mathbb{P}}_{\alpha}$ take the form $\langle \check{p}_{\alpha} : \alpha < \kappa \rangle$ for $\langle p_{\alpha} : \alpha < \kappa \rangle \in \prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$, the cartesian product. This means φ is surjective. φ is clearly injective since $x \neq y$ implies $\check{x} \neq \check{y}$.

To see that $\varphi(\vec{p}) \leq_{\kappa} \varphi(\vec{p}^*)$ iff $\vec{p} \leq^{\kappa} \vec{p}^*$ (writing \leq^{κ} as the product order on $\prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$). This is the only place

CH VI §34 G

ITERATED FORCING

where we use induction. For
$$\kappa = 0$$
, this is trivial. For $\kappa + 1$,

$$\vec{p}^{\frown}\langle p_{\kappa}\rangle \leq^{\kappa+1} \vec{p}^{*\frown}\langle p_{\kappa}^{*}\rangle \quad \text{iff} \quad \forall \alpha < \kappa+1 \ (p_{\alpha} \leq'_{\alpha} p_{\alpha}^{*}) \quad \text{iff} \quad \vec{p} \leq^{\kappa} \vec{p}^{*} \land p_{\kappa} \leq'_{\kappa} p_{\kappa}^{*} \\ \text{iff} \quad \varphi(\vec{p}) \leq_{\kappa} \varphi(\vec{p}^{*}) \land \mathbb{1}_{\kappa} \Vdash \quad \breve{p}_{\kappa} \leq'_{\kappa} \breve{p}_{\kappa}^{*} \\ \text{iff} \quad \varphi(\vec{p})^{\frown}\langle \check{p}_{\kappa} \rangle \leq_{\kappa+1} \varphi(\vec{p}^{*})^{\frown}\langle \check{p}_{\kappa}^{*} \rangle \quad \text{iff} \quad \varphi(\vec{p}^{\frown}\langle p_{\kappa} \rangle) \leq_{\kappa+1} \varphi(\vec{p}^{*\frown}\langle p_{\kappa}^{*} \rangle).$$

or limit κ the result follows immediately by induction

Fc

The nice thing about Lemma 34 G • 2 is that we can use all of our theorems about iterations when talking about products. In fact, several theorems become much easier to show for products, like factoring.

- 34 G • 3. Result (Factoring Product Forcing) Let \mathcal{P} be a collection of preorders. Let $I \subseteq \mathcal{P}(\mathcal{P})$ be an ideal or $\mathcal{P}(\mathcal{P})$ itself. Therefore, for any $X \subseteq \mathcal{P}$, $\prod \mathcal{P} \cong \prod X \times \prod (\mathcal{P} \setminus X),$

where $\prod X$ has support in $\{x \cap X : x \in I\}$ and $\prod (\mathcal{P} \setminus X)$ has support in $\{x \setminus X : x \in I\}$. In particular, $\prod_{\xi < \kappa} \mathbb{P}_{\xi} \cong \prod_{\xi < \alpha} \mathbb{P}_{\xi} \times \prod_{\alpha \le \xi < \kappa} \mathbb{P}_{\xi} \text{ for any } \alpha < \kappa \text{ and preorders } \mathbb{P}_{\alpha} \text{ for } \alpha < \kappa.$

Proof .:.

If $x, y \in I$ then $(x \cap \alpha) \cup (y \setminus \alpha) \in I$ since I is closed under subsets and unions. So if $p \in \prod_{P \in X} \mathbb{P}$ and $q \in \prod_{\mathbb{P} \in \mathcal{P} \setminus X} \mathbb{P}$, then $p \cup q$ is in $\prod \mathcal{P}$, and vice versa: if p is in the full product, $p \upharpoonright X \in \prod_{\mathbb{P} \in X} \mathbb{P}$ and $p \upharpoonright (\mathcal{P} \setminus X) \in \prod_{\mathbb{P} \in \mathcal{P} \setminus X} \mathbb{P}$. Because the order is entry-wise, this shows the map $\varphi(p) \mapsto \langle p \upharpoonright X, p \upharpoonright (\mathcal{P} \setminus X) \rangle$ is an isomorphism. \neg

Another nice property of product preorders is that order doesn't matter as much compared with iterations, where changing the order is largely unintelligible: $\mathbb{P} \times \mathbb{Q} \cong \mathbb{Q} \times \mathbb{P}$ since the order on the product is entry-wise. This tells us that generics for the product are *mutually* generic in the following sense, as a corollary of Two-Step Iterated Forcing (34 A • 6). Ostensibly, a G being $\mathbb{P} \times \mathbb{Q}$ -generic is weaker: the first-components, $G_{\mathbb{P}}$ should be generic over the ground model V, and the second-components $G_{\mathbb{Q}}$ should be generic over $V[G_{\mathbb{P}}]$. The ability to reverse the order allows us to say that $G_{\mathbb{P}}$ is actually generic over $V[G_{\mathbb{Q}}]$.

- 34 G • 4. Corollary (Finite Product Forcing)

Let $V \models ZFC$ be a transitive model we can force over. Let $\mathbb{P}, \mathbb{Q} \in V$ be appropriate for forcing. Let $G \subseteq \mathbb{P} \times \mathbb{Q}$. Therefore, G is $\mathbb{P} \times \mathbb{Q}$ -generic over V iff $G = G_{\mathbb{P}} \times G_{\mathbb{Q}}$ where

- $G_{\mathbb{P}}$ is \mathbb{P} -generic over $V[G_{\mathbb{Q}}]$; and
- $G_{\mathbb{Q}}$ is \mathbb{Q} -generic over $V[G_{\mathbb{P}}]$.

In this case, $G_{\mathbb{P}} = \{ p \in \mathbb{P} : (p, \mathbb{I}^{\mathbb{Q}}) \in G \}$, and similarly for $G_{\mathbb{Q}}$.

Proof .:.

 (\rightarrow) Let G be generic. Therefore, for any $(p,q) \in G$, by upward closure, $(p, \mathbb{1}^{\mathbb{Q}}) \in G$. So define $G_{\mathbb{P}} = \{p \in G\}$ $\mathbb{P}: \langle p, \mathbb{I}^{\mathbb{Q}} \rangle \in G$. We now want to show $G = G_{\mathbb{P}} \times G_{\mathbb{Q}}$. The argument above tells us $G \subseteq G_{\mathbb{P}} \times G_{\mathbb{Q}}$. So suppose $p \in G_{\mathbb{P}}$ and $q \in G_{\mathbb{Q}}$, i.e. $\langle p, \mathbb{I}^{\mathbb{Q}} \rangle, \langle \mathbb{I}^{\mathbb{P}}, q \rangle \in G$. Since G is a filter, there is a $\langle p^*, q^* \rangle \in G$ extending both, which means $p^* \leq^{\mathbb{P}} p$ and $q^* \leq^{\mathbb{Q}} q$. By upward closure, $\langle p^*, q^* \rangle \leq \langle p, q \rangle \in G$. Thus $G = G_{\mathbb{P}} \times G_{\mathbb{Q}}$.

It's easy to see that G^{-1} is $\mathbb{Q} \times \mathbb{P}$ -generic over V assuming G is $\mathbb{P} \times \mathbb{Q}$ -generic over V. Moreover, Lemma 34 G • 2 tells us we have an isomorphism $\varphi : \mathbb{P} \times \mathbb{Q} \to \mathbb{P} * \check{\mathbb{Q}}$ where $\varphi(\langle p, q \rangle) = \langle p, \check{q} \rangle$. Hence φ "G is $\mathbb{P} * \check{\mathbb{Q}}$ generic over V. We know that dom(φ "G) = dom(G) = G_P. Two-Step Iterated Forcing (34 A • 6) tells us that $\{\check{q}_{G_{\mathbb{P}}}:\check{q}\in G\}=G_{\mathbb{Q}}$ is $\check{\mathbb{Q}}_{G_{\mathbb{P}}}=\mathbb{Q}$ -generic over $V[G_{\mathbb{P}}]$. The same argument applies to G^{-1} to yield that $G_{\mathbb{P}}$ is \mathbb{Q} -generic over $V[G_{\mathbb{Q}}]$.

 (\leftarrow) By Two-Step Iterated Forcing (34 A • 6), $G_{\mathbb{P}} * G_{\mathbb{Q}} = \{\langle p, \check{q} \rangle : p \in G_{\mathbb{P}} \land \check{q}_{G_{\mathbb{P}}} = q \in G_{\mathbb{Q}}\}$ is $\mathbb{P} * \check{\mathbb{Q}}$ -generic over V. In other words, $G' = \{ \langle p, \check{q} \rangle : \langle p, q \rangle \in G_{\mathbb{P}} \times G_{\mathbb{Q}} \}$ is $\mathbb{P} * \check{\mathbb{Q}}$ -generic over V. Again, we have an isomorphism defined by $\varphi(\langle p, q \rangle) = \langle p, \check{q} \rangle$ so that $\varphi^{-1} G' = G$ is $\mathbb{P} \times \mathbb{Q}$ -generic over V. -

As discussed before, it's important to realize that not all properties of \mathbb{Q} are retained in the generic extension by \mathbb{P} : \mathbb{Q} being a preorder is absolute, but it being, say, ccc isn't absolute. So we cannot, for example, say that the product of ccc preorders is ccc in general, because \mathbb{P} may add an uncountable antichain to \mathbb{Q} . The results of Lemma 34 A • 7 and others like it still hold, it's just that one must be careful about the context for the hypotheses about \mathbb{Q} .

Note that $Add(\aleph_0, 1)$ can be factored as a product of two copies of $Add(\aleph_0, 1)$, which has the interesting consequence that adding just one cohen real adds infinitely many which are mutually generic over each other.

- **34G**•5. Result Add($\aleph_0, 1$) \cong Add($\aleph_0, 1$) \times Add($\aleph_0, 1$). As a result, if G is Add($\aleph_0, 1$)-generic over a transitive model $\mathbf{V} \models$ ZFC we can force over, there is an intermediate extension: some H Add($\aleph_0, 1$)-generic over V has $V \subsetneq V[H] \subsetneq V[G]$.

Proof .:.

Note that $\operatorname{Add}(\aleph_0, 1) = \langle \operatorname{Add}(\aleph_0, 1), \leq_0, \emptyset \rangle$ may be regarded as $\langle \{p : \omega \rightharpoonup \omega : |\operatorname{dom}(p)| < \aleph_0\}, \supseteq, \emptyset \rangle$. As a result, if we interlace $p, q \in \operatorname{Add}(\aleph_0, 1)$, we get an element of $\operatorname{Add}(\aleph_0, 1)$:

 $\varphi(p,q) = \{ \langle 2n, p(n) \rangle : n \in \operatorname{dom}(p) \} \cup \{ \langle 2n+1, q(n) \rangle : n \in \operatorname{dom}(q) \}.$

It's easy to see that φ : Add $(\aleph_0, 1) \times$ Add $(\aleph_0, 1) \rightarrow$ Add $(\aleph_0, 1)$ is a bijection. Moreover, $\langle p^*, q^* \rangle \leq \langle p, q \rangle$ iff $\varphi(p^*, q^*) \supseteq \varphi(p, q)$, meaning φ is an isomorphism.

Thus if G is $\operatorname{Add}(\aleph_0, 1)$ -generic over V, we get φ^{-1} " $G = G' = G_0 \times G_1$ for some $G_0 \operatorname{Add}(\aleph_0, 1)$ -generic over V and $G_1 \operatorname{Add}(\aleph_0, 1)$ -generic over $V[G_0]$ by Finite Product Forcing (34 G • 4) where therefore $V[G] = V[G_0][G_1]$. Hence $H = G_0$ yields $H \in V[G] \setminus V$ but $G \notin V[H]$ since this would mean $G_1 \in V[G_0]$, contradicting generics aren't in the ground model by Theorem 31 D • 5.

More generally, we have that $Add(\aleph_0, \kappa)$ is isomorphic to the finite support product $\prod_{\xi < \kappa} Add(\aleph_0, 1)$, which meshes with the intuition that $Add(\aleph_0, \kappa)$ adds κ -many cohen reals.

- **34 G • 6.** Theorem For any $\kappa \in \text{Ord}, \text{Add}(\aleph_0, \kappa)$ is isomorphic to the finite support product $\prod_{\xi < \kappa} \text{Add}(\aleph_0, 1)$.

Proof .:.

Write \mathbb{P} for $\prod_{\xi < \kappa} \operatorname{Add}(\aleph_0, 1)$. Conditions in $\operatorname{Add}(\aleph_0, \kappa)$ take the form $p : \kappa \times \omega \to 2$ where $|p| < \aleph_0$. For $\xi < \kappa$ and $p \in \operatorname{Add}(\aleph_0, \kappa)$, write $p_{\xi} = \{\langle n, y \rangle : \langle \xi, n, y \rangle \in p\}$ (which might be \emptyset). Conditions in the finite support product \mathbb{P} may be regarded are $p : \kappa \to \operatorname{Add}(\aleph_0, 1)$ such that $p(\xi) = \emptyset$ for all but finitely many ξ . So consider the map $f : \operatorname{Add}(\aleph_0, \kappa) \to \mathbb{P}$ defined by $f(p)(\xi) = p_{\xi}$ for $\xi < \kappa$ and $p \in \operatorname{Add}(\aleph_0, \kappa)$.

This map f will clearly be injective. Since $|p| < \aleph_0$, $p_{\xi} = \emptyset = \mathbb{1}^{\mathsf{Add}(\aleph_0,1)}$ for all but finitely many $\xi < \kappa$, meaning f(p) has finite support and so is indeed a condition in \mathbb{P} . f is also surjective, since for any $p \in \mathbb{P}$, we can define $q = \{\langle \xi, n, y \rangle : \langle n, y \rangle \in p(\xi)\}$ where then f(q) = p. It should be clear that this f maps $\mathbb{1}^{\mathsf{Add}(\aleph_0,\kappa)} = \emptyset$ to $\mathbb{1}^{\mathbb{P}} = \operatorname{const}_{\emptyset} \upharpoonright \kappa$, and $p \leq^{\mathsf{Add}(\aleph_0,\kappa)} q$ iff $f(p) \leq^{\mathbb{P}} f(q)$, meaning f is an isomorphism. \dashv

We also get a similar result for the lévy collapse, $Col(\kappa, < \lambda)$, which we define here. Recall $Col(\kappa, \lambda)$ is defined as $Fn_{<\kappa}(\kappa, \lambda)$, i.e. all partial functions from κ to λ of size $< \kappa$.

- **34G**•7. **Definition** For κ a regular cardinal and λ an ordinal, define the preorder $\operatorname{Col}(\kappa, < \lambda) = \langle \operatorname{Col}(\kappa, < \lambda), \supseteq, \emptyset \rangle$ by $\operatorname{Col}(\kappa, < \lambda) = \{ p \in \operatorname{Fn}_{<\kappa}(\lambda \times \kappa, \lambda) : \forall \langle \alpha, \xi \rangle \in \operatorname{dom}(p) \ (p(\alpha, \xi) < \alpha) \}.$

– 34 G•8. Theorem

For κ a regular cardinal and λ an ordinal, $Col(\kappa, < \lambda)$ is isomorphic to the product $\prod_{\alpha < \lambda} Col(\kappa, \alpha)$ with support in $I = \{X \subseteq \lambda : |X| < \kappa\}.$

Proof .:.

Write \mathbb{P} for $\prod_{\alpha < \lambda} \operatorname{Col}(\kappa, \alpha)$. Conditions in \mathbb{P} take the form $\langle p_{\alpha} : \alpha < \lambda \rangle$ such that $p_{\alpha}(\xi) < \alpha$ for each $\xi < \kappa$ in dom (p_{α}) and $\alpha \in \lambda \setminus \kappa$. This may be regarded instead as a single function $p : \lambda \times \kappa \rightarrow \lambda$ where $p(\alpha, \xi) = p_{\alpha}(\xi) < \alpha$ whenever $\xi \in \operatorname{dom}(p_{\alpha})$. So define

$$f(\vec{p}) = \{ \langle \alpha, \xi, \gamma \rangle \in \lambda \times \kappa \times \lambda : \langle \xi, \gamma \rangle \in \vec{p}(\alpha) \}.$$

We have that f indeed is a map from \mathbb{P} to $\operatorname{Col}(\kappa, < \lambda)$: since our support has size $< \kappa$, $\vec{p}(\alpha) = \emptyset$ for all but $< \kappa$ -many $\alpha < \lambda$ and therefore $|f(\vec{p})|$, as the union of $< \kappa$ -many sets each of size $< \kappa$, has size $< \kappa$ by the regularity of κ .

To see that f is an isomorphism, it's clearly injective. Surjectivity is also easy: if $q \in Col(\kappa, < \lambda)$, then take

$$\vec{p} = (\alpha \mapsto \{\langle \xi, \gamma \rangle : \langle \alpha, \xi, \gamma \rangle \in q\})$$

and it's easily seen that $f(\vec{p}) = q$. We also have $f(\emptyset) = \emptyset$, and $\vec{p}^* \supseteq \vec{p}$ iff $f(\vec{p}^*) \supseteq f(\vec{p})$, meaning $f : \mathbb{P} \to \operatorname{Col}(\kappa, < \lambda)$ is an isomorphism.

The main point of the lévy collapse is to have slightly more control over how things are collapsed in that we preserve that $|\kappa| = \kappa$ is regular and $|\lambda| = \lambda$ while ensuring $\lambda \le \kappa^+$ in the extension, rather than merely $|\lambda| = \kappa^+ < \lambda$ if we had forced merely with $Col(\kappa^+, \lambda)$. In fact, if κ is regular and λ is strongly inaccessible, then forcing with $Col(\kappa, < \lambda)$ yields $\kappa^+ = \lambda$ in the generic extension.

Getting back to Cohen forcing, Result 34 G • 5 tells us $Add(\aleph_0, n) \cong Add(\aleph_0, 1)$ for any $n < \omega$. Because we take finite support, the direct limit of these, $Add(\aleph_0, \aleph_0)$, is then isomorphic to $Add(\aleph_0, 1)$. In a sort of converse way, even if we consider $Add(\aleph_0, \kappa)$ for large κ , any new real added could have been added just by $Add(\aleph_0, 1)$.

- **34G**•9. Corollary - Suppose the following:

- $V \vDash$ ZFC is a transitive model.
- G is $\mathsf{Add}(\aleph_0, \kappa)^{\mathsf{V}}$ -generic over V for some $\kappa \in \mathrm{Ord} \cap V$.
- $x \in \mathcal{N}^{\mathcal{V}[G]} \setminus \mathcal{N}^{\mathcal{V}}$.

Therefore there's an $H \operatorname{Add}(\aleph_0, 1)^V$ -generic over V such that $x \in V[H]$. If in addition κ is uncountable in V, then V[G] = V[H][K] for some $K \operatorname{Add}(\aleph_0, \kappa)^V$ -generic over V[H].

Proof .:.

Work in V. Let \dot{x} be an Add (\aleph_0, κ) -name for x. Without loss of generality, by Result 32 E • 5, \dot{x} is a nice name in that it takes the form $\bigcup_{n,m<\omega} \{\langle \langle \check{n}, \check{m} \rangle \} \times A_{n,m}$ where each $A_{n,m}$ is either \emptyset or an antichain. Since Add (\aleph_0, κ) is ccc, each $A_{n,m}$ is countable in V. As a result,

$$A = \bigcup_{p \in ran(\dot{x})} \operatorname{sprt}(p) = \bigcup_{n,m < \omega} \bigcup_{p \in A_{n,m}} \operatorname{sprt}(p)$$

is countable (since sprt(*p*) is finite). By Factoring Product Forcing (34 G • 3), Add(\aleph_0, κ) $\cong \prod_{\alpha < \kappa} Add(\aleph_0, 1)$ can be factored into the product

$$\mathsf{Add}(\aleph_0,\kappa)\cong\prod_{\alpha\in A}\mathsf{Add}(\aleph_0,1)\times\prod_{\alpha\in\kappa\backslash A}\mathsf{Add}(\aleph_0,1).$$

But since $|A| \leq \aleph_0$, $\prod_{\alpha \in A} \operatorname{Add}(\aleph_0, 1) \cong \operatorname{Add}(\aleph_0, \aleph_0) \cong \operatorname{Add}(\aleph_0, 1)$ by the remark above the statement. Hence we can view V[G] = V[H][K] where H is $\operatorname{Add}(\aleph_0, 1)$ -generic over V, and K is $\prod_{\alpha \in \kappa \setminus A} \operatorname{Add}(\aleph_0, 1)$ -generic over V[H].

- Note that by definition of A, \dot{x} is actually a $\prod_{\alpha \in A} \operatorname{Add}(\aleph_0, 1)$ -name which can then be translated into a $\operatorname{Add}(\aleph_0, 1)$ -name τ such that $\dot{x}_G = \tau_H$ by Name Translation Theorem (33 C 8). This yields $x \in V[H]$.
- If κ is uncountable, then $|\kappa \setminus A| = |\kappa|$ and hence $\prod_{\alpha \in \kappa \setminus A} \operatorname{Add}(\aleph_0, 1) \cong \prod_{\alpha < \kappa} \operatorname{Add}(\aleph_0, 1) \cong \operatorname{Add}(\aleph_0, \kappa)$, and we may find a $K' \operatorname{Add}(\aleph_0, \kappa)$ -generic over V with V[G] = V[H][K] = V[H][K']. \dashv

Theorem 34 G • 6 also tells us that there's an infinite \subseteq -decreasing sequence of Add(\aleph_0 , 1)-generic extensions $V \subsetneq$

 $\dots \subsetneq V[G_2] \subsetneq V[G_1] \subsetneq V[G_0]$ whenever there's a single $G_0 \operatorname{\mathsf{Add}}(\aleph_0, 1)$ -generic over V. $\operatorname{\mathsf{Add}}(\aleph_0, 1)$ is important for other reasons as well, and the following motivates the notion of *weak homogeneity*^{xxxv} of a preorder.

- 34 G • 10. Definition

A preorder \mathbb{P} is *weakly homogeneous* iff for every $p, q \in \mathbb{P}$, there is a dense homomorphism $f_{p,q} : \mathbb{P} \to \mathbb{P}$ such that $f_{p,q}(p)$ is compatible with q.

It's not difficult to show that $Add(\aleph_0, 1)$ is weakly homogeneous, and the product of weakly homogeneous preorders (with support in some ideal) is weakly homogeneous. If we think of $Add(\aleph_0, 1)$ as $Col(\aleph_0, \aleph_0)$, we can generalize this to get that $Col(\kappa, \lambda)$ is weakly homogeneous for every infinite $\kappa \leq \lambda$ (recall $Col(\kappa, \lambda)$ consists of partial functions from κ to λ of size $< \kappa$).

– 34 G•11. Corollary -

Add(\aleph_0 , 1) is weakly homogeneous. In fact, Col(κ , λ) is weakly homogeneous for all ordinals $\kappa \leq \lambda$.

Proof .:.

Let \mathbb{P} be Add($\aleph_0, 1$) or Col(κ, λ), regarding Add($\aleph_0, 1$) = $\langle \text{Add}(\aleph_0, 1), \leq, \mathbb{I} \rangle$ as $\langle \{p : \omega \rightarrow \omega : | \text{dom}(p) | < \aleph_0\}, \supseteq, \emptyset \rangle$. Let $p, q \in \mathbb{P}$ be arbitrary. Define $f_{p,q} : \mathbb{P} \rightarrow \mathbb{P}$ by taking $f_{p,q}(r)$ to swap r(n) = p(n) with q(n), and r(n) = q(n) with p(n), and otherwise do nothing: for $n \in \text{dom}(r)$,

$$f_{p,q}(r)(n) = \begin{cases} p(n) & \text{if } r(n) = q(n) \land n \in \operatorname{dom}(p) \cap \operatorname{dom}(q) \\ q(n) & \text{if } r(n) = p(n) \land n \in \operatorname{dom}(p) \cap \operatorname{dom}(q) \\ r(n) & \text{otherwise.} \end{cases}$$

For example, if dom $(p) \cap \text{dom}(q) = \emptyset$, then $f_{p,q} = \text{id}$ and p and q are already compatible. If dom(p) = dom(q), then $f_{p,q}(p) = q$.

To see that $f_{p,q}$ is a dense homomorphism, clearly $f_{p,q}(1) = 1$. Suppose $r^* \leq r$. It's easy to see $f_{p,q}(r^*) \supseteq f_{p,q}(r)$, meaning $f_{p,q}$ is a homomorphism. Suppose $r_0 \perp r_1$ so that $r_0(n) \neq r_1(n)$ for some $n \in \text{dom}(r_0) \cap \text{dom}(r_1)$.

- If $r_0(n) = p(n)$ and $r_1(n) = q(n) \neq p(n)$, then clearly $f_{p,q}(r_0)(n) = q(n)$ while $f_{p,q}(r_1)(n) = p(n)$ are also not equal.
- If $r_0(n) = p(n)$ and $r_1(n) \neq q(n)$, then $f_{p,q}(r_0)(n) = q(n) \neq r_1(n) = f_{p,q}(r_1)(n)$.
- If $r_0(n), r_1(n) \notin \{p(n), q(n)\}$, then $f_{p,q}(r_0)(n) = r_0(n) \neq r_1(n) = f_{p,q}(r_1)(n)$.

The other cases work similarly, and in each, $f_{p,q}(r_0) \perp f_{p,q}(r_1)$. So $f_{p,q}$ is an incompatibility homomorphism. For density, $f_{p,q}$ is actually bijective: for any $r \in \mathbb{P}$, $f_{p,q}(f_{p,q}(r)) = r$.

Note that as an incompatibility homomorphism, if $p \perp q$, then $f_{p,q}(p)$ is compatible with q, but incompatible with $f_{p,q}(q)$. There are many other examples of weakly homogeneous preorders, and the following allows us to find even more examples.

- 34 G • 12. Result

Let \mathbb{P}_{α} be a weakly homogeneous preorder for each $\alpha < \kappa$. Let $I \subseteq \mathbb{P}(\kappa)$ be an ideal or $\mathbb{P}(\kappa)$ itself. Therefore the product with support in I, $\prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$, is weakly homogeneous.

Proof .:.

Let $p, q \in \prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$ be arbitrary. For each $\alpha \in \operatorname{sprt}(p) \cap \operatorname{sprt}(q) \in I$, let $f_{\alpha} : \mathbb{P}_{\alpha} \to \mathbb{P}_{\alpha}$ be a dense homomorphism such that $f_{\alpha}(p(\alpha))$ is compatible with $q(\alpha)$. For $\alpha \notin \operatorname{sprt}(p) \cap \operatorname{sprt}(q)$, let $f_{\alpha} = \operatorname{id}$. Define $f_{p,q} : \prod_{\alpha < \kappa} \mathbb{P}_{\alpha} \to \prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$ by

$$f_{p,q}(r) = \langle f_{\alpha}(r(\alpha)) : \alpha < \kappa \rangle.$$

^{xxxv} I believe this notion is "weak" homogeneity from the usual model theoretic notion of a model \mathbb{P} being homogeneous. One consequence of homogeneity is that for any $p, q \in \mathbb{P}$, there's an isomorphism $f : \mathbb{P} \to \mathbb{P}$ with f(p) = q (as long as \mathbb{P} satisfies the same formulas using either p or q). We have weakened this result by requiring only a dense homomorphism and f(p) to be compatible with q. Of course, we have also strengthened it slightly by removing the requirement of p and q having the same theory over \mathbb{P} .

Note that $\operatorname{sprt}(f_{p,q}(r)) = \operatorname{sprt}(r)$ so if r is a condition of the product, so is $f_{p,q}(r)$. It's easy to see that $f_{p,q}$ is an incompatibility homomorphism since each f_{α} is. To see that $f_{p,q} \, \prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$ is dense in $\prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$, let $r \in \prod_{\alpha < \kappa} \mathbb{P}_{\alpha}$ be arbitrary. For each $\alpha < \kappa$, there's a $r^*(\alpha) \in \mathbb{P}_{\alpha}$ where $f_{\alpha}(r^*(\alpha)) \leq^{\mathbb{P}_{\alpha}} r(\alpha)$. In particular, if $r(\alpha) = \mathbb{1}^{\mathbb{P}_{\alpha}}$, we can take $r^*(\alpha) = r(\alpha)$. As a result, $r^* = \langle r^*(\alpha) : \alpha < \kappa \rangle$ has $\operatorname{sprt}(r^*) = \operatorname{sprt}(r)$ and moreover, by construction, $f_{p,q}(r^*) \leq r$ since $f_{p,q}(r^*)(\alpha) = f_{\alpha}(r^*(\alpha)) \leq^{\mathbb{P}_{\alpha}} r(\alpha)$ for each $\alpha < \kappa$.

To see $f_{p,q}(p)$ is compatible with q, just note that each $f_{p,q}(p)(\alpha)$ is compatible with $q(\alpha)$ through some common extension $r(\alpha) \leq f_{p,q}(p)(\alpha), q(\alpha)$ which defines $r = \langle r(\alpha) : \alpha < \kappa \rangle$. By taking $r(\alpha) = \mathbb{1}^{\mathbb{P}_{\alpha}}$ whenever possible, we can ensure $\operatorname{sprt}(r) = \operatorname{sprt}(f_{p,q}(p)) \cup \operatorname{sprt}(q) = \operatorname{sprt}(p) \cup \operatorname{sprt}(q) \in I$ and therefore $r \leq f_{p,q}(p), q$. \dashv

In particular, Add(\aleph_0, κ) and similarly Col($\kappa, < \lambda$) are both weakly homogeneous for all $\kappa < \lambda$.

Part of the point of weak homogeneity is the ability to understand the theory of the generic extension independently of the chosen generic: these dense homomorphisms tell us $p \Vdash \varphi$ iff $\mathbb{1}^{\mathbb{P}} \Vdash \varphi$ whenever φ is a sentence. So if G, H are both \mathbb{P} -generic over V, then $\mathbf{V}[G] \models \varphi$ iff there's a $p \in G$ with $p \Vdash \varphi$ iff $\mathbb{1}^{\mathbb{P}} \Vdash \varphi$ which means $\mathbf{V}[H] \models \varphi$.

— 34 G•13. Theorem

Let \mathbb{P} be a weakly homogeneous preorder. Let φ be a FOL(\in)-sentence. Therefore for any $p \in \mathbb{P}$, $p \Vdash \varphi$ iff $\mathbb{1}^{\mathbb{P}} \Vdash \varphi$.

Proof .:.

One direction is clear. So suppose $p \Vdash \varphi$. If $\mathbb{1}^{\mathbb{P}} \not\Vdash \varphi$, then some $q \in \mathbb{P}$ forces $\neg \varphi$. Let $f_{p,q} : \mathbb{P} \to \mathbb{P}$ be a dense homomorphism with $f_{p,q}(p)$ compatible with q. By Name Translation Theorem (33 C • 8), $f_{p,q}(p) \Vdash \varphi$ but then any common extension $r \leq^{\mathbb{P}} f_{p,q}(p), q$ must force both φ and $\neg \varphi$, a contradiction. \dashv

We can also expand this notion to allow for parameters if we expand our notion of weak homogeneity to fix the interpretation of the relevant \mathbb{P} -names. The same idea of using Name Translation Theorem (33 C • 8) gives Theorem 34 G • 13 with those parameters.

— 34G•14. Definition -

Let \mathbb{P} be a preorder and $\vec{\tau}$ \mathbb{P} -names. \mathbb{P} is *weakly homogeneous with respect to* $\vec{\tau}$ iff for every $p, q \in \mathbb{P}$, there's a dense homomorphism $f_{p,q} : \mathbb{P} \to \mathbb{P}$ such that

• $f_{p,q}(p)$ is compatible with q;

• $T_{f_{p,q}}(\tau_n) = \tau_n$ for each entry τ_n of $\vec{\tau}$;

where as in Name Translation Theorem (33 C • 8), we iteratively define $T_f(\tau) = \{ \langle T_f(\sigma), f(p) \rangle : \langle \sigma, p \rangle \in \tau \}.$

It's easy to see that every weakly homogeneous preorder is weakly homogeneous with respect to check-names.

– 34 G•15. Corollary -

Let $\mathbb P$ be a weakly homogeneous preorder. Therefore $\mathbb P$ is weakly homogeneous with respect to all check-names.

Proof .:.

Let $p, q \in \mathbb{P}$ be arbitrary and let $f_{p,q}$ be the dense homomorphism witnessing the weak homogeneity of \mathbb{P} . Let x be arbitrary to show $T_{f_{p,q}}(\check{x}) = \check{x}$. As a dense homomorphism, $f_{p,q}(\mathbb{1}^{\mathbb{P}}) = \mathbb{1}^{\mathbb{P}}$ and therefore a simple induction yields $T_{f_{p,q}}(\check{x}) = \{\langle T_{f_{p,q}}(\check{y}), \mathbb{1}^{\mathbb{P}} \rangle : y \in x\} = \{\langle \check{y}, \mathbb{1}^{\mathbb{P}} \rangle : y \in x\} = \check{x}$, as desired. \dashv

– 34G•16. Corollary -

Let \mathbb{P} be a preorder and $\vec{\tau} \mathbb{P}$ -names. Suppose \mathbb{P} is weakly homogeneous with respect to $\vec{\tau}$. Therefore, for any $p \in \mathbb{P}$ and FOL(\in)-formula φ , $p \Vdash ``\varphi(\vec{\tau})$ '' iff $\mathbb{1}^{\mathbb{P}} \Vdash ``\varphi(\vec{\tau})$ ''.

Proof .:.

The same proof as Theorem 34 G • 13 works, just noting that the use of Name Translation Theorem (33 C • 8) doesn't change the parameters by weak homogeneity with respect to $\vec{\tau}$.

In particular, if we force with Cohen forcing a bunch, what the generic extension thinks of (finitely many) sets in the ground model is independent of the generic chosen. This has the nice side effect of telling us $HOD_V^{V[G]} \subseteq V$.

→ **34** G • 17. **Result** Let $V \models ZFC$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder weakly homogeneous with respect to $\vec{\tau}$ in $V^{\mathbb{P}}$ and let G be \mathbb{P} -generic over V. Suppose $x \in V[G]$ is FOLp(\in)-definable from $\vec{\tau}_G$ and $x \subseteq V$. Therefore $x \in V$.

Proof .:.

Let $x \subseteq V_{\alpha}$ for some $\alpha \in Ord \cap V$ (in particular, $x \subseteq V_{rank(x)} \cap V$). There is some $FOL(\epsilon)$ -formula φ where $V[G] \models "\forall y \ (y \in x \leftrightarrow \varphi(y, \vec{\tau}_G))$ " so some $p \in G$ forces this. By Corollary 34 G • 16, for any name \dot{x} for x, $\mathbb{1}^{\mathbb{P}} \Vdash "\forall y \ (y \in \dot{x} \leftrightarrow \varphi(y, \vec{\tau}))$ " and therefore $y \in x$ iff $\mathbb{1}^{\mathbb{P}} \Vdash "\varphi(\check{y}, \vec{\tau})$ ". This then defines x in V: $x = \{y \in V_{\alpha} \cap V : \mathbb{1}^{\mathbb{P}} \Vdash "\varphi(\check{y}, \vec{\tau})"\} \in V.$

Recall OD_A is the class of sets $FOLp(\in)$ -definable with parameters in Ord and A. HOD_A is then the class of sets whose transitive closure is contained in OD_A .

— 34 G•18. Corollary

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a weakly homogeneous preorder, and $G \mathbb{P}$ generic over V. Therefore if $x \subseteq V$ and $x \in V[G]$ is ordinal definable in $\mathbf{V}[G]$ from parameters in V, then $x \in V$.
In particular, $\mathrm{HOD}_{V}^{\mathbf{V}[G]} \subseteq V$.

Proof .:.

Let $x \subseteq V$ be defined from ordinals $\alpha_0, \dots, \alpha_n \in \text{Ord} \cap V = \text{Ord} \cap V[G]$ and sets $y_0, \dots, y_m \in V$, meaning $x \in \text{OD}_V^{V[G]}$. \mathbb{P} is weakly homogeneous with respect to $\check{\alpha}_0, \dots, \check{\alpha}_n, \check{y}_0, \dots, \check{y}_m$ by Corollary 34 G • 15. By Result 34 G • 17, $x \in V$. Now if $x \in \text{HOD}_V^{V[G]}$, then $x \in \text{OD}_V^{V[G]}$ and it suffices to show $x \subseteq V$ by induction on rank. Clearly this holds for $x = \emptyset$. Every $y \in x$ is also in $\text{HOD}_V^{V[G]}$ so inductively $y \subseteq V$ and thus $y \in V$ by the above. This means $x \subseteq V$ and so $x \in V$.

We know that for any set $A \in OD_A$, $HOD_A \models ZF$, and $HOD_A \models "x$ can be well-ordered" iff there's a well-order of x in OD_A . The point of this is that although $HOD_{\mathcal{N}}^{V[G]} \models ZF$, if G is $Add(\aleph_0, \alpha)$ -generic for an uncountable α , there will be no definable well-ordering of \mathcal{N} from real parameters, and hence $HOD_{\mathcal{N}}^{V[G]} \models \neg AC$.

- 34 G • 19. Theorem (Consistency of ZF + \neg AC) Let $V \models$ ZFC be a transitive model we can force over. Let G be Add $(\aleph_0, \kappa)^V$ -generic over V for some $\kappa \in$ Ord $\cap V$ uncountable in V. Therefore HOD $_{\kappa}^{V[G]} \models$ ZF + \neg AC.

Proof .:.

By absoluteness, $\mathsf{Add}(\aleph_0, \kappa)$ is the same in every transitive model with κ in it. So there's no worry about where this defined preorder is interpreted. Write $\mathbb{P} = \mathsf{Add}(\aleph_0, \kappa)$. Work in V[G]. It suffices to show there's no $OD_{\mathcal{N}}$ well-ordering of \mathcal{N} . So suppose otherwise. For the sake of notation, let $\varphi(x, y, a, \alpha)$ define a well-ordering \preccurlyeq_{φ} from just one parameter $a \in \mathcal{N}$, and one ordinal α . The proof easily generalizes to more parameters. Let $\dot{a} \in V^{\mathbb{P}}$ be a name for a. By Corollary 34 G • 9, let V[G] = V[H][K] where $a \in V[H]$ and K is \mathbb{P} -generic over V[H].

Thus going from V[H] to V[H][K] = V[G], every $x \in \mathcal{N}$ can therefore be FOLp(\in)-defined using $\alpha, a \in V[H]$, and its \preccurlyeq_{φ} -rank (another ordinal). So $\mathcal{N} \subseteq OD_{\mathcal{N} \cap V[H]}$. In fact $x \in \mathcal{N}$, as a subset of $\omega \times \omega \in HOD$, then has $x \in HOD_{\mathcal{N} \cap V[H]}$ and so \mathcal{N} is contained in this. By the weak homogeneity of \mathbb{P} , Corollary 34 G • 18 implies $HOD_{\mathcal{N} \cap V[H]} = HOD_{\mathcal{N} \cap V[H]}^{V[H][K]} \subseteq V[H]$ and therefore $\mathcal{N} \subseteq V[H]$. But \mathbb{P} adds reals from V[H] to V[H][K], meaning $\mathcal{N} \not\subseteq V[H]$, a contradiction. \dashv

This marks the start of a whole host of results related to weakenings of AC that might hold or fail [16]. Generally,

such results are studied with "symmetric" or "permutation" models, but these always take the form $HOD_{V\cup X}^{V[G]}$ for some generic extension V[G] and some $X \in V[G]$ [13]. That being said, such models are often much nicer to work with in these kinds of arguments than HOD-style models.

Section 35. Forcing and Elementary Embeddings

Many of the large large cardinals are phrased in terms of elementary embeddings. So it's natural to ask how these interact with forcing. Overall, the main tools used in the interaction between the two focus on lifting embeddings from the ground model to the generic extension, and (more importantly) doing this lifting in the generic extension.

§35A. Lifting embeddings

What we mean by "lifting" an embedding is the following.

- 35A • 1. Definition Let $V, W, N, M \vDash$ ZFC be transitive models with $V \subseteq W$ and $N \subseteq M$. Let $j : V \to N$ be an elementary embedding. A *lift-up of j to W, M* is an elementary embedding $j^+ : W \to M$ such that $j^+ \upharpoonright V = j$.

Note that the lift-up of an embedding has the same critical point. The main result we care about is the following, telling us that we can always lift-up embeddings so long as the image of the generic is compatible with a generic.

— 35 A • 2. Theorem (Generic Lifting)

- Let $V \vDash$ ZFC be a transitive model we can force over.
- Let $j: V \to M$ be an elementary embedding with M transitive (M need not be a subset of V).
- Let $\mathbb{P} \in V$ be a preorder and let G be \mathbb{P} -generic over V.

Therefore the following are equivalent for every H (which need not be in V[G]).

- 1. $j "G \subseteq H$ and H is $j(\mathbb{P})$ -generic over M.
- 2. There is an elementary embedding $j^+: V[G] \to M[H]$ such that $j^+ \upharpoonright V = j$ and $H = j^+(G)$.

Proof .:.

(1) → (2). To define j⁺, we'd like to work with names. For any τ ∈ V^P, take j⁺(τ_G) = j(τ)_H. First we must show that this definition makes sense: if τ_G = σ_G then j(τ)_H = j(σ)_H. Let p ∈ P force the equality of names in V: p ⊢ "τ = σ". We therefore have in M that j(p) ⊢ "j(τ) = j(σ)" and since j(p) ∈ H, M[H] ⊨ "j(τ)_H = j(σ)_H". So j⁺: V[G] → M[H] is well-defined.

To see that j^+ is elementary, note that this follows from the elementarity of j: if $V[G] \models "\varphi(\vec{\tau}_G)$ " for \mathbb{P} names $\vec{\tau}$ in V and FOL-formula φ , then there is some $p \in G$ such that $p \Vdash "\varphi(\vec{\tau})$ " and hence in M, $j(p) \Vdash$ " $\varphi(j(\vec{\tau}))$ ". Since $j"G \subseteq H$, this gives $M[H] \models "\varphi(j(\vec{\tau})_H)$ " in other words, $M[H] \models "\varphi(j^+(\vec{\tau}_G))$ ". And
clearly, the converse holds since $M[H] \models "\neg \psi$ " iff $M[H] \not\models \psi$ and then we use the contrapositive from the
above argument.

To see that $j^+(G) = H$, note that the canonical name for $G, \dot{G} = \{\langle \check{p}, p \rangle : p \in \mathbb{P}\}$ is moved by j to $j(\dot{G}) = \{\langle \check{p}, p \rangle : p \in j(\mathbb{P})\}$ which is clearly interpreted by H to be $j(\dot{G})_H = H$.

• (2) \rightarrow (1). Clearly since $G \subseteq \mathbb{P}$, for every $p \in G$, $j(p) \in j^+(G) = H \subseteq j(\mathbb{P})$ by elementarity. Moreover, since $\mathcal{P}(\mathbb{P}) \cap V \in V[G]$, V[G] knows G is \mathbb{P} -generic over V and hence by elementarity, M[H] knows $j^+(G)$ is $j(\mathbb{P})$ -generic over $\mathcal{P}(j(\mathbb{P})) \cap j^{"}V = \mathcal{P}(j(\mathbb{P})) \cap M$. But the notion of being generic is absolute between transitive models. Thus $j^+(G)$ is indeed generic over M and so $H = j^+(G)$ works. \dashv

The hard part of this theorem is really the existence of such an H. For example, if we collapse a measurable cardinal κ to, say, ω , then in the generic extension, κ is no longer measurable, and hence V[G] can't lift the canonical ultrapower

embedding of V to V[G]. But we can actually say something much stronger than this. It's not that V[G] simply doesn't have the proper information to lift^{xxxvi} it's that there *can't* be a lift.

– 35 A•3. Example -

• Let $V \vDash$ ZFC be a transitive model we can force over.

• Let κ be measurable in V as witnessed by a measure $U \in V$.

• Let $j: V \to \operatorname{cUlt}^{V}(V, U)$ be the canonical embedding j_U .

• Let $\mathbb{P} = \operatorname{Col}(\omega, \kappa)$.

• Let G be \mathbb{P} -generic over V.

Therefore, there is no H that is $j(\mathbb{P})$ -generic over $\text{cUlt}^{\vee}(V, U)$ with $j \, "G \subseteq H$.

Proof .:.

Note that \mathbb{P} collapses $|\kappa|$ to \aleph_0 by adding a surjection $g: \omega \to \kappa$. One useful property of \mathbb{P} is that $j"\mathbb{P} = \mathbb{P}$, because each $p \in \mathbb{P}$ has $p \subseteq \omega \times \alpha$ for some $\alpha < \kappa = \operatorname{cp}(j)$. Thus the conflict with lifting is that $G \subseteq H$ and G already determines that the added function has its image as κ not $j(\kappa)$. To be more explicit, for each $n < \omega$,

 $D_n = \{ p \in \mathbb{P} : n \in \operatorname{dom}(p) \} \subseteq \mathbb{P}$

is dense in \mathbb{P} and hence has an intersection with G. It should also be clear that any $p \in G$ has im $p \subseteq \kappa$ by definition. Working with $j(\mathbb{P}) = \text{Col}(\omega, j(\kappa))$, let $\kappa \leq \alpha < j(\kappa)$ also be arbitrary. We should have

 $E_{\alpha} = \{ p \in j(\mathbb{P}) : \exists n \in \omega \ (p(n) = \alpha) \}$

as dense in $j(\mathbb{P})$ and hence intersects H at some $h \in H \cap E_{\alpha}$ where then some $n \in \omega$ has $h(n) = \alpha$. But $G \cap D_n \neq \emptyset$ so there is some $g \in G \cap D_n$ where then $g(n) < \kappa \leq \alpha$. Thus j(g) = g and h are incompatible, contradicting that H is a filter. \dashv

As such, many arguments require a lot of effort to show we can find such generics. And more difficult is that we want to find such generics $H \in V[G]$. This is because most of the embeddings we want to consider will witness large cardinal properties. As such, we often have above that $M \subseteq V$ and j as a FOLp-definable class of V (with parameters being measures, extenders, and so forth). To use j^+ for large cardinal properties in V[G], we often want j^+ to be a class of V[G], which requires not just the existence of the $j(\mathbb{P})$ -generic H, but also that $H \in V[G]$.

A half-negative and half-positive example is the following, where we don't want to lift the embedding $j : V \to M$ to $j^+ : V[G] \to M[H]$ as a class of V[G]. Instead, the following example tells us that sometimes we must go to a larger model, namely V[G * H] to find a generic over M to apply Generic Lifting (35 A • 2).

— 35 A • 4. Example -

- Let $V \vDash$ ZFC be a transitive model we can force over.
- Let U be a measure on κ in V.
- Let $j_U: V \to \operatorname{cUlt}^{V}(V, U)$ be the canonical embedding.
- Let G be $\mathbb{P} = \operatorname{Col}^{V}(\omega, < \kappa)$ -generic over V.
- Let $j(\mathbb{P}) = \mathbb{P} \times \mathbb{Q}$, and let H be \mathbb{Q} -generic over V[G].

Therefore in V[G * H] we can lift j_U to $j_U^+ : V[G] \to \operatorname{cUlt}^{\mathsf{V}}(V, U)[G * H]$ with critical point $\omega_1^{\mathsf{V}[G]}$

Proof ...

Work in V. Using Theorem 34 G • 8, $\mathbb{P} = \text{Col}(\omega, < \kappa)$ is isomorphic to the finite support product $\prod_{\alpha < \kappa} \text{Col}(\omega, \alpha)$ so that $j(\mathbb{P})$ is isomorphic to

$$\prod_{\alpha \le j(\kappa)} \operatorname{Col}(\omega, \alpha) \cong \prod_{\alpha < \kappa} \operatorname{Col}(\omega, \alpha) \times \prod_{\kappa \le \alpha < j(\kappa)} \operatorname{Col}(\omega, \alpha)$$

So set $\mathbb{Q} = \prod_{\kappa \le \alpha < j(\kappa)} \operatorname{Col}(\omega, \alpha)$ so that $j(\mathbb{P}) \cong \mathbb{P} \times \mathbb{Q}$. Now work in $V[G * H]$ (or $V[G \times H]$). We may also

^{xxxvi}One might worry that we can't lift because if κ isn't measurable, then it can't be the critical point of an embedding, but this only applies to embeddings that are classes of the model. After all, in V[G], κ isn't measurable, but the embedding $j : V \to \text{cUlt}^V(V, U)$ still is a class there.

regard $G \times H$ as $j(\mathbb{P})$ -generic over V by considering

 $\{f \in j(\mathbb{P}) : f \upharpoonright (\omega \times \kappa) \in G \land j \upharpoonright (\omega \times [\kappa, j(\kappa))) \in H\} \in V[G * H].$

It should be clear also that for $p \in G$, $j_U(p) = p$ in the set above, meaning j_U "G is contained in the above set that is $j(\mathbb{P})$ -generic over V and therefore over $\operatorname{cUlt}^V(V, U) \subseteq V$. Thus in V[G * H], by Generic Lifting (35 A • 2), we can lift j_U to get $j_U^+ : V[G] \to \operatorname{cUlt}^V(V, U)[G * H]$.

Note the similarity in hypotheses to Example 35 A • 4 but the difference in conclusion. Loosely speaking, the reason why we have such different conclusions is that G in Example 35 A • 4 is comprised of parts which are all bounded in κ . So when we stretch G by an elementary embedding, j_U "G = G doesn't cover too much information like it did in Example 35 A • 3, it just covers one of the parts bounded below κ .

Generalizing Example 35 A • 4 is actually quite difficult, because in general we have the following easy fact about how iterations work with elementary embeddings. For most practical applications, iterations are built up with small preorders at first and larger preorders later.

— 35 A • 5. Result -

• Let $V \vDash$ ZFC be a transitive model we can force over.

- Let $j: V \to M$ be an embedding with $M \subseteq V$ a transitive class and $cp(j) = \kappa$.
- Let $\mathbb{P} = \mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ be a κ -length iteration in \mathbf{V} such that each $\dot{\mathbb{Q}}_{\alpha} \in \mathbf{V}_{\alpha}^{\mathbf{V}}$.

Therefore $j(\mathbb{P}) = \mathbf{*}_{\alpha < i(\kappa)} \dot{\mathbb{Q}}_{\alpha} \cong \mathbb{P} * \dot{\mathbb{Q}}$ where each $\dot{\mathbb{Q}}_{\alpha} \in V_{\alpha}^{\mathsf{M}}$, and $\dot{\mathbb{Q}}$ is the tail iteration.

Proof .:.

This just follows by elementarity and the fact that $j \upharpoonright V_{\kappa} = id$ by Result 12 A • 3.

-

This gives an alternative perspective on what is required to lift an embedding $j : V \to M$. The proof given for Example 35 A • 3 is essentially that we can't lift because no generic could be compatible with G. But more than that, V[G] doesn't have any $j(\mathbb{P})$ -generics over M since any such generics would witness that $j(\kappa)$ is countable in V[G], which isn't true. Similarly, with Example 35 A • 4, $j(\mathbb{P})$ collapses κ to be countable, meaning V[G] cannot find such a generic over M.

A very useful result in forcing is the following. This tells us that if $j : V[G] \to N$ is elementary and N is transitive, then N = M[j(G)] for some transitive $M \models ZFC$, an inner model of N.

- 35A.6. Result (Definability of the Ground Model) Let $V \models ZFC$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a preorder and let G be \mathbb{P} -generic over V. Therefore V is a FOLp-definable class of V[G], in particular using just the parameter $P = \mathcal{O}(|\mathbb{P}|)^V$.

In fact, this definition is uniform across all generic extensions: there is some FOL-formula φ such that for any $\dot{x} \in V^{\mathbb{P}}$, $\mathbf{V} \models "\exists p \exists y \ (p \Vdash "\dot{x} = \check{y}") \leftrightarrow \mathbb{1}^{\mathbb{P}} \Vdash "\varphi(\dot{x}, \check{P})"".$

The proof of this result is interesting, but is left through guided exercise Defining the Ground Model (35 • Ex22). The point is that V[G] knows it's the generic extension of some class V so that by elementarity of $j : V[G] \to N, N$ knows it's the generic extension of some class we call M. Explicitly, there is some formula φ and $P \in V[G]$ such that $V = \{x \in V[G] : V[G] \models "\varphi(x, P)"\}$. Hence by elementarity, we can define $M = \{x \in N : N \models "\varphi(x, j(P))"\}$. Note that we always get that the restriction $j \upharpoonright V : V \to M$ is elementary assuming V is a class of $V[G]^{xxxvii}$ but this alone doesn't tell us $M \subset V$, and in some cases, M may not be contained in V.

Let us now move on to thinking about how ultrapower embeddings are lifted, since ultrapowers are the primary source of embeddings for us.

^{xxxvii}To see this, just use the defining formula φ as a class so by elementarity, $V[G] \models "\varphi$ defines a proper class $\land \varphi(x)$ " iff $M[j(G)] \models$ " φ defines a proper class $\land \varphi(j(x))$ " so that φ defines a class M in M[j(G)] by the same formula and whenever $x \in V$, $j(x) \in M$. Elementarity of $j \upharpoonright V : V \to M$ then follows from the elementarity in V[G] but restricting quantifiers by the defining formula of V.

— 35 A • 7. Theorem (Lifting Ultrapowers)

- Let $V \models$ ZFC be a transitive model we can force over.
- Let $\mathbb{P} \in V$ be a preorder appropriate for forcing.
- Let G be \mathbb{P} -generic over V.
- Let *E* be a (short) (κ, λ) -extender in **V**.
- Suppose the canonical embedding $j : V \to M = \operatorname{cUlt}_E^V(V)$ lifts to $j^+ : V[G] \to M[j^+(G)]$ in that $j^+(\tau_G) = j(\tau)_{j^+(G)}$ for all $\tau \in V^{\mathbb{P}}$.

Therefore $M[j^+(G)] = \operatorname{cUlt}_{E^*}^{V[G]}(V[G])$ where $E^* = E_{j^+}^{\lambda}$ and j^+ is the canonical extender embedding. Moreover, $E = E^* \cap V$.

Proof .:.

Using Corollary 13 B • 8, we first aim to show $M[j^+(G)] = \{j^+(f)(r) : r \in [\lambda]^{<\omega} \land f : [\kappa]^{<\omega} \to V[G]\}$. Note that any $\tau \in M^{\mathbb{P}}$ can be represented by $\tau = j(f)(r)$ for some $f : [\kappa]^{<\omega} \to V$ and $r \in [\lambda]^{<\omega}$. In fact, without loss of generality, we may assume im $f \subseteq V^{\mathbb{P}}$. As a result, in V[G], we can consider the function f' defined by $f'(r) = f(r)_G$ for all r. By elementarity, $j^+(f')$ and $j^+(f)$ have the same relationship with $j^+(G)$: $j^+(f')(r) = j^+(f)(r)_{j^+(G)}$. It follows that

$$\tau_{j^+(G)} = j(f)(r)_{j^+(G)} = j^+(f)(r)_{j^+(G)} = j^+(f')(r).$$

Obviously, since im $j^+ \subseteq M[j^+(G)]$ we have the other containment, which means we have equality. So by Corollary 13 B • 8, $M[j^+(G)] = \text{cUlt}_{E^*}^{V[G]}(V[G])$ where $E^* = E_{j^+}^{\lambda}$. The fact that j^+ lifts $j = j_E$ shows us that $E^* \cap V = E$: for any $A \in \mathcal{O}([\lambda]^{<\omega}) \cap V$ and $r \in [\lambda]^{<\omega}$, $\langle r, A \rangle \in E$ iff $r \in j(A)$ iff $r \in j^+(A)$ iff $\langle r, A \rangle \in E^*$. \dashv

It's important to note that the converse of this need not hold in general. In fact, just because there's an embedding $j: V[G] \to M[j(G)]$ with $M[j(G)] \subseteq V[G]$, we don't even need that $M \subseteq V$.

Example 35 A • 3 is an example of when we can't find a lift-up, and generalizes to $Col(\alpha, \kappa)$ for any $\alpha < \kappa$. Example 35 A • 4 gives an example where we can lift $j : V \to M$, but only in a larger model that has access to a generic over M. But in many cases we can find lift-ups in the generic extension. The easiest examples of this are when the preorders we're using are sufficiently closed.

— 35 A • 8. Example

- Let $V \vDash$ ZFC be a transitive model we can force over.
- Let κ be measurable in **V** as witnessed by a measure $U \in V$.
- Let $j: V \to \operatorname{cUlt}^{V}(V, U)$ be the canonical embedding j_U .
- Let $\mathbb{P} \in V$ be $\leq \kappa$ -closed in \mathbf{V} .
- Let G be \mathbb{P} -generic over V.

Therefore U is still a measure in V[G] and $j_U^{V[G]}$ lifts j_U^V , meaning that $j_U^{V[G]} \upharpoonright V = j_U^V$ and $\operatorname{cUlt}^{V[G]}(V[G], U) = \operatorname{cUlt}^V(V, U)[j_U^{V[G]}(G)]$.

Proof .:.

Work in V. Since \mathbb{P} is $\leq \kappa$ -closed, it doesn't add any new subsets of κ and hence $\mathcal{P}(\kappa)^{V[G]}$. Thus U is still a measure in V[G]. To lift j_U^{V} , we consider $j_U^{V[G]}$, the ultrapower map of V[G] by U. We cannot make use of Lifting Ultrapowers (35 A • 7) until we know we can lift j_U^{V} to $j_U^{V[G]}$. But assuming we have this, it follows immediately from Lifting Ultrapowers (35 A • 7) that $\operatorname{cUlt}^{V[G]}(V[G], U) = \operatorname{cUlt}^{V}(V, U)[j_U^{V[G]}(G)]$. So we focus on showing $j_U^{V[G]} \upharpoonright V = j_U^{V}$.

Again, the $\leq \kappa$ -closure of \mathbb{P} means \mathbb{P} adds no new sequences from κ to V. So in particular, $[f]_U$ and $trcl([f]_U)$ are interpreted the same way in the ultrapowers $\mathsf{Ult}^{\mathsf{V}[G]}(\mathsf{V}[G], U)$, $\mathsf{Ult}^{\mathsf{V}[G]}(\mathsf{V}, U)$, and $\mathsf{Ult}^{\mathsf{V}}(\mathsf{V}, U)$ whenever f:

$$\kappa \to V$$
. In particular, $\mathsf{Ult}^{\mathsf{V}[G]}(\mathsf{V}, U) = \mathsf{Ult}^{\mathsf{V}}(\mathsf{V}, U)$ and the collapsing maps
 $\pi_0 : \mathsf{Ult}^{\mathsf{V}[G]}(V[G], U) \to \mathsf{cUlt}^{\mathsf{V}[G]}(V[G], U)$
 $\pi_1 : \mathsf{Ult}^{\mathsf{V}}(V, U) \to \mathsf{cUlt}^{\mathsf{V}}(V, U).$
satisfy $\pi_0 \upharpoonright \mathsf{Ult}^{\mathsf{V}[G]}(V, U) = \pi_1$ because of the uniqueness of collapsing maps. As a result, for $x \in V$,
 $j_U^{\mathsf{V}[G]}(x) = \pi_0([\mathsf{const}_x]_U) = \pi_1([\mathsf{const}_x]_U) = j_U^{\mathsf{V}}(x)$ and so $j_U^{\mathsf{V}[G]} \upharpoonright V = j_U^{\mathsf{V}}$.

Note that above, we might not have $G \subseteq j_U^{\mathsf{V}[G]}(G)$.

Closure isn't actually necessary through. The great part of $\leq \kappa$ -closure is that we don't really affect κ at all in the generic extension. A harder, but more *useful* example is the following. Here we again don't really affect κ , but in this case because our forcing is small enough to not rock the boat too much.

— 35 A • 9. Theorem (Lévy–Solovay) –

- Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over.
- Let κ be measurable in V as witnessed by a measure $U \in V$.
- Let $\mathbb{P} \in V$ be a preorder of size $|\mathbb{P}|^{V} < \kappa$.
- Let G be \mathbb{P} -generic over V.

Therefore in V[G],

$$U^+ = \{X \in \mathcal{P}(\kappa) : \exists Y \in U \ (Y \subseteq X)\} \in V[G]$$

is a measure on κ and the canonical embedding j_{U^+} is a lift-up of j_U , meaning both $j_{U^+} \upharpoonright V = j_U$ and also $\operatorname{cUlt}^{\mathsf{V}[G]}(V[G], U^+) = \operatorname{cUlt}^{\mathsf{V}}(V, U)[j_{U^+}(G)].$

Proof ...

First we show U^+ is a measure. That it's a filter should be clear: $\emptyset \notin U^+$ since $\emptyset \notin U$; it's trivially closed upwards; and if $X, Y \in U^+$ as witnessed by $X' \subseteq X, Y' \subseteq Y$ with $X', Y' \in U$, then $X \cap Y \supseteq X' \cap Y' \in U$. U^+ is also trivially non-principal because U is. Now for any $X \in \mathcal{O}(\kappa) \cap V[G]$ with name \dot{X} , for each $p \in \mathbb{P}$, let

$$X_p = \{ \alpha < \kappa : p \Vdash ``\check{\alpha} \in \dot{X}" \}$$

so that each $p \Vdash ``\check{X}_p \subseteq \dot{X} \cap \check{\kappa}"$.

— Claim 1 ——

 U^+ is an ultrafilter.

Proof .:.

Let $X \in \mathcal{P}(\kappa) \cap V[G]$ and $p \in \mathbb{P}$ be arbitrary. If $X_p \in U$, then $p \Vdash ``\dot{X} \cap \check{\kappa} \in \dot{U}^+$." On the other hand, if $X_p \notin U$ then $\kappa \setminus X_p \in U$. Note that

$$\kappa \setminus X_p = \{ \alpha < \kappa : p \not\Vdash ``\check{\alpha} \in \dot{X}"\} = \{ \alpha < \kappa : \exists q \leq p \ (q \Vdash ``\check{\alpha} \notin \dot{X}")\} = \bigcup_{q \leq p} (\kappa \setminus X)_q.$$

Since $|\mathbb{P}| < \kappa$, by the κ -completeness of U, this means some $(\kappa \setminus X)_q \in U$. But q forces that this is contained in $\kappa \setminus X$ so that $q \Vdash "\kappa \setminus X \in U^+$ ". This means for every $p \in \mathbb{P}$,

$$p \not\Vdash ``X \notin U^+ \land \check{\kappa} \setminus X \notin U^+$$

But this means $\mathbb{1}^{\mathbb{P}}$ forces the negation, i.e. U^+ is an ultrafilter.

So now we must show U^+ is κ -complete and normal. The κ -completeness of U^+ does not so trivially follow from the κ -completeness of U: $\{X_{\alpha} \in U^+ : \alpha < \gamma\}$ for $\gamma < \kappa$ should be witnessed by $\{Y_{\alpha} \in U : \alpha < \gamma\} \subseteq V$, but this set need not be in V.

— Claim 2 –

 U^+ is κ -complete.

 \neg

Proof .:.

Suppose not. Let $X = \{x_{\alpha} \in U^+ : \alpha < \gamma\}$ be a $< \kappa$ -sized family of sets in U^+ such that $\bigcap X \notin U^+$. Let \dot{U}^+ be a name for U^+ , let \dot{X} be a name for X, and let $p \in G$ be such that

 $p \Vdash ``|X| = \check{\gamma} < \check{\kappa} \land \dot{X} \subseteq \dot{U}^+ \land \bigcap \dot{X} \notin \dot{U}^+ ``.$

Consider now $\{(x_{\alpha})_p : \alpha < \gamma\}$. If any $(x_{\alpha})_p \notin U$, then the idea from Claim 1 tells us $(\kappa \setminus x_{\alpha})_q \in U$ for some $q \leq^{\mathbb{P}} p$, and thus below $q, \kappa \setminus x_{\alpha} \in U^+$. In other words, $q \Vdash ``\dot{x}_{\alpha} \notin \dot{U}^+$, contradicting that $q \leq^{\mathbb{P}} p \Vdash ``\dot{X} \subseteq \dot{U}^+$. Thus assume each $(x_{\alpha})_p \in U$ and therefore $\bigcap_{\alpha < \gamma} (x_{\alpha})_p \in U$ witnesses that $\bigcap X \in U^+$.

- Claim 3 - U^+ is normal.

Proof .:.

Suppose not. Let \dot{f} be a name for a function with $p \Vdash ``\dot{f} : \check{\kappa} \to \check{\kappa}$ is regressive" but is forced to never be U^+ -almost-constant: $p \Vdash ``\neg \exists \beta \forall^*_{\dot{U}^+} \alpha \ (\dot{f}(\alpha) = \beta)"$. For each $\beta < \kappa$, let $p_\beta \leq^{\mathbb{P}} p$ decide the values of \dot{f} on a set $X_\beta \in U$ so that

 $X_{\beta} = \{ \alpha < \kappa : \exists \gamma < \kappa \ p_{\beta} \Vdash ``\dot{f}(\check{\alpha}) = \check{\gamma}" \} \text{ and } p_{\beta} \Vdash ``\forall \alpha \in \check{X}_{\beta} \ (\dot{f}(\alpha) \neq \beta)".$ Since there are only $|\mathbb{P}| < \kappa$ many such $p_{\beta}s$, it follows that there are less than κ many $X_{\beta}s$ and so by Claim

2, $\bigcap_{\beta < \kappa} X_{\beta} \in U$. So define

$$F(\alpha) = \begin{cases} \beta & \text{if } \alpha \in \bigcap_{\xi < \kappa} X_{\xi} \land p_0 \Vdash ``\dot{f}(\check{\alpha}) = \check{\beta}``\\ 0 & \text{otherwise.} \end{cases}$$

It follows that $F : \kappa \to \kappa$ is regressive in V, but for each $\beta < \kappa$, $\{\alpha < \kappa : F(\alpha) \neq \beta\} \supseteq \bigcap_{\xi < \kappa} X_{\xi} \in U$, contradicting the normality of U.

Now we move on to showing that j_{U^+} is a lift-up of j_U . To do this, since j_{U^+} maps into $\text{cUlt}^{V[G]}(V[G], U^+)$ while j_U maps into $\text{cUlt}^{V}(V, U)$, we need to have some way of comparing both $[\text{const}_x]_U$ and $[\text{const}_x]_{U^+}$ and the collapses of these. We first show $\text{cUlt}^{V}(V, U) = \text{cUlt}^{V[G]}(V, U^+)$ and then proceed from there.

For every $f: \kappa \to V$ in V[G], there is some $g: \kappa \to V$ in V such that $f \approx_{U^+} g$.

Proof .:.

- Claim 4 -

Let $p \in \mathbb{P}$ and let \dot{f} be a \mathbb{P} -name for a κ -sequence such that for every $\alpha < \kappa$, $p \Vdash ``\dot{f}(\check{\alpha}) \in \check{V}$ '', meaning for each $\alpha < \kappa$,

 $D_{\alpha} = \{q \leq^{\mathbb{P}} p : \exists x \ q \Vdash ``\dot{f}(\check{\alpha}) = \check{x}"\}$ is dense in \mathbb{P} .

With that technicality out of the way, for each $\alpha < \kappa$, let $p_{\alpha} \in D_{\alpha}$. Most of these p_{α} s will be the same since $|\mathbb{P}| < \kappa$. More precisely to work with U, for each $p \in \mathbb{P}$, consider $\{\alpha < \kappa : p_{\alpha} = p\}$. These sets partition κ into $< \kappa$ -many sets so one of them must be in U. Hence there is some $p^* \in \mathbb{P}$ such that $p_{\alpha} = p^*$ for U-almost every α . Now define $g(\alpha) = x$ whenever $p^* \Vdash ``f(\alpha) = \check{x}``$ and $g(\alpha) = 0$ otherwise. In this case, it follows that $g \in V$ and $\forall_{U+}^* \alpha$ ($f(\alpha) = g(\alpha)$), as witnessed by A as desired. \dashv

Now we can show that $j_{U^+} \upharpoonright V$ is itself a canonical ultrapower embedding.

− Claim 5 j_{U^+} " $V \subseteq \text{cUlt}^{V[G]}(V, U^+)$. In fact, $j_{U^+} \upharpoonright V$ is the ultrapower map from V to $\text{cUlt}^{V[G]}(V, U^+)$.

Proof .:.

If $x \in V$, $[\operatorname{const}_x]_{U^+} \in \operatorname{Ult}^{V[G]}(V[G], U^+) \cap \operatorname{Ult}^{V[G]}(V, U^+)$. So it suffices to show that the collapses of these are the same. Note that for any $f : \kappa \to V[G]$ and $g : \kappa \to V$, if $\operatorname{Ult}^{V[G]}(V[G], U^+) \models$ " $[f]_{U^+} \in [g]_{U^+}$ " then without loss of generality, im $f \subseteq \operatorname{trcl}(\operatorname{im} g) \subseteq V$ and hence $f : \kappa \to V$ which witnesses that $\operatorname{Ult}^{V[G]}(V, U^+) \models$ " $[f]_{U^+} \in [g]_{U^+}$ ", and the converse of these clearly holds. As a result, $j_{U^+} \upharpoonright V$ maps into $\operatorname{cUlt}^{V[G]}(V, U^+)$.

It follows that taking the ultrapower of V by U^+ in V[G] doesn't change anything from the ultrapower in V in the following sense. Let π_{U^+} : Ult^{V[G]} $(V, U^+) \rightarrow \text{cUlt}^{V[G]}(V, U^+)$ and π_U : Ult^{V(V, U)} $\rightarrow \text{cUlt}^{V(V, U)}$ be the collapsing maps.

- $\mathsf{Ult}^{\mathsf{V}[G]}(\mathsf{V}, U^+) \cong \mathsf{Ult}^{\mathsf{V}}(\mathsf{V}, U)$ as witnessed by φ sending $[f]_{U^+} = [g]_{U^+}$ to $[g]_U$ for f, g as in Claim 4.
- $\pi_{U^+}([g]_{U^+}) = \pi_U([g]_U)$ whenever $g : \kappa \to V$ is in V, because $\pi_U \circ \varphi$ collapses $\text{Ult}^{V[G]}(V, U^+)$ and so does π_{U^+} . By the uniqueness of the collapsing map, $\pi_{U^+} = \pi_U \circ \varphi$.

As a result, whenever $x \in V$, $j_{U^+}(x) = \pi_{U^+}([\text{const}_x]_{U^+}) = \pi_U([\text{const}_x]_U) = j_U(x)$. So $j_{U^+} \upharpoonright V = j_U$. By Lifting Ultrapowers (35 A • 7), $\text{cUlt}^{\vee}(V, U)[j_{U^+}(G)] = \text{cUlt}^{\vee[G]}(V[G], U^+)$.

The original version of Lévy–Solovay (35 A • 9) in fact says much more [21]: every measure in V[G] on κ is generated from a measure in V as above. In a sense, Lévy–Solovay (35 A • 9) tells us that measurability and most large cardinal properties are unaffected by small forcing. The theorem also generalizes in various ways showing that small forcing is relatively harmless to large cardinals, as practically all large cardinal notions are preserved by small preorders. This will be shown later in the subsection on gap forcing.

What distinguishes the examples where we can lift, Lévy–Solovay (35 A • 9) and Example 35 A • 8, from the example where we can't, Example 35 A • 3, is the fact that the preorders used in the positive examples *avoid* κ . It is this sense of having a "gap" at κ that we will exploit in the subsection on gap forcing, dealing primarily with the non-introduction of large cardinals.

§ 35 B. Strategic closure and finding generics

Another way of finding generics is merely to *generate* them by transfering the original generic. This only occurs under special circumstances, but those circumstances are common enough by ultrapowers that its inclusion is warranted, generalizing Example 35 A • 8 and giving a better idea of what $j_U^{V[G]}(G)$ should be.

- 35 B•1. Theorem

- Let $V \vDash$ ZFC be a transitive model we can force over.
- Let E be a (κ, λ) -extender in V with $j_E : V \to \operatorname{cUlt}_E^V(V)$ the canonical embedding.
- Let $\mathbb{P} \in V$ be a preorder appropriate for forcing that is $\leq \kappa$ -distributive (e.g. $\leq \kappa$ -closed).
- Let G be \mathbb{P} -generic over V.

Therefore the filter generated by $j_E "G$, i.e. $H = \{p \in j_E(\mathbb{P}) : \exists q \in j_E "G \ (q \leq p)\}$, is $j_E(\mathbb{P})$ -generic over $\operatorname{cUlt}_E^{\mathsf{V}}(V)$, and thus j_E lifts to $j^+ : V[G] \to \operatorname{cUlt}_E^{\mathsf{V}}(V)[H]$.

Proof .:.

It's clear that *H* is a filter, so let $D \in M = \text{cUlt}_E^{\vee}(V)$ be open dense in $j(\mathbb{P})$. We can write D = j(f)(r) for some $r \in [\lambda]^{<\omega}$ and $f : [\kappa]^{<\omega} \to V$. Without loss of generality, f(s) is open dense in \mathbb{P} for each $s \in [\kappa]^{<\omega}$. Since \mathbb{P} is $\leq \kappa$ -distributive, $\bigcap_{s \in [\kappa]^{<\omega}} f(s) \neq \emptyset$ is open dense so there is some element $p \in G \cap \bigcap_{s \in [\kappa]^{<\omega}} f(s)$. Thus $j_E(p) \in j^{"}G$ and by elementarity, $\forall s \in [\lambda]^{<\omega} (j_E(p) \in j_E(f)(s))$. So in particular, $j_E(p) \in j_E(f)(r) = D$ and so $\emptyset \neq j^{"}G \cap D \subseteq H \cap D$. The fact that we can lift j_E to j^+ follows from Generic Lifting (35 A • 2). \dashv

We can go the opposite direction under less restrictive circumstances too.

─ 35B•2. Theorem

- Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over.
- Let $j: V \to M$ be an elementary embedding with M transitive and $cp(j) = \kappa$.
- Let $\mathbb{P} \in V$ be κ -cc in \mathbf{V} .
- Let H be $j(\mathbb{P})$ -generic over M.

Therefore j^{-1} "*H* is P-generic over *V*.

Proof .:.

That j^{-1} "*H* is a filter is easy: if $p^* \in j^{-1}$ "*H* and $p^* \leq^{\mathbb{P}} p$, then by elementarity, $j(p^*) \leq^{\mathbb{P}} j(p)$ so as a filter, $j(p) \in H$ and hence $p \in j^{-1}$ "*H*. Compatibility of elements also follows easily by elementarity. For genericity, we use antichains instead of dense sets as in Theorem 32 C • 5. Let $A \in V$ be a maximal antichain in \mathbb{P} so that $|A| < \kappa$ and hence j"A = j(A) is an antichain of $j(\mathbb{P})$. There is then some $p \in H \cap j$ "A and as $p \in j$ "A, p = j(q) for some $q \in \mathbb{P}$ so that $q \in j^{-1}$ " $H \cap A$ as desired.

These two results are useful for *transfering* generics one of two ways along an elementary embedding, but they merely reuse the same generic. The issue is that if $j(\mathbb{P})$ is fairly different from \mathbb{P} , by being a longer iteration for example, then the two ideas don't work because we aren't able to find generics over the later iterations. The main means of actually finding generics is the same idea as with Corollary 31 D • 3, but generalized to uncountable cardinals. All we require is a sufficient amount of closure and sufficiently few dense sets or antichains.

- 35 B • 3. Theorem

- Let $V \models$ ZFC be a transitive model we can force over.
- Let $\mathbb{P} \in V$ be a preorder appropriate for forcing.
- Let $\kappa \in \text{Ord.}$
- Suppose \mathbb{P} is $< \kappa$ -closed and $\{D \in V : D \text{ is dense}\}$ has size $\leq \kappa$ (ostensibly, not in V).

Therefore there is a G \mathbb{P} -generic over V, and in fact 2^{κ} -many such generics.

We will give a formal proof of a generalization of this later, but the idea is essentially the same as with Corollary 31 D • 3: just continually extend conditions $p_{\alpha} \in D_{\alpha}$ for $\alpha < \kappa$ for D_{α} open dense. By closure we can continue to do this, and since there are sufficiently few dense sets (in the real world), the resulting choices $\{p_{\alpha} : \alpha < \kappa\}$ generate a generic.

The easiest example of this type of reasoning is simply to collapse everything in sight down to just above the closure of a preorder.

— 35 B•4. Example -

- Let $V \models$ ZFC be a transitive model we can force over.
- Let $\mathbb{P} \in V$ be a < λ -closed preorder appropriate for forcing in V for some λ .
- Let *H* be $\mathbb{Q} = \operatorname{Col}(\lambda, 2^{|\mathbb{P}|})^{\mathsf{V}}$ -generic over *V*.

Therefore in V[H], there is a G \mathbb{P} -generic over V.

Proof .:.

Because \mathbb{Q} is $<\lambda$ -closed, it doesn't add any $<\lambda$ -sized sequences and thus \mathbb{P} remains $<\lambda$ -closed in V[H]. Since \mathbb{P} is appropriate for forcing, in $V, \lambda \leq |\mathbb{P}| \leq 2^{|\mathbb{P}|}$. In V we also trivially have at most $2^{|\mathbb{P}|}$ -many distinct dense sets of \mathbb{P} . Hence in V[H], \mathbb{P} has $<\lambda$ -closure with at λ -many dense sets of V. So we may apply Theorem 35 B • 3 to get a G \mathbb{P} -generic over V in V[H].

Generally, we use Theorem 35 B • 3 while working inside a generic extension of V, say V[G], and we use this to find an $H \in V[G]$ generic over M to lift; where $j : V \to M$ and $j"G \subseteq H$. Sometimes we cannot lift so directly, and instead must move to a larger model, lifting $j : V \to M$ to $j^+ : V[G] \to M[H]$ as a class of V[G * H] rather than V[G] as with Example 35 A • 4.

The generalization of Theorem $35 \text{ B} \cdot 3$ uses a concept of strategic closure of a poset, which is a kind of generalization of regular closure in a way that encompasses many more posets without compromising most of the arguments used with such posets.

Let \mathbb{P} be a preorder and κ, λ ordinals. The game $\mathscr{G}^{\lambda}_{\mathbb{P}}$ is the two person game of length $\leq \lambda$

I:
$$p_0 = \mathbb{1}^{\mathbb{P}}$$
 $p_2 \in \mathbb{P}$ \cdots p_{ω} $p_{\omega+2}$ \cdots
II: $p_1 \in \mathbb{P}$ p_3 \cdots $p_{\omega+1}$ \cdots ,

with the rules that $p_0 = \mathbb{1}^{\mathbb{P}}$ and that $p_{\alpha} \leq^{\mathbb{P}} p_{\beta}$ for $\beta \leq \alpha < \lambda$. Here I plays p_{α} for even $\alpha < \lambda$ (including limit α) and II plays $p_{\alpha} \in \mathbb{P}$ for odd $\alpha < \lambda$. The first player to break a rule loses, and if no one breaks a rule, then I wins. In other words, I wins iff the game can continue.

- \mathbb{P} is $\leq \kappa$ -strategically closed iff I has a winning strategy in $\mathcal{G}_{\mathbb{P}}^{\kappa+1}$.
- \mathbb{P} is κ -strategically closed iff I has a winning strategy in $\mathcal{G}_{\mathbb{P}}^{\kappa}$.
- \mathbb{P} is $< \kappa$ -strategically closed iff \mathbb{P} is $\leq \alpha$ -strategically closed for all ordinals $\alpha < \kappa$.
- \mathbb{P} is $\ll \kappa$ -strategically closed iff \mathbb{P} is $\leq \lambda$ -strategically closed for all cardinals $\lambda < \kappa$.

Strategic closure is a weakening of the usual closure of preorders:

- $\leq \kappa$ -closure corresponds to $\leq \kappa$ and $< \kappa^+$ -strategic closure.
- < κ -closure corresponds to $\ll \kappa$, < κ , and κ -strategic closure in increasingly stronger senses.

Basically, whereas $< \kappa$ -closure ensures we can always extend a $\leq^{\mathbb{P}}$ -decreasing sequence $\langle p_{\alpha} : \alpha < \lambda \rangle$ whenever $\lambda < \kappa$, κ -strategic closure only allows us to state this whenever the even entries of the sequence conform to a certain strategy by player I. Put in another sense, closure gives total freedom to choose a decreasing sequence (of appropriate length) and find something below it. Strategic closure only gives control over half of the sequence, relying on I's strategy to extend at limits.

Additionally, $\leq |\alpha|$ -strategic closure is ostensibly weaker than α -strategic closure since the game $\mathscr{G}_{\mathbb{P}}^{\alpha}$ could be much longer than $\mathscr{G}_{\mathbb{P}}^{|\alpha|+1}$. It's clear that we can play the game multiple times, which gives some additional strength: $\leq \kappa$ -strategic closure implies $\leq \kappa + \kappa$ -strategic closure and $< \kappa \cdot \omega$ -strategic closure. Going beyond this is more difficult, and it's unclear to me whether $\leq \kappa$ -strategic closure implies $\kappa \cdot \omega$ -strategic closure and more generally $< \kappa^+$ -strategic closure.

Usually, in proving things with strategic closure, we—the theorem provers—play the role of **II**, relying on **I**'s strategy to clean up our mess, especially at limit stages. Despite their differences, the similarities between $< \kappa$ -closed and κ -strategically closed preorders are enough to show to allow some arguments about $< \kappa$ -closure to go through about κ -strategic closure. For example, κ -strategically closed preorders are $\leq \kappa$ -distributive.

– 35 B•6. Corollary

```
For any infinite cardinal \kappa, if \mathbb{P} is \ll \kappa-strategically closed, then \mathbb{P} is < \kappa-distributive.
```

Proof .:.

Let \mathcal{D} be a collection of open, dense sets. Clearly $\bigcap \mathcal{D}$ is open, so it suffices to show it's dense. Let $p \in \mathbb{P}$ be arbitrary. Enumerate $\mathcal{D} = \{D_{\alpha} : \alpha < \lambda\}$ where $|\mathcal{D}| = \lambda < 2\lambda + 1 < \kappa$. Because we only have access to half of the extensions, we must work half as slow, accomplishing the same proof as Result 33 B • 3, but requiring a sequence of length 2λ with a condition beneath it instead of a sequence of length λ with a condition beneath it. In other words, to make space for **I**'s moves that we don't care about, we need $2\lambda + 1$ steps instead of $\lambda + 1$ steps.

Let τ be a strategy for **I** in $\mathscr{G}_{\mathbb{P}}^{2\lambda+1}$, and take $p_0 = \mathbb{1}^{\mathbb{P}}$. Assuming **I** plays according to τ , we prove by induction $p_{2\alpha} \in D_{\xi}$ for every $\xi < \alpha$. (*)

This is trivial for $\alpha = 0$. For limit α this is also easy: inductively, $p_{2\alpha} \leq p_{2\xi} \in D_{\xi}$ for $\xi < \alpha$ so $p_{2\alpha} \in \bigcap_{\xi < \alpha} D_{\xi}$

since each D_{ξ} is open. For $\alpha + 1$, let **II** play $p_{2\alpha+1} \in D_{\alpha} \cap \mathbb{P}_{\leq p_{2\alpha}} \cap \mathbb{P}_{\leq p}$ which exists by density of D_{α} . Thus $p_{2\alpha+1} \in \bigcap_{\xi \leq \alpha} D_{\xi}$. Since τ is a strategy and **I** plays according to τ , $\tau(\langle p_{\xi} : \xi \leq 2\alpha + 1 \rangle) = p_{2\alpha+2} \leq p_{2\alpha+1}$ so that $p_{2(\alpha+1)} \in \bigcap_{\xi < \alpha} D_{\xi}$ and thus (*) holds for $\alpha + 1$. This proves the induction that $p_{2\alpha} \in \bigcap_{\xi < \alpha} D_{\xi}$ for

every $\alpha \leq \lambda$. In particular, $p_{2\lambda} \in \bigcap_{\alpha < \lambda} D_{\alpha} \cap \mathbb{P}_{\leq p}$.

In paticular, considering $\kappa = \delta^+$, $\leq \delta$ -strategic closure implies $\leq \delta$ -distributivity.

– 35 B•7. Corollary –

For any limit cardinal κ , if \mathbb{P} is $\ll \kappa$ -strategically closed, then \mathbb{P} is $< \kappa$ -distributive.

Proof .:.

If κ is a limit, let \mathcal{D} be a collection of open, dense sets of size $|\mathcal{D}| < \kappa$. By $\ll \kappa$ -strategic closure, we have $\leq |\mathcal{D}|^+$ -strategic closure and hence $\bigcap \mathcal{D} \neq \emptyset$ is open, dense by $< |\mathcal{D}|^+$ -distributivity from Corollary 35 B • 6. \dashv

In general, $< \kappa$ -distributivity is indeed distinct from κ -strategic closure.^{xxxviii} But we continue to show its similarities with the usual notion of closure. In particular, we have the same sort of result with iterations as in Inverse Limit Closure (34 D • 7).

— 35 B • 8. Theorem (Inverse Limit Strategic Closure) –

Let κ, λ be ordinals. Let $\mathbf{*}_{\alpha < \lambda} \dot{\mathbb{Q}}_{\alpha}$ be a λ -stage iteration with support in $I \subseteq \mathcal{P}(\lambda)$, an ideal or $\mathcal{P}(\lambda)$ itself such that

- inverse limits are taken at every limit stage $\alpha \leq \lambda$ with $cof(\alpha) < \kappa$;
- we take either direct or inverse limits at all other limit stages;
- $\mathbb{1}_{\alpha} \Vdash ``\dot{\mathbb{Q}}_{\alpha}$ is $< \check{\kappa}$ -strategically closed" for each $\alpha < \lambda$; and

Therefore $\mathbf{a}_{\alpha < \lambda} \dot{\mathbf{Q}}_{\alpha}$ is $< \kappa$ -strategically closed.

Proof .:.

Proceed by induction on λ . $\lambda = 0$ is trivial as the only sequences are constant 1 sequences. We only show the successor case as the limit case is similar to Inverse Limit Closure (34 D • 7) but made merely more complicated by the introduction of (canonical) names of strategies. The successor case deals with this slight complication in a much more managable way. We proceed similarly to Lemma 34 D • 6. So work with $\mathbb{P} * \dot{\mathbb{Q}}$ which, to those attached to the induction, can be thought of inductively as $\mathbb{P} = \bigstar_{\alpha < \lambda} \dot{\mathbb{Q}}_{\alpha}$ and $\dot{\mathbb{Q}} = \dot{\mathbb{Q}}_{\lambda+1}$.

Let $\alpha < \kappa$ be fixed. We aim to define a strategy for I in $\mathscr{G}^{\alpha}_{\mathbb{P}*\dot{\mathbb{Q}}}$, so suppose II plays some sequence $\langle \langle p_{\xi}, \dot{q}_{\xi} \rangle : \xi < \gamma \rangle$ in this game, where γ is such that after including I's turns, we get a full α -length sequence. In the game $\mathscr{G}^{\alpha}_{\mathbb{P}}$, I wins with some strategy σ giving

$$\mathcal{G}^{\alpha}_{\mathbb{P}} \qquad \begin{array}{ccc} \mathbf{I}: & \mathbb{1}^{\mathbb{P}} & \sigma(\langle p_0 \rangle) & \sigma(\langle p_0, p_1 \rangle) \cdots & \sigma(\langle p_{\xi} : \xi < \gamma \rangle) = p^* \\ \mathbf{II}: & p_0 & p_1 & \cdots \end{array}$$

Thus in $\mathscr{G}^{\alpha}_{\mathbb{P}} * \dot{\mathbb{Q}}$, the strategy for I to choose the first coordinates will be just to use σ while ignoring $\dot{\mathbb{Q}}$. To choose the second coordinates for I, let $\dot{\tau} \in V^{\mathbb{P}}$ be such that $\mathbb{1}^{\mathbb{P}} \Vdash ``\dot{\tau}$ is winning for I in $\mathscr{G}^{\check{\alpha}}_{\check{Q}}$. Without loss of generality, since II loses otherwise, $\langle \langle p_{\xi}, \dot{q}_{\xi} \rangle : \xi < \gamma \rangle$ is $\leq^{\mathbb{P}*\dot{\mathbb{Q}}}$ -decreasing which means for $\xi < \zeta < \gamma$, $p_{\zeta} \Vdash ``\dot{q}_{\zeta} \leq^{\dot{\mathbb{Q}}} \dot{q}_{\xi}$. Hence I may play according to $\dot{\tau}$ as determined by the play by II thus far: there is some canonical name for $\dot{\tau} (\langle \dot{q}_{\xi} : \xi < \zeta \rangle)$ for each $\zeta \leq \gamma$, and since $\dot{\tau}$ is forced to win, the condition p^* below each p_{ξ} forces that this name is below each \dot{q}_{ξ} and below the previous plays by I.

So I plays these canonical names for the second coordinate. To sum up, I plays according to this strategy: using σ for the first coordinates, and canonical names for the outputs of \dot{t} (as determined by the first coordinate) for the second coordinates. This is clearly winning for I. Since $\alpha < \kappa$ was arbitrary, we get $< \kappa$ -strategic closure for $\mathbb{P} * \dot{\mathbb{Q}}$.

^{xxxviii}The example of this, which hasn't been defined here yet, is "shooting a club" in (a stationary, co-stationary subset of) ω_1 , which is $< \aleph_1$ distributive, but not \aleph_1 -strategically closed. Indeed, shooting a club isn't even $\leq \aleph_0$ -strategically closed: $\leq \aleph_0$ -strategically closed preorders
preserve stationary subsets of \aleph_1 with nearly the same proof as with $\leq \aleph_0$ -closed preorders. Trivially all separative preorders will be \aleph_0 -strategically
closed, so this is the worst we can possibly fail to have any of the above strategic closure properties. This also highlights the difference between $< \kappa$ -distributivity and $< \kappa$ -closure.

By the same proof, we also get the above result with κ -strategic closure replacing the occurrences of $< \kappa$ -strategic closure. This also gives, by the exact same proof, the analogue of Result 34 E • 8 for strategic closure.

- 35 B • 9. Corollary

Let $\mathbf{k}_{\xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ be a κ -stage iteration with support in some $I \subseteq \mathbb{P}(\kappa)$ —a non-principal ideal or $\mathbb{P}(\kappa)$ itself—such that for some $\alpha < \kappa$ and some λ ,

- inverse limits or direct limits are taken at every limit stage;
- inverse limits are taken at every limit stage $\geq \alpha$ of cofinality $< \lambda$;
- $\mathbb{1}_{\xi} \Vdash ``\dot{\mathbb{Q}}_{\xi}$ is $< \lambda$ -strategically closed" for every $\alpha \leq \xi < \kappa$; and
- $\mathbf{*}_{\xi < \alpha} \dot{\mathbf{Q}}_{\xi}$ doesn't collapse any cofinalities $\geq \lambda$ to be $< \lambda$.

Therefore $\mathbb{1}_{\alpha} \Vdash :: \mathbf{k}_{\alpha \leq \xi < \kappa} \dot{\mathbb{Q}}_{\xi}$ is $< \lambda$ -strategically closed".

Now we actually show the theorem we care about. Actually we could have shown this earlier, but the exploration of strategically closed preorders is useful for easton support iterations in a bit more generality than purely closed preorders.

- 35 B • 10. Theorem (Closure Giving Generics)

- Let $V \vDash$ ZFC be a transitive model we can force over.
- Let $\mathbb{P} \in V$ be a preorder appropriate for forcing.
- Suppose \mathbb{P} is κ -strategically closed and $\{D \in V : D \text{ is dense}\}$ has size $\leq \kappa$ (ostensibly, not in V) for some cardinal κ .

Therefore for each $p \in \mathbb{P}$, there is a G \mathbb{P} -generic over V with $p \in G$, and in fact 2^{k} -many such generics.

Proof .:.

Proceed as in Corollary 31 D • 3. Let $p \in \mathbb{P}$ be arbitrary. Enumerate $\{D \in V : D \text{ is dense}\} = \{D_{\alpha} : \alpha < \kappa\}$. As a cardinal, if $\alpha < \kappa$ then $2\alpha < \kappa$. This means that we can deal with every D_{α} in a game just of length κ while only having access to half the moves. More precisely, let τ be a strategy for I in $\mathscr{G}_{\mathbb{P}}^{\kappa}$. Let I play according to this strategy and for $p_{2\alpha}$ defined by I, let II play an arbitrary $p_{2\alpha+1} \in D_{\alpha} \cap \mathbb{P}_{\leq p_{2\alpha}} \cap \mathbb{P}_{\leq p}$. Since τ wins for I in $\mathscr{G}_{\mathbb{P}}^{\kappa}$, we can continue this to get a set $\{p_{\alpha} : \alpha < \kappa\} \subseteq \mathbb{P}$ which, by construction, intersects every dense set of \mathbb{P} in V and in which every two elements are compatible with a common extension (indeed, any two elements are *comparable*). Hence, the upward closure $\{q \in \mathbb{P} : \exists \alpha < \kappa \ (p_{\alpha} \leq q)\}$ is a filter \mathbb{P} -generic over V with $p \in G$.

To see that there are 2^{κ} -many such generics, note that for incompatible choices by **II** of $p_{2\alpha+1} \in D_{\alpha} \cap \mathbb{P}_{\leq p_{2\alpha}}$, we get different generics. Since \mathbb{P} is appropriate for forcing, any dense set has such incompatible elements. This gives at least 2 incompatible choices by **II** on each turn, and thus at least 2^{κ} -many choices by **II** overall, and so at least 2^{κ} -many generics.

We can also use the same idea with antichains as opposed to dense sets.

- 35 B • 11. Theorem

- Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over.
- Let $\mathbb{P} \in V$ be a preorder appropriate for forcing.

• Suppose \mathbb{P} is κ -strategically closed and $\{A \in V : A \text{ is a maximal antichain}\}$ has size $\leq \kappa$ for some κ .

Therefore for each $p \in \mathbb{P}$, there is a G \mathbb{P} -generic over V with $p \in G$, and in fact 2^{k} -many such generics.

Proof .:.

Proceed as in Corollary 32 C • 4. Let $p \in \mathbb{P}$ be arbitrary. Enumerate $\{A \in V : A \text{ is a maximal antichain}\} = \{A_{\alpha} : \alpha < \kappa\}$. Write $A \downarrow = \{p^* \in \mathbb{P} : \exists p \in A \ (p^* \leq p)\}$ which is dense if A is a maximal antichain. Thus we can consider $\{A_{\alpha} \downarrow : \alpha < \kappa\}$ as a κ -sized family of dense sets and find 2^{κ} -many G P-generic over this family as in Closure Giving Generics (35 B • 10). Each such G clearly intersects every maximal antichain and therefore is generic over V by Theorem 32 C • 5. \dashv

This idea is useful because were very rarely have combinatorial results related to dense sets, but chain conditions are very common and can give nice results related to counting the number of antichains: if a κ -sized preorder \mathbb{P} is λ -cc, then there are at most κ^{λ} -many antichains. As a result, combinatorial axioms that give nice characterizations of cardinal exponentiation, like GCH as per Theorem 5 E • 6, are useful in these contexts and are often assumed out of convenience.

§ 35 C. Easton support iterations

Easton support iterations are very useful with indestructibility axioms with large cardinals. A common idea with such iterations is to proceed in a trial by fire—not unlike The Consistency of MA without CH ($34 \text{ F} \cdot 10$)—where each preorder attempts to destroy some large cardinal property, and what remains is then indestructible by the kinds of preorders used in the iteration. But beyond this, easton support iterations can also be used with all sorts of large cardinals to produce a wide array of interesting results.

— 35C•1. Definition -

Let $\kappa \in \text{Ord.}$ The *easton ideal on* κ is the set $\{X \subseteq \kappa : \forall \delta \leq \kappa \ (\delta \text{ is weakly inaccessible} \rightarrow |X \cap \delta| < \delta)\}$. *Easton support iterations* of length κ are iterations with support in the easton ideal on κ .

More intelligibly, for κ -length iterations, we take bounded support at weakly inaccessible stages, and allow every other kind of support elsewhere as in Support of Inverse Limits (34 C • 12). Hence by Support of Direct Limits (34 C • 5), easton support corresponds to taking direct limits at weakly inaccessible stages (i.e. regular limit cardinals), and inverse limits elsewhere.

Easton support iterations have nice properties for their tail forcing, which is especially useful when thinking about how iterations are moved by elementary embeddings: if $j(\mathbb{P}) \cong \mathbb{P} * \dot{\mathbb{Q}}$, then $\dot{\mathbb{Q}}$ (is forced to) act like the how most of the tails of \mathbb{P} act. So if the tails of \mathbb{P} act nicely, so too does this tail iteration $\dot{\mathbb{Q}}$. So let's investigate what the tails of easton support iterations look like, assuming that we're forcing with sufficiently (strategically) closed, small preorders.

– 35 C•2. Lemma

- Let $\lambda < \kappa$ be *mahlo* in that λ is strongly inaccessible and $\{\delta < \lambda : \delta \text{ is strongly inaccessible}\}$ is stationary.
- Let $\mathbf{*}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ be a κ -length easton support iteration.
- Suppose each $\dot{\mathbb{Q}}_{\alpha}$ is forced to be $\ll \alpha$ -strategically closed, and $|\mathbf{x}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}| < \lambda$ for each $\alpha < \lambda$.
- Let $\dot{\mathbb{Q}} = \mathbf{k}_{\lambda < \alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ be the tail iteration after λ , and $\mathbb{P} = \mathbf{k}_{\alpha < \lambda} \dot{\mathbb{Q}}_{\alpha}$.

Therefore

1. \mathbb{P} is λ -cc; and

2. $\mathbb{1}^{\mathbb{P}} \Vdash ``\dot{\mathbb{Q}}$ is $< \dot{\lambda}$ -strategically closed".

Proof .:.

1. The result for \mathbb{P} follows from Corollary 34 C • 16 and more generally from Direct Limit Chain Conditions (34 C • 15) because mahlos, having a stationary set of inaccessibles below them and easton iterations taking direct limits at inaccessibles, means we satisfy Direct Limit Chain Conditions (34 C • 15) (3): we take the direct limit stationarily often. Direct Limit Chain Conditions (34 C • 15) (2) is trivial by the cardinality restriction, and (1) follows by λ again being inaccessible and easton iterations taking direct limits at inaccessibles.

2. For strategic closure, we must use Corollary 35 B • 9 and confirm the following four hypotheses:

- a. inverse limits or direct limits are taken at every limit stage;
- b. inverse limits are taken at every limit stage $\geq \lambda$ of cofinality $< \lambda$;
- c. $\mathbb{1}_{\alpha} \Vdash ``\dot{\mathbb{Q}}_{\alpha}$ is $\langle \dot{\lambda}$ -strategically closed'' for every $\alpha \in [\lambda, \kappa)$; and
- d. \mathbb{P} doesn't collapse any cofinalities $\geq \lambda$ to be $< \lambda$.

(a) is immediate by definition of easton support. For (b), note that we take inverse limits *everywhere* except weak inaccessibles. But those of cofinality $< \lambda$ are all dealt with before λ : if $\kappa = cof(\kappa)$ with $cof(\kappa) < \lambda$ then $\kappa < \lambda$. So (b) holds, and (c) holds by hypothesis. (d) follows from \mathbb{P} being λ -cc. \dashv

The fact that the heads are λ -cc is very useful when used in combination with Theorem 35 B • 11. The idea is that with a traditional $j : V \to M$, we often have $j(\mathbb{P}) = \mathbb{P} * \dot{\mathbb{Q}}$ where then $\dot{\mathbb{Q}}$ is $j(\kappa)$ -cc, and this can give a bound on the number of antichains, assuming some nice combinatorial properties and that $j(\kappa)$ is relatively small.

Here is one general example of this idea of counting antichains by using chain conditions and elementary embeddings. The idea behind the setup is that we have iteratively forced with preorders to change properties of some cardinals below κ using preorders below κ . For example, the reader can assume \dot{Q}_{α} is a name for $Add(\alpha, 1)$ whenever α is a successor cardinal and trivial otherwise. That particular preorder has some nice properties.

— 35 C·3. Example

- Let $V \vDash$ ZFC + GCH be a transitive model we can force over.
- Let κ be measurable in **V** with measure U.
- Let $j = j_U : V \to \text{cUlt}^V(V, U)$ be the canonical embedding.
- Let $\mathbb{P} = \mathbf{k}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ be an easton support iteration.
- Suppose each \dot{Q}_{α} is forced to be $\ll \alpha$ -strategically closed, and each \dot{Q}_{α} has rank (and thus cardinality) $< \kappa$.
- Suppose further that \dot{Q}_{α} is trivial for U-almost every $\alpha < \kappa$, and is non-trivial only at cardinals.
- Let G be \mathbb{P} -generic over V.

Therefore, in V[G],

- $V[G] \models "\forall \delta \ge \kappa (|\delta| = \delta \rightarrow \delta^+ = 2^{\delta})$ ", i.e. GCH holds above κ ; and
- there are κ^{++} -many $H_j(\mathbb{P})$ -generic over $\operatorname{cUlt}^{\mathsf{V}}(V, U)$ such that j lifts to $j^+: V[G] \to \operatorname{cUlt}^{\mathsf{V}}(V, U)[H]$.

Proof .:.

It's not too difficult to see that V[G] satisfies GCH above κ , but we really only need that $V[G] \models "2^{\kappa} = \kappa^+$ ". To see this, note $|\mathbb{P}| \leq \kappa$ so for any $\delta \geq \kappa$, using Result 32 E • 3: there are at most $(2^{|\delta \times \mathbb{P}|})^V = (2^{\delta})^V$ -many names for subsets of δ in V[G]. Since \mathbb{P} is κ -cc and thus δ -cc, it follows that $(\delta^+)^V = (\delta^+)^{V[G]}$ and thus $(\delta^+)^{V[G]} \leq (2^{\delta})^{V} = (\delta^+)^{V} = (\delta^+)^{V[G]}$ and so $V[G] \models "2^{\delta} = \delta^+$ " for every $\delta \geq \kappa$. In particular, $V[G] \models "\kappa^{++} = 2^{\kappa^+} = 2^{2^{\kappa}}$ ".

Note that $j(\mathbb{P})$ is an easton support iteration where for $\alpha < \kappa$, each preorder takes the form $j(\dot{\mathbb{Q}}_{\alpha}) = \dot{\mathbb{Q}}_{\alpha}$ since the rank of such preorders are below $\kappa = \operatorname{cp}(j)$. Hence $j(\mathbb{P})$ can be factored of the form $j(\mathbb{P}) \cong \mathbb{P} * \dot{\mathbb{Q}}$ where $\dot{\mathbb{Q}}$ is the tail forcing in $\mathbb{M} = \operatorname{cUlt}^{\mathbb{V}}(\mathbb{V}, U)$ from κ to $j(\kappa)$. Note that $\dot{\mathbb{Q}}_G$ is $j(\kappa)$ -cc in $\mathbb{M}[G]$ by Corollary 34 C • 16 (measurable cardinals are mahlo). Also note that $j"G \subseteq \mathbb{P}$ so we merely need to find a $\dot{\mathbb{Q}}_G$ -generic over M[G]in $\mathbb{V}[G]$ to lift.

Because of the restriction on rank, by elementarity, it's not hard to see that the hypotheses of Lemma 35 C • 2 hold in **M** where κ is the mahlo cardinal and $j(\kappa)$ is the length of the iteration. In particular, **P** is κ -cc in **M**. Since $\dot{\mathbb{Q}}_{\alpha}$ is trivial for *U*-almost every $\alpha < \kappa$, it follows that $\dot{\mathbb{Q}}_{\kappa}$ is trivial in **M**: the set of trivial stages is in *U* iff $\kappa \in j(\{\alpha < \kappa : \dot{\mathbb{Q}}_{\alpha} \text{ is trivial}\})$. Thus the first non-trivial preorder of the tail forcing $\dot{\mathbb{Q}}$ happens at some **M**-cardinal above κ , and so $\dot{\mathbb{Q}}_{G}$ must be κ^+ -strategically closed in **M**[*G*].

By Result 12 C • 1, *M* is closed under *V*'s κ -length sequences. Since \mathbb{P} is κ -cc, by Result 33 B • 7, *M*[*G*] is closed under *V*[*G*]'s κ -length sequences. In particular, $(\kappa^+)^{\mathsf{M}[G]} = (\kappa^+)^{\mathsf{V}[G]}$ and so $\dot{\mathbb{Q}}_G$ is still κ^+ -strategically closed in $\mathsf{V}[G]$. It's also not hard to see that in $\mathsf{V}[G]$, $|\dot{\mathbb{Q}}_G| = |j(\kappa)| = |(2^{\kappa})^{\mathsf{V}}| \le \kappa^+ = 2^{\kappa}$. In particular, in $\mathsf{V}[G]$, because $\dot{\mathbb{Q}}_G$ is $j(\kappa)$ -cc in $\mathsf{M}[G]$, there are at most $(2^{\kappa})^{\kappa} = 2^{\kappa} = \kappa^+$ -many maximal antichains of M[G]. By Theorem 35 B • 11, there is are 2^{κ^+} -many $H' \dot{\mathbb{Q}}_G$ -generic over M[G] in $\mathsf{V}[G]$ which means G * H' is $\mathbb{P} * \dot{\mathbb{Q}}$ generic over *M*. This gives $(2^{\kappa^+})^{\mathsf{V}[G]}$ -many $H \ j(\mathbb{P}) \cong \mathbb{P} * \dot{\mathbb{Q}}$ -generic over *M* with $j"G \subseteq H$. By Generic Lifting (35 A • 2) we can lift *j* in $\mathsf{V}[G]$.

An easy example of this is the following, showing from a single measurable cardinal that it's possible to have the maximal number of measures on it.

- 35 C·4. Example
 - Let $V \vDash$ ZFC + GCH be a transitive model we can force over.
 - Let κ be measurable in V with measure U and ultrapower embedding $j = j_U : V \to \text{cUlt}^V(V, U)$.
 - Let $\mathbb{P} = \mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ where $\dot{\mathbb{Q}}_{\alpha} = \dot{\mathsf{Add}}(\check{\alpha}, \check{1})$ whenever $\alpha \in \{\delta^+ : \delta = |\delta| < \kappa\}$, and otherwise $\dot{\mathbb{Q}}_{\alpha}$ is (forced to be) trivial.
 - Let G be \mathbb{P} -generic over V.

Therefore, in V[G], κ is measurable and has κ^{++} -many measures on it.

Proof .:.

Note that $\mathsf{Add}(\alpha, 1)$ has rank $< \alpha + \omega$ and is $< |\alpha|$ -closed and hence $\ll \alpha$ -strategically closed whenever nontrivial (and clearly Ord-strategically closed when trivial). Moreover, it's not hard to see that $A = \{\alpha < \kappa : \exists \lambda \ (\alpha = \lambda^+)\} \notin U$ since $\kappa \notin j(A)$ because κ is still a limit cardinal in $\mathbf{M} = \mathsf{cUlt}^{\mathsf{V}}(\mathsf{V}, U)$. Since measurable cardinals are mahlo, it follows that all hypotheses of Example $35 \, \mathbb{C} \cdot 3$ hold and thus there are κ^{++} -many $j(\mathbb{P})$ generics H over M such that j lifts to V[G], M[H]. Enumerate these generics H_{α} for $\alpha < \kappa^{++}$ in $\mathbf{V}[G]$. Let $j_{\alpha} : V[G] \to M[H_{\alpha}]$ be the corresponding liftup of j.

By Lifting Ultrapowers (35 A • 7), $M[H_{\alpha}] = \text{cUlt}^{V[G]}(V[G], U_{j_{\alpha}})$ for each $\alpha < \kappa^{++}$ where $U_{j_{\alpha}}$ is the derived measure (i.e. $(\kappa, \kappa + 1)$ -extender). And since each $M[H_{\alpha}]$ is distinct, each $U_{j_{\alpha}}$ is distinct for $\alpha < \kappa^{++}$, and thus V[G] has κ^{++} -many measures on κ .

A similar forcing demonstrates a common problem associated with long iterations: loosing control over the elementary embeddings in the generic extension. In particular, it's consistent relative to a measurable cardinal that we can add a measurable cardinal in a model with none.

- 35C•5. Example

Assume Con(ZFC + "there is a measurable cardinal"). Therefore, the following state of affairs is consistent:

- $V \models$ ZFC is a transitive model we can force over;
- $V \vDash$ "there is no measurable cardinal"; and
- there is a preorder $\mathbb{P} \in V$ such that $\mathbb{1}^{\mathbb{P}} \Vdash$ "there is a measurable cardinal".

Proof .:.

The V of the statement will actually be a generic extension V[G] of a ground model V with a measurable cardinal. The idea is to kill the measurable cardinal of V in V[G], and then resurrect it in another extension V[G][H]. Let's get on with the actual proof. Without loss of generality that we are working with a countable, transitive model of ZFC with a measurable cardinal κ . Also let κ be the least ordinal with a transitive model $V \models ZFC + GCH +$ " κ is measurable" + "2^{κ} = κ ⁺". And by cutting off V, we may assume κ is the only measurable of V (if there is another measurable above κ , then consider κ' the least above κ and consider instead the ground model as $V_{\nu'}^{V}$).

Let $\mathbb{Q} = \mathbf{*}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ be an easton support iteration that adds a cohen subset of α to each inaccessible below κ : $\dot{\mathbb{Q}}_{\alpha} = \dot{\mathsf{Add}}(\check{\alpha}, \check{1})$ whenever α is strongly inaccessible.

— Claim 1

Let G be Q-generic over V. Therefore $V[G] \models ZFC +$ "there are no measurable cardinals".

Proof .:.

That $\mathbf{V}[G] \models \text{GCH}$ isn't too difficult, but it is rather tedious, primarily using Corollary 33 A • 3 and the chain and closure properties of $\text{Add}(\alpha, 1)$ at each stage of the forcing. So suppose U is a measure on κ in $\mathbf{V}[G]$. Consider $j_U : V[G] \to N = \text{cUlt}^{\mathbf{V}[G]}(V[G], U)$ the canonical embedding. By Definability of the Ground Model (35 A • 6), N = M[j(G)] for some $M \subseteq N$ where j(G) is $j(\mathbb{P})$ -generic over M. Moreover, $j_U \upharpoonright V : V \to M$ is also elementary, although perhaps $M \not\subseteq V$. Nevertheless, we still get $\mathcal{P}(\kappa) \cap V \subseteq \mathcal{P}(\kappa) \cap M$. It's easy to see that $j(\dot{\mathbb{Q}}_{\alpha}) = \dot{\mathbb{Q}}_{\alpha}$ whenever $\alpha < \kappa$, and hence we can factor $j(\mathbb{Q}) \cong \mathbb{Q} * \bigstar_{\kappa \leq \alpha < j(\kappa)} \dot{\mathbb{Q}}_{\alpha}$ where $\dot{\mathbb{Q}}_{\alpha}$ is (forced to be) $\text{Add}(\alpha, 1)$ whenever $\alpha \in [\kappa, j(\kappa))$ is inaccessible in M.

Since $V_{\kappa+1}^{\mathbf{V}[G]} = V_{\kappa+1}^{\mathbf{M}[j(G)]}$, κ is still inaccessible in $\mathbf{M}[j(G)]$ and by downward absoluteness, κ is inaccessible in \mathbf{M} . Hence $j(\mathbb{Q})$ adds a subset $j(G)_{\kappa} \subseteq \kappa$ in $\mathbf{M}[j(G)]$. But then $j(G)_{\kappa} \in V[G]$ which is $\mathsf{Add}(\kappa, 1)^{\mathsf{M}[j(G)] \mathsf{k}]}$ -generic over $M[j(G) \upharpoonright \kappa] = M[G] \subseteq V[G]$. But since $|\mathsf{Add}(\kappa, 1)| = 2^{<\kappa} = \kappa$ in both $\mathbf{M}[G]$ and $\mathbf{V}[G]$, any dense subset of $\mathsf{Add}(\kappa, 1)^{\mathsf{V}[G]}$ in $\mathbf{V}[G]$ can be seen as a subset of κ and is thus in $\mathbf{M}[G]$ by the agreement between V and M given by $j_U \upharpoonright V$. In particular, $j(G)_{\kappa} \in M[j(G)] \subseteq V[G]$ is $\mathsf{Add}(\kappa, 1)^{\mathsf{V}[G]}$ -generic over V[G], a contradiction. Thus κ is not measurable in $\mathsf{V}[G]$.

There of course cannot be any measurables created above κ by Lévy–Solovay (35 A • 9): $|\mathbb{Q}| = \kappa$. No $\lambda < \kappa$ is measurable in V[G] by hypothesis of the minimality of κ (Add(λ , 1) forces that $2^{\lambda} = \lambda^+$). So there are no measurable cardinals in V[G].

Now consider $\dot{\mathbb{R}} = \dot{\mathsf{Add}}(\check{\kappa}, \check{1})$ and $\mathbb{P} = \mathbb{Q} \ast \dot{\mathbb{R}}$. To show that κ remains measurable after forcing with \mathbb{P} over V, we need to lift an embedding. Let U be a measure on κ and let $j = j_U : V \to M = \mathrm{cUlt}^V(V, U)$ be the canonical embedding. We can factor

$$j(\mathbb{P}) \cong \mathbb{Q} * \dot{\mathbb{R}} * \bigwedge_{\kappa < \alpha \le j(\kappa)} \dot{\mathbb{S}}_{\alpha} = \mathbb{P} * \bigwedge_{\kappa < \alpha \le j(\kappa)} \dot{\mathbb{S}}_{\alpha}$$

where $\dot{\mathbf{S}}_{\alpha}$ is defined like $\dot{\mathbf{Q}}_{\alpha}$ but for $\kappa < \alpha \leq j(\kappa)$ in **M**. Note that in **V**, the tail forcing $\bigstar_{\kappa < \alpha < j(\kappa)} \dot{\mathbf{S}}_{\alpha}$ is forced to be $(\kappa^+)^{\mathbf{V}}$ -closed since the first non-trivial stage of forcing happens at an inaccessible above κ . Also note that the tail forcing is forced to have size $j(\kappa)$ in **M**, which has size κ^+ in **V**. So let G * H be **P**-generic over *V*. Note that $(\kappa^+)^{\mathbf{V}} = (\kappa^+)^{\mathbf{V}[G*H]}$ because both **Q** and \mathbb{R}_G preserve cardinals and cofinalities $\geq \kappa$. As a result, $j(\kappa)$ still has size κ^+ in $\mathbf{V}[G * H]$, and the tail iteration is κ^+ -closed there. Moreover, since **P** is κ -cc in **V**, the tail iteration is forced to be $j(\kappa)$ -cc in **M**. Hence **M** thinks there are only $j(\kappa)^{<j(\kappa)} = j(\kappa) \cdot 2^{<j(\kappa)} = j(\kappa)$ -many antichains of the tail iteration. Hence in $\mathbf{V}[G * H]$, the tail iteration $(\bigstar_{\kappa < \alpha \le j(\kappa)} \dot{\mathbf{S}}_{\alpha})_{G*H}$ is κ^+ -closed and there are only $|j(\kappa)| = \kappa^+$ -many antichains of $\mathbf{M}[G * H]$. Thus by Theorem 35 B • 11, V[G * H] has a K generic over M for the tail iteration with $j"G * H = G * H \subseteq K$ so that $j : V \to M$ lifts to $j^+ : V[G * H] \to M[K]$ within V[G * H]. It follows that κ is measurable in V[G * H].

The next subsection attempts to give limits on when this can happen, which really is an attempt to gain control over the embeddings a preorder can add to the generic extension.

§ 35 D. Gap forcing

One topic in the theory of forcing is that of *indestructibility* axioms which state the preservation of large cardinals by forcing with various kinds of preorders. The general idea behind showing the consistency of these axioms is to force with everything bad until everything that remains is indestructible by such bad preorders. The issue with this approach is that one must take great care to check that no new large cardinals have been introduced by accident that might not have gone through the entire process. As a result, knowing that certain (common) preorders don't introduce any large cardinals is quite useful, and a quite general technique is shown here.

Recall Lévy–Solovay (35 A • 9), which states that small posets don't affect measurability. What's really going on, however, is two results: every measure U on κ in the ground model generates a measure in V[G], and every measure

U on κ in the generic extension is generated by a measure in the ground model. In this sense, small forcings don't introduce new large cardinals. While this is a nice result, it's not exactly all that useful if we want to deal with iterations of length $\geq \kappa$ or with larger preorders. Basically, we lose control over the embeddings added by more useful forcing notions in a large cardinal context. Gap forcing, a notion due to Joel David Hamkins [14], is a way to help with that.

35 D • 1. Definition Let κ be a cardinal. A preorder *admits a gap* at κ iff it's (forcing equivalent to) an iteration $\mathbb{P} * \dot{\mathbb{Q}}$ where

- \mathbb{P} is non-trivial of size $< \kappa$; and
- $\mathbb{1}^{\mathbb{P}} \Vdash ``\dot{\mathbb{Q}}$ is $\leq \check{\kappa}$ -strategically closed''.

For such a preorder, we also call κ *the gap*.

To help hint at why this concept will be useful, note that there are lots of examples of such preorders. In particular,

- \mathbb{P} admits a gap at every $\kappa \geq |\mathbb{P}|^+$ for any non-trivial \mathbb{P} ;
- $\mathbb{P} * \dot{\mathsf{Col}}(|\mathbb{P}|^+, \kappa)$ admits a gap at $|\mathbb{P}|^+$ for any non-trivial \mathbb{P} ;
- Add $(\omega, 1) * \dot{\mathbb{Q}}$ admits a gap at ω_1 for any countably closed $\dot{\mathbb{Q}}$; and so on.

The main theorem we're interested in is the following showing where embeddings in the generic extension come from.

— 35 D•2. Theorem (Gap Forcing) -

• Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over.

- Let $\mathbb{P} \in V$ admit a gap at $\delta < \kappa$ in V for some cardinals $\delta, \kappa \in V$.
- Let G be \mathbb{P} -generic over V.
- Suppose there is an elementary $j: V[G] \to M[j(G)]$ as a class of V[G] with $cp(j) = \kappa$.
- Suppose M[j(G)] is closed under δ -length sequences of V[G].

Therefore

- 1. $j \upharpoonright V : V \to M$ is a class of V with $M = V \cap M[j(G)]$.
- 2. $V[G] \models ``^{\lambda} M[j(G)] \subseteq M[j(G)]$ " implies $V \models ``^{\lambda} M \subseteq M$ " for all $\lambda \in Ord$.
- 3. $V_{\lambda}^{\vee} \subseteq M[j(G)]$ implies $V_{\lambda}^{\vee} \subseteq M$.

There are stronger versions of this theorem where it's not even assumed that j is amenable, but I believe the above is the simplest form to digest. Note that there are a few ways to view what exactly M and M[j(G)] are supposed to be. The first is merely to view **M** as some given transitive model with M[j(G)] a forcing extension that happens to have an elementary $j : V[G] \to M[j(G)]$. In this case, M is already given, and the result that $M = V \cap M[j(G)]$ is quite interesting. The other is to view M[j(G)] as just some model N, like an ultrapower of **V**[G]. In this case, M is less obvious, but we may view M either by definition as trcl(j"V), making it a less interesting result, or instead use elementarity and the definability of the ground model as per Definability of the Ground Model (35 A • 6).

Regardless of how M and M[j(G)] are viewed, Gap Forcing (35 D • 2) easily gives results like the previously unproven consequences of Lévy–Solovay (35 A • 9) discussed above.

— 35 D•3. Corollary -

- Let $V \vDash$ ZFC be a transitive model we can force over.
- Let $\mathbb{P} \in V$ admit a gap at $\delta < \kappa$ in V for some cardinals $\delta, \kappa \in V$ with κ measurable.
- Let G be \mathbb{P} -generic over V.

Therefore, for $\lambda \ge \kappa$, $V[G] \models ``\lambda$ is measurable" implies $V \models ``\lambda$ is measurable".

Proof .:.

Note that the ultrapower $j : V[G] \to \text{cUlt}^{V[G]}(V[G], U)$ by some measure U on κ is $\text{cp}(j) = \kappa > \delta$ -closed. Since the gap is below κ , by Gap Forcing (35 D • 2), $j \upharpoonright V$ is a class of V, and thus we can consider the derived measure as a measure on κ in V.

The tagline for this is that gap forcing creates no new measurable cardinals above its gap. By similarly easy reasoning, gap forcing creates no new strong cardinals above its gap. Note that strong embeddings can be difficult to work with

because ensuring the closure properties of Gap Forcing $(35 D \cdot 2)$ aren't immediate, but so long as we consider the canonical extender embeddings, we're fine.

– 35 D•4. Corollary

• Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model we can force over.

- Let $\mathbb{P} \in V$ admit a gap at $\delta < \kappa$ in **V**.
- Let G be \mathbb{P} -generic over V.

Therefore, for $\lambda \ge \kappa$ with $cof(\lambda) > \delta$, $V[G] \models "\kappa$ is λ -strong" implies $V \models "\kappa$ is λ -strong".

Proof .:.

If κ is λ -strong in V[G] as witnessed by some $E \in V[G]$, then $\operatorname{cUlt}_{E}^{V[G]}(V[G])$ is closed under δ -sequences in V[G]. Thus we can apply Gap Forcing (35 D • 2) to get that $j \upharpoonright V : V \to M$ is a class of V with $V_{\lambda}^{\mathsf{V}} \subseteq V_{\lambda}^{V[G]} \subseteq \operatorname{cUlt}_{E}^{V[G]}(V[G])$ and therefore $V_{\lambda}^{\mathsf{V}} \subseteq M \subseteq V \cap \operatorname{cUlt}_{E}^{V[G]}(V[G])$. This means, using the derived extender, κ is λ -strong in V .

Hence $V[G] \models "\kappa$ is strong" implies $V \models "\kappa$ is strong". Note that we needed $cof(\lambda) > \delta$, since otherwise the embedding witnessing the λ -strength of κ would only have $< \delta$ -closure and we couldn't apply Gap Forcing (35 D • 2). Note also that the converse of these need not hold: $\mathbb{P} * \dot{Col}(\check{\kappa}, \check{\lambda})$ would potentially move the strength of κ down, yet this iteration still admits a gap at $\delta < \kappa$ because $\dot{Col}(\check{\kappa}, \check{\lambda})$ has enough closure. And there are many, *many* more results like the above. The general idea is that gap forcing creates no new large cardinals above the gap. This is often useful in combination with easton support iterations by way of Lemma 35 C • 2.

— 35 D • 5. Corollary

- Let $\mathbf{k}_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ be a κ -length easton support iteration.
- Suppose that each $\dot{\mathbb{Q}}_{\alpha}$ is trivial unless α is mahlo.
- Suppose further that each \dot{Q}_{α} is forced to be $\ll \alpha$ -strategically closed.^{xxxix}
- Suppose that $|\mathbf{x}_{\xi < \alpha} \dot{\mathbb{Q}}_{\xi}| < |\alpha|$ for each $\alpha < \kappa$.

Therefore, $\mathbf{k}_{\alpha < \kappa} \dot{\mathbf{Q}}_{\alpha}$ admits a gap between any two stages of the iteration. More precisely, for each $\alpha < \kappa$, $\mathbf{k}_{\xi < \alpha} \dot{\mathbf{Q}}_{\alpha} * \mathbf{k}_{\alpha \leq \xi < \kappa} \dot{\mathbf{Q}}_{\alpha}$ admits a gap at $|\alpha|^+$.

Proof .:.

For $\alpha < \kappa$ arbitrary, note that the above break up is equivalent to breaking it up at the next mahlo cardinal. If there is no such cardinal, then the tail iteration is trivial and clearly we admit a gap at $|\alpha|^+$ because the start of the iteration only has size $|\alpha|$ by hypothesis. So without loss of generality, assume α is mahlo. By Lemma 35 C • 2, the tail iteration $\bigstar_{\alpha \le \xi < \kappa} \dot{\mathbb{Q}}_{\alpha}$ is forced to be $< \lambda$ -strategically closed where λ is the next mahlo above α . In particular, it is $\le \alpha^+$ -strategically closed. Again, by hypothesis, the start of the iteration $\bigstar_{\xi < \alpha} \dot{\mathbb{Q}}_{\alpha}$ has size $< \alpha^+$ and so the iteration admits a gap at α^+ .

The proof of Gap Forcing $(35 D \cdot 2)$, due to Hamkins, involves a series of lemmas, partially abstract theorems regarding *covering* properties as per Definition 33 B • 4. First we introduce a definition of the kinds of sequences gap forcings *can't* introduce.

- 35 D • 6. Definition

Let $\mathbf{V} \models \mathsf{ZFC}$ be a transitive model. A sequence $f : \alpha \to V$ is *fresh* over V iff $f \upharpoonright \xi \in V$ for every $\xi < \alpha$, but $f \notin V$.

The definition will be useful in a variety of places, especially in combination with covering properties. Of note is that forcing with a gap at δ doesn't introduce fresh sequences of cofinality > δ . That being said, any non-trivial preorder adds a fresh sequence over the ground model.

^{xxxix}We actually need much less than $\ll \alpha$ -strategic closure in most cases: we just need that \dot{Q}_{α} , when non-trivial, is $\leq \beta$ -strategically closed where β is the maximal mahlo below α . In this way, $\ll \alpha$ -strategic closure is extreme overkill, but it's simple to state.

— 35 D • 7. Result -

Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force over. Let $\mathbb{P} \in V$ be a non-trivial preorder in V. Let G be \mathbb{P} -generic over V. Therefore there is a fresh sequence over V of length $\leq |\mathbb{P}|^{V[G]}$ in V[G].

Proof .:.

Suppose $|\mathbb{P}|^{\mathbf{V}[G]} = \kappa$ as witnessed by a bijection $h : \kappa \to \mathbb{P}$ in $\mathbf{V}[G]$. Now consider the characteristic function of h^{-1} "G as a subset of $\kappa : \chi_G : \kappa \to 2$ has $\chi_G(\alpha) = 1$ iff $h(\alpha) \in G$. We can't have both $\chi_G \in V$ and $h \in V$ as otherwise $G = \{p \in \mathbb{P} : \chi_G(h^{-1}(p)) = 1\} \in V$. So just consider the least $\alpha \leq \kappa$ such that either $\chi_G \upharpoonright \alpha \notin V$ or $h \upharpoonright \alpha \notin V$. Such a sequence is therefore fresh over V of length $\leq \kappa = |\mathbb{P}|^{\mathbf{V}[G]}$.

With a bit of work, we can show that gap forcing adds very few fresh sequences. The idea being that \mathbb{P} adds some small $< \delta$ -length sequence and this should propagate through to longer sequences: no sequence of cofinality $\geq \delta$ escapes this because of the strategic closure of $\dot{\mathbb{Q}}$.

– 35 D•8. Lemma

- Let $V \vDash$ ZFC be a transitive model we can force over.
- Let $\mathbb{P} * \dot{\mathbb{Q}} \in V$ admit a gap at δ in V.
- Let G be $\mathbb{P} * \dot{\mathbb{Q}}$ -generic over V.
- Suppose $f \in V[G]$ is a θ -length sequence with $\operatorname{cof}(\theta)^{V[G]} \ge \delta$.

Therefore f is not fresh: if $f \upharpoonright \xi \in V$ for every $\xi < \text{dom}(f)$ then $f \in V$.

Proof .:.

It suffices to work with $f: \theta \to \text{Ord}$ and in fact binary sequences $f: \theta \to 2$.

— Claim 1 -

 $\mathbb{P} * \dot{\mathbb{Q}}$ adds no fresh sequences of a length θ with $cof(\theta)^{\vee} \ge \delta$ iff it adds no such binary sequences.

Proof .:.

Clearly the (\rightarrow) direction holds. For the other direction, suppose $f : \theta \to V$ with $\operatorname{cof}(\theta) \ge \delta$ and $f \upharpoonright \xi \in V$ for all $\xi < \theta$. We have $f : \theta \to V_{\alpha}^{\mathsf{V}}$ for some sufficiently large α . In V we have an injection $g : \theta \times V_{\alpha}^{\mathsf{V}} \to \kappa^+$ for some κ . Now consider the characteristic function not of f as a subset of $\theta \times V_{\alpha}^{\mathsf{V}}$ but instead of g''f as a subset of κ^+ : $\chi : \kappa^+ \to 2$. Since $f \upharpoonright \xi \in V$ for each $\xi < \theta$, it follows that $\chi \upharpoonright \xi \in V$ for each $\xi < \kappa^+$. Since κ^+ is regular (in V) with $\kappa \ge \theta \ge \delta$, it follows that $\chi \in V$ as otherwise it would be a fresh binary sequence. But then $f = \{x : \chi(g(x)) = 1\} \in V$, meaning f is not fresh.

So let $f : \theta \to 2$ have all of its (strict) initial segments in V. The general idea is to notice that \mathbb{P} should add a new sequence of length $< \delta$ that then propagates through to larger sequences like f in a coded way. The way we can code this is through a tree of conditions that decide more and more of \dot{f} .

— Claim 2 -

There is some $\langle p^*, \dot{q}^* \rangle \in \mathbb{P} * \dot{\mathbb{Q}}$ such that each set

 $D_{\xi} = \{ \langle p^*, \dot{q}^{**} \rangle \in \mathbb{P} \times \dot{\mathbb{Q}} : \langle p^*, \dot{q}^{**} \rangle \text{ decides every element of } \dot{f} \upharpoonright \dot{\xi} \}$

is dense below $\langle p^*, \dot{q}^* \rangle$ for $\xi < \theta$.

Proof .:.

Really we just need to ensure we can use the same p^* when strengthening. Note that for $\xi < \theta$, $f \upharpoonright \xi$ is decided by some $\langle p_{\xi}, \dot{q}_{\xi} \rangle \in G$ to be some element of the ground model: let \dot{f} be a name for f:

 $\mathbf{V} \vDash "\forall \boldsymbol{\xi} < \boldsymbol{\theta} \; \exists \boldsymbol{x} \; \exists \langle \boldsymbol{p}_{\boldsymbol{\xi}}, \dot{\boldsymbol{q}}_{\boldsymbol{\xi}} \rangle \in \mathbb{P} \ast \dot{\mathbb{Q}} \; (\langle \boldsymbol{p}_{\boldsymbol{\xi}}, \dot{\boldsymbol{q}}_{\boldsymbol{\xi}} \rangle \Vdash "\dot{f} \upharpoonright \dot{\boldsymbol{\xi}} = \check{\boldsymbol{x}}")".$

Since $\operatorname{cof}(\theta) \ge \delta$ and $|\mathbb{P}| < \delta$, unboundedly many of the p_{ξ} s must be the same $p^* \in \mathbb{P}$ so that in fact this p^* works for every $\xi < \theta$ since we're talking about initial segments. Using any \dot{q}^* such that $\langle p^*, \dot{q}^* \rangle \Vdash$ " \dot{f} is a function from $\check{\theta}$ to $\check{2}$ " then works.

Basically, this means that to decide more of f, we don't need to change the first coordinate, and can just deal

with the conditions of \dot{Q} in the tree of decisions we will form. We use then the $\leq \delta$ -strategic closure of \dot{Q} . We will consider a tree $T \in V$ of conditions of $\mathbb{P} * \dot{Q}$. An immediate corollary of this is the following.

– Claim 3

If f is fresh over V, then any condition below $\langle p^*, \dot{q}^* \rangle$ splits into two incompatible conditions that differ on initial segments of (what they decide of) \dot{f} .

Proof .:.

Since no condition decides every element of \dot{f} , any condition below $\langle p^*, \dot{q}^* \rangle$ splits into two incompatible conditions that differ on \dot{f} and by further extension, we can assume that they differ on initial segments of \dot{f} .

So assume f is fresh, i.e. $f \notin V$. To form T, let $\dot{\sigma}$ be a canonical P-name for a winning strategy in $\mathscr{G}_{\dot{\mathbb{Q}}}^{\delta+1}$ for I. The idea is now to pick the incompatible extensions as above while the other player plays according to $\dot{\sigma}$. It's important to note that P adds a fresh sequence P of length $\pi < \delta$ by Result 35 D • 7. Let \dot{P} be a P-name for P.

To form T, consider $\tau \in 2^{<\pi}$. We define \dot{q}_{τ} as a canonical \mathbb{P} -name for the last play by **II** in the game $\mathscr{G}_{\dot{\mathbb{Q}}}^{\delta+1}$ where **I** uses $\dot{\sigma}$. Firstly, have player **II** play $\dot{q}_{\tau \uparrow 0} = \dot{q}^*$ as in Claim 2. For any $\tau \in 2^{<\pi}$, let \dot{r}_{τ} be **I**'s response using $\dot{\sigma}$: a canonical name for $\dot{\sigma}(\langle \dot{q}_{\tau \uparrow \alpha} : \alpha < \ln(\tau) \rangle)$. Now suppose \dot{r}_{τ} has been defined by the following game (inductively p^* forces that this is all in accordance with $\dot{\sigma}$):

I:
$$\dot{\mathbb{I}}^{\mathbb{Q}}$$
 $\dot{r}_{\tau \uparrow 0}$ $\dot{r}_{\tau \uparrow 1}$ \cdots \dot{r}_{τ}
II: $\dot{q}^* = \dot{q}_{\tau \uparrow 0}$ $\dot{q}_{\tau \uparrow 1}$ $\dot{q}_{\tau \uparrow 2}$ \cdots

To define $q_{\tau^{\frown}(0)}$ and $q_{\tau^{\frown}(1)}$, use p^* and Claim 3. To define \dot{q}_{τ} for limit length τ , we may just consider any condition below \dot{r}_{τ} . The result is a map from $2^{<\pi}$ to $\dot{\mathbb{Q}}$ that, most importantly, lies in V.

Now we make use of $G = G_{\mathbb{P}} * G_{\mathbb{Q}}$ to finish off the proof where $G_{\mathbb{P}}$ is \mathbb{P} -generic over V and $G_{\mathbb{Q}}$ is $\dot{\mathbb{Q}}_{G_{\mathbb{P}}} = \mathbb{Q}$ generic over $V[G_{\mathbb{P}}]$. Work in $\mathbf{V}[G_{\mathbb{P}}]$. Let $P = \dot{P}_{G_{\mathbb{P}}}$, let $q_{\tau} = (\dot{q}_{\tau})_{G_{\mathbb{P}}}$, $r_{\tau} = (\dot{r}_{\tau})_{G_{\mathbb{P}}}$ for $\tau \in 2^{<\pi} \cap V$, and let $\sigma = (\dot{\sigma})_{G_{\mathbb{P}}}$. Consider $\langle q_{P \uparrow \xi} : \xi < \pi \rangle$ which therefore comes from an actual partial play of the game $\mathcal{G}_{\mathbb{Q}}^{\delta+1}$ where I uses σ . By the $\leq \delta$ -strategic closure of \mathbb{Q} , there is some $r_P = \sigma(\langle q_{P \uparrow \xi} : \xi < \pi \rangle)$:

I: $\dot{\mathbb{I}}^{\dot{\mathbb{Q}}}$ $r_{P \upharpoonright 0}$ $r_{P \upharpoonright 1}$ \cdots r_{P} II: $q_{P \upharpoonright 0}$ $q_{P \upharpoonright 1}$ \cdots

Note, however, that r_P , being below each $q_P \upharpoonright \xi$, decides the initial segments of \dot{f} that the $q_P \upharpoonright \xi$ s decide. But the initial segment of \dot{f} decided by r_P is in V since f was fresh. Yet from this initial segment $f \upharpoonright \xi$, we can reconstruct $P : \pi \to 2$ in V: for $\tau \lhd P$, if $\langle p^*, \dot{q}_{\tau} \frown \langle 0 \rangle \rangle$ decides \dot{f} in a way different from $f \upharpoonright \xi$, then $\tau \frown \langle 1 \rangle \lhd P$, and so we consider what $\langle p^*, \dot{q}_{\tau} \frown \langle 1, 0 \rangle \rangle$ decides, and so on. At limit stages we take unions and after π steps, this gives $P \in V$, a contradiction.

One might wonder about cases like $Add(\omega_1, 1)$ being factored as $1 * Add(\check{\omega}_1, \check{1})$. At first glance, this would seem to admit a gap at ω , but adds an ω_1 -length sequence with $cof(\omega_1) > \omega$. But by countable closure, every initial segment would be in the ground model, meaning the ω_1 -length sequence is fresh. The reason why this doesn't conflict with Lemma 35 D • 8 is because of the somewhat innocuous restriction in Definition 35 D • 1 that \mathbb{P} must be nontrivial in a gap forcing preorder $\mathbb{P} * \dot{\mathbb{Q}}$. This fact is of course used in the proof, but it's worth keeping in mind when trying to think about examples of gap forcing.

– 35 D•9. Lemma

- Consider the setup as in Gap Forcing $(35 D \cdot 2)$:
 - Let $V \vDash$ ZFC be a transitive model we can force over.
 - Let $\mathbb{P} * \dot{\mathbb{Q}} \in V$ admit a gap at some regular δ in V.
 - Let G be $\mathbb{P} * \dot{\mathbb{Q}}$ -generic over V.
 - Let $j: V[G] \to M[j(G)]$ be a class of V[G] with $cp(j) = \kappa > \delta$.
 - Suppose M[j(G)] is closed under δ -length sequences of V[G].

Therefore every set of ordinals in V[G] of size δ is covered by a set in $M \cap V$ of size δ .

Proof .:.

Let $x \in V[G]$ have size δ so, by $\leq \delta$ -strategic closure and thus $\leq \delta$ -distributivity via Corollary 35 B • 7, $\dot{\mathbb{Q}}$ didn't add it. So if we regard $G = G_{\mathbb{P}} * G_{\mathbb{Q}}$ where $G_{\mathbb{P}}$ is \mathbb{P} -generic over V, then $x \in V[G_{\mathbb{P}}]$. Note also that by closure under δ -length sequences of V[G], $x \in M[j(G)]$ and again since $j(\dot{\mathbb{Q}})$ didn't add it by $\leq j(\delta) = \delta$ -distributivity, $x \in M[j(G_{\mathbb{P}})]$. Since $|\mathbb{P}| < \delta < \kappa$, we might as well assume by a bijection that $\mathbb{P} = \delta$ and thus $\mathbb{P} \in V_{\delta+\omega}^{\mathsf{V}} \subseteq V_{\kappa}^{\mathsf{V}}$ so that $j(\mathbb{P}) = \mathbb{P}$ and thus $j(G_{\mathbb{P}}) = G_{\mathbb{P}}$. So all of this is just to say that $x \in V[G_{\mathbb{P}}] \cap M[G_{\mathbb{P}}]$. Note that $V[G_{\mathbb{P}}]$ and $M[G_{\mathbb{P}}]$ have the same δ -sized sets of ordinals by the closure properties of M[j(G)] in V[G] and the above argument about $\leq \delta$ -strategic closure not adding any such sequences.

Since \mathbb{P} is trivially δ -cc, by Chain Condition Covering (33 B • 5), we have the $\leq \delta$ -covering property between V and $V[G_{\mathbb{P}}]$, and similarly for \mathbb{M} and $\mathbb{M}[G_{\mathbb{P}}]$ so that δ is still regular in $V[G_{\mathbb{P}}]$ (Recall that the $\leq \delta$ -covering property for V, V[G] means that any $\leq \delta$ -sized set in $V[G_{\mathbb{P}}]$ is covered by a set in V of size $\leq \delta$.) As a result, we can cover $x \subseteq x_0 \in V$ with a set of size $|x_0|^V \leq \delta$. Since $x_0 \in V[G_{\mathbb{P}}] \cap M[G_{\mathbb{P}}]$, we can again cover $x_0 \subseteq x_1 \in M$ of size $|x_1|^{\mathbb{M}} \leq \delta$. Again we can cover $x_1 \subseteq x_2 \in V$ and so on. At limit stages $\gamma < \delta$, we have in $V[G_{\mathbb{P}}] \cap M[G_{\mathbb{P}}]$ the sequence $\langle x_{\xi} : \xi < \gamma \rangle$ and so can take the union $x_{\gamma} = \bigcup_{\xi < \gamma} x_{\xi}$ which, by regularity, has size $\leq \delta$ in $V[G_{\mathbb{P}}]$. This allows us to continue defining the sequence with $x_{\gamma+1} \in M$ and $x_{\gamma+2} \in V$ and so on. The result is a \subseteq -increasing sequence $\langle x_{\xi} : \xi < \delta \rangle$ with cofinally many in V and cofinally many in M.

So now consider $X = \bigcup_{\alpha < \delta} x_{\alpha} \in V[G_{\mathbb{P}}]$ which has sze $\leq \delta$. We'd like to show $X \in V \cap M$ since $x \subseteq X$. In $V[G_{\mathbb{P}}]$, let $p_{\alpha} \in G$ witness that in $V p_{\alpha} \Vdash ``\dot{x}_{\alpha} \in \check{V}$ '' where \dot{x}_{α} is a name for x_{α} . Since $|\mathbb{P}| < \delta$, there must be some $p^* \in G_{\mathbb{P}}$ such that $p_{\alpha} = p^*$ for unboundedly many α . But then this p^* decides all of X in V. And working in $M[G_{\mathbb{P}}]$, we can use the same idea for M to get some q^* deciding all of X in M. A common extension to p^* and q^* in $G_{\mathbb{P}}$ then shows $X \in V \cap M$.

It's not hard to see then that this gives agreement between M and V with respect to ordinal sequences.

- 35 D•10. Corollary

Consider the setup as in Gap Forcing (35 D \cdot 2). Therefore M and V have the same δ -sequences of ordinals.

Proof .:.

Without loss of generality, work with δ -sized sets of ordinals. Any such set in *V* can be enumerated and extended into $V \cap M$ by Lemma 35 D • 9. But by cutting off the enumeration at a point of length $< \delta^+ < \kappa$ in *M*, we get the set is in *M* and vice versa.

We will use a previous exercise Exercise $10 \cdot Ex29$ to simplify things slightly, which we prove here, giving a solution to the exercise.

- **35D**•11. Lemma — Let $\mathbf{M}, \mathbf{V} \models \mathsf{ZFC}$ be transitive models. Therefore $M \subseteq V$ iff every set of ordinals in M is in V.

Proof .:.

Let $x \in M$. Enumerate in M trcl($\{x\}$) and then code this enumeration as a subset of a single ordinal so that the code is in V and we can decode to get trcl($\{x\}$) $\in V$ and thus $x \in V$. More precisely, let $f : trcl(\{x\}) \to \kappa$ be a bijection in M. Let code : $\kappa \times \kappa \to \kappa$ be the definable bijection in M and V given in the proof of Lemma 5 D • 2. It follows that

 $A = \{ \operatorname{code}(f(a), f(b)) : a \in b \in \operatorname{trcl}(\{x\}) \} \in \mathcal{P}(\kappa) \cap M \cap V.$

V can then decode A with $\operatorname{code}^{-1} "A = \{ \langle f(a), f(b) \rangle : a \in b \in \operatorname{trcl}(\{x\}) \} \in V$. But $\langle \kappa, \operatorname{code}^{-1} "A \rangle \in V$ is isomorphic to $\langle \operatorname{trcl}(\{x\}), \in \rangle \in M$ by the mostowski collapse. Hence taking the mostowski collapse in V yields by uniqueness of the collapse that $\langle \operatorname{trcl}(\{x\}), \in \rangle \in V$. From this, $x \in V$ since x is the \in -maximal element of $\langle \operatorname{trcl}(\{x\}), \in \rangle \in V$.

Now we may actually get on with proving the main result.
- Proof of Gap Forcing $(35 D \cdot 2) \dots$

Note firstly that by Lemma 35 D•9, the statement " $cof(\theta) < \delta$ " is the same among V, V[G], M[j(G)] and M. This will be useful when using Lemma 35 D•8. Let $G = G_{\mathbb{P}} * G_{\mathbb{Q}}$ where $G_{\mathbb{P}}$ is \mathbb{P} -generic over V and $G_{\mathbb{Q}}$ is $\mathbb{Q} = \dot{\mathbb{Q}}_{G_{\mathbb{P}}}$ -generic over $V[G_{\mathbb{P}}]$. Recall that we need to show the following:

1. $M = V \cap M[j(G)].$

2. $j \upharpoonright V : V \to M$ is a class of V.

3. $\mathbf{V}[G] \models ``^{\lambda} M[j(G)] \subseteq M[j(G)]$ " implies $\mathbf{V} \models ``^{\lambda} M \subseteq M$ " for all $\lambda \in Ord$.

4. $V_{\lambda}^{\vee} \subseteq M[j(G)]$ implies $V_{\lambda}^{\vee} \subseteq M$.

Before we show any of these, it will be useful to show the following claim.

- Claim 1 $M \subseteq V.$

Proof .:.

By Lemma 35 D • 11, it suffices to show that every set A of ordinals in M is in V. We proceed by induction on $\theta = \sup A$. So suppose $A \in \mathcal{O}(\theta) \cap M$ with $A \cap \xi \in V$ for each $\xi < \theta$. If $\operatorname{cof}(\theta) \ge \delta$, then Lemma 35 D • 8 tells us $A \in V$ since otherwise $A \in V[G]$ would be fresh.

So assume $\operatorname{cof}(\theta) < \delta$. Write $G = G_{\mathbb{P}} * G_{\mathbb{Q}}$ for $G_{\mathbb{P}}$ P-generic over V. We know $A \in M \subseteq V[G]$ and by the $\leq \delta$ -distributivity of $\dot{\mathbb{Q}}_{G_{\mathbb{P}}} = \mathbb{Q}$ as a result of Corollary 35 B • 7, it follows that \mathbb{Q} didn't add A, i.e. $A \in V[G_{\mathbb{P}}]$. So let \dot{A} be a \mathbb{P} -name for A as a subset of θ . Let α be sufficiently large (e.g. such that $\mathbb{V}_{\alpha}^{\mathsf{V}}$ reflects the statements (1) " $p \Vdash$ " $\check{\alpha} \in \dot{A}$ " $\lor p \Vdash$ " $\check{\alpha} \notin \dot{A}$ "" and (2) $\forall \alpha < \theta$ ((1) holds) for $p \in \mathbb{P}$ and $\alpha < \theta$) and consider the (uncollapsed) hull

$$\mathsf{Hull} = \mathsf{Hull}^{\mathsf{V}_{\alpha}^{\mathsf{V}}}(\{\dot{A}, \mathbb{P}\} \cup \mathbb{P})$$

It should be clear that $G_{\mathbb{P}}$ is \mathbb{P} -generic over Hull and V_{α}^{\vee} . Moreover, Hull $\leq V_{\alpha}^{\vee}$. It's not hard to see that $|\operatorname{Hull}| = |\mathbb{P}| \cdot \aleph_0 \leq \delta$ so that Hull \cap Ord is a set of V and M by the previous argument. Since $A \in M$ already, it follows that $a = A \cap \operatorname{Hull} \subseteq$ Ord is also in M and as a small set, by Corollary 35 D • 10, $a \in V$. Hence there is some $p \in G_{\mathbb{P}}$ such that $p \Vdash ``\dot{A} = \check{a}$ '', meaning for each $\xi \in \operatorname{Hull}$,

$$p \Vdash ``\check{\xi} \in \dot{A}" \text{ or } p \Vdash ``\check{\xi} \notin \dot{A}". \tag{(*)}$$

By hypothesis on α being sufficiently large, (*) holds in V_{α}^{V} for each $\xi \in$ Hull and so by elementarity, (*) holds in Hull for each $\xi \in$ Hull, meaning Hull \models "*p* decides every element of \dot{A} " and thus the same holds in V_{α}^{V} and by hypothesis on α , this holds in V. Hence $A \in V$ defined by { $\alpha \in \theta : p \Vdash$ " $\check{\alpha} \in \check{\theta}$ "}. \dashv

As a corollary of Claim 1, using Corollary 35 D • 10, we have that ${}^{\delta}M \subseteq M$ in V. Now we show (1)–(4) above.

- Clearly M ⊆ M[j(G)] and M ⊆ V so it suffices to show V ∩ M[j(G)] ⊆ M. To show this, we defer to sets of ordinals. It's not immediate this suffices by Lemma 35 D 11, since it's not clear V∩M[j(G)] ⊨ ZFC yet. Proceed by induction: let x ∈ V ∩ M[j(G)] and assume inductively that x ⊆ M and so (since V, V[G], M[j(G)], and M all have the same ordinals) x ⊆ V^M_α for some α. In M, let g : λ → V^M_α be a bijection. Now consider A = {α < λ : g(α) ∈ x} as a set of ordinals. If A ∉ M, let θ ≤ λ be least such that A ∩ ξ ∈ M for all ξ < θ but A ∩ θ ∉ M. If cof(θ) ≥ δ then Lemma 35 D 8 tells us A ∩ θ ∈ M[j(G)] must be in M because it was added by a preorder admitting a gap at j(δ) = δ. So assume cof(θ) < δ as witnessed by θ = sup_{ξ<γ} θ_ξ in V for some γ < δ. Since ^δM ⊆ M in V, the sequence (A ∩ θ_ξ : ξ < γ) ∈ M and thus the union, A ∩ θ, is in M, a contradiction. Hence there is no such θ such that A ∩ θ ∉ M, i.e. A ∩ λ = A ∈ M. It follows that g⁻¹"A = x ∈ M since g ∈ M.
- 2. Showing $j \upharpoonright V$ is a class of V requires showing two things: that $j \upharpoonright V$ is amenable to V and that $j \upharpoonright V$ is FOLp-definable over V. First we show amenability, that if $x \in V$ then $j \cap x \in V$.

− Claim 2 − $j \upharpoonright V$ is amenable to *V* iff $j "θ \in V$ for every $θ \in Ord \cap V$.

Proof .:.

Clearly the (\rightarrow) direction holds: if $j \upharpoonright V$ is amenable to V then $j \cap (\theta \times V_{\alpha}^{V})$ for every $\theta, \alpha \in \text{Ord} \cap V$ and hence for $\alpha = \sup j "\theta, j "\theta = \operatorname{im}(j \cap (\theta \times V_{\alpha}^{V})) \in V$.

For the (\leftarrow) direction, assume $j "\theta \in V$ for every $\theta \in \text{Ord} \cap V$. Let $X \in V$ be arbitrary. We have $X \subseteq \text{dom}(X) \times V_{\alpha}^{V}$ for some α . Since $\text{dom}(X) \in V$, by enumerating $g : \lambda \text{ dom}(X)$, we have

$$j \cap X = \{ \langle a, b \rangle \in X : j(a) = b \} = \{ \langle g(\alpha), g(\beta) \rangle : \alpha \in \lambda \land j(\alpha) = \beta \}.$$

Note that $j "\lambda \in V$ and thus $j \upharpoonright \lambda \in V$ as the increasing enumeration of $j "\lambda$. Hence for $\alpha < \lambda$, $j(\alpha) = \beta$ is definable in V (via $\langle \alpha, \beta \rangle \in j \upharpoonright \lambda$) and so the above defines $j \cap X$ in V.

So let $\theta \in \text{Ord} \cap V$ be arbitrary and suppose inductively that $j"\xi \in V$ for each $\xi < \theta$. As a class of $\mathbf{V}[G]$, $j"\theta \in V[G]$. And as before, if $\operatorname{cof}(\theta)^{\mathbf{V}[G]} \ge \delta$ then $j"\theta \in V$ by Lemma 35 D • 8. So assume $\operatorname{cof}(\theta) < \delta$ as witnessed by $\langle \theta_{\xi} : \xi < \gamma \rangle \in V$ for $\gamma < \delta$. By $\le \delta$ -strategic closure and thus $\le \delta$ -distributivity of \mathbb{Q} , $\langle j"\theta_{\xi} : \xi < \gamma \rangle \in V[G_{\mathbb{P}}]$ and therefore $j"\theta$, as the supremum of these, is in $V[G_{\mathbb{P}}]$. So let τ be a \mathbb{P} -name for $j"\theta$. As with Claim 1, let α be sufficiently large (e.g. such that $\mathbf{V}_{\alpha}^{\mathsf{V}}$ reflects the statements (1) " $p \Vdash$ " $\check{\alpha} \in \dot{A}$ " $\lor p \Vdash$ " $\check{\alpha} \notin \dot{A}$ "" and (2) $\forall \alpha < \theta$ ((1) holds) for $p \in \mathbb{P}$ and $\alpha < \theta$) and consider again an uncollapsed hull

$$\mathsf{Hull} = \mathsf{Hull}^{\mathsf{V}_{\alpha}^{\mathsf{v}}}(\{\tau, \mathbb{P}\} \cup \mathbb{P}) \preccurlyeq \mathsf{V}_{\alpha}^{\mathsf{v}}.$$

Since $|\operatorname{Hull}|^{\vee} \leq \delta$, $j"\theta \cap \operatorname{Hull}$ has size $\leq \delta$ in \vee and so is in M by Corollary 35 D • 10. Furthermore, $j"\theta \cap \operatorname{Hull} = j"X$ for some $X \subseteq \theta$ in V[G] of size $\leq \delta$, and since we can enumerate $X = \{x_{\alpha} : \alpha < \delta\}$, $j(X) = \{j(x_{\alpha}) : \alpha < j(\delta) = \delta\} = j"X$ since $\delta < \kappa = \operatorname{cp}(j)$. By Lemma 35 D • 9, we can cover X by a set $Y \in V \cap M$ of size $\leq \delta$ and without loss of generality, $X \subseteq Y \subseteq \theta$. Thus we have $j"\theta \cap \operatorname{Hull} = j"X \subseteq j"Y \subseteq j"\theta \cap \operatorname{Hull}$ and so equality holds: $j"Y = j"\theta \cap \operatorname{Hull}$ and, as with X, j"Y = j(Y). But since $Y \in V, j(Y) \in M \subseteq V$. Thus \vee has access to all the information of $a = j"\theta \cap \operatorname{Hull}$ and we can find some $p \in G_{\mathbb{P}}$ such that $p \Vdash `\tau = \check{a}$ " meaning for each $\xi \in \operatorname{Hull}$,

$$p \Vdash ``\xi \in \tau$$
'' or $p \Vdash ``\xi \notin \tau$ '' *

By hypothesis on α being sufficiently large, (*) holds in V_{α}^{V} for each $\xi \in$ Hull and so by elementarity, (*) holds in Hull for each $\xi \in$ Hull, meaning Hull \models "p decides every element of τ " and thus the same holds in V_{α}^{V} . By hypothesis on α , this holds in V and thus j " θ is definable in V with j " $\theta = \{\alpha < \theta : p \Vdash$ " $\check{\alpha} \in \tau$ " $\} \in V$. By Claim 2, this finishes the proof that $j \upharpoonright V$ is amenable to V.

To show that $j \upharpoonright V$ is FOLp-definable over V, let φ be a formula that defines j over V[G]: j(x) = y iff $V[G] \models "\varphi(x, y, w)"$ for parameters w (which we write as a single parameter for notational convenience). Let \dot{w} be a $\mathbb{P} \ast \dot{\mathbb{Q}}$ -name for w. Thus for $x, y \in V$, j(x) = y iff $\exists p \in G$ ($p \Vdash "\varphi(\check{x}, \check{y}, \check{w})"$). For each $\theta \in \text{Ord} \cap V$, we then get a p_{θ} realizing this for $j \upharpoonright V_{\theta}^{\mathsf{V}}$ instead of just $j \upharpoonright V$, and so unboundedly many θ must have the same $p_{\theta} = p^*$. We can then use this p^* to define $j \upharpoonright V$ in V by j(x) = y iff $p^* \Vdash "\varphi(\check{x}, \check{y}, \check{w})"$.

- 3. This is easy since if ${}^{\lambda}M[j(G)] \subseteq M[j(G)]$ in V[G] then any λ -length sequence of $M \subseteq M[j(G)]$ in V is in V[G] and thus in M[j(G)]. So any such sequence is in $V \cap M[j(G)] = M$ by (1).
- 4. This is also easy since $V_{\lambda}^{\vee} \subseteq M[j(G)]$ implies $V_{\lambda}^{\vee} = V_{\lambda}^{\vee} \cap M[j(G)] \subseteq M$ by (1). \dashv

§ 35 E. Master conditions

One of the arguments used in the proof of Gap Forcing $(35 \text{ D} \cdot 2)$ above is to take a hull and use a single condition that gives a great deal of information about forcing over the hull. Such conditions are the last topic we focus on here. To motivate the idea, let's consider an extreme case. Suppose I have a traditional $j : V \to M$ and I stretch \mathbb{P} to $j(\mathbb{P})$ in such a way that there is some $p \in j(\mathbb{P})$ below every element of $j''\mathbb{P}$. Such a p has the property that if H is $j(\mathbb{P})$ -generic over M with $p \in H$, then $G = j^{-1}"H$ will be generic over \mathbb{P} and we can then clearly lift using Generic FORCING AND ELEMENTARY EMBEDDINGS

Lifting $(35 \text{ A} \cdot 2)$. This is an extreme idea, but the idea is that p provides a very nice way to lift the embedding, and more generally, p allows us to argue about forcing on the V side in M. This is the main motivation behind considering master conditions. First we begin with the easier idea of a *strong* master condition.

- **35E**•1. **Definition** Let V, M be transitive. Let $\mathbb{P} \in V$ be a preorder. Let $j : V \to M$ be elementary. Let $G \subseteq \mathbb{P}$. We call $p^* \in j(\mathbb{P})$ a strong master conditions for j and G iff $p^* \leq j(\mathbb{P})$ j(p) for all $p \in G$, i.e. p^* is below $j^{"}G$.

This easily allows us to lift up an embedding if this master condition is in a generic over M by Generic Lifting (35 A • 2).

— 35 E • 2. Corollary

- Let $j: V \to M$ be elementary between transitive models of ZFC we can force over.
- Let $\mathbb{P} \in V$ be a preorder appropriate for forcing.
- Let $G \subseteq \mathbb{P}$ be \mathbb{P} -generic over V.
- Let $p^* \in j(\mathbb{P})$ be a strong master condition for j and G.
- Let *H* be $j(\mathbb{P})$ -generic over *M* with $p^* \in H$.

Therefore, $G = j^{-1} H$ is P-generic over V, and so $j : V \to M$ lifts to $j^+ : V[G] \to M[H]$.

Proof .:.

Clearly $j"G \subseteq H$ since H is a filter and $p^* \leq^{j(\mathbb{P})} j(p)$ for any $p \in G$. That we can lift and $G = j^{-1}"H$ follows from Generic Lifting (35 A • 2) (2).

Now since H was an arbitrary generic containing the strong master condition, it follows that $G = j^{-1}$ "H for *any* such generic. One way to think about this is as a restriction on the kinds of generics we can find for M. But another way to think about this relationship is instead that our strong master condition is able to *generate* a generic. If we forget about being *given* G, such a condition p^* lets us generate a generic like G. So now we attempt to give a more general notion of a strong master condition that ignores G but still retains a similar property.

- 35 E • 3. Definition

Let V, M be transitive. Let $\mathbb{P} \in V$ be a preorder. Let $j : V \to M$ be elementary. We call $p^* \in j(\mathbb{P})$ a *strong master condition* for j and \mathbb{P} iff for every maximal antichain $A \subseteq \mathbb{P}$ in $V, j^{"}A$ has a unique element above p^* .

In essence, we may translate antichains from V and work with them in M. This gives Corollary $35 \text{ E} \cdot 2$ where we may not know what G looks like beforehand.

- 35 E • 4. Corollary

- Let $j: V \to M$ be elementary between transitive models of ZFC we can force over.
- Let $\mathbb{P} \in V$ be a preorder appropriate for forcing.
- Let $p^* \in j(\mathbb{P})$ be a strong master condition for j and \mathbb{P} .
- Let *H* be $j(\mathbb{P})$ -generic over *M* with $p^* \in H$.

Therefore, $G = j^{-1} H$ is P-generic over V, and so $j : V \to M$ lifts to $j^+ : V[G] \to M[H]$.

Proof ...

Clearly $j"G \subseteq H$. For genericity, we use the antichain characterization of Theorem 32 C • 5. Let $A \subseteq \mathbb{P}$ be a maximal antichain. Therefore j"A has a unique element $j(p) \in j"A$ such that $p \leq j(\mathbb{P})$ j(p). Hence $j(p) \in H \cap j"A$ so that $p \in G \cap A$. Since H is generic, $H \cap j"A$ has size at most 1 and therefore $|G \cap A| = 1$. Once we show G is a filter, this tells us G is \mathbb{P} -generic over M and we can lift by Generic Lifting (35 A • 2).

G is a filter because *H* is: if $p \leq \mathbb{P} q$ with $p \in G$ then $j(p) \leq \mathbb{P} j(q)$ with $j(p) \in H$. Since *H* is a filter, $j(q) \in H$ so that $q \in G$. For compatibility, clearly any two $p, q \in G$ are compatible in \mathbb{P} . Let *D* be the set of everything below both p, q or incompatible with one of them. By Result 32 C • 3, there is a maximal antichain $A \subseteq D$ where then $j"A \cap H$ is some j(r) with $p^* \leq j(r)$. It follows that j(r) can't be incompatible with j(p) or j(q) since all three are in *H* and *H* is a filter. Thus $j(r) \in H$ is below j(p), j(q) so that $r \in G$ is below p, q.

Forcing and Elementary Embeddings

This completes the proof that G is a filter and by the above remarks, we can lift j.

 \dashv

The general idea with strong master conditions is that they are able to not only generate a generic, but they tell us *exactly* what that generic should look like: p^* tells us where G should intersect antichains. Nevertheless, this is sometimes too strong of a condition to verify practically, and is more than what is necessary to make certain arguments go through. Loosely speaking, a strong master condition not only forces j^{-1} "H to be generic, it also tells us what that generic looks like.

A weak master condition can be defined just by forcing the preimage to be generic. The way to do this is by ensuring the preimage will always intersect maximal antichains, but it does not say *where* as in Definition 35 E • 3. This is equivalent to saying that j "*A* is *predense* below p^* for each such *A*. Such a notion is very useful in the context of PFA and related forcing axioms. Let us explain the common terminology.^{x1}

```
— 35 E • 5. Definition —
```

- Let \mathbb{P} be a preorder and $p^* \in \mathbb{P}$.
 - 1. A set *D* is *predense* iff every element of \mathbb{P} is compatible with some element of *D*.
 - 2. A set *D* is predense below p^* iff *D* is predense in $\mathbb{P}_{\leq p}$.

Such a set is *pre*dense in the sense that the downward closure of the set is dense in the same way that this is true of a maximal antichain. It's actually fairly clear that one can re-characterize a maximal antichain as a *predense* antichain. Moreover, every dense set is obviously predense. And this gives us the following.

- 35 E · 6. Corollary

Let $V \vDash ZFC$ be a transitive model we can force with $\mathbb{P} \in V$, a preorder, over. Therefore, for a filter $G \subseteq \mathbb{P}$, the following are equivalent:

1. G is \mathbb{P} -generic over V.

2. $G \cap A \neq \emptyset$ for every maximal antichain $A \in V$;

- 3. $G \cap D \neq \emptyset$ for every predense $D \in V$;
- 4. $G \cap D \neq \emptyset$ for every dense $D \in V$;

Proof .:.

The equivalence between (1), (2), and (4) is shown by Theorem 32 C • 5. (That $|G \cap A| = 1$ for A an antichain follows from G being a filter.) So suppose (3) holds. Since every dense set is predense, it follows that G is generic. For the other direction of $(1) \rightarrow (3)$, suppose D is predense. Let $D' \subseteq D$ be the downward closure of D so that D' is therefore dense. $G \cap D' \neq \emptyset$ and any element of this lies beneath some element $p \in D$. Since G is closed upward, $p \in G \cap D$ as desired.

And this motivates the idea of a weak master condition.

— 35 E • 7. Definition

Let $j: V \to M$ be elementary between transitive models of ZFC. Let $\mathbb{P} \in V$ be a preorder and $p^* \in j(\mathbb{P})$. We say p^* is a *weak master condition* for j, \mathbb{P} iff for every maximal antichain $A \subseteq \mathbb{P}$ in $V, j^{"}A$ is predense below p^* .

Predense sets and master conditions are often useful in the context of proper forcing where we often take skolem hulls.^{xli} But for now we will just cover the very basic ideas. Readers interested more in these topics should consult [17]. For now, what matters is that transforming critical information about the poset in V like maximal antichains into informative sets below a single p^* in M via $j : V \to M$ means that p^* has a lot of power in relation to the generics over V. This gives us Corollary 35 E • 4 for weak master conditions by basically the same proof.

xl This terminology was briefly brought up in Definition 33 D • 1, but we will actually look at the ideas here.

^{xli}Elsewhere in the literature, we would consider Hull^V(X) \subseteq H_K for some X and κ and consider j as the resulting (uncollapsing) map from cHull^V(X) to H_K. In this framework, a weak master condition for j, P is sometimes called (Hull^V(X), j(P))-generic [17]. The concept can also be thought of in terms of certain games and names for ordinals in a very nice way.

- 35 E • 8. Corollary -

- Let $j: V \to M$ be elementary between transitive models of ZFC we can force over.
- Let $\mathbb{P} \in V$ be a preorder appropriate for forcing.
- Let $p^* \in j(\mathbb{P})$ be a weak master condition for j and \mathbb{P} .
- Let *H* be $j(\mathbb{P})$ -generic over *M* with $p^* \in H$.

Therefore, $G = j^{-1} H$ is P-generic over V, and so $j : V \to M$ lifts to $j^+ : V[G] \to M[H]$.

Proof .:.

Clearly $j"G \subseteq H$. For genericity, we use the antichain characterization of Theorem 32 C • 5. Let $A \subseteq \mathbb{P}$ be a maximal antichain. Therefore j"A is predense below p^* . Since $p^* \in H$, $H \cap j"A \neq \emptyset$ by Corollary 35 E • 6 (3) and hence $G \cap A \neq \emptyset$. Once we show G is a filter, this tells us G is P-generic over M and we can lift by Generic Lifting (35 A • 2).

G is a filter because *H* is: if $p \leq^{\mathbb{P}} q$ with $p \in G$ then $j(p) \leq^{\mathbb{P}} j(q)$ with $j(p) \in H$. Since *H* is a filter, $j(q) \in H$ so that $q \in G$. For compatibility, clearly any two $p, q \in G$ are compatible in \mathbb{P} by elementarity and compatibility between $j(p), j(q) \in H$. Let *D* be the set of everything below both p, q or incompatible with one of them. By Result 32 C • 3, there is a maximal antichain $A \subseteq D$ where then j"A is predense below p^* and hence $j"A \cap H \neq \emptyset$ with some element j(r). It follows that j(r) can't be incompatible with j(p) or j(q) since all three are in *H* and *H* is a filter. Thus $j(r) \in H$ is below j(p), j(q) so that $r \in G$ is below p, q. This completes the proof that *G* is a filter and by the above remarks, we can lift *j*.

This actually provides a *characterization* of being a weak master condition.

– 35 E•9. Corollary -

- Let $j: V \to M$ be elementary between transitive models of ZFC we can force over.
- Let $\mathbb{P} \in V$ be a preorder appropriate for forcing.
- Suppose $p^* \in j(\mathbb{P})$ is such that $G = j^{-1}$ "H is \mathbb{P} -generic over V whenever H is $j(\mathbb{P})$ -generic over M with $p^* \in H$.

Therefore p^* is a weak master condition for j, \mathbb{P} .

Proof .:.

Suppose $A \subseteq \mathbb{P}$ is a maximal antichain in V. Let $q \leq j(\mathbb{P}) p^*$ be arbitrary. We need that there is some element of j "A compatible with q. Suppose not. Let H be $j(\mathbb{P})$ -generic over M such that $q \in H$ and therefore $p^* \in H$. It follows that $G = j^{-1}$ "H is \mathbb{P} -generic over V. As A is a maximal antichain, $G \cap A = \{r\}$ for some $r \in \mathbb{P}$ and therefore $j(r) \in H \cap j$ "A. Since H is a filter, $j(r) \in j$ "A is compatible with q. As $q \leq p^*$ was arbitrary, j"A must be predense below p^* .

Examples of master conditions can be found in, for example, the proof of Gap Forcing $(35 D \cdot 2)$, in particular the *p* such that (*) held. Other examples can be found with a common technique like the following.

35 E • 10. Result -

- Suppose $j: V \to M$ is traditional between transitive models of ZFC we can force over.
- Suppose $^{\lambda}M \subset M$.
- Suppose $\mathbb{P} \in V$ is such that
 - 1. $|\mathbb{P}|^{\mathsf{V}} \leq \lambda$; and
 - 2. $j(\mathbb{P}) \cong \mathbb{P} * \dot{\mathbb{Q}}$ where $\dot{\mathbb{Q}}$ is (forced to be) $\leq \lambda$ -directed closed in M.
 - A preorder is $\leq \lambda$ -directed closed iff for every set A such that $\forall p, q \in A \exists r \in A \ (r \leq p, q)$ (i.e. every *directed* set A), if $|A| \leq \lambda$ then there is some p^* below every element of A.
- Suppose *M* has only λ^+ -many dense sets (according to *V*) of $j(\mathbb{P})$.

Therefore we can lift j to $j^+: V[G] \to M[G * H]$.

Proof .:.

Let G be P-generic over V. By (1), $j \, "G$ has size $\leq \lambda$ in $\mathbf{V}[G]$ and is directed. $\dot{\mathbf{Q}}$ being forced to be $\leq \lambda$ -directed closed in \mathbf{M} implies $\dot{\mathbf{Q}}$ is also forced to be $\leq \lambda$ -directed closed in \mathbf{V} since $^{\lambda}M \subseteq M$. Thus $\dot{\mathbf{Q}}_G$ is also $\leq \lambda$ -directed closed in $\mathbf{V}[G]$ and so there is some p^* below every element of $j \, "G$, in other words, a strong master condition for j, G.

In V[G] we can enumerate the dense sets in M of $\dot{\mathbb{Q}}_G$ by $\{D_\alpha : \alpha < \lambda^+\}$. Now we construct a generic H below p^* as follows. Let $p_0 \in D_0$ be below p^* . Inductively, let $p_\alpha \in D_\alpha$ be below the directed set $\{p_\beta : \beta < \alpha\}$ (which, more than being directed is just linearly ordered). Such a p_α exists for $\alpha < \lambda$ by the $\leq \lambda$ -directed closure of $\dot{\mathbb{Q}}_G$. The result is $\{p_\alpha : \alpha < \lambda^+\}$ which intersects every dense set of $\dot{\mathbb{Q}}$ in $\mathbb{M}[G]$. It follows that the filter H generated by this set contains p^* and is generic over M[G]. Hence j^*G can be regarded as a subset of G * H and thus we can lift j to $j^+ : V[G] \to M[G * H]$.

It's not too difficult to see that M[G * H] remains closed under λ -length sequences in V[G]. These ideas and techniques are commonly used when dealing with *supercompact* cardinals, which assert the existence of a $j : V \to M$ with ${}^{\lambda}M \subseteq M$, $j(cp(j)) > \lambda$, for arbitrarily large λ . Such cardinals make calculating the number of dense sets of M relatively easy since such embeddings can be generated by a certain kind of ultrapower, which means one needs only to count the generating functions in V. And often, the \mathbb{P} of Result 35 E • 10 will be an iteration that starts with small preorders to ensure we can factor $j(\mathbb{P}) \cong \mathbb{P} * \dot{\mathbb{Q}}$ where the tail iteration has a certain amount of directed closure because the initial preorders in the iteration do too. Further reading in [17] and [4] is recommended.

Section 36. Exercises

In all of the following exercises, let $V \models ZFC$ be a transitive model we can force with $\mathbb{P} \in V$ over (for \mathbb{P} a preorder), and let G be \mathbb{P} -generic over V.

§36 A. Easier Exercises

- **36** Ex1. Exercise: Show that if \mathbb{P} has a minimal condition p, then $\{q \in \mathbb{P} : p \leq^{\mathbb{P}} q\}$ is \mathbb{P} -generic over V.
- **36 · Ex2.** Exercise: Let $S \subseteq \kappa$ be stationary in V, meaning $S \cap C \neq \emptyset$ for every club subset of $\kappa = cof(\kappa) > \aleph_0$ in V. Suppose \mathbb{P} is $< \kappa$ -closed. Show S is still stationary in V[G].
- **36**•**Ex3.** Exercise: Show $V_{\alpha}^{\vee[G]} = \{\tau_G : \tau \in V_{\alpha}^{\mathbb{P}}\}$ for any $\alpha \in \text{Ord.}$
- **36** Ex4. Exercise: Show that if \mathbb{P} is $< \kappa$ -distributive in **V** for some regular $\kappa \in \text{Ord then } V_{\kappa}^{\mathsf{V}} = V_{\kappa}^{\mathsf{V}[G]}$.
- **36 · Ex5. Exercise:** Suppose $j: V \to M$ is traditional and lifts to $j^+: V[G] \to M[H]$. Suppose every element of $V_{\alpha}^{\mathbb{P}}$ has a name in $V_{\alpha}^{\mathbb{V}}$, and $V_{\alpha}^{\mathbb{V}} = V_{\alpha}^{\mathbb{M}}$. Show $V_{\alpha}^{\mathbb{V}[G]} = V_{\alpha}^{\mathbb{M}[H]}$.
- **36 · Ex6.** Exercise: Show a countable support iteration of length ω_1 takes inverse limits at every limit stage except stage ω_1 where it takes the direct limit of previous iterations.
- 36 Ex7. Exercise: Show the results of Theorem 33 A 1 using previous proofs in Section 32.
- **36 Ex8. Exercise:** Show the results of The Generalized Δ-System Lemma (33 A 2) using the proof of The Δ-System Lemma (32 D 2).

— 36 A • 1. Definition -

Define $\mathbb{T} = \langle \mathbb{T}, \leq \rangle$ to be a *normal suslin tree* iff the following hold:

- 1. \mathbb{T} is a tree with height ω_1 .
- 2. All branches of \mathbb{T} have length $< \omega_1$.
- 3. All levels of \mathbb{T} have size $\langle \aleph_1$.
- 4. All antichains of \mathbb{T} have size $\langle \aleph_1$.
- 5. For all $t \in \mathbb{T}$, if $\alpha > \operatorname{rank}^{\mathbb{T}}(t)$, then there is an *s* of \mathbb{T} -rank (i.e. height) α with $t \leq s$ (i.e. every element can be extended to an element of arbitrary height).
- 6. For all $t_0, t_1 \in \mathbb{T}$ and limit $\alpha < \omega_1$, if $\operatorname{pred}_{\leq}(t_0) = \operatorname{pred}_{\leq}(t_1)$ then $t_0 = t_1$ (i.e. no splitting at limit stages).
- 7. For all $t \in \mathbb{T}$, the set $\{s \in \mathbb{T} : t \leq s \land \neg \exists r \ (t \leq r \leq s)\}$ is infinite (i.e. every element has infinitely many direct successors).

A suslin tree is just a tree satisfying (1)–(4).

Suslin trees are important in relation to the standard linear order on \mathbb{R} . We know $\langle \mathbb{R}, < \rangle$ is uniquely defined by three properties: it's a dense linear order without endpoints, it's a complete linear order, and it's separable. It's unclear whether separability can be weakened to merely being ccc (where an *antichain* for a linear order is a set of disjoint open intervals). A *suslin line* is a linear order witnessing that we cannot weaken separability to being ccc: a suslin line is not isomorphic to $\langle \mathbb{R}, < \rangle$ but it is a complete, ccc, dense linear order without endpoints. The existence of suslin lines is equivalent to the existence of suslin trees. *Suslin's hypothesis*, SH, is that there are no suslin lines. In the end, we have SH is independent of ZFC, of ZFC + CH, and of ZFC + \neg CH. In particular, we have that ZFC + MA + \neg CH \models SH while ZFC + "V = L" $\models \neg$ SH + CH. Using techniques of forcing, we can use this last relative consistency of ZFC + \neg SH

to force \neg CH while preserving \neg SH, basically showing \neg SH $\not\rightarrow$ CH. One can also show the relative consistency of ZFC + CH + SH so that SH and CH are completely independent of each other relative to ZFC.

36 · Ex9. Exercise:

- a. Show the existence of normal suslin trees is equivalent to the existence of suslin trees. (*Hint*: remove the points not satisfying (5))
- b. Show $MA + \neg CH$ implies there are no normal suslin trees.

Suslin trees are an easy counter-example to the nice looking claim that the product of ccc posets is ccc.

36 • Ex10. Exercise: Show that if $\mathbb{T} = \langle \mathbb{T}, \leq \rangle$ is a normal suslin tree as in Definition 36 A • 1, then $\mathbb{T}' = \langle \mathbb{T}, \geq \rangle$ is a ccc poset such that $\mathbb{T}' \times \mathbb{T}'$ is not ccc.

36 • Ex11. Exercise: Let $G = G_0 \times G_1$ be $\mathsf{Add}(\aleph_0, \aleph_2) \times \mathsf{Add}(\aleph_1, \aleph_3)$ in V. Show $V[G] \models "2^{\aleph_0} = \aleph_2 \wedge 2^{\aleph_1} = \aleph_3$ ".

§36B. Medium Exercises

- **36** Ex12. Exercise: Assume MA. Let $\kappa < |\mathbb{R}|$ be a cardinal and suppose X_{α} has lebesgue measure 0 for each $\alpha < \kappa$. Show $\bigcup_{\alpha < \kappa} X_{\alpha}$ has lebesgue measure 0. *Hint*: consider the preorder $\mathbb{P} = \{p \subseteq \mathbb{R} : p \text{ is open } \land \mu(p) <^{\mathsf{R}} \varepsilon\}$ where μ is lebesgue measure and $<^{\mathsf{R}}$ is the usual ordering on reals.
- **36** Ex13. Exercise: Let $\mathbb{P} = \bigstar_{n < \omega} \hat{\mathbb{Q}}_n$ be a finite support iteration with each $\hat{\mathbb{Q}}_{\alpha}$ appropriate for forcing. Show that \mathbb{P} adds a cohen real in the sense that there's a $g \in V[G]$ that is $\mathsf{Add}(\omega, 1)$ -generic over V.
- **36** Ex14. Exercise: Suppose $\mathbb{P} = \mathbb{Q} \times \mathbb{Q}$ is κ -cc for some $\mathbb{Q} \in V$. Suppose κ is measurable in V[G]. Show that κ is measurable in V.
- **36** Ex15. Exercise: Let κ be regular. Show that Add $(\kappa, 1)$ forces $2^{<\kappa} = \kappa$.
- **36 · Ex16. Exercise:** Consider the iteration $\mathbb{P} = \bigstar_{\alpha < \kappa} \dot{\mathbb{Q}}_{\alpha}$ where $\dot{\mathbb{Q}}_{\alpha}$ is (forced to be) Add($\aleph_{\alpha+1}, 1$), using easton support. Assume $2^{\lambda} = \lambda^+$ for all $\lambda \ge \kappa$ and show $\mathbb{P} \Vdash \mathsf{GCH}$.
- **36 Ex17.** Exercise: Suppose T is a suslin tree as in Definition 36 A 1. Let $\kappa > 2^{\aleph_0}$ be a regular cardinal. Show that Add (ω, κ) forces that T is still a suslin tree (and hence the consistency \neg CH with the existence of a suslin tree.)

36 • Ex18. Exercise: Let $S \subseteq \omega_1$ be stationary in V. Consider $\mathbb{P} = \langle \mathbb{P}, \leq \rangle$ defined by (in V)

$$\mathbb{P} = \{ X \subseteq S : \sup(X) < \omega_1 \land \forall Y \subseteq X \ (\sup(Y) \in X) \}.$$

ordered by $p \leq q$ iff $p \cap \max(q) = q$. Show that

- a. \mathbb{P} is $\leq \aleph_0$ -distributive in V.
- b. \mathbb{P} is not $\leq \aleph_0$ -closed in V.
- c. $V[G] \models " \bigcup G \subseteq S$ is a club of ω_1 ".

Conclude there is no FOLp(\in)-formula φ such that ZFC \vdash "{ $\alpha < \omega_1 : \mathbf{V}_{\omega_1} \models$ " $\varphi(\alpha)$ "} is stationary but doesn't contain a club".

§ 36 C. Harder Exercises

36 • Ex19. Exercise: Assuming whatever reasonable/standard large cardinal assumptions you'd like, find an example of an embedding $j : V[G] \to M[j(G)]$ such that $M \not\subseteq V$.

Exercises

- **36** Ex20. Exercise: For regular $\kappa \in \text{Ord}$, \Diamond_{κ} is the statement that there is a sequence $\langle A_{\alpha} \subseteq \alpha : \alpha < \kappa \rangle$ such that for every $X \subseteq \kappa$, $\{\alpha < \kappa : X \cap \alpha = A_{\alpha}\}$ is stationary in κ . We often write \Diamond for \Diamond_{\aleph_1} . We can use this statement to show the existence of a suslin tree in L, thus showing the independence of the existence of suslin trees from ZFC when coupled with Exercise 36 Ex9.
 - a. Show $Add(\kappa, 1)$ forces \Diamond_{κ} whenever κ is regular.
 - b. Show $L \models \Diamond_{\kappa}$ for every regular κ , where \Diamond_{κ} is as in Exercise 36 Ex20.
 - c. Show \Diamond_{κ} implies $2^{<\kappa} = \kappa$. Conclude $L \models GCH$.
 - d. Show that if \mathbb{T} is a tree satisfying (1)–(3) of Definition 36 A 1, any maximal antichain A of \mathbb{T} has

 $C = \{ \alpha < \omega 1 : A \cap \mathbb{T}_{\alpha} \text{ is a maximal antichain of } \mathbb{T}_{\alpha} \}$

as a club. Here \mathbb{T}_{α} refers to the subtree of \mathbb{T} consisting of elements of rank $< \alpha$.

- e. Show that $ZFC + \Diamond \vDash$ "there is a suslin tree", defined in Definition 36 A 1. (*Hint*: choose which points to extend in $\mathbb{T}_{\alpha} \subseteq \alpha$ according to whether A_{α} in the \Diamond -sequence is a maximal antichain of \mathbb{T}_{α} . Use this with (d).)
- **36** Ex21. Exercise (Easton Forcing): Let $F \in V$ be a function from regular cardinals to cardinals such that for all $\kappa \leq \lambda \in \text{dom}(F), \kappa < \text{cof}(F(\kappa))$ and $F(\kappa) \leq F(\lambda)$. Therefore there is a poset \mathbb{P} that forces $2^{\kappa} = F(\kappa)$ for all κ . We now walk through a proof.
 - a. Suppose \mathbb{P} is $\leq \kappa$ -closed and \mathbb{Q} is κ^+ -cc. Let $G = G_{\mathbb{P}} \times G_{\mathbb{Q}}$ be $\mathbb{P} \times \mathbb{Q}$ -generic over \mathbb{V} . Show every function $f : \kappa \to V$ in V[G] is in $V[G_{\mathbb{Q}}]$.
 - b. Define \mathbb{E}_F to be the product forcing $\prod_{\kappa \in \text{dom}(F)} \text{Add}(\kappa, F(\kappa))$ with Easton support, i.e. direct limits at weakly inaccessible stages and inverse limits elsewhere. Define $\mathbb{E}_F^{\leq \kappa}$ to be $\mathbb{E}_{F \upharpoonright \kappa^+}$, and similarly for the tail $\mathbb{E}_F^{>\kappa}$. Assuming GCH, show for all $\kappa \in \text{dom}(F)$, $\mathbb{E}_F^{\leq \kappa}$ is κ^+ -cc and $\mathbb{E}_F^{>\kappa}$ is $\leq \kappa$ -closed. (*Hint*: to show it's κ^+ -cc, use GCH with The Generalized Δ -System Lemma (33 A 2).)
 - c. Assume $V \models \text{GCH}$, and let $F \in V$ be as above. Let G be \mathbb{E}_F -generic over V. Show $V[G] \models "\forall \kappa \in \text{dom}(F) \ (2^{\kappa} = F(\kappa)) + \forall \kappa > \sup(\text{dom}(F)) \ (2^{\kappa} = \kappa^+)"$.

The next exercise about defining the ground model (using parameters) will make use of the following definitions. For the sake of reference, we include the concept of covering from Definition $33 B \cdot 4$.

- Let $V \subseteq W$ be transitive models of (sufficiently large fragments of) ZFC. Let δ be a cardinal of W.
 - 1. **V**, **W** have the $<\delta$ -covering property iff for all $A \in W$, if $A \subseteq V$ with $|A|^{\mathsf{W}} < \delta$ then there is some $A' \in V$ such that $A \subseteq A'$ and $|A'|^{\mathsf{V}} < \delta$.
 - 2. V, W have the $< \delta$ -approximation property iff for all $A \in W$, if $A \subseteq V$ and $\forall x \in V (|x|^V < \delta \rightarrow A \cap x \in V)$

then $A \in V$.

Succinctly, $<\delta$ -covering says that we can cover $<\delta$ -sized sets of **W** with $<\delta$ -sized sets in **V**, and $<\delta$ -approximation ssays that if every $<\delta$ -sized subset of A is in V then $A \in V$ (even if A is very large).

- **36** Ex22. Exercise (Defining the Ground Model): Let V be a transitive model of set theory we can force with \mathbb{P} over and let G be \mathbb{P} -generic over V.
 - a. Suppose δ is a regular cardinal of V. Suppose M, N \subseteq V are two inner models such that
 - i. $\mathcal{P}(\delta)^{\mathsf{M}} = \mathcal{P}(\delta)^{\mathsf{N}};$
 - ii. M, V and N, V both have the $< \delta$ -covering and $< \delta$ -approximation properties;
 - iii. $(\delta^+)^{\mathsf{M}} = (\delta^+)^{\mathsf{V}} = (\delta^+)^{\mathsf{N}}$.

Show that every set of ordinals in V of size $< \delta$ is contained in a set in $M \cap N$ of size $\leq \delta$.

- b. With the same setup, show that M and N have the same sets of ordinals. Conclude M = N.
- c. Suppose $\mathbb{P} \cong \mathbb{R} * \dot{\mathbb{Q}}$ admits a gap at δ^+ . Show V, V[G] have the $< \delta^+$ -covering property.
- d. Suppose $\mathbb{P} \cong \mathbb{R} * \dot{\mathbb{Q}}$ admits a gap at δ^+ . Show V, V[G] have the $< \delta^+$ -approximation property by the following line of reasoning.
 - i. Let A ∈ V[G] with A ⊆ Ord be arbitrary such that A ∩ X ∈ V for any δ-sized subset X ⊆ sup(A) ≤ κ in V. Let A be a P-name for A and consider M = Hull^{H^V_θ} (ℝ ∪ {κ, ℙ, A}) for some sufficiently large θ. Show that if ⟨p, q⟩ doesn't decide "α̃ ∈ A" then there are extensions p₀^{*}, p₁^{*} ≤^ℝ p and 1^ℙ ⊢ "q^{*} ≤^Q q" such that

 $\langle p_0^*, \dot{q}^* \rangle \Vdash ``\check{\alpha} \in \dot{A}''$ and $\langle p_1^*, \dot{q}^* \rangle \Vdash ``\check{\alpha} \notin \dot{A}''$.

ii. Show there is a $\dot{q} \in M$ where for every $p \in \mathbb{R}$, there are extensions $p_0^*, p_1^* \leq p$ and an ordinal $\alpha \in M$ such that

 $\langle p_0^*, \dot{q}^* \rangle \Vdash ``\check{\alpha} \in \dot{A}`` \text{ and } \langle p_1^*, \dot{q}^* \rangle \Vdash ``\check{\alpha} \notin \dot{A}``.$

- iii. Let $a = A \cap M \in V$ and extend $\langle \mathbb{1}^{\mathbb{R}}, \dot{q} \rangle$ to some $\langle p^*, \dot{q}^* \rangle \Vdash ``\dot{A} \cap \check{M} = \check{a}''$. Show $\langle p^*, \dot{q}^* \rangle$ decides all of A in V.
- e. Show V is FOLp(\in)-definable in V[G] from just the parameter $\mathcal{O}(|\mathbb{P}|)^{V}$. *Hint*: show the $<\delta^+$ -covering and $<\delta^+$ -approximation properties for $V_{\lambda}^{V}, V_{\lambda}^{V[G]}$, and use (b) to define V_{λ}^{V} as a certain inner model of $V_{\lambda}^{V[G]}$ (of sufficiently enough set theory) with certain properties that agrees with V on $\mathcal{O}(|\mathbb{P}|^+)$. Then reduce the need for $\mathcal{O}(|\mathbb{P}|^+)^{V}$ to $\mathcal{O}(|\mathbb{P}|)^{V}$.

Appendix A. Computability Theory

Computability is one of the most boring, tedious aspects of logic, akin to estimating infinite sums for traditional mathematics in that it mostly consists of looking at the fine details of long, hideous formulas and calculations. That said, it is unfortunately one of the most important aspects of the study, as it has close connections with definability and absoluteness. Fortunately, it develops into a study with more interesting questions with interesting methods through descriptive set theory and inner model theory. For now, we focus on the basics of computability theory.ⁱ

The most basic question that motivates the study is the following: "what can we compute?" This can be said in a slightly more formal manner of "What sets $A \subseteq \omega$ are such that for any x, we can compute whether $x \in A$ or not?" Of course, to answer such questions, we must first fix a notion of what it means to be able to compute something.

There is a distinction between the meta-theory and the formal theory similar to that of proof with first-order logic: we have our meta-theoretic sense of computation, and our more formal sense of it. Setting things up correctly, we should have the analog of "soundness": that if something is computable in the formal sense, then it's computable in the meta-theoretic sense. A priori, the converse, analogous to "completeness", can't be proven, as "computable" is a vague meta-theoretic concept that isn't as precise as truth for FOLp-formulas.

In light of the equivalence between all sufficiently strong methods of computationⁱⁱ, we adopt the Church–Turing thesis which states that if something is computable in the meta-theoretic sense, then it is computable in the formal as well. Often we will appeal to this philosophical stance as a form of laziness, avoiding having to give the precise details of an algorithm.

With all notions we present, we will start by defining which *functions* are considered computable by means of assuming that we can carry out certain processes. From this, we can say that a *set*, *relation*, and so on are computable according to whether there is a computable function outputting "true" or "false" (coded by 1 and 0 respectively) in accordance to whether the element is in the set, or the relation holds, and so on. In this way, we will expand the notion of computability from functions to sets, relations, propositions, and so on.

In doing so, we will often encounter the issue of where an algorithm does not give a final answer: it just keeps going through the instructions forever, never able to arrive anywhere. Hence we will often be dealing with *partial functions* rather than full fledged functions. In essence, a partial function is just a function where we state a larger domain than it actually has, similar to stating a larger range of a function than its actual image.

_ ^ · 1		inition —									
A -	i. Dei	muon									
A func	ction f	is a partia	l function from	n A to B,	written f	$: A \rightarrow$	B, iff there is	is some D	$\subseteq A$ where	$f: D \to B.$	

The only purpose of this is to say that the algorithm that determines $f: \omega \to \omega$ accepts inputs from ω , but perhaps doesn't return an answer for all of them. Moreover, it might not be clear what dom(f) will be just from its complicated definition. The use of partial functions is a nice concept that moves past these issues. Fortunately, we will not need to deal with the issue at first, since our first notion of computability yields only total functions.

ⁱComputability theory historically was called "recursion theory", and has since been re-branded by those within the field. Many old-guard set theorists still refer to it as recursion theory, and this is especially so when used in connection to set theory. To avoid a possible mix-up of terminology between recursive functions and functions defined by recursion, "computable functions" is used here. Unfortunately, *primitive recursive* functions have no such re-branding, and we use this term which no longer has a connection in name to computable functions.

ⁱⁱThere have been many different, complex systems of computation given over time, all of which have been proven to be equivalent in that something is computable in one sense iff it is in the other. This first started with Alonzo Church's λ -calculus and Alan Turing's turing machines, which were proven to have the same notion of computability by Turing.

Section A1. Primitive Recursion

We begin with the simplest notion of computability. We allow essentially two operations: adding 1, and the ability to repeat this. Unsurprisingly, we will need slightly more to flesh out this system, but nothing more computationally difficult than adding 1. Firstly, analogous to Recursion on ω (3 B • 2), we have the following definition.

A function f is defined by recursion iff dom $(f) \subseteq \omega^{n+1}$ for some $n < \omega$ and there are functions g and h where $f(0, \vec{y}) = g(0, \vec{y})$, and $f(x + 1, \vec{y}) = h(f(x), x, \vec{y})$, for all $x \in \omega$ and $\vec{y} \in {}^{n}\omega$

for all $x \in \omega$ and $y \in \omega$

This allows us to make precise the notion of "repeating".

– A1•2. Definition –

The set of primitive recursive functions is the \subseteq -smallest subset of partial functions from $\omega^{<\omega}$ to ω that is closed under composition, and recursion as in the scheme Definition A1 • 1, and the functions

• $\langle x_0, \cdots, x_n \rangle \mapsto x_i$ for each $i \leq n$ and each $n < \omega$;

- $x \mapsto 0$; and
- $x \mapsto x + 1$.

- A1•3. Corollary –

Every primitive recursive function is a "total" function, meaning its domain is all of ω^n for some $n < \omega$.

Proof .:.

Clearly each $\langle x_0, \dots, x_n \rangle \mapsto x_i, x \mapsto 0$, and $x \mapsto x + 1$ is total. By the inductive hypothesis, for total functions g and h as in Definition A1 • 1, the resulting f is total. And clearly the composition of total functions is total. Hence by structural induction, every primitive recursive function is total. \dashv

Primitive recursive functions are also FOLp-definable over $\langle \omega, +, \cdot, 0, 1 \rangle$. In fact, the axioms of peano arithmetic, PA, are able to define f on all of the *actual* natural numbers \mathbb{N} .

- A1.4. Result Let $f: \omega^n \to \omega$ be primitive recursive. Therefore f is definable over $\mathbf{N} = \langle \omega, +, \cdot, 0, 1 \rangle$.

The proof of this fact is delayed until later, since it makes use of some coding of finite sequences. In particular, to deal with definitions by recursion, we need to have a number that codes a list of previous calculations obeying the definition. Much of the following subsection will be to prove that such a coding is primitive recursive.

§ A1 a. Examples of primitive recursion

We immediately get by recursion that addition is primitive recursive. In particular, we proceed as follows:

$$f(0, y) = y$$

 $f(n + 1, y) = f(n, y) + 1.$

Here $g(x_0, x_1)$ is the map sending $\langle x_0, x_1 \rangle$ to x_1 as in Definition A1 • 2. $h(x_0, x_1, x_2) = x_0 + 1$ is then given by taking the first component of $\langle x_0, x_1, x_2 \rangle$ and adding 1: the composition of two functions from Definition A1 • 2. One can prove by induction that f(x, y) = x + y.

The same sort of idea easily gives the following as primitive recursive.

• The constant map $x \mapsto n$ is primitive recursive for each $n \in \omega$ as seen by repeated composition of 0 and $x \mapsto x + 1$.

PRIMITIVE RECURSION

- $(x, y) \mapsto x \cdot y$ is primitive recursive as seen by g(x, y) = x + y and h(z, x, y) = z + x.
- $\langle x, y \rangle \mapsto x^y$ is primitive recursive as seen by considering the map f' where $f'(x, y) = y^x$ as defined by g(x, y) = 1 and $h(z, x, y) = z \cdot y$.
- $x \mapsto x!$ is primitive recursive as seen by g(x) = 1 and $h(z, x) = z \cdot x$.
- $x \mapsto pd(x)$, the predecessor of x, is primitive recursive where pd(x + 1) = x while pd(0) = 0.
- $\langle x, y \rangle \mapsto x y$ is primitive recursive, where x y is 0 if $y \ge x$ and is x y otherwise. In the context of \mathbb{Z} rather than $\omega, x y = \max(0, x y)$.
- $\langle x, y \rangle \mapsto \max(x, y)$ is primitive recursive.
- $\langle x, y \rangle \mapsto \min(x, y)$ is primitive recursive.

To work out the details more precisely, we have the following slightly more difficult consequences.

– A1 a • 1. Corollary

Let f be primitive recursive. Therefore the maps

$$\sigma(\vec{x}, z) = \sum_{y < z} f(\vec{x}, y)$$
 and $\pi(\vec{x}, z) = \prod_{y < z} f(\vec{x}, y)$

are primitive recursive.

Proof .:.

We define each by recursion: take $g_{\sigma}(\vec{x}, y) = 0$ and $h_{\sigma}(t, y, \vec{x}) = t + f(\vec{x}, y)$ so that σ obeys

$$\sigma(\vec{x}, 0) = 0$$
 and $\sigma(\vec{x}, n+1) = h(\sigma(\vec{x}, n), n, \vec{x}) = \sigma(\vec{x}, n) + f(\vec{x}, n).$

By induction, it's easy to see σ defined in this way satisfies the requirements: inductively, $\sigma(\vec{x}, 0) = 0 = \sum_{y < 0} f(\vec{x}, y)$; and at successor stages, $\sigma(\vec{x}, n) = \sum_{y < n} f(\vec{x}, y)$ so that the above equality gives that

$$\sigma(\vec{x}, n+1) = \left(\sum_{y < n} f(\vec{x}, y)\right) + f(\vec{x}, n) = \sum_{y < n+1} f(\vec{x}, y).$$

A similar definition and idea holds for π :

$$\pi(\vec{x}, 0) = 1$$
 and $\pi(\vec{x}, n+1) = \pi(\vec{x}, n) \cdot f(\vec{x}, n)$ \dashv

By considering characteristic functions, we can also regard relations and sets as primitive recursive. Recall that for a set $A \subseteq {}^{<\omega}\omega$, the characteristic function is defined by

$$\chi_A(\vec{x}) = \begin{cases} 1 & \text{if } \vec{x} \in A \\ 0 & \text{otherwise.} \end{cases}$$

This demonstrates the usefulness of Corollary A1 a • 1 as the product $\prod_{y < z} \chi_A(\vec{x}, y)$ is 1 iff $\forall y < z \ (\langle \vec{x}, y \rangle \in A)$.

A1a•2. Definition

A set or relation $A \subseteq {}^{<\omega}\omega$ is *primitive recursive* iff the characteristic function χ_A is primitive recursive.

— A1 a • 3. Corollary

Let f be a function. Therefore f is primitive recursive as a function as in Definition A1 \cdot 2 implies it is primitive recursive as a relation as in Definition A1 $a \cdot$ 2.

Proof .:.

Suppose f is primitive recursive as a function. Then as a set, consider $\chi(\vec{x}, y) = 0^{y - f(\vec{x})} \cdot 0^{f(\vec{x}) - y}$. Recall that $0^0 = 1$ while $0^n = 0$ for $n \neq 0$. So if $f(\vec{x}) \neq y$, then either $f(\vec{x}) - y$ or $f(\vec{x}) - y$ will be non-zero (and in fact, the other will be 0), and hence $\chi_f(\vec{x}, y) = 1 \cdot 0 = 0$. Similarly, if $f(\vec{x}) = y$, then both $f(\vec{x}) - y = y - f(\vec{x}) = 0$ and hence $\chi_f(\vec{x}, y) = 0^0 \cdot 0^0 = 1 \cdot 1 = 1$. Therefore $\chi = \chi_f$ is primitive recursive, given by the composition of primitive recursive functions: -, multiplication, exponentiation, and f itself. \dashv

The converse to Corollary A1 a • 3 is not true in general. As we will see, primitive recursive functions do not encompass all of what is computable, and requires a kind of "bounded search" for an answer. Ostensibly, from a characteristic

function χ_f , we could just take the least y such that $\chi_f(\vec{x}, y) = 1$ and then get $f(\vec{x})$. But this has the issue of potentially being unbounded in a precise sense. So although it's computable from χ_f , that process isn't primitive recursive.

For now, the ability to talk about relations in a primitive recursive way allows for more complicated definitions. In particular, we have the following.

is primitive recursive.

Proof .:.

Note that $f(\vec{x}) = \chi_A(\vec{x}) \cdot g(\vec{x}) + (1 - \chi_A(\vec{x})) \cdot h(\vec{x})$ is clearly primitive recursive since χ_A , g, and h are. \dashv

This allows for some relations to be easily seen as primitive recursive, as it says that the set of primitive recursive functions is closed under definitions by cases (where the cases are primitive recursive). In particular, we have some further closure properties.

— A1 a • 5. Result — The set of primitive recursive relations is closed under

1. relative complements;

- 2. intersections; and
- 3. unions; and
- 4. bounded quantification (where the bound is primitive recursive).

Proof .:.

Let *A* and *B* be primitive recursive relations and *g* a primitive recursive function. Without loss of generality, *A* and *B* have the same arity, as we are more focused on χ_A and χ_B , and if one has fewer variables, just consider instead the expansion $f_A(\vec{x}, \vec{y}) = \chi_A(\vec{x})$ where we ignore the extra variables.

1. $\chi_{A \setminus B}(\vec{x}) = \chi_A(\vec{x}) - \chi_B(\vec{x}).$

2. $\chi_{A\cap B}(\vec{x}) = \chi_A(\vec{x}) \cdot \chi_B(\vec{x}).$

3. $\chi_{A\cup B}(\vec{x}) = \max(\chi_A(\vec{x}), \chi_B(\vec{x})).$

4. The last is bounded quantification, and by this, we mean the sets

$$A' = \{\vec{x} : \exists y \le g(\vec{x}) \ (\langle \vec{x}, y \rangle \in A)\} \text{ and } A'' = \{\vec{x} : \forall y \le g(\vec{x}) \ (\langle \vec{x}, y \rangle \in A)\}.$$

As one follows easily from the other, we will show A'' is primitive recursive. To do this, note that

$$\chi_{A''}(\vec{x}) = \min\{\chi_A(\vec{x}, y) : y \le g(\vec{x})\}.$$

From this, $\chi_{A''}(\vec{x}) = \pi(\vec{x}, g(\vec{x}))$ for π as in Corollary A1 a • 1 with f as χ_A , meaning $\chi_{A''}(\vec{x}) = \prod_{y \le g(\vec{x})} \chi_A(\vec{x}, y)$. Hence A'' is primitive recursive. A' being primitive recursive follows in the same sort of way: $\chi_{A'}(\vec{x}) = 1 - \prod_{y \le g(\vec{x})} (1 - \chi_A(\vec{x}, y))$ witnesses that A' is primitive recursive. \dashv

This is perhaps more appropriately stated in terms of *logic* rather than sets. In particular, intersections correspond to conjunctions, unions to disjunctions, relative complements to negations, and bounded quantification now more directly says what it means. In this way, we can restate Result A1 a \cdot 5 as follows.ⁱⁱⁱ

— A1 a · 6. Result

The set of primitive recursive relations is closed under conjunctions, disjunctions, unions, and bounded quantification (where the bound is primitive recursive). More formally, for $P(\vec{x})$ and $Q(\vec{x})$ two primitive recursive relations and $g(\vec{v})$ a primitive recursive function, then

ⁱⁱⁱNote the similarity here with Σ_0 -formulas in the Lévy hierarchy.

PRIMITIVE RECURSION

- $P(\vec{x}) \wedge Q(\vec{x})$ and $P(\vec{x}) \vee Q(\vec{x})$ are primitive recursive;
- $\neg P(\vec{x})$ is primitive recursive; and
- $\forall y \leq g(\vec{v}) \ P(\vec{v}, y)$ and $\exists y \leq g(\vec{v}) \ P(\vec{v}, y)$ are primitive recursive.

The benefit of bounded quantification in particular is that it allows for many simpler definitions, and for minimalization.

- A1 a • 7. Definition Let $P(\vec{x}, y)$ be a relation over ω . Write $\mu y < z P(\vec{x}, y)$ for the function $f(\vec{x}, z) = \begin{cases} \text{the least } y < z \text{ such that } P(\vec{x}, y) & \text{if there is one} \\ z & \text{otherwise} \end{cases}$

Write $\mu y P(\vec{x}, y)$ for the least y such that $P(\vec{x}, y)$, and leave it undefined if there is no such y.

Using this, we can show that primitive recursive functions are closed under bounded minimalization.

– A1a•8. Theorem

Let $P(\vec{x}, y)$ be a primitive recursive relation. Therefore $\mu y < z P(\vec{x}, y)$ is primitive recursive.

Proof .:.

Define by recursion $f(\vec{x}, z)$ so that it satisfies

$$f(\vec{x}, 0) = 0$$

$$f(\vec{x}, z + 1) = \begin{cases} f(\vec{x}, z) & \text{if } f(\vec{x}, z) \neq z \\ z & \text{if } P(\vec{x}, z) \\ z + 1 & \text{otherwise.} \end{cases}$$

It is easily seen that this is primitive recursive, and by induction that this is the same as $\mu y < z P(\vec{x}, y)$. \dashv

This is useful as it will generalize nicely to general computable functions. For now, we turn our attention to codings.

A "code" of a set A into a set B is just an injection $f : A \to B$ that ideally allows us to recover elements from A according to the encoded elements in B. We have already seen many examples of this, like associating a set $A \in \mathcal{O}(\omega)$ with its characteristic function $\chi_A \in {}^{\omega}2$. The first coding we will consider is encoding $\omega^{<\omega}$ into ω . We say that this coding is primitive recursive in the sense that $\operatorname{code}_n = \operatorname{code} \upharpoonright \omega^n$ is primitive recursive for each $n < \omega$.

– A1 a•9. Result

For each $n < \omega$, there is a primitive recursive, injective function $\operatorname{code}_n : \omega^n \to \omega$ where $\operatorname{im}(\operatorname{code}_n) \cap \operatorname{im}(\operatorname{code}_m) = \emptyset$ for $n \neq m < \omega$. Hence $\operatorname{code} = \bigcup_{n < \omega} \operatorname{code}_n$ is a primitive recursive injection from $\omega^{<\omega}$ to ω .

Proof .:.

We first prove two things are primitive recursive.

- Being a prime is primitive recursive. To see this, x is prime iff $\forall y < x \ \forall z < x \ (y \cdot z \neq x)$. By Result A1 a 6, this is primitive recursive as multiplication and equality of primitive recursive functions is primitive recursive.
- The map from *i* to the *i*th prime number is primitive recursive. We know already that there are infinitely many prime numbers. Thanks to the widely known proof of Euclid, we can put a bound on where to search for the next prime. In particular, if we have a list of all the primes below a number *x*, there must be another prime below the product of these primes +1 (because none of the primes in our list divide this number). In particular, there is one below x!+1. Therefore we can say that the *i*th prime number is defined by recursion where $p_0 = 2$ and $p_{i+1} = \mu p \le (p_i! + 1)$ (*p* is a prime). This shows that $i \mapsto p_i$ is primitive recursive.

As a result, defining $\operatorname{code}_n(\vec{x}) = \prod_{i < n} p_i^{x_i+1}$ (with $\operatorname{code}_0(\emptyset) = 1$) yields a function which is injective by the fundamental theorem of arithmetic. Moreover, this is primitive recursive, and $\operatorname{im}(\operatorname{code}_n) \cap \operatorname{im}(\operatorname{code}_m) = \emptyset$ for $n \neq m$, since for n < m, every $y \in \operatorname{im}(\operatorname{code}_m)$ is divisible by p_{m-1} (explaining the +1 in the exponent of each p_i) whereas no element of $\operatorname{im}(\operatorname{code}_n)$ is.

Such a coding is easy to work with in that we can also primitive recursively say whether a given natural number is a code of a sequence, what the length is, and what the *n*th entry in the sequence is.^{iv}

— A1 a•10. Result

The following are primitive recursive relations or functions from ω to ω : for x and y the codes of sequences in ${}^{<\omega}\omega$ and $n \in \omega$,

1. seq(n) iff *n* is the code of a finite sequence (of natural numbers).

2. $x \mapsto h(x)$, the length of $code^{-1}(x)$.

- 3. $x \mapsto x_n$, the *n*th entry of code⁻¹(*x*).
- 4. $\langle x, y \rangle \mapsto x^{\gamma} y$ which is the concatenation $\operatorname{code}^{-1}(x)^{\gamma} \operatorname{code}^{-1}(y)$.
- 5. $\langle x, i \rangle \mapsto x \upharpoonright i$ which is the restriction $\operatorname{code}^{-1}(x) \upharpoonright i$.

If x and y aren't the codes of sequences, define the functions to be 0.

Proof .:.

From these instructions or directions of algorithms, it should be clear that they are primitive recursive.

- 1. x is the code of a sequence iff x is of the form $p_0^{x_0+1} \cdots p_n^{x_n+1}$ for some $n < \omega$ and some $x_0, \cdots, x_n \in \omega$. So it's not hard to see that x is the code of a sequence iff for some n, every p_i divides x for i < n. This is easily seen as primitive recursive as $a \mid b$ is primitive recursive as witnessed by $a \mid b$ iff $\exists c \leq b \ (a \cdot c = b)$.
- 2. $\ln(x)$ is the least n < x such that $p_n \mid x$ while $p_{n+1} \nmid x$. And if $\neg seq(x)$, then set $\ln(x) = 0$.
- 3. Similar to h(x), if h(x) < n, we take x_n to be 0 and otherwise we take x_n to be the least e < x such that $p_n^{e+1} \mid x$ while $p_n^{e+2} \nmid x$.
- 4. Here we just set $x \uparrow y$ to be the least $m < \prod_{i < x+y} p_i^{x+y}$ (this is a bad bound, but it works) such that $\forall i < \ln(x) \ (p_i^{x_i+1} \mid m \land p_i^{x_i+2} \nmid m)$ and $\forall i < \ln(y) \ (p_{\ln(x)+i}^{y_i+1} \mid m \land p_{\ln(x)+i}^{y_i+2} \nmid m)$. And if x or y isn't (the code of) a sequence, we set it equal to 0.
- 5. Here we just set $x \upharpoonright i$ to be the least m < x such that $\forall k < i \ (p_k^{x_k+1} \mid x \land p_k^{x_k+2} \nmid x)$. If x isn't (the code of) a sequence, we say $x \upharpoonright i = 0$.

In essence, this allows us to *decode* natural numbers into sequences. The usefulness of this idea will be to talk about natural numbers as both sequences and as natural numbers, and hence talk about codes of codes and so forth. Ultimately, this will lead to the theorems of Tarski and Gödel, which have implications for logic and set theory in the form of the undefinability of truth and Gödel's incompleteness theorems.

For now, the idea of coding and decoding sequences allows us to prove Result A1 • 4. The proof of this does not make use of the fact that the coding above is primitive recursive, but instead just the idea of coding finite sequences of natural numbers as natural numbers themselves. The only part of the coding we care about is really the ability to *decode* natural numbers.

🖵 A1a•11. Lemma -

There is a definable coding (and decoding) of finite sequences over $\mathbf{N} = \langle \omega, 0, 1, +, \cdot \rangle$.

Proof .:.

To shorten the notation, we can clearly define \leq over N: $x \leq y$ iff $\exists z (y + z = x)$. Therefore x < y iff $x \leq y \land x \neq y$. Similarly, $x \mid y$ iff $\exists z (x \cdot z = y)$. Hence x is a prime number of N iff N \models " $x \neq 1 \land \forall y < x (y \mid x \rightarrow y = 1)$ ". To ease up notation even more, we have $x \equiv y \pmod{z}$ iff $\exists q (z = q \cdot x + y)$.

We unfortunately need to make use of a simpler (for N) to define coding mechanism, since we have no way to reference general exponentiation (and thus the usual coding) directly. This simpler to define coding is due to

^{iv}As with the discussions of HOD and L, the model of ZFC we're working in might have misinterpreted ω and hence might contain elements of " $\omega^{<\omega}$ " that aren't finite sequences of natural numbers. This is not a problem for our results, but it is something to keep in mind: the interpretation of ω might not be the actual N. So when I write "x is the code of a finite sequence", I really mean that $x \in \text{im}(\text{code})$, however "code" is interpreted in the model.

Gödel, and the actual function is called Gödel's β function. First, we require some number theory in the form of the Chinese Remainder Theorem. Note that this is really a theorem scheme for N: for each k, we get a new theorem.

— Claim 1 (The Chinese Remainder Theorem) ——

Let $\langle x_0, \dots, x_k \rangle \in \omega^k$ be given. Let $\langle n_0, \dots, n_k \rangle \in \omega^k$ be such that there is no prime p with $p \mid n_i$ and $p \mid n_j$ for $i \neq j$. Therefore there is exactly one $m < n_0 \cdot \ldots \cdot n_k$ such that for each $i \leq k, m \equiv x_i \pmod{n_i}$.

Proof .:.

For uniqueness, suppose *m* and *m'* both have $m, m' \equiv x_i \pmod{n_i}$ for all $i \leq k$. For the sake of definiteness, suppose m > m'. Therefore $m-m' \equiv x_i - x_i \pmod{n_i}$, meaning $n_i \mid m-m'$ for $i \leq k$. But then m-m' > 0 must have $n_0 \cdots n_k \mid m-m'$, meaning $m \geq n_0 \cdots n_k + m'$, contradicting the hypothesis that $m < n_0 \cdots n_k$.

For existence, the map M being $x \mapsto \langle a_0, \dots, a_k \rangle$ —where $a_i \in \omega$ is the least such that $x \equiv a_i \pmod{n_i}$ (and thus $a_i < n_i$)—is injective by the argument above. Looking at $M \upharpoonright n_0 \cdots n_k$ yields an injective function from $n_0 \cdots n_k \subseteq \omega$ to $n_0 \times \cdots \times n_k$. The Pigeonhole Principle (5 B • 8) then implies the map is surjective. Hence $M^{-1}(\langle x_0, \dots, x_k \rangle)$ witnesses the result.

This suggests we can code sequences of natural numbers just by looking at a few natural numbers. In particular, we define Gödel's β function:

 $\beta(x, y, i) = b$ where b is the least such that $x \equiv b \pmod{y \cdot (i+1) + 1}$.

Clearly β is definable over **N** as $\beta(x, y, i) = b$ iff $\exists q < x \ (x = q \cdot (y \cdot (i + 1) + 1) + b \land b < y \cdot (i + 1) + 1)$. The point of the function is the following claim.

Claim 2 Let $k \in \omega$ and let $\langle x_0, \dots, x_k \rangle \in \omega^{k+1}$. Therefore there are $x, y < \omega$ such that $\beta(x, y, i) = x_i$ for every $i \leq k$.

Proof .:.

Set $y = (\max(x_0, \dots, x_k, k) + 1)!$ and $n_i = y \cdot (i + 1) + 1$.

The n_i s are relatively prime: there can be no prime $p \mid n_i$ and $p \mid n_j$ for $n \neq j$ as otherwise for i < j, $p \mid n_i - n_j = y \cdot (j - i)$. As a prime, $p \mid j - i$ or $p \mid y$. p can't divide y since $p \mid 1 + (i + 1)y$ which would imply $p \mid 1$. Thus $p \mid j - i$. But p must be greater than $\max(x_0, \dots, x_k, k) + 1$ as otherwise $p \mid y$. Hence $p > k \geq j - i$, contradicting that $p \mid j - i \neq 0$ implies $p \leq j - i$.

The Chinese Remainder Theorem then states there is some $x < \prod_{i \le k} (y \cdot (i+1) + 1)$ such that $x \equiv x_i \pmod{n_i}$ for $i \le k$ and so $\beta(x, y, i) = x_i$ for $i \le k$.

This claim not only tells us that we can code all finite sequences of ω by triplets (x, y, and the length of the sequence), it also tells us we can *decode* them with the β function.

Now at this point, we could choose to define the usual coding $\vec{x} \mapsto \prod_{i < \ln(\vec{x})} p_i^{x_i+1}$ over **N**. But the process for doing so would take up more space than the goal for introducing it: proving Result A1 • 4. So because this would take up even more space for a topic already overstaying its welcome, we will work with the clumsy β function to show that primitive recursive functions are definable over **N**, and more generally, they are definable over PA: $f(\vec{x}) = y$ iff PA $\vdash \varphi(\#\vec{x}, \#y)$ where #y writes out the number y as " $0 + 1 + \cdots + 1$ " where there are y 1s. This generalization requires a bit more work in confirming that PA can prove and define all the necessary background material used thus far.

- Proof of Result A1 • 4 ...

Suppose f is defined by composition of primitive recursive functions: $f(\vec{x}) = g(h_1(\vec{x}), \dots, h_n(\vec{x}))$. Let φ_g define g and φ_{h_i} define h_i for each i as in the statement of the result. Therefore, for $\vec{x} \in \mathbb{N}^m$ and $y \in \mathbb{N}$, $f(\vec{x}) = y$ iff

 $\mathbf{N} \models ``\exists v_1 \cdots \exists v_n (\varphi_{h_1}(\vec{x}, v_1) \land \cdots \land \varphi_{h_n}(\vec{x}, v_n) \land \varphi_g(v_1, \cdots, v_n, y))''.$

PRIMITIVE RECURSION

Suppose f is defined by recursion: $f(0, \vec{w}) = g(\vec{w})$ and $f(x + 1, \vec{w}) = h(f(x), x, \vec{w})$. Let φ_g define g and φ_h define h as in the statement of the result. But because the following formula is so long, rather than writing " $\varphi_g(\vec{w}, y)$ ", write " $g(\vec{w}) = y$ ", and do similarly for h. To define f, we code the steps of computation as a finite sequence up to x, and then declare y is the last computation. More precisely, for $\vec{w} \in \mathbb{N}^m$ and $x, y \in \mathbb{N}$, $f(x, \vec{w}) = y$ iff

$$\mathbf{N} \models ``\exists a \exists b \exists y_0 \left(\begin{array}{c} g(\vec{w}) = y_0 \land \beta(a, b, 0) = y_0 \land \beta(a, b, x) = y \\ \land \forall i < y \exists y_1 \exists y_2 (\beta(a, b, i) = y_1 \land h(y_1, i, \vec{w}) = y_2 \land \beta(a, b, i + 1) = y_2) \end{array} \right) ".$$

That this is equivalent to $f(x, \vec{w}) = y$ can be easily checked inductively for $x, y \in \omega$ and $\vec{w} \in \omega^{<\omega}$. Hence f is definable over **N** and therefore by induction, all primitive recursive functions are definable over **N**.

§ A1 b. Bounds and primitive recursion

At this point, one might wonder what total, computable functions could exist that *aren t* primitive recursive. Of course, simply through combinatorial means, one can show that most total functions aren't primitive recursive.

– A1 b•1. Result –

There are only countably many primitive recursive functions.

Proof .:.

The set of functions containing $x \mapsto 0$, $\langle x_0, \dots, x_n \rangle \mapsto x_i$, and $x \mapsto x + 1$ for $i \le n < \omega$ is clearly countable. The closure of this set under the operations of composition and definitions by recursion is thus computable, and this yields all primitive recursive functions.

As a result, we can enumerate these functions $\{f : \omega^{<\omega} \rightarrow \omega : f \text{ is primitive recursive}\} = \{f_n : n < \omega\}$ and diagonalize: set $f : \omega \rightarrow \omega$ to be where $f(n) = \max\{f_i(n) : i \leq n\} + 1$. Hence we have an (eventual) bound on all primitive recursive functions. One might think that f is computable, as we just go through the list of primitive recursive functions and add one to the max up to that point each time. But this is dependent on the coding of primitive recursive functions: how n is associated to f_n , and more precisely how $\langle n, x \rangle$ is associated to $f_n(x)$. In particular, this association needs to be computable in order for the diagonalizing element f to be computable. We can show that such an association is possible later.

We now give an explicit example of an intuitively computable function that is not primitive recursive, but which (eventually) bounds every primitive recursive function. This famous function is referred to as the Ackermann function named after Wilhelm Ackermann.

- A1b•2. Definition The Ackermann function Ack : $\omega^2 \rightarrow \omega$ is defined by "double recursion" in that it satisfies Ack(0, x) = x + 1Ack(n + 1, 0) = Ack(n, 1)Ack(n + 1, x + 1) = Ack(n, Ack(n + 1, x)).

So to compute Ack(n, x), we need to compute Ack(n, y) for each y < x, eventually leading down to Ack(n, 0) and thus requiring Ack(n - 1, 1) and so on. It's not hard to see that such a function is well-defined. In particular, for each n, Ack_n being the map $x \mapsto Ack(n, x)$ satisfying the above can be defined by recursion and thus Ack, being the map $\langle n, x \rangle \mapsto Ack_n(x)$, is well-defined. The function grows very quickly, as Ack(3, 2) = 29 while $Ack(4, 2) = 2^{65536} - 3$. To show this in a semi-compact manner, firstly note that Ack(1, x + 1) = Ack(0, Ack(1, x)) = Ack(1, x) + 1. So inductively,

Ack(1, x) = Ack(1, 0) + x = Ack(0, 1) + x = x + 2.Therefore Ack(2, x + 1) = Ack(1, Ack(2, x)) = Ack(2, x) + 2. So inductively, Ack(2, x) = Ack(2, 0) + 2x = 2x + Ack(1, 1) = 2x + 3. PRIMITIVE RECURSION

Therefore Ack(3, x + 1) = Ack(2, Ack(3, x)) = 2Ack(3, x) + 3. So inductively,

$$\operatorname{Ack}(3, x) = 2^{x}\operatorname{Ack}(3, 0) + \sum_{k < x} 3 \cdot 2^{k}.$$

Since $Ack(3, 0) = Ack(2, 1) = 2 \cdot 1 + 3 = 5$, it follows that

Ack(3, x) =
$$5 \cdot 2^{x} + \sum_{k \in \mathbb{Z}^{x}} 3 \cdot 2^{k} = 5 \cdot 2^{x} + 3 \cdot 2^{x} - 3 = 8 \cdot 2^{x} - 3 = 2^{x+3} - 3$$

In particular, $Ack(3, 2) = 5 \cdot 4 + 3 \cdot 4 - 3 = 20 + 12 - 3 = 29$, and we can compute

$$Ack(4, 2) = Ack(3, Ack(4, 1))$$

= Ack(3, Ack(3, Ack(4, 0)))
= Ack(3, Ack(3, Ack(3, 1)))
= Ack(3, Ack(3, 2¹⁺³ - 3)) = Ack(3, Ack(3, 13))
= Ack(3, 2¹⁶ - 3)
= 2^{2¹⁶} - 3 = 2⁶⁵⁵³⁶ - 3.

As a result, despite using a relatively small number like 100, Ack(100, 100) is a truly enormous (finite) number. But the point of the Ackermann function (for our purposes anyway) isn't to generate large numbers, but to demonstrate that it isn't primitive recursive. To do this, we first should note the following properties of the Ackermann function.

- A1b•3. Lemma Write Ack_n for the map $x \mapsto Ack(n, x)$. Therefore for every $n < \omega$,

- 1. Ack_n is increasing: $x < y \in \omega$ implies Ack_n(x) < Ack_n(y);
- 2. $x < \operatorname{Ack}_n(x)$ for all $x \in \omega$;

3. $n < m < \omega$ implies $\forall x \in \omega (\operatorname{Ack}_{m}(x) < \operatorname{Ack}_{m}(x));$

- 4. $\operatorname{Ack}_n(\operatorname{Ack}_n(x)) < \operatorname{Ack}_{n+2}(x)$ for all $x \in \omega$.
- 5. Ack $_n$ is primitive recursive.

Proof .:.

1, 2. It suffices to show $x < Ack_n(x) < Ack_n(x + 1)$ for every $x < \omega$. Proceed by induction on *n*. For n = 0, this is clear: $x < Ack_0(x) = x + 1 < x + 2 = Ack_0(x + 1)$. For n + 1, proceed by induction on *x*.

For x = 0, $Ack_{n+1}(0) = Ack_n(1) > 1 > 0$ by the inductive hypothesis on n, which also yields that $Ack_{n+1}(1) = Ack_n(Ack_{n+1}(0)) > Ack_{n+1}(0)$. Hence the result holds for n + 1 when x = 0.

For x + 1, by the inductive hypothesis on n,

$$Ack_{n+1}(x+2) = Ack_n(Ack_{n+1}(x+1)) > Ack_{n+1}(x+1)$$

By the inductive hypothesis on *x*,

$$\operatorname{Ack}_{n+1}(x+1) = \operatorname{Ack}_n(\operatorname{Ack}_{n+1}(x)) > \operatorname{Ack}_{n+1}(x) > x.$$

Thus the result holds for all $n, x < \omega$.

3. Proceed by induction on n then by induction on m. For n = 0 and m = 1, Ack_n(x) = x + 1 while we already calculated above that Ack_m(x) = x + 2, which is clearly always larger. For m + 1 assuming ∀x ∈ ω(Ack_n(x) < Ack_m(x)), again proceed by induction on x. For x = 0, Ack_{m+1}(0) = Ack_m(1) > 1 = Ack_n(0) by (1). For x + 1,

$$Ack_{m+1}(x + 1) > Ack_{m+1}(x) > Ack_m(x) > Ack_n(x) = x + 1.$$

Hence $\operatorname{Ack}_{m+1}(x+1) > (x+1) + 1 = \operatorname{Ack}_n(x+1)$ so that the result holds for all *m* and *x* when n = 0.

For n + 1 < m, proceed by induction on x. For x = 0, we of course have

 $Ack_{n+1}(0) = Ack_n(1) < Ack_m(1) = Ack_{m+1}(0).$

Appx §A1 b

 \neg

For x + 1, by the inductive hypothesis on n and x,

 $\operatorname{Ack}_{n+1}(x+1) = \operatorname{Ack}_n(\operatorname{Ack}_{n+1}(x)) < \operatorname{Ack}_m(\operatorname{Ack}_{n+1}(x)) < \operatorname{Ack}_m(\operatorname{Ack}_{m+1}(x)) = \operatorname{Ack}_{m+1}(x+1).$ Thus the result holds for all $n < m < \omega$ and $x \in \omega$.

4. Proceed by induction on n. For n = 0,

$$\operatorname{Ack}_n(\operatorname{Ack}_n(x)) = (x+1) + 1 = x + 2 < 2x + 3 = \operatorname{Ack}_{n+2}(x).$$

For n + 1 and x = 0,

$$\begin{aligned} Ack_{n+3}(x) &= Ack_{n+2}(1) \\ &> Ack_n(Ack_n(1)) = Ack_n(Ack_{n+1}(0)) \\ &> Ack_n(Ack_{n+1}(0) - 1) = Ack_{n+1}(Ack_{n+1}(0)). \end{aligned}$$

For n + 1 and x + 1,

$$\begin{aligned} \operatorname{Ack}_{n+1}(\operatorname{Ack}_{n+1}(x+1)) &= \operatorname{Ack}_{n+1}(\operatorname{Ack}_n(\operatorname{Ack}_{n+1}(x))) \\ &< \operatorname{Ack}_{n+1}(\operatorname{Ack}_{n+1}(\operatorname{Ack}_{n+1}(x))) \\ &< \operatorname{Ack}_{n+1}(\operatorname{Ack}_{n+3}(x)) \\ &< \operatorname{Ack}_{n+2}(\operatorname{Ack}_{n+3}(x)) = \operatorname{Ack}_{n+3}(x+1). \end{aligned}$$

5. Clearly $Ack_0 = x \mapsto x + 1$ is primitive recursive. So assume $Ack_n(x)$ is primitive recursive. Therefore the function f defined by

$$f(0) = \operatorname{Ack}_n(1)$$
$$f(x+1) = \operatorname{Ack}_n(f(x))$$

is primitive recursive and is equal to $Ack_{n+1}(x)$.

- A1b•4. Theorem

For every primitive recursive function $f: \omega^m \to \omega$, there is an $n < \omega$ such that $f(\vec{x}) < \operatorname{Ack}_n(\max(\vec{x}))$ for all $\vec{x} \in \omega^m$.

Proof ...

We proceed by structural induction on f. If f is one of the basic, given functions $x \mapsto 0, \vec{x} \mapsto x_i$, or $x \mapsto x+1$; then clearly $Ack_1(max(\vec{x})) = max(\vec{x}) + 2 > f(\vec{x})$ for all \vec{x} .

Suppose *f* is defined by composition: $f(\vec{x}) = g(h_1(\vec{x}), \dots, h_k(\vec{x}))$. By the hypothesis, $g(\vec{x}) < \operatorname{Ack}_{n_0}(\max(\vec{x}))$ and $h_i(\vec{x}) < \operatorname{Ack}_{n_i}(\max(\vec{x}))$ for all \vec{x} and 1 < i < k. Therefore for $n = \max\{n_i : i < k\}$, by Lemma A1 b • 3,

 $f(\vec{x}) < \operatorname{Ack}_{n_0}(\max(h_1(\vec{x}), \cdots, h_k(\vec{x}))) \le \operatorname{Ack}_n(\operatorname{Ack}_n(\max(\vec{x}))) < \operatorname{Ack}_{n+2}(\max(\vec{x})).$

So suppose f is defined by recursion:

$$f(0, \vec{x}) = g(\vec{x}) f(n + 1, \vec{x}) = h(f(n, \vec{x}), n, \vec{x}).$$

Let *g* and *h* both be bounded by Ack_{m-1} .

— Claim 1 -

 $f(n, \vec{x}) < \operatorname{Ack}_m(n + \max(\vec{x}))$ for all $n < \omega$ and \vec{x} .

PRIMITIVE RECURSION

Proof ::. For n = 0, we know this holds since $f(0, \vec{x}) = g(\vec{x}) < \operatorname{Ack}_{m-1}(\max(\vec{x})) < \operatorname{Ack}_m(\max(\vec{x})) < \operatorname{Ack}_m(n + \max(\vec{x})).$ For n + 1, we have that $f(n + 1, \vec{x}) = h(f(n, \vec{x}), n, \vec{x}) < \operatorname{Ack}_{m-1}(\max(f(n, \vec{x}), n, \vec{x}))$ $< \operatorname{Ack}_{m-1}(\max(\operatorname{Ack}_m(n + \max(\vec{x})), x, \vec{x})))$ $= \operatorname{Ack}_m(n + \max(\vec{x}) + 1) \dashv$

Note that then

$$f(n, \vec{x}) < \operatorname{Ack}_{m}(n + \max(\vec{x})) \leq \operatorname{Ack}_{m}(2\max(n, \vec{x}))$$

$$< \operatorname{Ack}_{m}(2\max(n, \vec{x}) + 3) = \operatorname{Ack}_{m}(\operatorname{Ack}_{2}(\max(n, \vec{x})))$$

$$< \operatorname{Ack}_{m+2}(\operatorname{Ack}_{m+2}(\max(n, \vec{x}))) < \operatorname{Ack}_{m+4}(\max(n, \vec{x})).$$

Hence we have the result for f defined by recursion, and therefore, by induction, for all primitive recursive f.

- A1 b • 5. Corollary

The Ackermann function is not primitive recursive.

Proof .:.

If Ack were primitive recursive, then it would need to bound itself: there'd be an $n \in \omega$ with Ack $(x, y) < Ack_n(\max(x, y))$ for all x, y. In particular, for x = n and y = n, Ack $(n, n < Ack_n(\max(n, n)) = Ack_n(n) = Ack(n, n)$, a contradiction.

We also now have a counter-example to the converse of Corollary A1 a • 3.

— A1 b • 6. Corollary

The Ackermann function is primitive recursive as a relation (meaning the characteristic function for the graph is primitive recursive) but not as a function.

Proof .:.

We know the Ackermann function isn't primitive recursive as a function. As a relation, however, we know Ack(x, y) = z implies both x and y are less than z by Lemma A1 b • 3 (2) and (3). In particular, define χ as just whether a function obeying the Ackermann function up to that point would give z as the output to x and y. To talk about this, however, we need to do a bit of coding. What this means is we will calculate a certain number of Ack values and write them in a table and code this into a number. Note that since

$$\operatorname{Ack}(x, y) < \operatorname{Ack}(x - 1, z) = \operatorname{Ack}(x - 1, \operatorname{Ack}(x, y)) = \operatorname{Ack}(x, y + 1)$$

we need to check at most all values Ack(x', y') with $x' \le x$ and $y' \le z$. In particular, the number of outputs we need to keep track of is $z \cdot x$ and so using the usual primitive recursive coding, we need at most $z \cdot x$ primes. Since these values are bounded by Ack(x, y) = z, the number coding the calculation of Ack(x, y) is bounded by $\prod_{i < z \cdot x+1} p_i^{z+1} \le p_{z \cdot x+1}^{(z+1)}$, and thus we will work within this. As a primer, Ack(x', y') will be the $x' \cdot z + y'$ th entry of the coded table.

So define $\chi(x, y, z)$ to be 1 iff there is a number $c \le p_{z \cdot x+1}^{(z \cdot x+1) \cdot (z+1)}$ such that

- $\operatorname{seq}(c)$ with $\operatorname{lh}(c) = z \cdot x + 1$;
- for every x' < x and y' < z,

$$-c_{0\cdot z+0}=1,$$

 $-c_{(x'+1)\cdot z+0} = c_{x'\cdot z+1}$, and

PRIMITIVE RECURSION

$$- c_{(x'+1)\cdot z + (y'+1)} = \begin{cases} c_{x'\cdot z + c_{(x'+1)\cdot z + y'}} & \text{if } x' \cdot z + c_{(x'+1)\cdot z + y'} < \ln(c) \\ z + 1 & \text{otherwise;} \end{cases}$$

• $c_{z \cdot x + y} = z$.

Otherwise, $\chi(x, y, z) = 0$. It should be clear that this χ is primitive recursive as all these operations with the code are primitive recursive, and we're using bounded existential quantification. Moreover, χ is the characteristic function for the Ackermann function as a relation. Showing it's primitive recursive as a relation.

Ostensibly, we could then just take $Ack(x, y) = \mu z (\chi(x, y, z) = 1)$ for χ as above and get that Ack(x, y) is primitive recursive. This, of course, doesn't work, as we would need *bounded* minimalization. So the reason the Ackermann function isn't primitive recursive is that we can't get a bound on what Ack(x, y) should be just based on (in a primitive recursive way) x and y.

The existence of the Ackermann function not only tells us that is there a function which bounds all of the countably many primitive recursive functions, but also that there's one that is actually computable in an intuitive, meta-theoretic sense and is easy to compute in the sense above: it's graph is primitive recursive. Yet to have that Ack is computable in a formal sense, we need to expand our notion of computability from primitive recursive functions to computable functions.

Section A2. Computablility

As stated before, we will adopt the Church–Turing thesis, which in essence states that the only (partial) functions that are computable from $\omega^{<\omega}$ to ω are those that are computable via a turing machine, via a register machine, via an expression of λ -calculus, or via any sort of computer program given unlimited time and memory.

Now with primitive recursive functions, all the functions had domain ω^n for some $n < \omega$. This will not be the case with all computable functions, yet we will still be working with natural numbers. So recall the following definition.

```
A2.1. Definition —
A function f is a partial function from A to B, written f : A \rightarrow B iff f : D \rightarrow B for some D \subseteq A.
```

To deal with the fact that not all inputs have an output, we will introduce a new notion of equality, $\stackrel{\circ}{=}$. We will also introduce the notation that an input is not in the domain, leaving equality of functions as the usual set equality: f = g iff $\forall x \ (f(x) \stackrel{\circ}{=} g(x))$ iff $\forall x \ ((x \notin \text{dom}(f) \land x \notin \text{dom}(g)) \lor (x \in \text{dom}(f) \land x \in \text{dom}(g) \land f(x) = g(x)))$.

A2•2. Definition
Let f: A → B be a partial function. Let a ∈ A. We say
f converges at a, written f(a) ↓, iff a ∈ dom(f); and
f diverges at a, written f(a) ↑, iff a ∉ dom(f).
For g another partial function, we write f(a) = g(b) iff both diverge, or both converge and f(a) = g(b).

The empty function, \emptyset , for example, has $\emptyset(x) \uparrow$ for every x. There is no (set) function which converges everywhere^v although class sized functions obviously can, the identity being an obvious example. Now in the context of evaluating partial functions through algorithms, we may encounter the situation where $f(a) \uparrow$ but we need to evaluate $g(\vec{x}, f(a))$. In these cases, we say that $g(\vec{x}, f(a))$ diverges as well. Hence in definitions by recursion and composition, we often will use \triangleq instead of equality.

§ A2 a. μ -recursive functions

The next notion of computability we will investigate will be so-called μ -recursive functions in that they allow the operations of primitive recursive functions in addition to the minimalization operator μ as in Definition A1 a • 7. We restate this definition as a reminder since previously, we in practice only dealt with *bounded* minimalization, where $\mu x < z(P(x)) = z$ if there is no x < z where P(x). In the unbounded case, we leave the function as undefined if there is no such x. Since we are dealing with partial functions now, we should be a little more precise about how we do this. As we are interested in computation, the idea to find the minimal witness is just to proceed through the numbers, calculating one at a time until we reach a witness. This has some consequences for when this process diverges.

- A2a • 1. Definition Let $f: \omega^n \to \omega$ be a partial function. $\mu n \ge y$ $(f(n, \vec{x}) = 1)$ is the function *m* defined by $m(y, \vec{x}) \stackrel{\circ}{=} \begin{cases} \text{the least (and only) } n \text{ such that } f(n, \vec{x}) = 1 \\ \text{and } \forall i \ge y(i < n \to f(i, \vec{x}) \downarrow \land f(i, \vec{x}) \ne 1) & \text{if } \exists y \ (f(y, \vec{x}) = 1) \\ \text{undefined} & \text{otherwise.} \end{cases}$ We will just write " μn " for " $\mu n \ge 0$ ".

In particular, μn ($f(n, \vec{x}) = 1$) \uparrow iff

• there is no *n* with $f(n, \vec{x}) = 1$; or

votherwise replacement would yield that V is a set, contradicting Russell's Paradox (2 • 6).

• $f(n, \vec{x}) \uparrow$ with no $i \leq n$ where $f(i, \vec{x}) = 1$.

So if $f(0, \vec{x}) \uparrow$, then μy $(f(y, \vec{x}) = 1) \uparrow$, even if $f(1, \vec{x}) = 1$.

A2a•2. Definition

The set of μ -recursive functions is the \subseteq -least set of partial functions closed under composition, definitions by recursion, and minimalization (meaning $\vec{y} \mapsto \mu x$ ($f(\vec{y}, x) = 1$) for $f \mu$ -recursive) and containing all the primitive recursive functions.

– A2a•3. Corollary -

The Ackermann function is μ -recursive.

Proof .:.

By Corollary A1 b • 6, the Ackermann function as a set is primitive recursive. In other words, the graph of the Ackermann function is primitive recursive, as witnessed by the primitive recursive characteristic function $\chi(x, y, z)$ which is 1 iff Ack(x, y) = z. Therefore Ack $(x, y) = \mu z \chi(x, y, z)$ witnesses that the Ackermann function is μ -recursive.

We also have the same sort of closure conditions as in Result A1 a • 6.

— A2a·4. Corollary

The set of μ -recursive relations is closed under \wedge , \neg , and bounded quantification.

Another corollary of Definition A2 a • 2, as a consequence of Result A1 • 4, is that every μ -recursive function is definable over $\mathbf{N} = \langle \omega, 0, 1, +, \cdot \rangle$.

- A2a • 5. Corollary Every μ -recursive function is FOL-definable over $\mathbf{N} = \langle \omega, 0, 1, +, \cdot \rangle$.

Proof .:.

We have already proven all the primitive recursive functions are definable, and that the functions definable over **N** is closed under composition and definitions by recursion. Hence it suffices to show that they are closed under minimalization. In particular, suppose $f(\vec{x}, v) = y$ is defined by $\varphi_f(\vec{x}, v, y)$. Therefore the map sending \vec{x} to μv ($f(\vec{x}, v) = 1$) can be defined as follows. In particular, $\mu v f(\vec{x}, v) = 1$) = w iff

 $\mathbf{N} \vDash "\varphi(\vec{x}, w, 1) \land \forall v < w \ (\neg \varphi(\vec{x}, v, 1))".$

Hence the definable functions of N are closed under minimalization and therefore all μ -recursive functions are definable.

By the absoluteness of ω , 0, 1, and natural number addition and multiplication, it follows that the definitions of these computable functions are absolute between transitive models of ZF – P.

An example^{vi} of μ -recursive functions are those gotten by if-then-else statements.

– A2a•6. Result

Let χ , g, and h be partial functions. Define $f(\vec{x}) \stackrel{\circ}{=} \text{if } (\chi(\vec{x}) = 0)$ then $g(\vec{x})$ else $h(\vec{x})$ as follows. Firstly, $f(\vec{x}) \uparrow \text{if } \chi(\vec{x}) \uparrow$. And otherwise, for $\chi(\vec{x}) \downarrow$,

$$f(\vec{x}) = \begin{cases} g(\vec{x}) & \text{if } \chi(\vec{x}) = 0\\ h(\vec{x}) & \text{if } \chi(\vec{x}) \neq 0. \end{cases}$$

Therefore, if χ , g, and h are all μ -recursive (primitive recursive), then f is μ -recursive (primitive recursive).

Proof .:.

Ostensibly, we'd like to proceed by recursion, taking $f'(0, \vec{x}) \stackrel{\circ}{=} g(\vec{x})$ and $f'(n + 1, \vec{x}) \stackrel{\circ}{=} h(\vec{x})$, and then set $f(\vec{x}) \stackrel{\circ}{=} f'(\chi(\vec{x}), \vec{x})$. The problem with this approach is that to compute $f'(1, \vec{x})$, we necessarily

^{vi}We could have introduced this earlier, but there would be no point, as primitive recursive functions are all total (rather than mere partial) functions. So the worry demonstrated in the proof wouldn't have applied.

need to compute $g(\vec{x})$: more formally, we're taking $f'(n + 1, \vec{x}) \stackrel{\circ}{=} h'(f'(n, \vec{x}), n, \vec{x})$ where $h'(y, n, \vec{x}) \stackrel{\circ}{=} h(p_3(y, n, \vec{x}), p_4(y, n, \vec{x}), \cdots)$ where p_i is the map $\vec{y} \mapsto y_i$. The point is that if $g(\vec{x}) \uparrow$, then $h'(g(\vec{x}), 0, \vec{x}) \uparrow$ and hence $f(1, \vec{x}) \uparrow$ whereas the definition of the problem statement yields $f(1, \vec{x}) \stackrel{\circ}{=} h(\vec{x})$.

So instead, we define two functions by recursion:

$f_g(0, \vec{x}) = 0$	$f_h(0, \vec{x}) = 0$
$f_g(n+1,\vec{x}) \stackrel{\circ}{=} g(\vec{x})$	$f_h(n+1,\vec{x}) \stackrel{\circ}{=} h(\vec{x})$

Now we can define $f(\vec{x}) \stackrel{\circ}{=} f_g(1 - \chi(\vec{x}), \vec{x}) + f_h(\chi(\vec{x}), \vec{x})$. The result then follows immediately as both primitive recursive and μ -recursive functions are closed under composition and definitions by recursion. \dashv

The point of if-then-else is both because it is an intuitively computable operation, and it can be used as a word of caution when working with *partial* functions. To combine these two purposes, we have the following result that nicely ties together the programming or algorithm side of computability with the semantic side.

The function
$$m(y, \vec{x}) = \mu n \ge y$$
 $(f(n, \vec{x}) = 1)$ is the \subseteq -least partial function $g : \omega^n \to \omega, n < \omega$, satisfying
 $g(y, \vec{x}) \stackrel{\circ}{=}$ if $(f(y, \vec{x}) = 1)$ then y else $g(y + 1, \vec{x})$, (*)

for all y, \vec{x} .

Proof .:.

This consists of showing two parts: firstly, that *m* satisfies (*); and secondly, that $m \subseteq g$ for any *g* satisfying (*). That *m* satisfies (*) is pretty clear, but to show this formally, we break down into three cases. Write $m'(y, \vec{x})$ for if $(f(y, \vec{x}) = 1)$ then *y* else $m(y + 1, \vec{x})$ so that the goal is to show m = m'.

If $m(y, \vec{x}) \uparrow$, then one of the following holds:

• There is no y with $f(y, \vec{x}) = 1$ and thus

$$m'(y, \vec{x}) \stackrel{\circ}{=}$$
if $(f(y, \vec{x}) = 1)$ then y else $m(y + 1, \vec{x}) \stackrel{\circ}{=} m(y + 1, \vec{x})$,

which diverges for the same reason that $m(y, \vec{x}) \uparrow$.

- $f(y, \vec{x}) \uparrow$ and thus $m(y, \vec{x}) \stackrel{\circ}{=} m'(y, \vec{x}) \uparrow$.
- For some n > y, $f(n, \vec{x}) \uparrow$, where then $m'(y, \vec{x}) \stackrel{\circ}{=} m(y+1, \vec{x})$. Since $y+1 \le n$, $m(y+1, \vec{x}) \uparrow$ and thus $m'(y, \vec{x}) \stackrel{\circ}{=} m(y, \vec{x})$.

If $m(y, \vec{x}) = y$, then $f(y, \vec{x}) = 1$ and thus $m'(y, \vec{x}) = y$. So suppose $m(y, \vec{x}) = n > y$. Therefore for every *i* with $y \le i < n, 1 \ne f(y, \vec{x}) \downarrow$ and hence $m'(i, \vec{x}) = m(i + 1, \vec{x})$. So inductively, $m'(y, \vec{x}) \stackrel{\circ}{=} m((n-1)+1, \vec{x}) \stackrel{\circ}{=} m(n, \vec{x}) = n = m(y, \vec{x})$. Hence m = m' satisfies (*).

Now suppose g satisfies (*). Suppose $n = m(y, \vec{x}) \downarrow$ and thus $f(n, \vec{x}) = 1$ and for every *i* with $y \le i < n$, $1 \ne f(i, \vec{x}) \downarrow$. It follows for these *i* that $g(i, \vec{x}) \triangleq$ if $(f(i, \vec{x}) = 1)$ then *i* else $g(i + 1, \vec{x}) \triangleq g(i + 1, \vec{x})$ and so, $g(y, \vec{x}) \triangleq g(y + 1, \vec{x}) \triangleq \cdots \triangleq g(n - 1, \vec{x}) \triangleq g(n, \vec{x}) = n = m(y, \vec{x}) \dashv$

It's important to realize that there are g satisfying (*) that aren't equal to m. In particular, if $\neg \exists y \ (f(y, \vec{x}) = 1)$, then $g = m \cup \{\langle y, \vec{x}, 17 \rangle : y \in \omega\}$ satisfies (*) (in addition each function with any particular constant in place of 17).

§A2b. Computation

Definitions as in Result A2 a • 7 are important in that they introduce the idea of minimal solutions to equations involving partial functions as variables rather than merely natural numbers. Moreover, the natural way to find these *minimal* solutions is just to continually expand their definitions: a minimal solution will be undefined as much as possible unless forced to be a value by the definition. This leads to the idea of a *computer*.

The idea behind a computer is that it has a collection of inputs, and a collection of outputs, and starting from inputs,

COMPUTABLILITY

one can proceed along according to how the computer allows one to transition, and once one cannot go any further meaning one is at a *terminal* position—the computer outputs this, and this is *supposed* to be the output of the function with these inputs.

In set theoretic terms, a computer is just a set of trees each of height $\leq \omega$ with various rules dictating how one node transitions to the next. A branch of finite height yields either a computation (or else a defective input). To formalize all of this, we need to specify what these transitions are and what these nodes look like. The idea is to have nodes of the form " $\vec{f} \mid \vec{x}$ " where \vec{x} denotes a series of numbers or inputs, and \vec{f} is a series of functions that these inputs should be applied to, or some things to be simplified or expanded.

— A2b•1. Definition -

Let F be a set of partial functions over ω . For $f \in F$, let ' \dot{f} ' be a symbol intended to represent f. The *language* of COM(F) consists of $\stackrel{\circ}{=}$ between terms. The *terms of* COM(F) is the \subseteq -least set T of expressions containing the variables v_i for $i < \omega$ as well as the constants n for $n < \omega$ such that T is closed under the following:

- " $\dot{f}(\vec{t})$ " for $f \in F$ and \vec{t} the proper number of COM(F)-terms;
- " $\mathbf{p}_i^n(t_1, \dots, t_n)$ " for COM(F)-terms t_1 through $t_n, i, n \in \omega$, and " \mathbf{p}_i^n " a symbol playing the role of an *n*-place function *variable*.
- "if $(t_1 = 0)$ then t_2 else t_3 " for COM(F)-terms t_1, t_2 , and t_3 .

The point of these new function variables is to make precise the kind of thinking as in Result A2 a • 7. There, we're saying $\mu n \ge y$ ($f(n, \vec{x}) = 1$) is the \subseteq -least solution to the COM({ $f, x \mapsto x + 1$ })-equation

" $p(y, \vec{x}) \stackrel{\circ}{=}$ if $(f(y, \vec{x}) = 1)$ then y else $p(y + 1, \vec{x})$ ".

We have of course seen many examples of definitions of this form just through our usual definitions: " $S(x) \stackrel{\circ}{=} x + 1$ " and "proj_iⁿ $(x_1, \dots, x_n) \stackrel{\circ}{=} x_i$ ", for example. But generally the idea with a COM-equation " $p(\vec{x}) \stackrel{\circ}{=} t$ " is that t might involve p. For example, we can represent definitions by recursion: f defined by recursion using g and h is the \subseteq -least solution \hat{p}_0 where

"
$$p_0(y, \vec{x}) = \text{if } (y = 0) \text{ then } g(\vec{x}) \text{ else } h(p_0(y - 1, \vec{x}), y - 1, \vec{x})$$
".

If g and h are themselves defined by recursion or composition, we can then easily fully describe f according to a series of COM-equations. This idea naturally leads us to form programs (and Result A2 a • 7 tells us that we only need the simpler if-then-else in our language instead of minimalization if we want to arrive at all μ -recursive functions).

- A2b•2. Definition

Let F be a set of partial functions over ω . A COM(F)-program $P \neq \emptyset$ is a (finite) sequence of COM(F)-equations $p_0(\vec{x}) \stackrel{\circ}{=} t_0, \quad \dots, \quad p_n(\vec{x}) \stackrel{\circ}{=} t_n,$

where the only function variables occurring in the t_i s are the p_j s, and moreover, ' p_i ' \neq ' p_j ' for $i \neq j$.

Using this computer language, we can define a computer that computes the least solution of p_0 —call this function \hat{p}_0 —to an COM(*F*)-program *P*. In particular, from a state " $\vec{f} \mid \vec{x}$ ", we transition by certain rules to another state. If we start with input " $p_0 \mid \vec{x}$ " and end up in the state " $\mid n$ " for some $n \in \omega$, we output *n*, and write "that the *P*-computer computes $p_0(\vec{x}) = n$ ". More explicitly, we have the following transition rules, which clearly are deterministic: there's only ever at most one state to transition to.

A2b·3. Definition (Computer Transitions) -

Let F a set of partial functions over ω . Let P be a COM(F)-program. The following are allowable transitions.

• "x |" transitions to "| x" for $x \in M$.

• " $\dot{f}_i \mid \vec{x}$ " transitions to " $\mid f_i(\vec{x})$ ", where $f_i \in F$ and \vec{x} is of the appropriate length.

• " $\mathbf{p}_i \mid \vec{x}$ " transitions to " $t_i(\vec{x}) \mid$ ", where " $\mathbf{p}_i(\vec{v}) \stackrel{\circ}{=} t_i$ " is a COM(F)-equation of P.

In addition, we also have some transitions that make things go a bit smoother.

• " $h(\tau_1, \cdots, \tau_n)$ |" transitions to " $h \tau_1 \cdots \tau_n$ |".

- "(if $(\tau_1 = 1)$ then τ_2 else τ_3) |" transitions to " $\tau_2 \tau_3$? τ_1 |".
- " $\tau_2 \tau_3$? | 1" transitions to " τ_2 |".
- " $\tau_2 \tau_3$? | *n*" transitions to " τ_3 |" whenever $n \neq 1$.

And we don't care about surrounding material: if X transitions to Y, then $\sigma \gamma \gamma \zeta$ transitions to $\sigma \gamma \gamma \zeta$.

COMPUTABLILITY

For example, we can compute $x \mapsto 2x + 5$ when x = 1 with the COM($\{+, \cdot\}$)-program

$$p_0(x) \stackrel{\circ}{=} p_1(x) + 5$$
, $p_1(x) \stackrel{\circ}{=} 2 \cdot x$.

If we were to try to compute this at x = 1, we start as always with input \vec{x} —which in this case is just x being 1—on the node "p₀ | \vec{x} ", since the idea is that we always are trying to compute p₀. The computation—which of course, is not the most optimal way of calculating 2x + 5—then proceeds as follows:^{vii}

```
\begin{array}{c|c} p_{0} & | \ 1 \\ +(p_{1}(1), 5) & | \\ +p_{1}(1) & 5 & | \\ +p_{1}(1) & | \ 5 \\ +p_{1} & | \ 1 & 5 \\ +p_{1} & | \ 1 & 5 \\ + & (2, 1) & | \ 5 \\ + & \cdot & 21 & | \ 5 \\ + & \cdot & 2 & | \ 1 & 5 \\ + & \cdot & | \ 2 & 1 & 5 \\ + & \cdot & | \ 2 & 1 & 5 \\ + & | \ 2 & 5 \\ + & | \ 2 & 5 \\ & | \ 7. \end{array}
```

The idea is that if we end at "| n" for some $n \in \omega$, then we have computed $\hat{p}_0(\vec{x}) = n$. Of course, not all programs need to end at such a terminal state. The easiest such example would be " $p_0(x) \stackrel{\circ}{=} p_0(x)$ " which then has a computation for $n \in \omega$ as follows:

$$p_0 \mid n$$

$$p_0(n) \mid$$

$$p_0 \mid n \mid$$

$$p_0 \mid n$$

$$.$$

and this goes on forever. In this case, there is nothing computed by these transitions. This reflected in that the \subseteq -least solution to the equation " $p_0(x) \stackrel{\circ}{=} p_0(x)$ " is just $\hat{p}_0 = \emptyset$.

We now collect these ideas together in the form of a "computer". Of course, we still need to prove that such a computer will actually compute the least solution to a program.

Let F be a set of partial functions over ω . Let P be a COM(F)-program. For $\vec{x} \in \omega^{<\omega}$ and $n \in \omega$;

- We say the *P*-computer computes $p_0(\vec{x}) = n$ iff starting from " $p_0 | \vec{x}$ ", and transitioning as in Computer Transitions (A2 b 3), we end at "| n".
- If the *P*-computer does not compute any value of $p_0(\vec{x})$, then we say it computes $p_0(\vec{x}) \uparrow$.

We define the resulting function $\check{p}_0^P = \{\langle \vec{x}, n \rangle \in \omega^{<\omega} : \text{the } P \text{-computer computes } p_0(\vec{x}) = n\}.$ The \subseteq -least partial functions satisfying P are the functions $\hat{p}_0^P, ..., \hat{p}_n^P$ for P involving the function variables $p_0, ..., p_n^P$.

The \subseteq -least partial functions satisfying P are the functions p_0^r , ..., p_n^r for P involving the function variables p_0 , ..., p_n . (And it turns out that $\hat{p}_0^P = \check{p}_0^P$.)

Through reordering the equations of P, this also easily provides definitions for \check{p}_k^P for k other than 0 (if there are any in P).

We should check two things of this sort of computer: firstly that it computes correctly, and secondly that it computes minimal solutions.^{viii} Showing soundness is relatively easy, although long and tedious. We skip the proof for the sake

^{vii}Note that we are abusing notation by writing, for example, " $2 \cdot x$ " above, which really should be regarded as " $\cdot (2, x)$ ", and similarly for +. This is reflected in the computation.

viiiThis is analogous to soundness and completeness in first-order logic.

of clarity.

Let F be a set of partial functions over ω , and let P be a COM(F)-program involving $p_0, ..., p_n$. Therefore $\check{p}_0^P, ..., \check{p}_n^P$

are the \subseteq -least partial functions that satisfy P, meaning they satisfy P and any sequence of partial functions $f_0, ..., f_n$ f_n satisfying P have $\check{p}_i^P \subseteq f_i$ for $i \leq n$.

Proof .:.

By soundness of the *P*-computer, these \check{p}_n^P satisfy *P*. Now suppose $f_0, ..., f_n$ also satisfy *P*. To see that the \check{p}_i^P are the least solutions, we need to check that if $\check{p}_i^P(\vec{x}) \downarrow$, then $f_i(\vec{x}) \downarrow$ and $\check{p}_i^P(\vec{x}) = f_i(\vec{x})$ for each $i \leq n$.

Suppose the *P*-computer computes $\check{p}_i^P(\vec{x}) = n_i$. This means that starting with " $p_i \mid \vec{x}$ " and following the transitions of Computer Transitions (A2 b \cdot 3), we arrive at "| n_i ". We now proceed by induction on COM(F)-terms τ involving no variables (excluding the p_i^k s).

- Claim 1

Let τ be a COM(F)-term involving no variables except possibly function variables. Suppose " τ |" eventually transitions to "| y". Therefore, the interpretation of τ —meaning where we replace \dot{f} with $f \in F$, the p_is with the f_i s, and then evaluate as normal—is y.

Proof .:.

Let τ be a counter-example such that the computation from " τ |" to "| y" is of minimal length. Clearly τ can't be a constant, since otherwise " τ |" transitions to "| τ " and so the result clearly holds.

Suppose τ is of the form " $f(\vec{\sigma})$ " where $\vec{\sigma}$ are COM(F)-terms. In computing τ , we also compute each σ_i and thus their computations are shorter than τ . Say the *P*-computer computes that σ_i is y_i . Inductively, the interpretation of each σ_i is y_i so that the interpretation of τ is $f(\vec{y})$. Note that " τ " transitions as follows: " $\dot{f}(\vec{\sigma})$]" goes to " $\dot{f} \sigma_0 \cdots \sigma_m$]", which eventually transitions to " $\dot{f} \sigma_0 \cdots \sigma_{m-1}$ | y_m ", and by repetition to " $\dot{f} \mid y_0 \cdots y_m$ " which goes to " $\mid f(y_0, \cdots, y_m)$ " which is the interpretation of τ .

If τ is of the form "(if $(\sigma_1 = 1)$ then σ_2 else σ_3)", then the interpretation of τ is y_2 if $y_1 = 1$, and otherwise y_3 . Note that " τ |" transitions to " $\tau_2 \tau_3$? τ_1 |" and thus eventually to " $\tau_2 \tau_3$? | y_1 ". If $y_1 = 1$, this transition sitions to " τ_2 |" and thus to "| y_2 ", the same as the interpretation. If $y_1 \neq 1$, then this transitions to " τ_3 |" and thus eventually to " $| y_3$ ", which is the interpretation.

If τ is of the form " $\mathbf{p}_i(\vec{\sigma})$ ", then we proceed as follows: " τ |" transitions to " $\mathbf{p}_i \sigma_1 \cdots \sigma_m$ |" and so eventually to " $\mathbf{p}_i \mid y_1 \cdots y_m$ " and then to " $t_i(\vec{y}) \mid$ " where P contains the equation " $\mathbf{p}_i(\vec{v}) = t_i$ ". From here, we somehow eventually arrive at "| y" and thus since the computation from " $t_i(\vec{y})$ |" is shorter than from " τ |", it follows that the interpretation of $t_i(\vec{y})$ is y. But the interpretation of $t_i(\vec{y})$ is $f_i(\vec{y})$ since the f_i 's satisfy P. Moreover, $y = f_i(\vec{y})$ is the interpretation of " $p_i(\vec{\sigma})$ ", which is τ . Hence the result holds for τ as well.

For each $i \leq n$, we know " $\mathbf{p}_i \mid \vec{x}$ " transitions to " $t_i(\vec{x}) \mid$ ". Since $t_i(\vec{x})$ involves no variables other than function variables, its interpretation under the f_i s is the same as what the P-computer computes, which is n_i . Hence $n_i = f_i(\vec{x}) \downarrow$, which is by definition the same as $\check{p}_i^P(\vec{x})$. Therefore $\check{p}_i^P \subseteq f_i$, and the result is proven. +

Another way to phrase the result above is that $\check{p}_i^P = \hat{p}_i^P$ for each *i*. With this, we can then represent each μ -recursive partial function by a program. In particular, we have the following definition.

- A2b•6. Definition

Write s for the map $x \mapsto x + 1$ and pd for the map $x \mapsto x - 1$. Say a function f is computable iff there is a $COM(\{s, pd\})$ -program P such that $f = \check{p}_0^P$.

A2b•7. Corollary

Every μ -recursive function is computable.

Proof .:.

The programs

$$v_{p_0}(v_0, \dots, v_n) = v_i$$
", and " $v_0(v) = 0$ "

clearly compute projections and the constant 0 map. Since $x \mapsto x + 1$ is given, to show that all primitive recursive functions are computable, it suffices to show the set of computable functions is closed under definitions by recursion. But this was shown just below Definition A2 b • 1: if f is defined by recursion using g and h, which are in term computed by programs P_g (with first variable p_{g_0}) and P_h (with first variable p_{h_0} ,) then (adjusting the names of the function variables as needed) f is the least solution \hat{p}_0^P to the program

" $p_0(n, \vec{x}) \stackrel{\circ}{=} \text{if } (n = 0) \text{ then } p_{g_0}(\vec{x}) \text{ else } p_{h_0}(p_0(\text{pd}(y), \vec{x}), \text{pd}(y), \vec{x})$ " $P_g \stackrel{\frown}{=} P_h$.

It should be clear that this program computes f and thus the set of computable functions is closed under definitions by recursion, implying that all primitive recursive functions are computable. Moreover, by Result A2 a • 7, the set of computable functions is closed under minimalization. Hence all the μ -recursive functions are computable. \neg

We can also show that the reverse is true: all computable functions are μ -recursive just by the coding idea introduced to deal with the Ackermann function. In particular, we code the symbols of COM({s, pd}) as numbers as follows:

- the code of 's' is 0;
- the code of 'pd' is 1;
- the code of v_i is 2^i ;
- the code of ' p_i^n ' is p_{i+1}^n where p_i is the *i*th prime;
- the code of *n* for $n \in \omega$ is 6^n ;
- the code of '(' is 10 and ')' is 12;
- the code of ',' is 14;
- the code of '=' is 15;
- the code of 'if' is 17;
- the code of 'then' is 18; and
- the code of 'else' is 20.
- the code of '|' is 21.

From these codes, sequences of symbols are coded by the classical coding $(x_0, \dots, x_n) \mapsto p_0^{x_0+1} \dots p_n^{x_n+1}$. We appeal to laziness to convince the reader that the property of a number $n \in \omega$ being the code of a COM({s, pd})-term is primitive recursive. Mostly this just means translating the rules governing term formation into the corresponding rules in terms of the corresponding symbol codes.

This also allows us to code nodes in computations. Then sequences of these nodes correspond to computations. It's also not hard—although very tedious—to show that the property of a number being the code of a computation following the rules of Computer Transitions (A2 b • 3) is primitive recursive. And as a result, the output of a computation of a COM({s, pd})-program P can be seen as the least pair (or rather, the first element of the least number coding a pair) $\langle y, c \rangle$ where c codes a computation from P that ends with code("| y"). Thus every partial function given by a program P can be represented in this way and is thus μ -recursive.

- A2 b \cdot 8. Corollary - Every computable function is μ -recursive.

This—as with the equivalence between μ -recursive functions, turing machines, and so on—lends credence to the idea of the Church–Turing thesis, which states the equivalence between an intuitive, meta-theoretic notion of computability, and a formal notion of computability like that of μ -recursive or computable as in the sense of Definition A2 b • 6.

§A2c. Normal form

The idea of coding a program is quite important. Although the explicit details are skipped here for the sake of space, they can in principle be worked out. But the main ideas do not rely too much on these details, since there are many notions of computability that all turn out to be equivalent.

Let's first recall the ideas informally stated above. The details of the proof of this lemma are left to any reader that is interested enough^{ix} given the discussion above Corollary $A2b \cdot 8$ as a hint.

— A2c·1. Lemma

The following are primitive recursive relations and functions:

- The property Prog(e) stating "e is the code of a COM({s, pd})-program".
- For each $1 \le n < \omega$, the relation CompCode_n(e, x_1, \dots, x_n, y) stating "e is the code of a program and y is the code of a computation with "p₀ | $x_1 \cdots x_n$ " at the start".
- The function Output(y) being $\mu n < y$ ("| n" is the last entry of the computation coded by y).

For every $f : \omega^n \to \omega$, where $n < \omega$, f is computable iff there is some e such that $f(\vec{x}) \stackrel{\circ}{=} \text{Output}(\mu y \text{ (CompCode}_n(e, \vec{x}, y)).$

In particular, every μ -recursive function only needs to use (unbounded) minimalization *once* with all other relevant functions and relations being primitive recursive. Another corollary of this is the identification of computable functions with programs as well as the extremely important—although a bit odd looking—result below. In essence, it's saying that given the first *m* inputs in an *m* + *n*-ary partial function given by a program *e*, we can find a new program *e'* that outputs the same partial function but fixing those first *m* inputs. So S_n^m takes the first *m* inputs, and leaves the last *n* inputs.

— A2 c · 3. Theorem (The S_n^m -Theorem) — — —

For each $1 \le n, m < \omega$, there is an injective, primitive recursive function $S_n^m : \omega^{m+1} \to \omega$ such that for any $e \in \omega$, $\vec{a} \in \omega^m$, and $\vec{x} \in \omega^n$,

Output(μy (CompCode_{*m*+*n*}(e, \vec{a}, \vec{x}, y)) $\stackrel{\circ}{=}$ Output(μy (CompCode_{*n*}($S_n^m(e, \vec{a}), \vec{x}, y$)).

Let us introduce some notation. Because we can identify computable functions with codes of programs and the number of inputs, we write $[e]^n$ for the *n*-ary partial function mapping \vec{x} to Output(μy (CompCode_n(e, \vec{x}, y))^x. Hence

```
\begin{bmatrix} [0]^{1}, [[1]]^{1}, [[2]]^{1}, [[3]]^{1}, \cdots \\ \begin{bmatrix} [0]^{2}, [[1]]^{2}, [[2]]^{2}, [[3]]^{2}, \cdots \\ \begin{bmatrix} [0]]^{3}, [[1]]^{3}, [[2]]^{3}, [[3]]^{3}, \cdots \\ \vdots \end{bmatrix}
```

lists all of the computable functions. Moreover, this notation also allows us to more easily state The S_n^m -Theorem (A2 c • 3) as saying $[\![e]\!]^{m+n}(\vec{a}, \vec{x}) \triangleq [\![S_n^m(e, \vec{a})]\!](\vec{x})$ for all $\vec{a} \in \omega^m$.

It's important to realize that the list above is well defined in that $[x]^n$ is a (computable) partial function even if x doesn't code a COM({s, pd})-program with n inputs. So "most" of these functions will merely be \emptyset , since if e doesn't code a program, CompCode_n(e, \vec{x}, y) is always false. This means μy (CompCode_n(e, \vec{x}, y)) diverges, implying $[e](\vec{x}) \stackrel{\circ}{=} \text{Output}(\mu y \text{ (CompCode}_n(e, \vec{x}, y))) \uparrow$.

§ A2 d. Noncomputable sets

As can be easily seen by Normal Form (A2 $c \cdot 2$), there are only countably many computable functions. Hence most partial functions and relations over ω must be noncomputable. The easiest example is the so-called halting problem.

^{ix}or a masochist

^xMany other texts will write φ_e for **[[e]**], but to make it easier to read the index (which is arguably more important than the input), we adopt the latter notation.

Let us begin with a slight modification.

Note that all primitive recursive functions can be enumerated in a similar way as with all computable functions. If we diagonalize through these functions, we end up with a computable function that is not primitive recursive. Doing the same for the computable *partial* functions leaves us with a function that *is* still computable. Let us work out the details to see where the usual argument goes astray.

– A2d•1. Theorem

There is a *universal* computable partial function, meaning a map $f: \omega^2 \to \omega$ where for every computable $g: \omega^n \to \omega$, there is some $e \in \omega$ where for all $\vec{x} \in \omega^n$, $g(\vec{x}) \stackrel{\circ}{=} f(e, \operatorname{code}(\vec{x}))$. In other words, $f(e, x) \stackrel{\circ}{=} \llbracket e \rrbracket^1(x)$.

Proof .:.

Just take f to be the map taking $\langle e, x \rangle$ to

 $Output(\mu y CompCode_1(e, x, y)).$

This f is clearly able to compute all the partial functions $g : \omega \to \omega$: just note that $g = \llbracket e \rrbracket^1$ for some e and thus $\llbracket e \rrbracket^1(x) = f(e, x)$ for all $x \in \omega$ by Normal Form (A2 c • 2).

Now for each $g: \omega^n \to \omega$, consider the map h being $x \mapsto g(x_0, x_1, \dots, x_n)$. Note that this map is computable as we are (in a primitive recursive way) decoding x (and if x isn't a code, we have taken $x_i = 0$). Thus $h: \omega \to \omega$ satisfies $h(\operatorname{code}(\vec{x})) \stackrel{\circ}{=} g(\vec{x})$ for every $\vec{x} \in \omega^n$. Therefore the above gives us some e where f(e, x) = h(x) for all $x \in \omega$, and in particular for all $x = \operatorname{code}(\vec{x})$ with $\vec{x} \in \omega^n$.

If we try to diagonalize as we can with the primitive recursive functions, taking $f'(e, x) \triangleq \llbracket e \rrbracket^1(x) + 1$, we don't necessarily have a problem. The natural argument would be that there must be some e' with $f' = \llbracket e' \rrbracket^1$ and thus we should have f'(e', x) = f'(e', x) + 1. But in reality, we only get $f'(e', x) \triangleq f'(e', x) + 1$ and thus we must have $f'(e', x) \uparrow$. The problem is that this argument only works on total functions. Hence this argument actually shows that there is no way to computably enumerate the computable *total* functions.

This suggests a fundamental issue related to computable (partial) functions is their domains. And this gives us our first example of a non-computable relation.

- A2 d · 2. Definition Define Halt to be the set $\{\langle e, x \rangle \in \omega^2 : x \in \text{dom} [\![e]\!]^1\}$.

- A2 d • 3. Corollary Halt is not computable.

Proof .:.

Suppose Halt were computable. Therefore the (total) function

$$f(x) = \begin{cases} \llbracket e \rrbracket^1(e) + 1 & \text{if } \langle e, e \rangle \in \text{Halt} \\ 0 & \text{otherwise} \end{cases}$$

is computable and thus can be represented as $f = \llbracket e_f \rrbracket^1$ for some $e_f \in \omega$. But since f is total, $\langle e_f, e_f \rangle \in$ Halt so that $f(e_f) = \llbracket e_f \rrbracket^1(e_f) + 1 = f(e_f) + 1$, a contradiction.

We actually have a much more general theorem of this sort. The idea is that we can work with codes of programs syntactically to get partial functions, but to do the reverse—i.e. to work with functions to understand their programs—is always noncomputable. Put in another way, there's no computable way to answer the question "what the hell does this program do?", meaning for any given collection of functions F, whether $[e]^n \in F$ or not for $e, n < \omega$ isn't computable.

— A2d•4. Theorem (Rice's Theorem) —

For $n < \omega$, let $F \subseteq \{f \subseteq \omega^n \times \omega : f \text{ is a function}\}$ be an arbitrary set of partial functions over ω , and let $E = \{e \in \omega : [\![e]\!]^n \in F\}$ be the set of codes of these. Therefore, E is either \emptyset, ω , or noncomputable.

In particular, as we noted before, the codes that give total functions is noncomputable: there is no way to tell beforehand whether any given program's code will always return an output.

To prove this theorem, it will be useful to know the Second Recursion Theorem^{xi} of Kleene, which has a similar flavor to The S_n^m -Theorem (A2 c • 3) but is a more remarkable fixed-point theorem. Unfortunately, the proof of the theorem isn't as impressive as the result itself, and feels more like symbol pushing.

- A2 d • 5. Lemma (The Second Recursion Theorem) For every computable $f : \omega^{n+1} \rightharpoonup \omega$, there is an *e* such that for all \vec{x} , $[\![e]\!]^n(\vec{x}) \stackrel{\circ}{=} f(e, \vec{x})$.

Proof .:.

Firstly, define $g(e, \vec{x}) \stackrel{\circ}{=} f(S_n^1(e, e), \vec{x})$ so that g is clearly computable and thus there is some e_g where $\llbracket e_g \rrbracket^{n+1} = g$. But then setting $e = S_n^1(e_g, e_g)$ yields by The S_n^m -Theorem (A2 c • 3) that $\llbracket e \rrbracket^n(\vec{x}) \stackrel{\circ}{=} \llbracket e_g \rrbracket^{n+1}(e_g, \vec{x}) \stackrel{\circ}{=} f(S_n^1(e_g, e_g), \vec{x}) \stackrel{\circ}{=} f(e, \vec{x})$

This fixed-point theorem has an enormous number of counter-intuitive consequences. For example, there is an *e* where $[\![e]\!]^1(x) = x + e$, and another where $[\![e]\!]^1$ is just the constant *e* function. Such weird statements show that the codes of programs can have as little or as much correspondence with their defined function as we'd like. In particular, we can never determine in general which codes yield functions with certain properties—formally stated via Rice's Theorem (A2 d • 4)—unless it trivially applies to all or to no partial functions over ω . For now, we repress the *n* superscript as it doesn't really add any understanding.

- Proof of Rice's Theorem $(A2 d \cdot 4) \dots$

Suppose $E \neq \emptyset$ and $E \neq \omega$ so that there exist $e_0 \notin E$ and $e_1 \in E$. Define a partial function f where

 $f(x, \vec{y}) \stackrel{\circ}{=} \begin{cases} \llbracket e_0 \rrbracket (\vec{y}) & \text{if } x \in E \\ \llbracket e_1 \rrbracket (\vec{y}) & \text{otherwise.} \end{cases}$

Note that for $x \in E$ fixed, $\vec{y} \mapsto f(x, \vec{y})$ is the map $\llbracket e_0 \rrbracket \notin F$. And for $x \notin E$ fixed, $\vec{y} \mapsto f(x, \vec{y})$ is the map $\llbracket e_0 \rrbracket \notin F$. If membership in E were computable, then clearly f would be computable so that by The Second Recursion Theorem (A2 d • 5), there's an $e \in \omega$ where for all $\vec{y}, f(e, \vec{y}) \triangleq \llbracket e \rrbracket (\vec{y})$. But this is impossible: $e \in E$ implies $\llbracket e \rrbracket = \llbracket e_0 \rrbracket \notin F$, contradicting that $e \in E$. Similarly, $e \notin E$ implies $\llbracket e \rrbracket = \llbracket e_1 \rrbracket \in F$, contradicting that $e \notin E$. Hence membership in E is not computable.

xiThe First Recursion Theorem is a kind of fixed-point theorem about so-called *effective* operations.

Section A3. The Arithmetical Hierarchy

What theorems like Rice's Theorem $(A2 d \cdot 4)$ tell us is that we should develop a theory of complexity for noncomputable sets. Our first step into this topic will be with examining complexity in terms of *form*. More specifically, we begin with the primitive recursive relations, and then continually consider relations defined by existential quantification or universal quantification over lower levels. The result is a hierarchy of complexity not unlike the Lévy hierarchy for formulas.



A3 • 1. Figure: The arithmetical hierarchy

These sets share a close connection with the Lévy hierarchy on $\mathbf{N} = \langle \omega, 0, 1, +, \cdot \rangle$ in that a set $A \subseteq \omega$ is Σ_n^0 iff A is defined by a Σ_n -formula. Here a Δ_0 -formula is a formula with only bounded quantifiers where the "bound" is using the defined natural order: x < y iff $\exists z \ (x + z = y)$ so a quantifier is bounded iff it is of the form $\exists x < y$ or $\forall x < y$, each of which is shorthand for a larger FOL($\{0, 1, +, \cdot\}$)-formula.

Note that elsewhere in this document, we will refer to the sets in Figure A3 • 1 as $\Sigma_n^{0,\omega}$, $\Delta_n^{0,\omega}$, or $\Pi_n^{0,\omega}$ to distinguish them from analogous notions for subsets of ${}^{\omega}\omega$ rather than just ω . But since we're dealing just with ω here, we will ignore the " ω " in the superscript. One might suggest also doing away with the superscript of '0', but then we might get confused with the lévy hierarchy, which is intimately related. This 0 can also be generalized to larger numbers, though we will not investigate that in this appendix.

§ A3 a. First levels: Δ_0^0 , Δ_1^0 , and Σ_1^0

The first kind of not-necessarily-computable sets we will look at are the Σ_1^0 sets.^{xii} Recall that the halting problem is not computable just because we couldn't tell whether a given input was in the domain of a partial function or not: it's not computable in general whether $x \in \text{dom}(\llbracket e \rrbracket)$. But this relation is very close to being computable in that $x \in \text{dom}(\llbracket e \rrbracket)$ iff we will *eventually* reach an output when computing. More precisely, $x \in \text{dom}(\llbracket e \rrbracket)$ iff $\exists y$ (CompCode(e, x, y)). If we know beforehand that there is such a y, we can use minimalization to say so; but if there is no such y, we won't be able know beforehand in a computable manner (in general).

If we examine the form of the above equivalence, this has the form of one existential quantifier in front of a primitive recursive relation. We can *semi-compute* this in that we can yield a "yes" answer if there is one, but we don't necessarily have an algorithm to determine if there is not. This motivates the definition of Σ_1^0 sets, sometimes called *semicomputable* or *computably enumerable*. The 0 on the top references that we are quantifying only over ω . Σ_1^1 , for example, allows quantification over ${}^{\omega}\omega$.

A3a • 1. Definition Let $R \subseteq \omega^n$, $n < \omega$, be a relation. • R is $\Pi_0^0 = \Sigma_0^0 = \Delta_0^0$ iff it is primitive recursive. • R is Σ_1^0 iff for some Δ_0^0 -relation Q, $\vec{x} \in R \leftrightarrow \exists y \ (\langle \vec{x}, y \rangle \in Q)$. • R is Δ_1^0 iff R and $\omega^n \setminus R$ are Σ_1^0 .

^{xii}Note that Σ_1^0 sets are different from the Σ_1^0 sets, which are the open sets of real numbers.

 \neg

Note that just by adding dummy variables, any Δ_0^0 relation R is also Σ_1^0 : $Q = R \times \{0\}$ witnesses this. Hence $\Delta_0^0 \subseteq \Sigma_1^0$, and as a result of Corollary A2 d • 3, this containment is strict: $\Delta_0^0 \subsetneq \Sigma_1^0$. We will eventually see that the Δ_1^0 sets encompass all of the computable relations, meaning that all of the containments are strict: $\Delta_0^0 \subsetneq \Delta_1^0 \subsetneq \Sigma_1^0$.

The definition of Σ_1^0 sets is equivalent to a great number of other statements. In particular, we can see the connection with Halt in that a relation is Σ_1^0 iff it is the domain of a computable partial function (also called *semirecursive* or *semicomputable*).

A3a•2. Theorem

Let $R \subseteq \omega^n$. Therefore the following are equivalent:

- 1. *R* is Σ_{1}^{0} ;
- 2. the ostensibly weaker statement that $\vec{x} \in R$ iff for some *computable* $Q, \exists y \ (\langle \vec{x}, y \rangle \in Q);$
- 3. R = dom(f) for some computable partial function f over ω ;

Proof .:.

Clearly (1) implies (2) as all primitive recursive relations are computable. Assume (2), working towards (3). Consider the partial function f defined by $f(\vec{x}) \stackrel{\circ}{=} \mu y \chi_Q(\vec{x}, y)$. Note that f is computable since χ_Q is. But $\vec{x} \in \text{dom } f$ iff $\exists y (\langle \vec{x}, y \rangle \in Q)$ iff $\vec{x} \in R$. Hence R = dom f.

So assume (3), working towards (1), with R = dom(f), meaning that $\vec{x} \in R$ iff $f(\vec{x}) \downarrow$. By Normal Form (A2 c • 2), there is some $e \in \omega$ where $f(\vec{x}) \stackrel{\circ}{=} \text{Output}(\mu y \text{ CompCode}_n(e, \vec{x}, y))$. Hence $f(\vec{x}) \downarrow$ iff the function on the right converges. Since Output is total (because it's primitive recursive), it follows that $f(\vec{x}) \downarrow$ iff $\mu y \text{ CompCode}_n(e, \vec{x}, y) \downarrow$. And since CompCode_n is primitive recursive and therefore total, this happens iff there is such a y:

 $\vec{x} \in R$ iff $f(\vec{x}) \downarrow$ iff $\exists y \operatorname{CompCode}_n(e, \vec{x}, y)$,

As with primitive recursive functions and computable ones, Σ_1^0 (being the set of Σ_1^0 sets) is closed under conjunctions, disjunctions, and so on. The main difference to notice about this is that Σ_1^0 is *not* closed under negations, nor under universal quantification.

- A3 a • 3. Result

which is then Σ_1^0 .

 Σ_1^0 is closed under conjunctions, disjunctions, bounded quantification, and existential quantification.

Proof .:.

Let $R(\vec{x})$ be equivalent to $\exists y \ P_R(\vec{x}, y)$, and similarly let $Q(\vec{x}')$ be equivalent to $\exists z \ P_Q(\vec{x}', z)$, where P_R and P_Q are primitive recursive.

- $R(\vec{x}) \wedge Q(\vec{x}')$ is equivalent to $\exists c(\operatorname{seq}(c) \wedge \operatorname{lh}(c) = 2 \wedge P_R(\vec{x}, c_0) \wedge P_Q(\vec{x}', c_1))$, which is clearly Σ_1^0 as seq and lh are both primitive recursive, and the collection of primitive recursive relations is closed under conjunction. Disjunction follows more easily: $R(\vec{x}) \vee Q(\vec{x}')$ is equivalent to $\exists y (P_R(\vec{x}, y) \vee P_Q(\vec{x}', y))$.
- $\exists x \leq y \ R(\vec{x}, x)$ iff $\exists c \ (seq(c) \land lh(c) = 2 \land c_0 \leq y \land P_R(\vec{x}, c_0, c_1))$, which is Σ_1^0 . Bounded universal quantification is a little more complicated: $\forall x \leq y \ \exists z \ P_R(\vec{x}, x, z)$ is equivalent to saying that there is a coding of a function from y to these z:

$$\exists c \; (\operatorname{seq}(c) \land \operatorname{lh}(c) > y \land \forall n \leq y \; P_R(\vec{x}, n, c_n)).$$

•
$$\exists z \exists y \ P_R(\vec{x}, z, y)$$
 is equivalent to the existence of a pair $\langle y, x \rangle$ with $P_R(\vec{x}, z, y)$:
 $\exists z \ R(\vec{x}, z) \leftrightarrow \exists c(\operatorname{seq}(c) \land \operatorname{lh}(c) = 2 \land P_R(\vec{x}, c_0, c_1)) \dashv$

The main reason why Σ_1^0 sets are not closed under negations (and hence not under universal quantification) is the following very nice result restating computability in terms of Σ_1^0 sets. The idea behind this is that generally a Σ_1^0 -set only gives a "yes" answer to membership in a finite amount of steps, but might go on searching forever for a "no".

 \neg

But if both the set and its complement have this property, then we get a bound on how far we need to search for a "no" answer too. In particular, we just simultaneously search for a "yes" answer to membership in the set and to membership in its complement. Either way, we reach the result in a finite number of steps.

- A3 a • 4. Theorem - A relation $R \subseteq \omega^n$ is computable iff R and $\omega^n \setminus R$ are Σ_1^0 , in other words R and $\neg R$ are Σ_1^0 . Hence R is computable iff R is Δ_1^0 .

Proof .:.

If both R and $\neg R$ were Σ_1^0 , then for some primitive recursive P_R and $P_{\neg R}$, $R(\vec{x}) \leftrightarrow \exists y P_R(\vec{x}, y)$ and $\neg R(\vec{x}) \leftrightarrow \exists y P_{\neg R}(\vec{x}, y)$. But then $R(\vec{x}) \leftrightarrow \forall y P_{\neg R}(\vec{x}, y)$. So to say whether $R(\vec{x})$ holds or not, we just need to search for an example of P_R or one to $P_{\neg R}$: consider the computable function $f(\vec{x}) = \mu y (P_R(\vec{x}, y) \vee P_{\neg R}(\vec{x}, y))$. Note that f is total, because $R(\vec{x})$ is either true or false, implying there is always a y where $P_R(\vec{x}, y) \vee P_{\neg R}(\vec{x}, y)$. Now consider

$$\chi(\vec{x}) = \begin{cases} 1 & \text{if } P_R(\vec{x}, f(\vec{x})) \\ 0 & \text{if } P_{\neg R}(\vec{x}, f(\vec{x})). \end{cases}$$

It should be clear that then $\chi_R = \chi$, which is clearly computable.

Another characterization of computability in terms of Σ_1^0 is the following result, acting similar to the converse of (1) implying (3) in Theorem A3 a • 2.

– A3a•5. Result –

A partial function f over ω is computable iff f as a relation is Σ_1^0 .

Proof .:.

If f is computable, then f as a relation is Σ_1^0 as $\langle \vec{x}, y \rangle \in f$ iff $\vec{x} \in \text{dom}(f)$, which is Σ_1^0 by Theorem A3 a • 2. So suppose f as a relation is Σ_1^0 . Therefore there is some primitive recursive P where $\langle \vec{x}, y \rangle \in f$ iff $\exists z (P(\vec{x}, y, z))$. So take $f(\vec{x}) \triangleq (\mu y P(\vec{x}, y_0, y_1))_0$. As P is primitive recursive, f is computable.

In some sense, the above result should seem counter intuitive, since $\chi_{=}(x, y)$ is primitive recursive, and so by composition, $\langle x, y \rangle \mapsto \chi_{=}(f(x), y)$ should also be computable if f is. The issue with this argument is that $\chi_{f}(x, y) = \chi_{=}(f(x), y)$ when $f(x) \downarrow$, but there are issues when $f(x) \uparrow$: $\chi_{f}(x, y) = 0$ while $\chi_{=}(f(x), y) \uparrow$. This is why f as a relation may only be "semi-computable" or Σ_{1}^{0} . But in the case of total functions, this idea works.

— A3 a • 6. Corollary

A *total* function f over ω is computable iff f as a relation is computable.

Now if Σ_1^0 sets can be seen as domains of computable functions, they can also be seen as *images* of computable functions. This motivates the names "recursively enumerable" and "computable enumerable" sometimes used in the literature.

— A3 a • 7. Definition -

A set $A \subseteq \omega$ is *computably enumerable* iff $A = \emptyset$ or some total, computable $f : \omega \to \omega$ has $A = \operatorname{im} f$.

In principle, we could expand the definition of A being computably enumerable to allow for $A \subseteq \omega^n$ for $n < \omega$ or for $f : \omega^n \to \omega$, but this doesn't grant any additional generality by the next theorem. Indeed, since we can code tuples in a Δ_0^0 way, we can easily translate between the two notions.

— A3 a • 8. Result

A set $A \subseteq \omega$ is computably enumerable iff it is Σ_1^0 .

Proof .:.

If $A = \emptyset$, this is obvious. So assume $A \neq \emptyset$. If A is Σ_1^0 , then for some Δ_0^0 -relation $R, x \in A$ iff $\exists y \ R(x, y)$. Let

 $a \in A$ be fixed and consider the partial function

$$f(u) = \begin{cases} a & \text{if } \neg R(u_0, u_1) \\ u_0 & \text{if } R(u_0, u_1). \end{cases}$$

Clearly f is computable and has $f''\omega = A$, since if $x \in A$ as witnessed by R(x, y), then $f(2^{x+1} \cdot 3^{y+1}) = x$ so that $A \subseteq f''\omega$. And by definition, $f(u) \in A$ for each $u \in \omega$. Hence $f''\omega = A$, so A is computably enumerable.

So assume A is computably enumerable with $A = f''\omega$ for $f : \omega \to \omega$ computable. Therefore $y \in A$ iff $\exists x (f(x) = y)$. As f is computable, by Result A3 a • 5, f(x) = y is Σ_1^0 , and since Σ_1^0 is closed under existential quantification by Result A3 a • 3, it follows that $\exists x (f(x) = y)$ is Σ_1^0 and thus $y \in A$ is Σ_1^0 . \dashv

We may actually strengthen Definition A3 a • 7 to $A \subseteq \omega$ requiring a total, *injective*, computable $f : \omega \to \omega$ with $A = f''\omega$. We cannot, however, require that f is *increasing*, as this would require A to be computable. The basic idea why is that f being increasing allows us to place a bound on where we need to search and thus give a "no" answer in a finite amount of steps.

A **3** a • 9. Result
A set
$$A \subseteq \omega$$
 is computably enumerable iff $|A| < \aleph_0$ or $A = \text{im } f$ for some injective, total, computable $f : \omega \to \omega$.

Proof .:.

One direction is immediate, so assume A is computably enumerable as witnessed by the (non-injective) $g: \omega \to \omega$. The idea is to define f is just to skip the repeated values: define f by the (shorthand for the) COM({s, pd})-program

" $p_0(x) = y_0$ for the least $y = \text{code}(y_0, y_1)$ where $y_0 = g(y_1) \land y_1 = \mu z < x \ (\forall x' < x \ p_0(x') \neq g(z))$ ". Since g is total, y = g(x) is computable by Corollary A3 a • 6 which means we can expand the above into an actual COM({s, pd})-program to yield a computable $\hat{p}_0 = f$. In essence, this is saying that we choose a y_0 and y_1 such that

• $g(y_1)$ differs from the previous values of p_0 ; and

• $y_1 \le x$ is the least such that the above holds (recall $\mu z < x(\dots) = x$ if there are no z < x such that \dots). Since A is infinite, there will always be such a y, meaning that $f(x) = \hat{p}_0(x)$ is always defined, and thus f is total. Moreover f is injective by the defining program, and it should be clear that if g(x) = y, then f(x') = y for some $x' \le x$. In particular, $A = f''\omega$.

The result below uses "increasing" in the sense of *strictly* increasing $(x < y \rightarrow f(x) < f(y))$ rather than the weaker notion $(x < y \rightarrow f(x) \le f(y))$ that would encompass constant functions. The result, along with Result A3 a • 9, shows that while we may delete repeated values in a computable way, we cannot reorder them.

— A3a•10. Result

A set $A \subseteq \omega$ is computable iff $|A| < \aleph_0$ or $A = \operatorname{im} f$ for some increasing, total, computable $f : \omega \to \omega$.

Proof .:.

If $A \subseteq \omega$ is computable, then clearly either $|A| < \aleph_0$, or $|A| = \aleph_0$ in which case, setting $a = \min A$, we can define by recursion f(0) = a, $f(n + 1) = \mu x$ ($x \in A \land x > f(n)$). Since membership in A is computable, this is computable, and clearly $f''\omega = A$.

If $A = f''\omega$ for some increasing, total, computable $f: \omega \to \omega, n \le f(n)$ for every $n \in \omega$ and thus $x \in A$ iff there is some $n \le x$ with f(n) = x. So setting $\chi(x) = \exists n \le x$ (f(n) = x) yields that χ is computable (by Corollary A3 a • 6 since f is total) and equal to χ_A .

These few results allow us to conclude the following useful results about pointwise images and preimages. Note that the above result unfortuantely precludes $f''X \in \Delta_1^0$ even if f and X are both computable: take $X = \omega$ and f such that $f''\omega =$ Halt as an example.
- A3 a • 11. Result — Let $f : \omega \to \omega$ be a (total) computable function, and let $X \subseteq \omega$. Therefore,

- If X is Σ_1^0 , then f''X is Σ_1^0 ;
- If X is Σ_1^0 , then f^{-1} "X is Σ_1^0 ; and
- If X is Δ_1^0 , then f^{-1} "X is Δ_1^0 .

Proof .:.

Given that f is computable as a function, it is Σ_1^0 as a subset of ω^2 by Result A3 a • 5. Therefore, we have the following by the closure conditions of Result A3 a • 3 and Corollary A2 a • 4.

- $y \in f^{"}X$ iff $\exists x \ (x \in X \land y = f(x))$. So $f^{"}X$ is Σ_{1}^{0} if X is.
- $x \in f^{-1}$ "X iff $\exists y (y \in X \land f(x) = y)$, which is again Σ_1^0 if X is.
- $x \in f^{-1}$ "X iff $f(x) \in X$ iff $\chi_X \circ f(x) = 1$. Therefore, if $X \in \Delta_1^0$, then $\chi_X \circ f$ is computable. Since it's also the characteristic function for f^{-1} "X (we need totality for this), it follows that f^{-1} "X is Δ_1^0 . \dashv

§ A3 b. The rest of the arithmetical hierarchy

Thus far, we have categorized subsets of ω (and subsets of $\omega^{<\omega}$ through a primitive recursive coding and decoding) into three distinct classes: Δ_0^0 , Δ_1^0 , and Σ_1^0 with $\Delta_0^0 \subseteq \Delta_1^0 \subseteq \Sigma_1^0$ and all three containments as strict.^{xiii} We can generalize this hierarchy in a way analogous to the Lévy hierarchy that motivates the notation used in the previous section.

- A3 b • 1. Definition Let $R \subseteq \omega^n$ be a relation over ω .

- R is Σ_0^0 iff it is primitive recursive.
- R is Σ_{m+1}^{0} iff for some \prod_{m}^{0} -relation $Q, \vec{x} \in R$ iff $\exists y \ (\langle \vec{x}, y \rangle \in Q)$.
- R is Π^0_m iff $\omega^n \setminus R$ is Σ^0_m .

With these definitions, we set Σ_m^0 to be the set of all Σ_m^0 -relations, and $\Delta_m^0 = \Sigma_m^0 \cap \Pi_m^0$.

With such a technical definition, we should do some justification for it, in addition to alternative characterizations. To help gain some intuition, $\{\langle e, x \rangle \in \omega^2 : [\![e]\!](x) \uparrow\}$ is Π_1^0 , being the negation of the Σ_1^0 -relation Halt. We can also arrive at standard ways of representing the relations in the higher classes. For example, R is Π_3^0 iff

 $\vec{x} \in R \quad \text{iff} \quad \neg(\vec{x} \in P) \quad \text{for } P \text{ some } \Sigma_3^0 \text{-relation} \\ \text{iff} \quad \neg(\exists y \ (\langle \vec{x}, y \rangle \in Q)) \quad \text{for } Q \text{ some } \Pi_2^0 \text{-relation} \\ \text{iff} \quad \neg \exists y \ \neg(\langle \vec{x}, y \rangle \in S) \quad \text{for } S \text{ some } \Sigma_2^0 \text{-relation} \\ \text{iff} \quad \neg \exists y \ \neg \exists z \ (\langle \vec{x}, y, z \rangle \in T) \quad \text{for } T \text{ some } \Pi_1^0 \text{-relation} \\ \text{iff} \quad \neg \exists y \ \neg \exists z \ \neg(\langle \vec{x}, y, z \rangle \in U) \quad \text{for } U \text{ some } \Sigma_1^0 \text{-relation} \\ \text{iff} \quad \neg \exists y \ \neg \exists z \ \neg \exists t \ (\langle \vec{x}, y, z, t \rangle \in V) \quad \text{for } V \text{ some } \Pi_0^0 \text{-relation} \\ \text{iff} \quad \neg \exists y \ \neg \exists z \ \neg \exists t \ \neg(\langle \vec{x}, y, z, t \rangle \in W) \quad \text{for } W \text{ some } \Sigma_0^0 \text{-relation} \\ \text{iff} \quad \forall y \ \exists z \ \forall t \ (\langle \vec{x}, y, z, t \rangle \in W) \quad \text{for } W \text{ some primitive recursive relation.} \end{cases}$

The idea is that a Δ_0^0 relation (analogous to the bounded quantification of Δ_0 -formulas in the Lévy hierarchy) is the primitive recursive relations; and for the larger classes, we alternate existential quantifiers and universal quantifiers. The following are some basic results.

- Σ_1^0 -relations are the form $\exists x \ R$ for R primitive recursive.
- Π_1^0 -relations are of the form $\forall x \ R$ for R primitive recursive.
- Σ_2^0 -relations are of the form $\exists x \forall y R$ for R as above;
- Σ_m^0 -relations in general alternate the quantifiers starting with \exists : they are of the form $\exists x_1 \ \forall x_2 \ \exists x_3 \ \cdots \ Q x_m \ R$ where Q is either \exists or \forall to ensure the alternating pattern.

xiiiIn particular, the coded Ackermann function code"Ack $\in \Delta_1^0 \setminus \Delta_0^0$, and the coded halting set code"Halt $\in \Sigma_1^0 \setminus \Delta_1^0$.

THE ARITHMETICAL HIERARCHY

- The same holds for Π_m^0 -relations but starting with \forall .
- R is Σ_m^0 iff $\neg R$ is Π_m^0 .

We also get the following theorem which will be often used without referring back to it, allowing us to consider just sets in $\mathcal{P}(\omega)$ rather than in $\bigcup_{n < \omega} \mathcal{P}(\omega^n)$. This also easily holds for the other classes of \prod_n^0 and Δ_n^0 too.

```
- A3 b • 2. Theorem
Let R \subseteq \omega^m with m < \omega. Therefore R is \Sigma_n^0 iff code"R = \{ \text{code}(\vec{x}) : \vec{x} \in R \} is \Sigma_n^0.
```

Proof .:.

Proceed by induction on *n*. For n = 0, this is clear, because the characteristic function χ_R is primitive recursive iff (by our primitive recursive coding) $\chi(x) = seq(x) \wedge lh(x) = m \wedge \chi_R(x_0, \dots, x_{m-1})$ is primitive recursive. Note that χ is the characteristic function of code" *R*.

For n + 1, suppose R is \sum_{n+1}^{0} so that R satisfies $\vec{x} \in R$ iff $\exists y (\langle \vec{x}, y \rangle \in P')$ for some $P' \in \prod_{n=1}^{0}^{0}$. Equivalently, R satisfies $\vec{x} \in R$ iff $\exists y (\langle \vec{x}, y \rangle \notin P)$ for some $P \in \sum_{n=1}^{0}^{0}$. By the inductive hypothesis (really a slight modification of it to allow for some inputs to not be squashed together in a single, coded input), this is equivalent to $\vec{x} \in R$ iff $\exists y (\langle \operatorname{code}(\vec{x}), y \rangle \notin P_0)$ where P_0 is $\sum_{n=1}^{0}^{0}$. So this gives a $\sum_{n=1}^{0}^{0}$ -definition for code" R. The reverse direction follows similarly. Hence by induction, the result holds.

Now recall Figure A3 • 1, presenting the containments. We still, of course, need to prove these facts, but by introducing irrelevant variables with either quantifier, we know thus far that $\Delta_n^0 \subseteq \Sigma_n^0 \subseteq \Delta_{n+1}^0$ for each $n \in \omega$ (and similarly $\Delta_n^0 \subseteq \prod_{n=1}^{\infty} \Delta_{n+1}^0$). Given the above definitions, placing relations in the hierarchy is relatively easy just based on

A3 b • 3. Figure: The arithmetical hierarchy

their defining form, just as with the Lévy hierarchy. The issue, however, is that showing this placement is optimal—showing that $R \in \Sigma_n^0$ or that $R \in \Pi_n^0$ but not Δ_n^0 —is far more difficult.

One might also notice that Definition A3 b • 1 offers different definitions for the classes of Δ_0^0 , Δ_1^0 , and Σ_1^0 than Definition A3 a • 1. That said, we can still show that the two are equivalent.

- A3b•4. Result $\Sigma_0^0 = \Delta_0^0 = \Pi_0^0$ is the set of primitive recursive relations. Hence Σ_1^0 is the same as in Definition A3 a•1, and thus Δ_1^0 is the same as in Theorem A3 a•4.

Proof .:.

Firstly, note that $\Sigma_0^0 = \Delta_0^0 = \Pi_0^0$ since every Σ_0^0 (i.e. primitive recursive) relation is the negation of a primitive recursive relation, $R \in \Sigma_0^0$ has $\neg R \in \Sigma_0^0$ and thus $\neg \neg R = R \in \Pi_0^0$ implying $\Sigma_0^0 = \Delta_0^0 = \Pi_0^0$. This implies that the definitions of Σ_1^0 is equivalent as before. So that by Theorem A3 a • 4, *R* is computable iff $R, \neg R \in \Sigma_1^0$, meaning $R \in \Sigma_1^0$ and $R = \neg \neg R \in \Pi_1^0$, which is equivalent to $R \in \Delta_1^0$. So the computable relations are again the Δ_1^0 -relations and so Δ_0^0, Δ_1^0 , and Σ_1^0 are the same collections as before.

We now begin work towards identifying Σ_n^0 with sets defined by Σ_n -formulas over **N**, so long as n > 0 (Δ_0 -definable sets are Δ_0^0 , but the reverse isn't necessarily true). Firstly, recall the Lévy hierarchy of formulas.

— A3 b • 5. Definition —

Let **N** be the model $\langle \omega, 0, 1, +, \cdot \rangle$. Write "x < y" as shorthand for the formula " $\exists z \ (x + z = y \land z \neq 0)$ ". A quantifier is called *bounded* iff it is of the form " $\exists x < t(y) \ (\cdots)$ " for some FOL($\{0, 1, +, \}$)-term t. Let φ be a

Appx §A3 b

 $FOL(\{0, 1, +, \cdot\})$ -formula.

- φ is $\Delta_0^{\mathsf{N}} = \Sigma_0^{\mathsf{N}} = \Pi_0^{\mathsf{N}}$ iff all quantifiers in φ are bounded.
- φ is $\sum_{n=1}^{N}$ iff φ is of the form $\exists x \ \psi$ where ψ is $\prod_{n=1}^{N}$.
- φ is Π_n^{N} iff φ is $\neg \psi$ for some $\Sigma_n^{\mathsf{N}} \psi$.

Write $\varphi(\mathbf{N})$ for the set $\{\vec{x} \in \omega^{<\omega} : \mathbf{N} \models ``\varphi(\vec{x})"\}$. We also say φ is $\Sigma_n^{\mathbf{N}}$ (or any of the other classes) iff $\varphi(\mathbf{N}) = \psi(\mathbf{N})$ for some $\Sigma_n^{\mathbf{N}}$ -formula ψ as above. And finally, φ is $\Delta_n^{\mathbf{N}}$ iff φ is $\Sigma_n^{\mathbf{N}}$ and $\Pi_n^{\mathbf{N}}$.

By again adding dummy variables, we get the same inequalities as in Figure A3 b • 3.

A3 b • 6. Figure: The elementary relations of N

It is more intuitive, however, that all the containments are strict for this diagram compared to the arithmetical hierarchy. Note that we have the expected closure properties. As a small remark, note that "x < y", despite having the form $\exists z \ (x + z = y)$ can be easily recast as a Δ_0^N formula, since obviously x < y iff $\exists z < y \ (x = z)$.

— A3b•7. Lemma ·

For each $n < \omega$, the Σ_n^N , Π_n^N , and Δ_n^N classes are closed under bounded quantification, conjunction, and disjunction. Moreover,

- Σ_n^{N} is closed under existential quantification for n > 0;
- Π_n^{N} is closed under universal quantification for n > 0; and
- Δ_n^{N} is closed under negation.

Proof .:.

Proceed by induction on *n*. For n = 0, the result clearly holds: $\Sigma_0^N = \Delta_0^N = \Pi_0^N$ consists of all formulas that have only bounded quantifiers and thus is closed under bounded quantification, conjunction, and disjunction.

For n + 1, formulas that are \sum_{n+1}^{N} have the form " $\exists x \varphi$ " where φ is $\prod_{n=1}^{N}$. So let φ and ψ be $\prod_{n=1}^{N}$.

- For disjunction, " $\exists x \ \varphi \lor \exists y \ \psi$ " is equivalent to " $\exists x \ (\varphi \lor \psi)$ ". Since, inductively, " $\varphi \lor \psi$ " is $\prod_{n=1}^{N}$, it follows that $\exists x \ \varphi \lor \exists y \ \psi$ is $\sum_{n=1}^{N}$ and so $\sum_{n=1}^{N}$ is closed under disjunction.
- For conjunction, " $\exists x \ \varphi \land \exists y \ \psi$ " is the same over **N** as " $\exists c \ \exists x < c \ \exists y < c \ (\varphi \land \psi)$ ". As \prod_n^N is closed under conjunction and bounded quantification, it follows that this is \sum_{n+1}^N and thus \sum_{n+1}^N is closed under conjunction.
- For existential quantification, " $\exists x \exists y \varphi$ " is equivalent to " $\exists c \exists x < c \exists y < c \varphi$ ", and inductively, since $\prod_{n=1}^{N} |v_{n+1}|$ is closed under bounded quantification, this is $\sum_{n=1}^{N}$. Hence $\sum_{n=1}^{N} |v_{n+1}|$ is closed under existential quantification.
- For bounded existential quantification, " $\exists x < y \exists z \varphi$ " is just shorthand for " $\exists x (x < y \land \exists z \varphi)$ ". Since "x < y" is $\Delta_0^{\mathsf{N}} \subseteq \Sigma_{n+1}^{\mathsf{N}}$, and $\Sigma_{n+1}^{\mathsf{N}}$ is closed under conjunction and existential quantification, it follows that this is $\Sigma_{n+1}^{\mathsf{N}}$, and thus $\Sigma_{n+1}^{\mathsf{N}}$ is closed under bounded existential quantification.

To proceed further, we should think back to coding. Thanks to previous work, we have a Δ_0^N -definable function that is able to code finite sequences: Gödel's β -function as per Lemma A1 a • 11. The following is clearly Δ_0^N , and defines the β -function:

 $\beta(x, y, i) = b \leftrightarrow \mathbf{N} \models ``\exists q < x \exists z < y \cdot (i+1) + 1 (x = q \cdot (y \cdot (i+1) + 1) + z \land b = z)`'.$

From this, decoding only adds two existential quantifiers: for a k-tuple \vec{x} , there is some a and some b where $\beta(a, b, i) = x_i$ for all i < k. Hence we can show that $\sum_{n=1}^{N} i$ is closed under bounded universal quantification.

• For bounded universal quantification, " $\forall x < y \exists z \varphi$ " is equivalent to the existence of a sequence $\langle c_n : n < y \rangle$ where x = n and $z = c_n$ witness φ for $n < 2 \cdot y$. Using Gödel's β -function, this says

 $\exists a \exists b \forall n < y \forall x \forall z (\beta(a, b, 2n) = n = x \land \beta(a, b, 2n + 1) = z \to \varphi).$

Since φ is Π_n^N , and $P \to Q$ is shorthand for $\neg P \lor Q$ (and Π_n^N is closed under those operations), we can calculate that this is Σ_{n+1}^N :



It then follows that $\sum_{n=1}^{N}$ is closed under bounded universal quantification.

Thus we have proven the result for each Σ_n^N , $n < \omega$, and the result for the Π_n^N s follows easily just by distributing negations through the Σ_n^N s: φ is Π_n^N iff $\neg \varphi$ is Σ_n^N . Similarly, the result for the Δ_n^N s follows easily from the result for Σ_n^N and Π_n^N .

More important for our purposes is that Δ_0^{N} -formulas, while not able to encompass all Δ_0^{O} -relations, always define Δ_0^{O} -relations.

- A3b·8. Lemma Let φ be a Δ_0^{N} -formula. Therefore $\varphi(\mathsf{N})$ is Δ_0^0 (or at least code" $\varphi(\mathsf{N})$ is Δ_0^0).

Proof .:.

Firstly, note that for any FOL($\{0, 1, +, \cdot\}$)-term t, the map $\vec{x} \mapsto t^{N}(\vec{x})$ is primitive recursive, as all such terms are given by the composition of $+, \cdot, 0, 1$, and projections $\vec{x} \mapsto x_i$.

Now we proceed by induction on φ . The only atomic formulas of FOL({0, 1, +, ·}) are those that state equality of terms: $\varphi(\vec{x})$ is " $t(\vec{x}) = t'(\vec{x})$ ", which then has $\varphi(\mathbf{N})$ coinciding with the characteristic function $\chi_{=}(t^{\mathbf{N}}(\vec{x}), t'^{\mathbf{N}}(\vec{x}))$, which is primitive recursive. The rest of the induction follows by Result A1 a • 6.

The converse to Lemma A3 b • 8 is unfortunately false. Nevertheless, we do get the induction holding for the higher levels of the arithmetical hierarchy, and this is partly due to our work in Lemma A3 b • 7 and the fact that we have a Δ_0^{N} -definable coding of finite sequences.

- A3b.9. Theorem Let
$$0 < n < \omega$$
 and $X \subseteq \omega^{<\omega}$. Therefore, X is Σ_n^0 iff X is Σ_n^N -definable, meaning $X = \varphi(\mathbf{N})$ for a Σ_n^N -formula φ .

Proof .:.

Proceed by induction on n. For n = 1, we have the following.

 (\rightarrow) Because we have a Δ_0^{N} -definable function capable of coding of finite sequences of natural numbers, it follows by closely analyzing Result A1 • 4, or more precisely the proof of it (along with Lemma A3 b • 7), that every primitive recursive function has a Σ_1^{N} -definition:

$$f(\vec{x}) = y$$
 iff $\mathbf{N} \models ``\exists z \varphi_f(\vec{x}, y, z)`',$

where φ_f is Δ_0^N . In particular, since X is Σ_1^0 , for some primitive recursive $R, \vec{x} \in X$ iff $\exists y (\langle \vec{x}, y \rangle \in R)$ which is then clearly equivalent to $N \models \exists c (\varphi_R(\vec{x}, c_0, c_1))$ which witnesses the result.

(\leftarrow) If $X = \varphi(\mathbf{N})$ for φ a $\Sigma_1^{\mathbf{N}}$ -formula, then for some $\Delta_0^{\mathbf{N}}$ -formula $\psi, \vec{x} \in X$ iff $\mathbf{N} \models \exists y \ \psi(\vec{x}, y)$. By Lemma A3 b • 8, $\psi(\mathbf{N}) = R$ is primitive recursive. Hence $\vec{x} \in X$ iff $\exists y \ (\langle \vec{x}, y \rangle \in R)$ shows that X is Σ_1^0 .

For n + 1, the inductive hypothesis holding at n implies by the definition of \prod_n^{N} and \prod_n^0 (just being negations applied to the corresponding Σ_n^{N} and Σ_n^0 sets) that the result holds there. But then it's obvious that the result holds for n + 1: $\vec{x} \in X$ iff $\exists y \ (\langle \vec{x}, y \rangle \in R)$ for $R \ \prod_n^0$ iff $\mathsf{N} \models ``\exists y \ \varphi_R(\vec{x}, y)$ '' for $\varphi_R \ a \ \prod_n^{\mathsf{N}}$ -formula.

So it's easy to conclude the result for all levels of the arithmetical hierarchy, excluding Δ_0^0 .

– A3b•10. Corollary -

- For $0 < n < \omega$ and $X \subseteq \omega^{<\omega}$,

 - X is Σ_n⁰ iff X is Σ_n^N-definable;
 X is Π_n⁰ iff X is Π_n^N-definable; and
 X is Δ_n⁰ iff X is Δ_n^N-definable.

Therefore every N-definable relation can be placed in the arithmetical hierarchy, which is a difficult condition to satisfy, given that of the 2^{\aleph_0} relations, only \aleph_0 -many are arithmetical: Σ_0^0 is countable, and each Σ_{n+1}^0 adds only (by induction) $|\Pi_n^0| = |\Sigma_n^0| = \aleph_0$ -many relations and is thus also countable, meaning $\bigcup_{n \in \omega} \Sigma_n^0$ (also written Σ_{ω}^0) is countable.

But this countability gives hope that we can talk about their elements coded as natural numbers themselves. This has already been done with Δ_1^0 where we consider codes of programs that are used to compute functions. This resulted in a universal computable function, meaning a computable function that is able to compute every other computable function: $f(e, x) = [e]^1(x)$. We have this same result for the higher classes in the arithmetical hierarchy.

This idea also allows us to show that all the containments of Figure A3 b • 3 and Figure A3 b • 6 are strict. The main idea behind the proof is that there is a *universal* element in each class that is able to compute the others. Such an element can be diagonalized in a way that produces a relation that only moderately increases complexity.

Section A4. Reducibility and Logic

Indeed, in some cases, a problem A can be reduced to another problem B in the sense that knowing the answer to B implies knowing the answer to A through some computable means. In general, we won't be able to compare complexity in a linear fashion, but we can still have meaningful statements that will generalize to our ideas about descriptive set theory.

The primary idea behind these ideas is a kind of relative computability in that we take the same notion of computability as before, but allow certain special functions and relations as given. The role these special functions play is like that of an oracle of myth: a human that serves as a mouthpiece for gods. So the outputs of these functions, although noncomputable by us, are simply given to us like how an oracle speaks the will of divinity, giving partial access to the noncomputable.

- A4.1. Definition Let F be a set of partial functions over ω . A partial function f over ω is F-computable iff it is computed by a $COM(F \cup \{s, pd\})$ -program.

Equivalently, f is F-computable iff it is F- μ -recursive in the sense that it is in the \subseteq -least collection of partial functions over ω containing the μ -recursive functions and F that is closed under the usual operations. One can show by the same idea as before that these two definitions of F-computability coincide: each program and computation uses only finitely many functions from F, and then doing the appropriate coding we can code a computation from these special functions. Given that we assume that these special functions are F- μ -recursive, we can decode such this in an F- μ -recursive way, and get that the resulting function is F- μ -recursive. Similarly, the converse holds analogous to Corollary A2 b • 7. Moreover, so long as F is countable—which it will be in each case we consider—we can consider a single coding of these functions and do away with the "finitely many used" idea.

§ A4 a. Reducibility in the arithmetical hierarchy

There are many types of reducibility that one can consider. The fundamental idea behind all of them is to embed one relation into another. The result is that if one is too complex, the other is forced to be by this reduction. Consider the following statement which motivates the idea.

- A4a·1. Theorem Consider the set $K = \{x \in \omega : [[x]]^1(x) \downarrow\}$. Therefore, for every $R \in \Sigma_1^0$, there is some computable, injective, total function f where $\vec{x} \in R$ iff $f(\vec{x}) \in K$.

Proof .:.

Note that any $R \in \Sigma_1^0$ is the domain of some computable partial function $r : \omega^n \to \omega$, $n < \omega$. Now add a variable that does nothing: $r' : \omega^{n+1} \to \omega$ defined by $r'(\vec{x}, y) \stackrel{\circ}{=} r(\vec{x})$. We can let $[\rho'] = r'$ so that by The S_n^m -Theorem (A2 $\mathfrak{c} \cdot 3$), for each fixed $\vec{x} \in \omega^n$, and any $y \in \omega$,

 $\vec{x} \in R$ iff $r'(\vec{x}, y) \downarrow$ iff $\llbracket \rho' \rrbracket (\vec{x}, y) \downarrow$ iff $\llbracket S_1^n(\rho', \vec{x}) \rrbracket (y) \downarrow$.

Write f for the map $\vec{v} \mapsto S_1^n(\rho', \vec{v})$. Since y is arbitrary, we can consider $y = f(\vec{x})$ so that $\vec{x} \in R$ iff $[\![f(\vec{x})]\!](f(\vec{x})) \downarrow$. In other words, we've shown $\vec{x} \in R$ iff $f(\vec{x}) \in K$. Note that $f : \omega^n \to \omega$ is computable, injective, and total for each $n < \omega$.

In terminology we will introduce, this says that every Σ_1^0 -relation is *reducible* to the set K above. Given that K, just by examining its definition, is Σ_1^0 too, this gives an example of a relation that is, in some sense, the most complex a Σ_1^0 set can be.

The simplest example of this is "many-to-one" reducibility. Arguably this is harder to deal with than turing reducibility, but for our purposes, it serves as a nice introduction.

- A4a·2. Definition $R \subseteq \omega$ is many-to-one reducible to $Q \subseteq \omega$, written $R \leq_{\mathrm{m}} Q$ iff for some computable (total) function $f : \omega \to \omega$, $\vec{x} \in R \Leftrightarrow f(\vec{x}) \in Q$. If, in addition, f is injective, we say R is one-to-one reducible to Q, written $R \leq_1 Q$. And if f is bijective, write $R \equiv_1 Q$.

Note the following immediate properties of \leq_m :

- many-to-one reducibility is reflexive and transitive;
- $R \leq_{\mathrm{m}} Q$ iff $\neg R \leq_{\mathrm{m}} \neg Q$;
- $R \leq_{\mathrm{m}} \omega$ iff $R = \omega$;
- $R \leq_{\mathrm{m}} \emptyset$ iff $R = \emptyset$.

We also get that \leq_m is able to help place a bound on where a given relation is in the arithmetical hierarchy. In particular, each Σ_n^0 class is closed downward under \leq_m .

- A4a·3. Theorem Let $0 < n < \omega$, and $R, Q \subseteq \omega$. Therefore $Q \leq_m R \in \Sigma_n^0$ implies $Q \in \Sigma_n^0$.

Proof .:.

This is really a generalization of Result A3 a • 11: let $f : \omega \to \omega$ witness the reduction so that $x \in Q$ iff $f(x) \in R$, meaning $Q = f^{-1} R$. Thus, since Σ_n^0 is closed under existential quantification for n > 0, $x \in Q$ iff

$$\exists y \ (\underbrace{f(x) = y \land y \in R}_{\Sigma_1^0 \subseteq \Sigma_n^0} \land \underbrace{y \in R}_{\Sigma_n^0}_{\Sigma_n^0}$$

This then yields a Σ_n^0 -definition for Q.

- A4a·4. Corollary — If $Q \leq_{m} R$ with $Q \notin \Sigma_{n}^{0}$, then $R \notin \Sigma_{n}^{0}$. The same also holds for the Π_{n}^{0} s.

As one would expect from the notation, a theorem of Myhill tells us that $R \leq_1 Q$ and $Q \leq_1 R$ implies $R \equiv_1 Q$, a kind of computable version of Cantor–Bernstein (5 C • 4). Proving this isn't especially useful nor interesting, basically being a more refined and careful checking of the proof of Cantor–Bernstein (5 C • 4). Instead we will be more interested in the question of relations which are maximal in the sense of Theorem A4 a • 1, which can be restated as saying $R \leq_1 K$ for each $R \in \Sigma_1^0$.

— A4 a • 5. Definition

Let $X \subseteq \mathcal{O}(\omega^{<\omega})$. A relation $R \in X$ is \leq_{m} -complete in X iff every $Q \in X$ has $Q \leq_{\mathrm{m}} R$. We have similar definitions for \leq_1 , and any other reduction we introduce.

Thus Theorem A4 a • 1 gives an example of a \leq_1 -complete in Σ_1^0 -set. Note that Halt (or code"Halt) is another example that is \leq_1 -complete in Σ_1^0 for the same sort of reason as *K*. Other classes also exhibit \leq_1 -complete sets.

- A4a • 6. Result $R \subseteq \omega^{<\omega}$ is \leq_{m} -complete in Σ_{n}^{0} iff $\neg R$ is \leq_{m} -complete in Π_{n}^{0} .

Proof .:.

Note that $Q \leq_{\mathrm{m}} R$ iff $\neg Q \leq_{\mathrm{m}} \neg R$ since any f witnessing the reduction has $x \in Q$ iff $f(x) \in R$, which is equivalent to $x \notin Q$ iff $f(x) \notin R$. So if R is \leq_{m} -complete in Σ_n^0 , then every $Q \in \Pi_n^0$ has $\neg Q \in \Sigma_n^0$ so that $\neg Q \leq_{\mathrm{m}} R$ iff $\neg \neg Q = Q \leq_{\mathrm{m}} R$ so that R is \leq_{m} -complete in Π_n^0 . The reverse direction follows similarly. \dashv

Hence Π_1^0 has $\omega \setminus K$ as \leq_1 -complete. The benefit of complete relations is for showing something is noncomputable.

 \neg

-

For example, K is not computable for the same reason as Corollary A2 d \cdot 3: the defined f is computable if K is.

- A4a·7. Corollary If R is \leq_m -complete in Σ_1^0 (or in Π_1^0), then $R \notin \Delta_1^0$.

Proof .:.

By Corollary A4 a • 4, since $K \leq_1 R$ with $K \notin \Pi_1^0$, it follows that $R \notin \Pi_1^0$ and therefore $R \notin \Delta_1^0$.

This yields the following, showing there is complexity even within the classes of the arithmetical hierarchy.



A4 a • 8. Figure: Σ_1^0 -relations

This can also be generalized to the larger classes, meaning that to show the strict containments of $\Sigma_n^0 \subseteq \Sigma_{n+1}^0$ for $n < \omega$, it suffices to find \leq_m -complete sets. To show this, we need to introduce kind of relative versions of the arithmetic hierarchy corresponding to the relativized notion of computable in Definition A4 • 1. The general roadmap to do show the strict containments is as follows:

- Show $\sum_{n=1}^{0}$ sets are $\sum_{n=1}^{0}$ over $\sum_{n=1}^{0}$ -sets.
- Analogous to K and Halt for Σ_1^0 , produce a sequence of sets \emptyset' , \emptyset'' , \cdots such that $\emptyset^{(n+1)}$ is \leq_1 -complete in $\Sigma_1^0(\emptyset^{(n)})$ and thus \leq_1 -complete in Σ_{n+1}^0 .
- Conclude $\emptyset^{(n+1)} \in \Sigma_{n+1}^0 \setminus \Sigma_n^0$, analogous to Corollary A4 a 7.
- Use the $\emptyset^{(n)}$ s to show there are sets in $\Delta_n^0 \setminus \Sigma_n^0$.

To show the first, we need to make precise what it means to be Σ_1^0 over a set $Q \in \Sigma_n^0$.

— A4a•9. Definition

Let *F* be a set of total functions over ω .

- The set of *partial functions recursive in* F is the \subseteq -least set containing the primitive recursive functions and F that is closed under composition and definitions by recursion.
- $\Sigma_0^0(F)$ is the set of relations that are primitive recursive in F;
- $\Sigma_{n+1}^{0}(F)$ is the set of relations R satisfying $\vec{x} \in R$ iff $\exists y \ (\langle \vec{x}, y \rangle \in Q)$ for some $\Pi_{n}^{0}(F)$ -relation Q;
- $\Pi_n^0(F)$ is the set of relations R of the form $\neg R$ where $R \in \Sigma_n^0$.

•
$$\Delta_n^0(F) = \Sigma_n^0(F) \cap \Pi_n^0(F)$$

We also say X is Σ_n^0 over F for $X \in \Sigma_n^0(F)$.

By the same sort of reasoning as before, $\Sigma_0^0(F)$ is closed under negations and thus $\Sigma_0^0(F) = \Delta_0^0(F) = \Pi_0^0(F)$. By the same arguments as before, we have the same closure conditions. Thus the proof that Δ_1^0 consists of all computable relations also shows that $\Delta_1^0(F)$ consists of all *F*-computable relations.

– A4a•10. Corollary —

Let *F* be a set of (total) functions over ω . Therefore, $R \in \Delta_1^0(F)$ iff *R* is *F*-computable.

We also have (again, just from the same proofs as before) the following easy results as examples generalized from Section A3. For all of the results below, we will be assuming F is finite so we can more easily code things. So if F is finite, we have the following:

• Each $R \in \Sigma_n^0(F)$ satisfies

 $\vec{x} \in R$ iff $\exists y_1 \forall y_2 \cdots Qy_n (\langle \vec{x}, \vec{y} \rangle \in P)$

for some $P \subseteq \omega^n$ that is F-computable, where Q is a quantifier symbol that keeps the alternating pattern.

- The *F*-relativized version of Normal Form $(A2 c \cdot 2)$ holds with Output and each CompCode_n now being primitive recursive in *F* (and using a slightly different coding to account for *F*).
- The *F*-relativized version of The S_n^m -Theorem (A2 c 3) holds.
- Every *F*-computable function is thus coded by an $e \in \omega$ with the function then being $[\![e]\!]_F^n : \omega^n \rightharpoonup \omega$.
- $\Sigma_n^0(F)$ for $0 < n < \omega$ is closed under \land , \lor , bounded quantification, and existential quantification.
- $f: \omega^n \to \omega, n < \omega$, is *F*-computable iff *f* as a relation is $\Sigma_1^0(F)$.
- $R \in \Sigma_1^0(F)$ iff R = dom(f) for some *F*-computable partial function *f*.
- Halt(F) = { $\langle e, x \rangle \in \omega^2 : x \in \text{dom}[\![e]\!]_F^1$ } and $K_F = \{e \in \omega : e \in \text{dom}[\![e]\!]_F^1\}$ are not F-computable and are both \leq_1 -complete in $\Sigma_1^0(F)$.

Knowing a fair amount about the *start* of the *F*-arithmetical hierarchy is very important, because taking *R* to be a relation \leq_1 -complete in Σ_n^0 allows us to say that $\Sigma_1^0(R)$ -relations are precisely the Σ_{n+1}^0 -relations. Note that for $Q \subseteq \omega^{<\omega}$, we generally write " $\Sigma_1^0(Q)$ " instead of more correct but more cumbersome " $\Sigma_1^0(\{\chi_Q\})$ ".

– A4a•11. Lemma –

Let $R \subseteq \omega$ and $n < \omega$. Therefore R is Σ_{n+1}^0 iff R is $\Sigma_1^0(Q)$ for some $Q \in \Sigma_n^0 \cup \Pi_n^0$.

Proof .:.

Suppose R is Σ_{n+1}^0 . Thus R satisfies $x \in R$ iff $\exists y \ (\langle x, y \rangle \in Q)$ for some $Q \in \Pi_n^0$. Since χ_Q is clearly $\Delta_0^0(Q)$, it follows that R is $\Sigma_1^0(Q)$.

Now suppose R is $\Sigma_1^0(Q)$ for some $Q \in \Sigma_n^0 \cup \prod_n^0$. As with the proof that (1) implies (3) in Theorem A3 a • 2, $R = \operatorname{dom}(f)$ for some Q-computable $f = \llbracket e \rrbracket_F^1$. But then $x \in R$ iff $\exists y (\llbracket e \rrbracket_F^1(x) = y)$ iff there is a sequence of computations that obey the rules of Computer Transitions (A2 b • 3). Given that Q is Σ_n^0 or Π_n^0 , it's easy to see that a natural number c being the code of a sequence of computations allowing Q is Δ_{n+1}^0 (we need to make use of both Q and $\neg Q$ to state whether " $\chi_Q \mid x$ " transitions to " \mid 1" or to " \mid 0", increasing the complexity slightly from whichever Q is). This means $x \in R$ iff there is a c where some Δ_{n+1}^0 -relation holds, which is then a Σ_{n+1}^0 -relation.

So to generate more complete sets, we can use the same idea as with K and Halt.

- A4a·12. Definition Let $R \subseteq \omega^n$, $n < \omega$. Write R' for the *jump* of R: $\{e \in \omega : [[e]]_R^1(e) \downarrow\}$. For $m < \omega$, write $R^{(m)}$ for $R''' \stackrel{\dots}{\longrightarrow} ', m$ applications of the jump operator.

In particular, the definition of K before is just \emptyset' . So we can write \emptyset'' , \emptyset''' , and so on to get $\emptyset^{(n)}$ for any $n < \omega$.

The point of these, if you recall, is to show the strict containments of Figure A3 • 1 and Figure A3 b • 3. Recalling the plan from above Definition A4 a • 9, we have shown the first step: Σ_{n+1}^0 sets are Σ_1^0 over Σ_n^0 -sets. We now want to show that we can really just consider sets that are Σ_1^0 over $\emptyset^{(n)}$ so that $\Sigma_{n+1}^0 = \Sigma_1^0(\emptyset^{(n)})$.

- A4a·13. Lemma Let $X \subseteq \omega$. Therefore X' is \leq_1 -complete in $\Sigma_1^0(X)$. Moreover, $X' \notin \Delta_1^0(X)$

Proof .:.

That X' is \leq_1 -complete in $\Sigma_1^0(X)$ follows from the same proof as Theorem A4 a • 1 with the X-relativized forms of Normal Form (A2 c • 2) and The S_n^m -Theorem (A2 c • 3).

To see that X' is not X-computable, we use the same idea as before: otherwise consider f defined by $f(n) = [n]_X^1(n) + 1$ if $n \in X'$ and f(n) = 0, which would be X-computable if X' were X-computable. This means

 $f = \llbracket m \rrbracket_X^1$ for some $m < \omega$. Since f is total, this yields $m \in X'$ and thus $f(m) = \llbracket m \rrbracket_X^1(m) + 1 = f(m) + 1$, a contradiction.

– A4 a • 14. Theorem (Post's Theorem) —

 $\emptyset^{(n)}$ is \leq_1 -complete in Σ_n^0 , and $\Sigma_{n+1}^0 = \Sigma_1^0(\emptyset^{(n)})$ for each $n < \omega$. And the same hold for the Π_n^0 s, $n < \omega$.

Proof .:.

Proceed by induction on n. For n = 0, both are clear:

• $\Sigma_1^0(\emptyset^{(0)}) = \Sigma_1^0(\emptyset) = \Sigma_1^0$; and • $\emptyset^{(1)} = \emptyset' = K$ is \leq_1 -complete in Σ_1^0 by Theorem A4 a • 1. For n + 1, $\emptyset^{(n+2)}$ is \leq_1 -complete in $\Sigma_1^0(\emptyset^{(n+1)})$ by Lemma A4 a • 13. It then suffices to show $\Sigma_{n+2}^0 = 0$ $\Sigma_1^0(\emptyset^{(n+1)}).$

(\subseteq) Suppose $R \in \Sigma_{n+2}^{0}$. Thus for some $Q \in \Pi_{n+1}^{0}$, R satisfies $\vec{x} \in R$ iff $\exists y (\langle \vec{x}, y \rangle \in Q)$. Since $\emptyset^{(n+1)}$ is \leq_1 -complete in $\Sigma_1^0(\emptyset^{(n)}) = \Sigma_{n+1}^0$, $\neg \emptyset^{(n+1)}$ is \leq_1 -complete in Π_{n+1}^0 . So there is some total, injective, computable function f where $\vec{y} \in Q$ iff $f(\vec{y}) \notin \emptyset^{(n+1)}$. Hence

$$\vec{x} \in R \quad \text{iff} \quad \exists y \ (f(\langle \vec{x}, y \rangle) \notin \emptyset^{(n+1)}) \quad \text{iff} \quad \exists y \ \exists z \ (\underbrace{f(\vec{x}, z) = y}_{\Sigma_1^0 \subseteq \Sigma_1^0(\emptyset^{(n+1)})} \land \underbrace{y \notin \emptyset^{(n+1)}}_{\Sigma_1^0(\emptyset^{(n+1)})})$$

shows that $R \in \Sigma_1^0(\emptyset^{(n+1)})$, and therefore $\Sigma_{n+2}^0 \subseteq \Sigma_1^0(\emptyset^{(n+1)})$.

 (\supseteq) Conversely, suppose $R \in \Sigma_1^0(\emptyset^{(n+1)})$ and thus for some $\emptyset^{(n+1)}$ -computable relation $Q, \vec{x} \in R$ iff $\exists y \ (\langle \vec{x}, y \rangle \in Q)$. Note that being a $\emptyset^{(n+1)}$ -computable relation means that there exists the code of a program $e \in \omega$ where $\vec{y} \in Q$ iff there is a code of computations following the program e and Computer Transitions (A2 b • 3) using $\tilde{\emptyset}^{(n+1)}$ and $\neg \hat{\emptyset}^{(n+1)}$ that ends in "| 1". Because $\hat{\emptyset}^{(n+1)}$ is $\Sigma_1^0(\hat{\emptyset}^{(n)}) = \tilde{\Sigma}_{n+1}^0$ and $\neg \emptyset^{(n+1)}$ is Π^0_{n+1} , it follows being a code of computations is Δ^0_{n+2} , and thus Q—stating the existence of such a code—is Σ^0_{n+2} (in fact, Q will be Δ^0_{n+2} , but this isn't important for us). So R is Σ^0_{n+2} . \dashv

So completeness is really quite useful, showing that $\Sigma_n^0(R)$ —with $R \leq_1$ -complete in Σ_m^0 —is just Σ_{n+m}^0 . This allows us to conclude all the strict containments of the arithmetical hierarchy through the following figure.





REDUCIBILITY AND LOGIC

Proof .:.

Post's Theorem (A4 a • 14) tells us by Lemma A4 a • 13 that $\emptyset^{(n+1)} \in \Sigma_1^0(\emptyset^{(n)}) \setminus \Delta_1^0(\emptyset^{(n)}) = \Sigma_{n+1}^0 \setminus \Delta_{n+1}^0$ demonstrating $\Delta_{n+1}^0 \subsetneq \Sigma_{n+1}^0$.

To show $\Sigma_n^0 \subsetneq \Delta_{n+1}^0$, we instead note that $\Pi_n^0 \subsetneq \Delta_{n+1}^0$, and the result for Σ_n^0 follows similarly. Note that $\emptyset^{(n)}$ is \leq_1 -complete in Σ_n^0 . Since n > 0, we thus know $\emptyset^{(n)} \in \Sigma_n^0 \setminus \Delta_n^0 = \Sigma_n^0 \setminus \Pi_n^0$ by Post's Theorem (A4 a • 14) and Lemma A4 a • 13. Since $\emptyset^{(n)} \in \Delta_{n+1}^0$ it follows that $\Pi_n^0 \subsetneq \Delta_{n+1}^0$ and similarly $\Sigma_n^0 \subsetneq \Delta_{n+1}^0$ by considering $\neg \emptyset^{(n)}$ in $\Pi_n^0 \setminus \Sigma_n^0$.

What Post's Theorem (A4 $a \cdot 14$), Lemma A4 $a \cdot 13$, and Corollary A4 $a \cdot 16$ should signal is that reducibility in many of its forms is really necessary to study ever more complex sets, and by Corollary A3 $b \cdot 10$, it also has consequences for definability. Let us turn to some more consequences for this topic that we have encountered thus far.

§A4b. Consequences for logic

Corollary A4 a • 16 essentially tells us that there are ever more complicated sets of natural numbers. And, in particular, there are ever more complicated *definable* sets of natural numbers through identifying Σ_n^0 -relations with Σ_n^N -definable sets. In particular, we have the following, showing that we can't actually define a truth relation over **N**.

- A4 b · 1. Theorem (Tarski's Nondefinability of Truth) There is no FOL($\{+, \cdot, 0, 1\}$)-formula φ where for all FOL-sentences σ , N $\vDash \sigma$ iff N \vDash " φ (code(" σ "))".

Proof .:.

Suppose otherwise, and let φ be a $\sum_{n=1}^{N}$ (note that every formula can be placed in this hierarchy by placing it in prenex normal form and then reducing blocks of quantifiers to a single quantifier coding a sequence). Without loss of generality, n > 0 so that $\sum_{n=1}^{0}$ is closed under existential quantification. Therefore the property of x being a true FOL-sentence is $\sum_{n=1}^{0}$. Write True $\in \sum_{n=1}^{0}$ for the set of (codes of) true sentences.

Note that for a formula of one variable, $\theta(x)$, the map $\langle \theta, n \rangle \mapsto \operatorname{code}("\theta(n)")$ is still primitive recursive (since the map $n \mapsto \operatorname{code}(n)$ is primitive recursive and we are just replacing every free occurrence of x in the sequence of symbols " $\theta(x)$ " with " $1 + \cdots + 1$ "). Call this map f. Thus for every relation R defined by a formula σ_R coded by a number $\operatorname{code}(\sigma_R) = \rho_R \in \omega$, and every $x \in \omega$,

 $x \in R \Leftrightarrow \operatorname{True}(f(\rho_R, n)) \Leftrightarrow \exists y \ (y = f(\rho_R, n) \land \operatorname{True}(y)),$

which is Σ_n^0 since the graph of f (i.e. f as a relation) is $\Sigma_1^0 \subseteq \Sigma_n^0$ and True is Σ_n^0 by hypothesis. Thus every definable relation is Σ_n^0 . But this contradicts Corollary A4 a • 16 and Corollary A3 b • 10. In particular, $\emptyset^{(n+1)}$ is Σ_{n+1}^0 and Σ_{n+1}^N but not Σ_n^0 .

As a corollary, we can't reduce truth to any particular arithmetical relation. This yields Gödel's incompleteness theorem in that we cn't reduce truth to whether there is a proof from a certain set of axioms that we can computably code.

- A4 b \cdot 2. Corollary (Gödel's First Incompleteness Theorem) - There is no theory T such that all the following hold:

- The set of (codes of) formulas in *T* is computable;
- T is sound: $\mathbf{N} \models T$ (this implies consistency); and
- *T* is complete: every φ has $T \vdash \varphi$ or $T \vdash \neg \varphi$ ".

Proof ...

Suppose such a T exists. Therefore, for any formula φ , $\mathbf{N} \models \varphi$ iff $T \vdash \varphi$, since by soundness, $T \vdash \varphi$ implies $\mathbf{N} \models \varphi$. For the converse, by completeness, if $\mathbf{N} \models \varphi$, then either $T \vdash \varphi$ or $T \vdash "\neg \varphi$ ". If $T \vdash \varphi$, we're done, so

suppose $T \vdash \neg \varphi$ so that $T \cup \{\varphi\}$ is inconsistent and thus has no model, contradicting that $N \models \{\varphi\} \cup T$.

But then any $n \in \omega$ has $n \in$ True iff there is a natural number coding a proof with all the axioms of the proof in T. Since T is computable, and it's not hard (although, again, extremely tedious) to show that being the code of a proof from T is Δ_1^0 . Hence True is Σ_1^0 , contradicting that True isn't arithmetical by Tarski's Nondefinability of Truth (A4 b • 1).

More generally, and by the same proof, there can be *no* intelligible proof system that is sound and complete.

— A4b•3. Corollary -

A system for proofs is a set \mathcal{P} where a proof is a sequence of formulas and the system for proofs specifies which sequences are acceptable. For a formula φ , we say $\vdash_{\mathcal{P}} \varphi$ iff there is a proof of \mathcal{P} that ends with φ .

Therefore, there is no proof system \mathcal{P} such that all the following hold:

- being the code of a proof of \mathcal{P} is computable;
- for every $FOL(\{0, 1, +, \cdot\})$ -sentence σ , if $\vdash_{\mathscr{P}} \sigma$, then $N \models \sigma$; and
- for every FOL($\{0, 1, +, \cdot\}$)-sentence σ , either $\vdash_{\mathscr{P}} \sigma$ or $\vdash_{\mathscr{P}} "\neg \sigma"$.

The idea behind this generalization is that it goes beyond the usual ideas of proof in terms of first order logic or the restrictions of only using 0, 1, +, and \cdot over ω . We could add in whatever additional relations, functions, or constants we want. The proof system doesn't need to behave in the same sort of way or directly make reference to **N**. And at worst, we can always take \mathcal{P} to just be the set of theorems of whatever weird system one desires. The result is always the same: no matter what attempt we make at codifying the basic principles of logic and arithmetic, we always will have sentences that are left unproven and unrefuted by the system or else it's impossible to tell whether a given proof is acceptable.

Note that the hypothesis that things be computable cannot be dropped: from the theory $T = \{\varphi : \mathbf{N} \models \varphi\}$, we obviously get that T is sound and complete. The issue is that this T isn't computable, and really is just the truth set that Tarski's Nondefinability of Truth (A4b • 1) showed isn't N-definable. Indeed, it's hard to imagine the usefulness of a system where we can't even fact check whether a proposed proof is actually a proof.

One concept that comes up frequently for us is the idea of a theory being encoded by ω just by coding the symbols of the signature and using the standard coding mechanisms to code formulas as finite sequences of these.

- A4b•4. Definition

A FOL(σ)-theory *T* is *encodable* iff there is a coding of FOL(σ)-symbols into ω such that the set of coded sequences of symbols {code(φ) : $\varphi \in T$ } is computable.

A4b•5. Corollary

If T is encodable, then

• "x is the code of a proof of y from T" is a computable relation; and

• so is "x is the code of a proof of $code(\neg)^{\gamma} y$ from T".

Hence there are FOL({0, 1, +, ·})-formulas proof_T(x, y) and disproof_T(x, y) that represent these over N. Thus there is a sentence Con(T) where $N \models Con(T)$ iff T is consistent:

 $\neg \exists x \left[\operatorname{proof}_T \left(x, \operatorname{code} \left((v_0 = v_0 \land \neg v_0 = v_0) \right) \right) \right]$

A remarkable version of Gödel's incompleteness theorem has us we can come up with a specific sentence for it. And moreover, we can work with systems that aren't necessarily sound, and whose consistency is a question mark. This is especially useful with various strengthenings of ZF and ZFC.

A4b•6. Theorem (Gödel's Second Incompleteness Theorem)

Let T be an encodable theory that can interpret PA. Therefore, $T \nvDash "Con(T)"$ (or rather, the translation of Con(T)).

Meta-theoretically, this means that we will never have a justification that our foundations for mathematics are actually correct from the foundations themselves: we must always go beyond to justify their correctness. Some common responses to this would be to consider a sort-of "consistent closure" of our theory T, taking $T_0 = T$ and $T_1 = T$

 $T_0 \cup \{\text{``Con}(T_0)\text{''}\}, T_2 = T_1 \cup \{\text{``Con}(T_1)\text{''}\}, \text{ and so on. The resulting theory, } T_\omega = \bigcup_{n < \omega} T_n \text{ would seem to be consistent, right? Well, the issue with this approach is that } T_\omega \vdash \text{``Con}(T_n)\text{''} \text{ for each } n < \omega, \text{ but we don't necessarily have } T_\omega \vdash \text{``Con}(T_\omega)\text{''}. \text{ This is important, because if we are to continue this approach, we need to be able to encode limit steps like } T_\omega \text{ into } \mathbf{N} \text{ and proceed to define } T_{\omega+1}, T_{\omega+2}, \cdots, T_{\omega+\omega}, \text{ and so on. The result is that eventually we cannot actually define our theory and what formulas are axioms becomes unintelligible.xiv$

To prove Gödel's Second Incompleteness Theorem $(A4b \cdot 6)$, we need to make precise what it means to interpret another theory in possibly a different signature. In essence, the idea is similar to how we can interpret \mathbb{N} and \mathbb{N} in ZFC: relativize the operations and statements to ω : instead of saying $\neg \exists x \forall y (y < x \rightarrow y + 1 < x)$, something false under ZFC by the axiom of infinity, we say $\neg \exists x \in \omega \forall y \in \omega (y < x \rightarrow y + 1 < x)$, something true under ZFC as ω is the least limit ordinal. Note that this translation is easy to carry out through codings: if we code the FOL(\in) and FOL($0, 1, +, \cdot$)-formulas, it's easy to transform the FOL($0, 1, +, \cdot$)-formulas into the FOL(\in)-formulas in a primitive recursive way: just replace each occurrence of + with its defining FOL(\in)-formula and similarly for the other symbols.

```
– A4b•7. Definition
```

Let σ and τ be two signatures. Let Σ be a FOL(σ)-theory and T a FOL(τ)-theory.

We say T can *interpret* Σ via iff there is a primitive recursive function tolk—called a *translation*—from FOL(σ)-sentences to FOL(τ)-sentences such that for all FOL(σ)-sentences φ and ψ ,

- $\Sigma \vdash \varphi$ implies $T \vdash tolk(\varphi)$; and
- tolk preserves \neg and \land : $\top \vdash$ "tolk($\neg \varphi$) $\Leftrightarrow \neg$ tolk(φ)", and $\top \vdash$ "tolk($\varphi \land \psi$) \Leftrightarrow tolk(φ) \land tolk(ψ)".

We call tolk above primitive recursive in the sense that the corresponding map code \circ tolk \circ code⁻¹ : $\omega \rightarrow \omega$ is primitive recursive. It should be noted that this complexity requirement on tolk will be satisfied in every theory we will consider. In particular, ZF – P, ZF, ZFC, and any other strengthenings of these can interpret PA by the following corollary.

— A4b•8. Corollary -

ZF - P (and many weaker set theories too) can interpret PA.

Up to this point, we've only proven equivalences over N. So it should be noted that all of the equivalences used thus far are provable in just PA rather than the entirety of Th(N). And if we wanted to be very careful, we could a lot more work and show that similar results hold over much weaker axioms that suffice to prove Gödel's work. Note that we abandon the use of the translation τ below, since it doesn't add much. In particular, the following lemma is all we need, although it is left unproven here. Mostly it just consists in proving Corollary A2 a • 5 from the proof of Result A1 • 4 over PA.

— A4b•9. Lemma

Let $f : \omega \to \omega$ be a computable (total) function. For $n \in \omega$ write #n for " $0 + 1 + \cdots + 1$ ", adding n '1's. Therefore, there is some FOL($\{0, 1, +, \cdot\}$)-formula φ such that

- φ defines f over N: N $\vDash \varphi(x, y)$ iff f(x) = y; and
- φ that also numeralwise represents f over PA: for all $x, y \in \omega$, PA $\vdash \varphi(\#x, \#y)$ iff f(x) = y.

To prove Gödel's Second Incompleteness Theorem (A4 b • 6), we will consider another variant of Gödel's First Incompleteness Theorem (A4 b • 2) where we explicitly construct a sentence σ that has $T \not\vdash \sigma$ and $T \not\vdash "\neg \sigma$ ".

— A4b•10. Definition -

Let T be an encodable theory and that interprets PA via translation τ . A rosser sentence for τ and T is a FOL($\{0, 1, +, \cdot\}$)-sentence ρ such that, for $r = \text{code}(\tau(\rho))$,

$$\mathsf{PA} \vdash ``\rho \leftrightarrow \forall y \ (\mathrm{proof}_T(y, \#r) \rightarrow \exists z \leq y \ \mathrm{disproof}_T(z, \#r))"$$

Clearly a rosser sentence ρ is unprovable from a consistent theory *T*, as any proof of it has a code $n \in \omega$ where then PA shows there is a (code for a) proof $m \le n < \omega$ of " $\neg \rho$ " from *T*, meaning *T* would be inconsistent. So this is Rosser's form^{xv} of Gödel's First Incompleteness Theorem (A4b • 2).

 x^{iv} More precisely, we could only carry out the process to define at most T_{α} for each α less than a countable ordinal known as $\omega_1^{CK} < \omega_1$, but we cannot deal with $T_{\omega_1^{CK}}$ itself.

xvnamed after John Rosser, and frequently referred to as "Rosser's trick"

A4b•11. Lemma (Rosser's Form of Gödel's First Incompleteness Theorem) –

Let T be a consistent, encodable theory that can interpret PA. Therefore there is a rosser sentence for T.

Proof .:.

Since T is encodable, proof and disproof can be represented over N and PA is strong enough to prove the equivalence when we restrict ourselves to talking about actual natural numbers. More precisely, if φ represents the a computable relation, then $N \models \varphi(x)$ iff $PA \vdash "\varphi(\#x)"$.

Write Rosser(v) for a more general attempt at the rosser sentence: Rosser(v) is

$$\forall y \; (\operatorname{proof}_T(y, \tau(v)) \to \exists z \leq y \; \operatorname{disproof}_T(z, \tau(v)))^{\circ}.$$

So we just need to find a fixed point of the map $x \mapsto \text{code}(\text{"Rosser}(x)\text{"})$. We can find an explicit example (well, explicit from whatever fixed coding we chose).

Write sub(x, n) for the (computable) function that substitutes each free occurrence of " v_0 " in the sequence coded by x with the number n written as " $1 + \cdots + 1$ ". More concretely, we have

$$ub(code(\varphi(y)), n) = code(\varphi(\#n)). \tag{(\star)}$$

With this function at our disposal and representable over PA in the sense above, let $\varphi(x)$ be the formula

$$\exists z \ (z = \operatorname{sub}(x, x) \land \operatorname{Rosser}(z))".$$

Let $e = \operatorname{code}(\varphi)$ and let θ be the sentence $\varphi(\#e)$, which is just to say $\exists z \ (z = \operatorname{sub}(\#e, \#e) \land \operatorname{Rosser}(z))$. Note by (\star) that $\operatorname{sub}(\operatorname{code}(\varphi(v_0)), \#e) = \operatorname{code}(\varphi(\#e)) = \operatorname{code}(\theta)$. Hence θ states $\operatorname{Rosser}(\operatorname{code}(\theta))$, and thus θ is a rosser sentence for T.

As a corollary, we can prove the second incompleteness theorem. The idea behind the proof is that the rosser sentence for T is equivalent to the consistency of T.

- Proof of Gödel's Second Incompleteness Theorem (A4 b • 6) ...

Let T interpret PA via the translation τ . Let ρ be a rosser sentence for T by Rosser's Form of Gödel's First Incompleteness Theorem (A4b • 11). It suffices to show $\tau(\rho)$ is *independent* of T and ρ equivalence to Con(T):

1. $T \not\vdash \tau(\rho)$; and

2. $\mathsf{PA} \vdash \mathsf{``Con}(T) \rightarrow \rho$ ''.

If we show these, then $T \vdash \tau(\operatorname{Con}(T)) \to \tau(\rho)^n$. So if $T \vdash \tau(\operatorname{"Con}(T))^n$, then $T \vdash \tau(\rho)$, contradicting (1) and showing that $T \not\vdash \tau(\operatorname{"Con}(T))^n$.

 $T \not\vdash \tau(\rho)$ since this would imply a disproof, contradicting the consistency of T. Explicitly, suppose $y \in \omega$ is (the code of) a proof of $\tau(\rho)$ from T so that $\mathsf{PA} \vdash "\mathsf{proof}_T(\#y, \#r)"$ and thus $\mathsf{PA} \vdash "\exists z \leq y(\mathsf{disproof}_T(z, \#r))"$. In particular, for $\mathsf{N} \models \mathsf{PA}$,

$$\mathsf{N} \models ``\exists y \exists z (\operatorname{proof}_T(y, r) \land \operatorname{disproof}_T(z, r))",$$

contradicting that T is consistent.

So suffices to show $\mathsf{PA} \vdash "\operatorname{Con}(T) \to \rho$ ". The argument above can be pretty easily formalized in PA (where $T = \mathsf{PA}$ and τ is just the identity) and this shows that $\mathsf{PA} \vdash "\operatorname{Con}(T) \to \neg \exists y (\operatorname{proof}_T(y, \#r))$ ". And clearly, if there is no proof of the rosser sentence, then ρ vacuously holds: every proof has a shorter disproof, just because there are no proofs. So $\mathsf{PA} \vdash "\operatorname{Con}(T) \to \rho$ ".

The same idea tells us that a formula of first-order logic being valid (meaning $M \vDash \varphi$ for every M) is Σ_1^0 but not Δ_1^0 . A sketch of the proof of this is given below.

A4 b • 12. Result (Church's Theorem)

For any fixed coding of formulas into ω , the set of *valid* FOL($\{0, 1, +, \cdot\}$)-*formulas* is not computable, i.e. Valid = $\{e \in \omega : e \text{ is the code of a formula and there is a proof of <math>e$ without using any axioms} \in \Sigma_1^0 \setminus \Delta_1^0. Proof Sketch .:.

Firstly, one can show that a weakened version of Lemma A4b•9 holds. In particular, we only need that Robinson's arithmetic axioms, denoted Q, can numeralwise represent primitive recursive functions (Q is a weakening of PA in the sense that PA \models Q). The important point of this weakening is that Q is finitely axiomatizable. In particular, there is some particular sentence Q that is the conjunction of all the sentences of Q. So Q $\vdash \varphi$ iff $\vdash "Q \rightarrow \varphi"$. So if Valid were Δ_1^0 , then so would the theorems of Q—call this set Th(Q) $\subseteq \omega$ —be, as seen just by asking whether " $Q \rightarrow \varphi$ " is vaid or not.

Clearly Th(Q) is Σ_1^0 since $x \in \text{Th}(Q)$ iff there is the code of a proof of $\text{code}("Q \to ")^x$. Being the code of a proof is Δ_1^0 so that the existence is Σ_1^0 . To show that this is not Δ_1^0 , we note that Q can numeralwise represent the CompCode₁ of Lemma A2 c • 1 and Normal Form (A2 c • 2). In particular, consider the halting problem again: $\langle e, x \rangle \in \text{Halt iff } [\![e]\!]^1(x) \downarrow$. This is equivalent, since both CompCode₁ and Output are total, to the existence of a y where CompCode₁ (e, x, y). By numeralwise representation in Q, we get a FOL($\{0, 1, +, \cdot\}$)-formula representing CompCode₁ where then

 $\langle e, x \rangle \in \text{Halt}$ iff $N \models \exists y \text{ CompCode}_1(e, x, y)$ iff $Q \vdash \exists y \text{ CompCode}_1(\#e, \#x, y)$.

The last equivalence can be seen as follows: if Q proves this, then clearly $N \models Q$ models it too. If $N \models CompCode_1(e, x, y)$, then by numeralwise representation, $Q \vdash "CompCode_1(#e, #x, #y)"$ and hence the existence of such a y. Hence if Th(Q) were Δ_1^0 , then so would Halt be, contradicting Corollary A2 d • 3. \dashv

The above proof actually shows Valid is \leq_1 -complete in Σ_1^0 as witnessed by a \leq_1 -reduction from Halt to Valid witnessed by the map f sending $\langle e, x \rangle$ to the code of " $Q \to \exists y$ CompCode₁(#e, #x, y)". Hence while if something is always true we can find a proof of it, we won't be able to tell beforehand whether there is a proof. This is in contrast to propositional logic where validity is computable just by examining the corresponding truth table.

Ideas used in Gödel's Second Incompleteness Theorem (A4b•6) also lead to the idea of consistency strength, since PA $\not\vdash$ Con(PA), but ZFC \vdash Con(PA) with ZFC $\not\vdash$ Con(ZFC). So ZFC has a stronger consistency strength than PA: Con(ZFC) implies Con(PA), but the reverse doesn't hold.

§A4 c. Turing reducibility and the turing degrees

The idea of many-to-one and one-to-one reduction is a fairly *fine* means of showing something is "reducible" to something else. A less subtle, more coarse notion of reducibility is that of turing reducibility.

- A4c·1. Definition Let $R, Q \subseteq \omega$ be relations. We say R is *turing reducible* to Q, written $R \leq_T Q$, iff χ_R is Q-computable. We say R and Q are *turing equivalent*, written $R \equiv_T Q$, iff $R \leq_T Q$ and $Q \leq_T R$.

Note the following basic facts about turing reducibility.

- \leq_{T} is transitive: $P \leq_{\mathrm{T}} Q \leq_{\mathrm{T}} R$ implies $P \leq_{\mathrm{T}} R$;
- \equiv_{T} is an equivalence relation over $\mathcal{P}(\omega)$;
- $R \equiv_{\mathrm{T}} Q$ for all $R, Q \in \Delta_1^0$; and
- $R \leq_{\mathrm{m}} Q$ implies $R \leq_{\mathrm{T}} Q$ (the reverse might not hold).
- $R \equiv_{\mathrm{T}} \neg R$ for all $R \subseteq \omega$.

As a result of this last item, being \leq_m or \leq_1 -complete implies being \leq_T -complete (and the reverse may not hold). In particular, we have the following characterization of the Δ_n^0 sets of the arithmetical hierarchy: they are exactly those computable from $\emptyset^{(n-1)}$. This actually tells us that there's a difference between many-to-one reducibility and turing reducibility, which makes sense, since many-to-one reducibility requires just a single, already computable function to translate between the two sets. Turing reducibility, on the other hand, allows all means of computation by way of using one as an oracle.

— A4c·2. Corollary

For $0 < n < \omega$ and $R \subseteq \omega$, R is Δ_n^0 iff $R \leq_T \emptyset^{(n-1)}$. Hence $\emptyset^{(n-1)}$ is \leq_T -complete in Δ_n^0 but \leq_m -complete merely in Σ_{n-1}^0 (so long as $n-2 \ge 0$). In particular, there is some Δ_2^0 -relation R where $R \leq_T \emptyset''$ but $R \not\leq_m \emptyset''$.

REDUCIBILITY AND LOGIC

Proof .:.

Let $n < \omega$. By Post's Theorem (A4 a • 14), R is Σ_{n+1}^0 iff R is $\Sigma_1^0(\emptyset^{(n)})$, and similarly for Π_{n+1}^0 . It follows that R is Δ_{n+1}^0 iff R is $\Delta_1^0(\emptyset^{(n)})$, but these are precisely the sets computable from $\emptyset^{(n)}$ by the same reasoning as in Theorem A3 a • 4. This proves the first statement, and clearly implies that $\emptyset^{(n)}$ is \leq_{T} -complete in Δ_{n+1}^0 .

So for n = 2, \emptyset'' is \leq_{T} -complete in Δ_3^0 . Suppose towards a contradiction that every $R \in \Delta_3^0$ not only has $R \leq_{\mathrm{T}} \emptyset''$ but also $R \leq_{\mathrm{m}} \emptyset''$. Since \emptyset'' is Σ_2^0 , it follows that $\neg \emptyset'' \in \Pi_2^0 \subseteq \Delta_3^0$ has $\emptyset'' \leq_{\mathrm{m}} \emptyset'' \in \Sigma_2^0$, implying $\neg \emptyset''$ is Σ_2^0 by Theorem A4 a • 3, contradicting that it is \leq_1 -complete in Π_2^0 . Therefore there must be some $R \leq_{\mathrm{T}} \emptyset''$ with $R \not\leq_{\mathrm{m}} \emptyset''$.

So turing reducibility gives rise to sets complete in $\Delta_n^{0, xvi}$ Results like the above tell us that there is a substantial difference between the equivalence classes of turing equivalence compared to many-to-one equivalence. These equivalence classes can be understood as degrees of computation, and are of fundamental importance to computability theory, which often cares about how sets are classified by their complexity.

- A4 c • 3. Definition Let $X \subseteq \omega$. The *turing degree* of X is the equivalence class of X up to turing equivalence: $[X]_{\equiv_{T}} = \{A \subseteq \omega : A \equiv_{T} X\}.$

We will often refer to turing degrees with boldface variables: **a**, **b**, **c**, and so on. Note that turing reducibility can be abstracted away from the subsets of ω to instead yield an order on the turing degrees themselves: **a** \leq_T **b** iff for some (equivalently *any*) $X \in \mathbf{a}$ and $Y \in \mathbf{b}$, $X \leq_T Y$. By the above remarks, we have the following.

− A4 c • 4. Definition Write 0 for the turing degree of 0. For a a turing degree with $X \in a$, write a' for the turing degree of X'.

In particular, $\mathbf{0} = \Delta_1^0$, and similarly, $\mathbf{0}^{(n)} = \Delta_{n+1}^0$ for $n < \omega$ by Corollary A4 c • 2.

— A4c•5. Result –

For any turing degree **a**, **a**' is well-defined: if $X \equiv_T Y$ then $X' \equiv_T Y'$. Moreover, **a** <_T **a**' for all turing degrees **a**.

Proof Sketch .:.

That \mathbf{a}' is well defined follows easily from Definition A4 $\mathbf{a} \cdot \mathbf{12}$: using the fact that $Y \leq_T X$, one can mimic the computations of $[\![e]\!]_Y^1(e)$ in X. This process converges iff the function itself converges meaning it's in Y'. But this process converging can be carried out in X' by considering the analogous Halt for X-computable functions and seeing that this is turing equivalent to X'.

Thus for any turing degree **a**, since $X \in \mathbf{a}$ has $X <_T X'$, it follows that $\mathbf{a} = [\mathbf{X}]_{\equiv_T} < [\mathbf{X}']_{\equiv_T} = \mathbf{a}'$.

This gives one method of getting more turing degrees from previous. Another method is to take the join: the least turing degree that sits above two given turing degrees.

A4c•6. Result

Let $X, Y \subseteq \omega$. Set $X \oplus Y = \{2n : n \in X\} \cup \{2n + 1 : n \in Y\}$. Therefore $X, Y \leq_{T} X \oplus Y$, and for any Z, $X, Y \leq_{T} Z$ implies $X \oplus Y \leq_{T} Z$.

Proof .:.

Clearly $X, Y \leq_T X \oplus Y$ since $X, Y \leq_1 X \oplus Y : x \in X$ iff $2x \in X \oplus Y$ and similarly for Y. And clearly, $X, Y \leq_T Z$ implies we can compute from Z that $n \in X \oplus Y$ iff $\exists m < n \ ((n = 2m \land m \in X) \lor (n = 2m + 1 \land m \in Y))$, which yields that $X \oplus Y \leq_T Z$.

Stated in terms of orders, this gives the following.

^{xvi}This is not the case with many-to-one reducibility, which has no \leq_m -complete sets in Δ_n^0 for n > 1.

— A4c•7. Corollary

Let **a** and **b** be turing degrees. Therefore there is some **c** such that $\mathbf{a} <_{\mathrm{T}} \mathbf{c}$ and $\mathbf{b} <_{\mathrm{T}} \mathbf{c}$. In particular, the turing degrees under \leq_{T} form an *upper semi-lattice*:

- \leq_{T} is transitive and anti-reflexive;
- For any two turing degrees, there is a \leq_T -least upper bound—namely the join $\mathbf{a} \oplus \mathbf{b}$; and
- for any turing degree $\mathbf{a}, \mathbf{a} <_{\mathrm{T}} \mathbf{a}'$.

Many-to-one equivalence is a relatively *fine* equivalence relation in that it has smaller equivalence classes than turing equivalence. That said, all of the equivalence classes (for either) are still only countable.

A4 c • 8. Result Each turing degree $\mathbf{a} \subseteq \mathcal{P}(!)$ has $|\mathbf{a}| = \aleph_0$. Hence there are 2^{\aleph_0} -many turing degrees.

Proof .:.

For each $X \subseteq \omega$, there are only \aleph_0 -many programs that make use of X as an oracle (or rather, its characteristic function just as a given relation symbol). As each $Y \leq_T X$ has an associated program, this implies there are at most \aleph_0 -many $Y \in [X]_{\equiv_T}$, and so each turing degree is countable.

To see that each turing degree **a** has $|\mathbf{a}| = \aleph_0$, note that for every $X \in \mathbf{0} = \Delta_1^0$ and $Y \in \mathbf{a}$, Result A4c•6 tells us $X \oplus Y \in \mathbf{a}$ and since there are \aleph_0 -many elements in **0** (the finite subsets of ω , for instance), it follows that $|\mathbf{a}| \ge \aleph_0$ and thus $|\mathbf{a}| = \aleph_0$.

A topic in the study of turing degrees that we will not pursue here is the existence of certain kinds of turing degrees. In particular, there are *minimal* turing degrees in the sense that they are $<_{\rm T}$ -minimal over the set of non-computable turing degrees: $\mathbf{0} <_{\rm T} \mathbf{a}$ but there is no **b** with $\mathbf{0} <_{\rm T} \mathbf{b} <_{\rm T} \mathbf{a}$. Moreover, the structure of the turing degrees differs significantly from the arithmetical hierarchy. In fact, a turing degree **a** is called an *r.e. degree*^{xvii} in the sense that **a** contains a Σ_1^0 -set. The study of r.e. turing degrees is long and complicated, involving many arguments using the so-called *priority method*, but the results regarding these degrees are very nice, and we state some of these below.

– A4 c • 9. Result

- 1. (Sacks' Density Theorem) For any two r.e. degrees $\mathbf{a} <_{\mathrm{T}} \mathbf{b}$, there is a \mathbf{c} where $\mathbf{a} <_{\mathrm{T}} \mathbf{c} <_{\mathrm{T}} \mathbf{b}$. This means the r.e. degrees are *dense* as an order.
- 2. Hence no r.e. degree is $<_{\rm T}$ -minimal in the non-computable degrees.
- 3. Moreover, there is then a solution to Post's problem: there is a turing degree between 0 and 0'.
- 4. For any r.e. degree $\mathbf{a} >_T \mathbf{0}$, there is an r.e. degree \mathbf{b} such that \mathbf{a} and \mathbf{b} are not \leq_T -comparable: $\mathbf{a} \not\leq_T \mathbf{b}$ and $\mathbf{b} \not\leq_T \mathbf{a}$.
- 5. (Sacks' Splitting Theorem) For any r.e. degree $\mathbf{a} >_{T} \mathbf{0}$, there are r.e. degrees $\mathbf{b}, \mathbf{c} <_{T} \mathbf{a}$ such that $\mathbf{a} = \mathbf{b} \oplus \mathbf{c}$.
- 6. There are r.e. degrees with no greatest lower bound, i.e. r.e. degrees **a**, **b** such that $\forall \mathbf{c} \leq_T \mathbf{a}, \mathbf{b} \exists \mathbf{d} (\mathbf{c} \leq_T \mathbf{d} \leq_T \mathbf{a}, \mathbf{b})$. This is why the turing degrees form an upper semilattice rather than an actual lattice.

It's also of interest that there is a solution to Post's problem with regard to many-to-one reducibility as well: a set X that is not Δ_1^0 nor \leq_m -complete in Σ_1^0 . While the proofs of these results are not of interest to us, the resulting picture of the turing degrees *is*, which is mostly just that their structure is extremely complicated in a precise way we will not pursue here. Some examples of the rich structure are the following.

– A4c•10. Result

- 1. For **a** a turing degree, **a** is the jump of some other degree $\mathbf{a} = \mathbf{b}'$ iff $\mathbf{a} \ge_T \mathbf{0}'$.
- 2. The map $\mathbf{a} \mapsto \mathbf{a}'$ is not injective: in fact, for any \mathbf{a} , there is a $\mathbf{b} >_{\mathrm{T}} \mathbf{a}$ with $\mathbf{a}' = \mathbf{b}'$.
- 3. \leq_{T} is ill-founded. In fact, we can choose a sequence $\langle \mathbf{a}_{\mathbf{n}} : \mathbf{n} < ! \rangle$ not only such that $\mathbf{a}_{\mathbf{n+1}} <_{\mathrm{T}} \mathbf{a}_{\mathbf{n}}$ but also that $\mathbf{a}'_{\mathbf{n+1}} <_{\mathrm{T}} \mathbf{a}_{\mathbf{n}}$ for all $n < \omega$.
- 4. Every countable poset can be embedded into the turing degrees ordered under \leq_{T} .

^{xvii}This comes from the terminology of a set being "recursively enumerable", which is the same as being Σ_1^0 .

For the purposes of set theory, we have the following interesting result about the turing degrees, coming from the results about computability and the natural numbers. One might worry that some inner models have a distinct $\mathcal{P}(\omega)$ —and this is certainly possible—but this doesn't pose an issue, since these new subsets yield only new turing degrees.

— A4c•11. Result

Being a turing degree is absolute between models of ZF - P.

Proof Sketch .:.

For a given function f, being a COM({f, s, pd})-program is absolute as it is just a statement about natural numbers. This implies downward absoluteness: if **a** is a turing degree in $W \supseteq U$ with $W, U \models ZF - P$, then every set Y in $\mathbf{a} = [\mathbf{X}]_{\equiv_{\mathrm{T}}}^{\mathsf{W}}$ is given by such a program in W. Since such a program is just a natural number, it follows that it's in U, and since the transition system of computation is absolute, it follows that this set Y is also computed by this program and is thus in $[X]_{\equiv_{\mathrm{T}}}^{\mathsf{W}}$. Hence $[X]_{\equiv_{\mathrm{T}}}^{\mathsf{W}} = [X]_{\equiv_{\mathrm{T}}}^{\mathsf{U}}$.

Now we show upward absoluteness. The idea here is that if $W \supseteq U$ has a larger $\mathcal{P}(\omega)$, then those new subsets have different turing degrees. To show this, again just note that the turing degree of U, $\mathbf{a} = [\mathbf{X}]_{\equiv_T}^U$, will be closed under the computations of all COM($\{\chi_X, s, pd\}$)-programs coded by natural numbers. But since W and U agree on what these programs are—being natural numbers—and they also agree on how these programs are computed, it follows that every subset of W X-computable is also in U. The only point of ZF - P is just to have enough set theory to carry out the required operations on ω .

Appendix B. Recursion Theory*

The point of this appendix is to consider further the hyperarithmetical hierarchy, and more generally the interaction of computability with set theoretic concepts, especially ordinals and ideas from descriptive set theory. As stated in the previous appendix, what is currently called "computability theory" was previously called "recursion theory". Many old-guard set theorists still refer to the field as recursion theory, giving "recursion theory" a more set theoretic tone. This is why I'm referring to the topic in this way. Regardless of the label, what will be presented here will be some of the basics of a field sometimes referred to as recursion theory, higher recursion theory, (to a lesser extent) effective descriptive set theory, or something else looking at computability with set theory. To make the distinction less meaningless, I will often use the label "recursive" instead of "computable" here, in line with the fact that computable functions over ω were once called recursive functions.

The goal of this appendix is to do four main things:

- 1. define the hyperarithmetical hierarchy;
- 2. show two characterizations of $\omega_1^{\rm CK}$ are equivalent;
- 3. show the length of the hyperarithmetical hierarchy is ω_1^{CK} ; and
- 4. investigate admissible sets and the theory KP's relation to recursion theory.

Doing any of these requires introducing many more new definitions and technology. (3) is just to say that for every $\alpha < \omega_1^{\text{CK}}, \Delta_{\alpha}^0 \subsetneq \Sigma_{\alpha}^0 \subsetneq \Delta_{\alpha+1}^0$, and that $\bigcup_{\xi < \alpha} \Delta_{\xi}^0 \subsetneq \Delta_{\alpha}^0$, meaning that at both successor and at limit stages we add new sets.

It should also be noted that this appendix relies on Appendix A in addition to the first several sections of [MISSING "Chapter III. Descriptive Set Theory"], particularly Section 24. So recall some notation from Appendix A which will be used *extensively*:

- $\llbracket e \rrbracket^n$ refers to the partial function from ω^n to ω computed by the (code of the) program e. $\llbracket e \rrbracket^n_x$ refers to this where x is used as an oracle in the (code of the) program e. We often leave the superscript out leaving the number of arguments implicit.
- As we refer to partial functions, we say f(x) [≜] g(x) iff x ∉ dom(f) ∪ dom(g), or else x ∈ dom(f) ∩ dom(g) ∧ f(g) = g(x). In other words, f(x) [≜] g(x) iff both are undefined, or both are defined and equal.

Note that the map $\langle e, n \rangle \mapsto [\![e]\!](n)$ is computable as a function from ω^2 to ω . Note that every computable map over \mathcal{N} is of the form $x \mapsto [\![e]\!]_x$ for some $e \in \omega$.

Section B1. The Hyperarithmetical Hierarchy

Everything we state in this section can be generalized to other polish spaces with great effort and onerous notation. To help with readability, this extra weight is dropped, and we state things only for \mathcal{N} , leaving the generalizations to the reader, and often using the unproven generalizations to products of copies of \mathcal{N} and ω . That said, the study of the hyperarithmetical hierarchy for ω , i.e. the study of hyperarithmetic reals, is itself an important topic although not one we will investigate too much here.

Recall that we have defined Σ_n^0 for $n < \omega$. To go further than Σ_α^0 for $\alpha = 1$, especially at limit stages, we need a more substantial coding. In essence, $x \in \mathcal{N}$ witnessing $X \in \Sigma_1^0(x)$ is supposed to code the construction of $X \in \Sigma_1^0$ from basic open sets. We can then easily code constructions of sets in $\Pi_1^0(x)$ and then in $\Sigma_2^0(x)$ and so on. But to make this intelligible at later stages, we want a uniform coding mechanism, which we call a borel code, which is a function that basically codes the construction of some $X \in \Sigma_\alpha^0$ from basic open sets.

B1 • 1. Definition

The set BC $\subseteq \mathcal{N}$ of *borel codes* is defined recursively in a hierarchy. Through simple coding, we may think of $x \in BC$ as an element of $\omega \times {}^{\omega}\mathcal{N}$, writing $x = \langle x(0) \in \omega \rangle^{\frown} \langle x_i \in \mathcal{N} : i < \omega \rangle$. For $\alpha \in Ord$,

- $x \text{ is } c\Sigma_1^0 \text{ iff } x(0) \notin \{0, 1\};$
- $x \text{ is } c \prod_{\alpha}^{0} \text{ iff } x(0) = 0 \text{ and } x_0 \in c \Sigma_{\alpha}^{0};$ $x \text{ is } c \Sigma_{\alpha}^{0} \text{ for } \alpha > 1 \text{ iff } x(0) = 1 \text{ and } x_i \in \bigcup_{\xi < \alpha} c \Sigma_{\xi}^{0} \cup c \prod_{\xi}^{0} \text{ for each } i < \omega.$

We set BC = $\bigcup_{\alpha < \omega_1} c\Sigma_{\alpha}^0 = \bigcup_{\alpha < \omega_1} c\Pi_{\alpha}^0$. To make things simpler, we also say x is $c\Pi_{\alpha}^0$ or $c\Sigma_{\alpha}^0$ if x is in $\bigcup_{\xi<\alpha} c\Sigma^0_{\xi} \cup c\Pi^0_{\xi}.$

These borel codes define borel sets in the following way. Note that this lines up with the notion of Σ_1^0 as before: a set X is Σ_1^0 iff there is some computable $x \in \mathcal{N}$ (viewed as the characteristic function of a computable set $A \subseteq \omega$) where $X = \bigcup_{\text{code}(\tau) \in A} \mathcal{N}_{\tau} = \bigcup_{x(\text{code}(\tau))=1} \mathcal{N}_{\tau}$. The following definition and corollary are the only places where we don't drop the added generality, simply to make it clear what the hyperarithmetical hierarchy is on products of \mathcal{N} and ω .

B1•2. Definition

Let \mathcal{M} be polish with basic open sets $\{\mathcal{M}_n : n < \omega\}$ and $A \in \mathcal{M}$. Let $x \in BC$. We define the borel set coded by $x = \langle x(0) \in \omega \rangle^{\frown} \langle x_i \in \mathcal{N} : i < \omega \rangle$, here $B_x^{\mathcal{M}}$, as follows.

- If x is $c\Sigma_1^0$, then $B_x^{\mathcal{M}} = \bigcup_{x \in \{n\}=1} \mathcal{M}_n$;
- If x is $c\Pi^0_{\alpha}$ and x(0) = 0, then $B^{\mathcal{M}}_x = \mathcal{M} \setminus B^{\mathcal{M}}_{x_0}$
- If x is $c\Sigma^0_{\alpha}$ for $\alpha > 1$ and x(0) = 1, then $B^{\mathcal{M}}_x = \bigcup_{i \in \alpha} B^{\mathcal{M}}_{x_i}$.

For $\mathcal{M} = \mathcal{N}$, we just write B_x for $B_x^{\mathcal{N}}$.

One can easily show by induction that these are borel, and in fact give all of the borel sets.

– B1•3. Corollary – For each $\alpha < \omega_1$ and polish $\mathcal{M}, \Sigma_{\alpha}^{0,\mathcal{M}} = \{B_x^{\mathcal{M}} : x \in c\Sigma_{\alpha}^0\}$, and similarly for $\prod_{\alpha=0}^{0}$ and $c\Pi_{\alpha}^0$. In particular, $\mathcal{B} = \{B_x : x \in \mathrm{BC}\}.$

Proof .:.

It should be clear that B_x is borel for every $x \in BC$ by induction. Fix an enumeration of basic open sets of \mathcal{M} , $\{\mathcal{M}_n : n < \omega\}$. It suffices to show that for every $\alpha < \omega_1, \sum_{\alpha}^{0,\mathcal{M}} = \{B_x^{\mathcal{M}} : x \in c\Sigma_{\alpha}^0\}$. Proceed by induction on α . The result on $\sum_{\alpha}^{0,\mathcal{M}}$ yields the result for $\prod_{\alpha}^{0,\mathcal{M}}$ since if $X \in \sum_{\alpha}^{0,\mathcal{M}}$ has borel code $x \in c\Sigma_{\alpha}^{0}$ then $\mathcal{M} \setminus X \in \prod_{\alpha}^{0,\mathcal{M}}$ has borel code $(0, x) \in c \prod_{\alpha}^{0}$

As noted above, this holds for $\alpha = 1$: clearly every $x \in c\Sigma_1^0$ has $B_x^{\mathcal{M}}$ as open, and every open set is $\bigcup_{n \in A} \mathcal{M}_n$ for some $A \subseteq \omega$ and thus is coded by $x = \langle 2 \rangle^{\frown} \langle A : i < \omega \rangle \in c \Sigma_1^0$.

For $\alpha > 1$, let $x \in c\Sigma_{\alpha}^{0}$. If $x(0) \neq 1$ then $x \in c\Sigma_{\xi}^{0} \cup c\Pi_{\xi}^{0}$ for some $\xi < \alpha$, where we may then appeal to the inductive hypothesis and the containments of the borel hierarchy. So we may assume x(0) = 1 and $x_{i} \in \bigcup_{\xi < \alpha} c\Sigma_{\xi}^{0} \cup c\Pi_{\xi}^{0}$ for each $i < \omega$. Inductively, each $B_{x_{i}}^{\mathcal{M}} \in \bigcup_{\xi < \alpha} \Sigma_{\xi}^{0,\mathcal{M}} \cup \Pi_{\xi}^{0,\mathcal{M}}$ and thus by properties of the borel hierarchy, each $B_{x_{i}} \in \Sigma_{\alpha}^{0,\mathcal{M}}$. Note that $B_{x}^{\mathcal{M}} = \bigcup_{i < \omega} B_{x_{i}}^{\mathcal{M}}$ is then in $\Sigma_{\alpha}^{0,\mathcal{M}}$ as this pointclass is closed under countable unions. This shows $\{B_{x}^{\mathcal{M}} : x \in c\Sigma_{\alpha}^{0}\} \subseteq \Sigma_{\alpha}^{0,\mathcal{M}}$.

Now suppose $X \in \sum_{\alpha}^{0,\mathcal{M}}$ as witnesses by $X = \bigcup_{n < \omega} X_n$ where for each $n < \omega, X_n \in \bigcup_{\xi < \alpha} \prod_{\xi_n}^{0,\mathcal{M}}$ for some $\xi_n < \alpha$. Inductively, each $X_n = B_{x_n}^{\mathcal{M}}$ for some code $x_n \in c \prod_{\xi_n}^{0,\mathcal{M}}$. Therefore $\langle 1 \rangle^{\frown} \langle x_n : n < \omega \rangle \in c \Sigma_{\alpha}^{0}$ is a borel code for X. This shows $\sum_{\alpha}^{0,\mathcal{M}} \subseteq \{B_x^{\mathcal{M}} : x \in c\Sigma_{\alpha}^0\}$ and hence we have equality for all α . \dashv

So whereas \sum_{α}^{0} -sets are coded by $c \sum_{\alpha}^{0}$ codes, we can form $\sum_{\alpha}^{0}(A)$ by restricting ourselves to only A-computable codes. This allows us to form the lightface borel hierarchy and its relativizations as follows.

— B1•4. Definition

For $A \in \mathcal{N}$, the (*relativized*, or A-) *hyperarithmetical* hierarchy, also known as the *lightface borel* hierarchy, consists of sets $\{B_x^{\mathcal{M}} : x \in BC \text{ is } A\text{-computable}\}$. This forms a (relativized) hierarchy as follows: for $X \subseteq \mathcal{N}$, and $\alpha < \omega_1$,

- X is $\Sigma^0_{\alpha}(A)$ iff X has an A-computable borel code $x \in c\Sigma^0_{\alpha}$;
- X is $\Pi^0_{\alpha}(A)$ iff X has an A-computable borel code $x \in c\Pi^0_{\alpha}$; and
- X is $\Delta^0_{\alpha}(A)$ iff X is $\Sigma^0_{\alpha}(A)$ and $\Pi^0_{\alpha}(A)$.

The hyperarithmetical hierarchy is this hierarchy for $A = \emptyset$.

This definition gives another hierarchy mirroring the argyle picture with the borel hierarchy in Figure 22 A • 3. But this time the length of the hierarchy is shortened significantly: to a countable ordinal $\omega_1^{CK} < \omega_1$. This ordinal will be defined later, but it suffices to think of it as the supremum of the ordinals reached by computable operations.



B1 • 5. Figure: The hyperarithmetical hierarchy

We immediately get the containments from the fact that $c\Sigma_{\alpha}^{0} \cup c\Pi_{\alpha}^{0} \subseteq c\Sigma_{\beta}^{0}$ for $\alpha < \beta$, and similarly for $c\Pi_{\beta}^{0}$. We can also re-characterize these pointclasses.

- **B1**•6. Corollary – Let
$$A \in \mathcal{N}$$
. Therefore,

1. For $\beta < \alpha$,

$$\Delta^{0}_{\beta}(A) \subseteq \Sigma^{0}_{\beta}(A) \subseteq \Sigma^{0}_{\alpha}(A) \subseteq \Delta^{0}_{\alpha+1}(A), \text{ and } \Delta^{0}_{\theta}(A) \subseteq \Pi^{0}_{\theta}(A) \subseteq \Pi^{0}_{\theta}(A) \subseteq \Delta^{0}_{\alpha+1}(A).$$

2. X is $\Sigma^0_{\alpha}(A)$ iff $\mathcal{N} \setminus X$ is $\Pi^0_{\alpha}(A)$; and

3. For $\alpha > 1$, $X \in \Sigma^0_{\alpha}(A)$ iff X is the A-computable union of sets in $\bigcup_{\xi < \alpha} \Pi^0_{\xi}(A)$, meaning $X = \bigcup_{n < \omega} B_{x_n}$ for some A-computable sequence $\langle x_n \in BC : n < \omega \rangle$.

Proof .:.

We only prove the result for $A = \emptyset$, but the idea easily generalizes. All of the following easily hold for $\alpha = 1$, so we assume $\alpha > 1$.

- 1. This is obvious from the fact that $\Sigma^0_\beta(A) \cup \Pi^0_\beta(A) \subseteq \Delta^0_\alpha(A)$ for all $\beta \leq \alpha$ by Definition B1 1.
- 2. If *X* has a computable borel code $x \in c\Sigma_{\alpha}^{0}$, then $\mathcal{N} \setminus X$ has a computable borel code $\langle 0, x \rangle \in c\Pi_{\alpha}^{0}$. Similarly if $\mathcal{N} \setminus X \in \Pi_{\alpha}^{0}$, then there is some computable borel code for $\mathcal{N} \setminus X$ in $c\Pi_{\alpha}^{0}$, $x = \langle x(0) \rangle^{\frown} \langle x_{i} : i < \omega \rangle \in \omega \times {}^{\omega} \mathcal{N}$.
 - If $x(0) \notin \{0, 1\}$, then $\mathcal{N} \setminus X \in \Sigma_1^0$ so that $\mathcal{N} \setminus (\mathcal{N} \setminus X) = X \in \Pi_1^0 \subseteq \Sigma_0^0$
 - If x(0) = 0, then $X = \mathcal{N} \setminus (\mathcal{N} \setminus X) = B_{x_0}$ with $x_0 \in c\Pi^0_{\xi} \cup c\Sigma^0_{\xi} \subseteq c\Sigma^0_{\alpha}$ for some $\xi < \alpha$;
 - If x(0) = 1, then $x \in c\Pi_{\xi}^{0} \cup c\Sigma_{\xi}^{0}$ for some $\xi < \alpha$. So inductively, $\mathcal{N} \setminus X \in \Sigma_{\xi}^{0} \cup \Pi_{\xi}^{0}$ implies $X \in \Sigma_{\xi}^{0} \cup \Pi_{\xi}^{0} \subseteq \Sigma_{\alpha}^{0}$.
- 3. Σ^0_{α} consists of sets whose borel codes are computable and either
 - a. in $c\Sigma_{\xi}^{0}$ or in $c\Pi_{\xi}^{0}$ for some $\xi < \alpha$; or
 - b. of the form $x = \langle 1 \rangle^{\frown} \langle x_i \in \bigcup_{\xi < \alpha} c\Sigma^0_{\xi} \cup c\Pi^0_{\xi} : i < \omega \rangle$.

For (a), if $X \in \Sigma_{\xi}^{0}$ for some $\xi < \alpha$, then inductively X is the computable union of sets. If $X \in \Pi_{\xi}^{0}$ for some $\xi < \alpha$, then there is some computable borel code $x \in c\Pi_{\xi}^{0}$ for X. In this case $X = \bigcup_{n < \omega} B_x$. For

(b), each B_{x_i} by the reasoning just given can be represented as a computable union (possibly just as a single set) of sets in $\bigcup_{\xi < \alpha} \prod_{\xi}^{0}$. In other words, write $x_n = \langle 1 \rangle^{\frown} \langle x_i^n : i < \omega \rangle$ so that $x = \langle 1 \rangle^{\frown} \langle x_{k_1}^{k_0} : k < \omega \rangle$ is computable and in $c \Sigma_{\alpha}^{0}$ (we regard $k \in \omega$ as a coded pair $\langle k_0, k_1 \rangle$) with $B_x = \bigcup_{k < \omega} B_{x_{k_1}^{k_0}} = X$. \dashv

We also get some nice closure properties for these pointclasses, as one would expect. Mostly this just comes the fact that the constructions for the usual borel pointclasses can be done in a computable manner.

— B1•7. Result —

Let $1 \leq \alpha < \omega_1$ and $A \in \mathcal{N}$. Therefore

- 1. $\Sigma^0_{\alpha}(A)$ is closed under A-computable unions, and finite intersections;
- 2. $\Pi^0_{\alpha}(A)$ is closed under finite unions, and A-computable intersections;
- 3. $\Delta^0_{\alpha}(A)$ is closed under finite unions, finite intersections, and complements.

Proof .:.

We again only prove the result for $A = \emptyset$ for the sake of notation.

1. If $\{X_n : n < \omega\} \subseteq \Sigma_{\alpha}^0$, then without loss of generality, each X_n can be realized as the union of sets in $\bigcup_{\xi < \alpha} \Sigma_{\xi}^0 \cup \Pi_{\xi}^0$. So let $X_n = \bigcup_{m < \omega} X_{n,m}$ where for each $n < \omega$, $\{X_{n,m} : m < \omega\} \subseteq \bigcup_{\xi < \alpha} \Sigma_{\xi}^0 \cup \Pi_{\xi}^0$. Let $x_{n,m} \in \bigcup_{\xi < \alpha} c\Sigma_{\xi}^0 \cup c\Pi_{\xi}^0$ be a computable borel code for $X_{n,m}$. As we can decode coded pairs in a computable way, $\langle 1 \rangle^{\frown} \langle x_{k_0,k_1} : k < \omega \rangle \in c\Sigma_{\alpha}^0$ is a computable borel code for $\bigcup_{k < \omega} X_{k_0,k_1} = \bigcup_{n,m < \omega} X_{n,m} = \bigcup_{n < \omega} X_n \in \Sigma_{\alpha}^0$. Hence Σ_{α}^0 is closed under computable unions.

For finite intersections, suppose X is the computable union $\bigcup_{n < \omega} X_n \in \Sigma_{\alpha}^0$ and Y is the computable union $\bigcup_{n < \omega} Y_n \in \Sigma_{\alpha}^0$ and Y is the computable union $\bigcup_{n < \omega} Y_n \in \Sigma_{\alpha}^0$ where $X_n, Y_n \in \bigcup_{\xi < \alpha} \Sigma_{\xi}^0 \cup \Pi_{\xi}^0$ for each $n < \omega$. Note that $X \cap Y = \bigcup_{n,m < \omega} X_m \cap Y_n$. For each m, n, there is some $\xi < \alpha$ with $X_m, Y_n \in \Sigma_{\xi}^0 \cup \Pi_{\xi}^0$ where then $X_m \cap Y_n \in \Sigma_{\xi}^0 \cup \Pi_{\xi}^0$ inductively. It follows that $\bigcup_{n,m < \omega} X_m \cap Y_n = X \cap Y$, and this is clearly a computable union so that $X \cap Y \in \Sigma_{\alpha}^0$.

- 2. Suppose $\{X_n : n < \omega\} \subseteq \Pi^0_{\alpha}$ with x_n a computable borel code for X_n for each $n < \omega$. If $\langle x_n : n < \omega \rangle$ is computable, then the set of codes of the complements is computable: $\langle \langle 0, x_n \rangle : n < \omega \rangle$. Since $\mathcal{N} \setminus X_n \in \Sigma^0_{\alpha}$ and there is a computable set of codes for these, by (1) the computable union $\bigcup_{n < \omega} \mathcal{N} \setminus X_n \in \Sigma^0_{\alpha}$ and therefore the complement $\mathcal{N} \setminus \bigcup_{n < \omega} \mathcal{N} \setminus X_n = \bigcap_{n < \omega} X_n \in \Pi^0_{\alpha}$. Hence Π^0_{α} is closed under computable intersections. For finite unions, suppose $X, Y \in \Pi^0_{\alpha}$. Therefore, $(\mathcal{N} \setminus X) \cap (\mathcal{N} \setminus Y) \in \Sigma^0_{\alpha}$ by (1). Hence the complement of this $X \cup Y \in \Pi^0_{\alpha}$.
- 3. Closure under finite unions and intersections follows just from the fact that both Σ^0_{α} and Π^0_{α} are closed under these. For complements, $X \in \Delta^0_{\alpha}$ implies $X \in \Sigma^0_{\alpha}$ so that $\mathcal{N} \setminus X \in \Pi^0_{\alpha}$. But we also have $X \in \Pi^0_{\alpha}$, implying $\mathcal{N} \setminus X \in \Sigma^0_{\alpha}$. Hence $X, \mathcal{N} \setminus X \in \Sigma^0_{\alpha} \cap \Pi^0_{\alpha} = \Delta^0_{\alpha}$.

As with the borel pointclasses, we also get closure under continuous preimages restricted in the sense of Lemma $24 \text{ A} \cdot 10$: closure under *computable* preimages, even from other spaces.

— B1•8. Result

Let \mathcal{M} and \mathcal{W} be topologies with basic open sets $\{\mathcal{M}_n : n < \omega\}$ and $\{\mathcal{W}_n : n < \omega\}$ respectively. Let $f : \mathcal{M} \to \mathcal{W}$ be such that f is A-computable for some $A \in \mathcal{N}$ in the sense that

$$\mathrm{NG}_f = \{ \langle x, n \rangle \in \mathcal{M} \times \omega : f(x) \in \mathcal{W}_n \} \in \Sigma_1^{0, \mathcal{M} \times \omega}(A).$$

Therefore for any $\alpha > 0$, $X \in \Sigma^{0,\mathcal{W}}_{\alpha}(A)$ implies $f^{-1} X \in \Sigma^{0,\mathcal{M}}_{\alpha}(A)$, and similarly for $\Pi^{0,\mathcal{W}}_{\alpha}(A)$, $\Pi^{0,\mathcal{M}}_{\alpha}(A)$ and $\Delta^{0,\mathcal{W}}_{\alpha}(A)$, $\Delta^{0,\mathcal{M}}_{\alpha}(A)$.

Proof .:.

Work with $A = \emptyset$. Firstly, we need a way of translating between the basic open sets in a computable way. We don't need to worry about empty preimages, since trivially $\emptyset \in \Sigma_{\alpha}^{0,\mathcal{M}}(A)$.

THE HYPERARITHMETICAL HIERARCHY

f

 \dashv

- Claim 1 If f is computable, then there is a computable function $h: \omega \times \omega \to \omega$ such that $f^{-1} W_n = \bigcup_{m < \omega} \mathcal{M}_{h(n,m)}$ if $f^{-1} W_n \neq \emptyset$.

Proof .:.

Since f is computable, $NG_f \in \Sigma_1^{0,\mathcal{M}\times\omega}$ meaning for some computable $p: \omega \to \omega$ (where we regard $p(n) = \operatorname{code}(p_0(n), p_1(n))$), $NG_f = \bigcup_{n < \omega} \mathcal{M}_{p_0(n)} \times \{p_1(n)\}$. So $f(x) \in \mathcal{W}_n$ iff there is some $m < \omega$ such that $x \in \mathcal{M}_{p_0(m)}$ and $n = p_1(m)$. So set

$$h(n,m) = \begin{cases} p_0(m) & \text{if } p_1(m) = n \\ p_0(m') & \text{for some fixed } m' \text{ with } p_1(m') = n \text{ otherwise.} \end{cases}$$

Then $f(x) \in \mathcal{W}_n$ iff $x \in \bigcup_{m < \omega} \mathcal{M}_{h(n,m)}$.

This provides a basis to define a function transforming codes. Let $X \in \Sigma_1^{0, W}$ have $f^{-1} X \neq \emptyset$. Therefore $X = \bigcup_{n < \omega} W_{k(n)}$ for some computable $k : \omega \to \omega$ where then

$${}^{-1}"X = \bigcup_{n < \omega} f^{-1}"\mathcal{W}_{k(n)} = \bigcup_{n < \omega} \bigcup_{m < \omega} \mathcal{M}_{h(k(n),m)} = \bigcup_{n < \omega} \mathcal{M}_{h(k(n_0),n_1)} \in \Sigma_1^{0,\mathcal{M}}$$

This establishes the result for $\alpha = 1$, and we proceed by induction from here. Note that complements work nicely with preimages, yielding the result for $\Pi_1^{0,\mathcal{W}}$ and $\Pi_1^{0,\mathcal{M}}$ as well. Going beyond $\Sigma_1^{0,\mathcal{W}}$, we need to translate between more complex borel codes.

— Claim 2 -

There is a computable g where for $x \in BC$, $B_{g(x)}^{\mathcal{M}} = f^{-1} {}^{"}B_{x}^{\mathcal{W}}$.

Proof .:.

Proceed by induction on β for $x \in c\Sigma_{\beta}^{0} \cup c\Pi_{\beta}^{0}$ to define g(x). The inductive definition is immediate: take $g(x) = \langle x(0) \rangle^{\frown} \langle g(x_{n}) : n < \omega \rangle$ for $x \notin c\Sigma_{\beta}^{0}$, preserving the operations and merely applying them to the hereditarily transformed codes. So suppose $x \in c\Sigma_{\beta}^{0}$ (i.e. suppose $x(0) \ge 2$) then $B_{x}^{W} = \bigcup_{x_{0}(n)=1} W_{n}$ where $x = \langle x(0) \rangle^{\frown} \langle x_{n} : n < \omega \rangle$, in which case

$$f^{-1} \, {}^{\mathsf{w}} B_x^{\mathfrak{w}} = \bigcup_{x_0(n)=1} f^{-1} \, {}^{\mathsf{w}} \mathcal{W}_n = \bigcup_{x_0(n)=1} \bigcup_{m < \omega} \mathcal{M}_{h(n,m)}$$

for the *h* from Claim 1. So taking $g(x) = \langle x(0) \rangle^{\frown} \langle h(x_n, n) : n < \omega \rangle$ completes the definition of *g*, which is clearly computable:

$$g(\langle x(0) \rangle^{\frown} \langle x_n : n < \omega \rangle) = \begin{cases} \langle x(0) \rangle^{\frown} \langle h(x_n, n) : n < \omega \rangle & \text{if } x(0) \ge 2, \\ \langle x(0) \rangle^{\frown} \langle g(x_n) : n < \omega \rangle & \text{otherwise.} \end{cases}$$

Note that a simple induction shows $g(x) \in c\Sigma_{\beta}^{0}$ if $x \in c\Sigma_{\beta}^{0}$ for all β , and similarly for the other levels of the hierarchy. In particular, for $x \in c\Sigma_{\alpha}^{0}$, $f^{-1}{}^{"}B_{x}^{\mathcal{W}} = B_{g(x)}^{\mathcal{M}}$. Note that $g(x) \in \mathcal{N}$ is still computable because x is and $x \mapsto g(x)$ is computable over \mathcal{N} . Hence $B_{g(x)}^{\mathcal{M}} \in \Sigma_{\alpha}^{0,\mathcal{M}}$ with a computable code $g(x) \in c\Sigma_{\beta}^{0}$, i.e. $B_{g(x)}^{\mathcal{M}} = f^{-1}{}^{"}B_{x}^{\mathcal{W}} \in \Sigma_{\alpha}^{0}$.

In particular, if we are able to code pairs in a computable way, then there's no worry about conflating the hierarchies on \mathcal{N} and on $\mathcal{N} \times \mathcal{N}$, for example, as we just consider computable preimages. More succinctly, this gives the following for \mathcal{N} .

- B1 • 9. Corollary
Let
$$1 \le \alpha < \omega_1$$
 and $A \in \mathcal{N}$. Therefore $\Sigma^0_{\alpha}(A)$, $\Pi^0_{\alpha}(A)$, and $\Delta^0_{\alpha}(A)$ are closed under A-computable preimages.

For the most part we will be concerned with the first few pointclasses of the hyperarithmetical hierarchy, mostly Σ_1^0 and

 Π_1^0 . So it's useful to have a couple theorems about these pointclasses, mostly being the computable analogs of results around Σ_1^0 and Π_1^0 . These pointclasses will be important for us as they form the basis of the analytical hierarchy, which is arguably much more important than the hyperarithmetical hierarchy, especially because the family of hyperarithmetic sets, $\Sigma_{\omega_{CK}}^0$, is exactly the set Δ_1^1 .

Section B2. Ordinals and Computability

To determine how long the hyperarithmetical hierarchy can go, we need to think about how far computation can take us, and this requires thinking about what ordinals are computable in the following sense.

An ordinal α is *recursive* iff there is some computable $R \subseteq \omega \times \omega$ such that R has order type α , i.e. $\langle \alpha, \in \rangle \cong \langle \operatorname{dom}(R) \cup \operatorname{ran}(R), R \rangle$. The *church–kleene ordinal* $\omega_1^{\operatorname{CK}}$ is sup{ $\alpha \in \operatorname{Ord} : \alpha$ is recursive}.

It's not immediately obvious that $\alpha < \beta$ with β recursive implies α is recursive, but this is the case and the basis for the following result.

– B2•2. Corollary –

 $\omega_1^{CK} < \omega_1$, and moreover every $\alpha < \omega_1^{CK}$ is recursive, but ω_1^{CK} isn't recursive.

Proof .:.

To see that $\omega_1^{CK} < \omega_1$, just note that there are only countably many computable subsets of $\omega \times \omega$. Clearly every recursive ordinal is countable as $\langle \alpha, \in \rangle \cong \langle X, R \rangle$ for some $X \subseteq \omega$ implies $|\alpha| \leq \aleph_0$. In particular, ω_1^{CK} is the supremum of countably many countable ordinals and is thus countable.

Now suppose β is recursive. It suffices to show any $\alpha < \beta$ is recursive. If $\langle \beta, \epsilon \rangle \cong \langle X, R \rangle$ with *R* computable, then any $\alpha < \beta$ is isomorphic to an initial segment of *R*. In particular, for some $n_{\alpha} < \omega$, $\langle \alpha, \epsilon \rangle$ is isomorphic to $\langle \operatorname{dom}(P) \cup \operatorname{ran}(P), P \rangle$ where n P m iff $n R m \wedge n R n_{\alpha}$. As $n_{\alpha} = 0 + 1 + \dots + 1$ is some fixed number, *P* is computable, and therefore α is recursive. Hence the set of recursive ordinals is closed downward, and therefore $\alpha < \sup\{\beta : \beta \text{ is recursive}\} = \omega_1^{CK}$ implies α is recursive.

Now suppose ω_1^{CK} is recursive. As the supremum of recursive ordinals, this implies $\omega_1^{CK} + 1$ isn't recursive, but this makes no sense: if α is recursive, so is $\alpha + 1$ as follows. If $\langle \alpha, \epsilon \rangle \cong \langle X, R \rangle$ then consider (to free up space) $R' = \{\langle 2n, 2m \rangle : \langle n, m \rangle \in R\}$. Then say n P m iff $n R' m \lor m = 1$, taking R and then adding a point at the end. This clearly has $\langle \alpha + 1, \epsilon \rangle \cong \langle \operatorname{dom}(P) \cup \operatorname{ran}(P), P \rangle$ with P computable. Thus ω_1^{CK} is recursive implies $\omega_1^{CK} + 1$ is too, a contradiction.

It is this sense that ω_1^{CK} is analogous to ω_1 for recursion. And in a similar way, we can define $\omega_1^{CK}(A)$ as the supremum of *A*-recursive ordinals and get that $\omega_1 = \sup_{A \in \mathcal{N}} \omega_1^{CK}(A)$.ⁱ

It's useful to develop codes for such ordinals along the same lines as the borel codes. This gives representations for these $\alpha < \omega_1^{CK}$ as well as a means of carrying out effective definitions and inductions.

The benefit of this will be to give an alternative characterization of the hyperarithmetical sets analogous to Post's theorem for the arithmetical hierarchy on ω . Recall that Post's theorem says $\Sigma_{n+1}^{0,\omega} = \Sigma_1^{0,\omega}(\emptyset^{(n)})$ where $\emptyset^{(n)} \in \mathcal{N}$ is the *n*th jump of \emptyset , defined recursively by $\emptyset^{(0)} = \emptyset$ and

 $\emptyset^{(n+1)} = \{e \in \omega : e \text{ is a (code of an) } \emptyset^{(n)} \text{-computable program that halts on input } e\}.$

If we are to have characterization for $\Sigma_{\alpha}^{0,\omega}$ for $\alpha \geq \omega$, we need another way to define $\emptyset^{(\alpha)}$, especially for limit α .

[–] B2•1. Definition

ⁱTo see this, for $\alpha < \omega_1$, for $f : \omega \to \alpha$ a bijection, $R = \{\langle a, b \rangle : f(a) < f(b)\}$ has $\langle \omega, R \rangle \cong \langle \alpha, \epsilon \rangle$ where then α is *R*-recursive.

Appx §B2

– B2•3. Definition

The set of *ordinal notations*, also called *Kleene's* \mathcal{O} , is a set of natural numbers ordered by $<_{\mathcal{O}}$, both built up by recursion:

- (base case) $0 \in \mathcal{O}$;
- (successor case) if $n \in \mathcal{O}$ then $\operatorname{code}(0, n) \in \mathcal{O}$ and $\forall m \leq_{\mathcal{O}} n \ (m \leq_{\mathcal{O}} \operatorname{code}(0, n));$
- (limit case) if e computes $\operatorname{im}(\llbracket e \rrbracket) = \{n_k : k < \omega\} \subseteq \mathcal{O} \text{ and } \forall k < \omega \ (n_k <_{\mathcal{O}} n_{k+1}), \text{ then } \operatorname{code}(1, e) \in \mathcal{O} \text{ and } \forall k < \omega \ \forall m \leq_{\mathcal{O}} n_k (m \leq_{\mathcal{O}} \operatorname{code}(1, e)).$

The codes for ordinals can be decoded as follows, similar to how borel codes can be decoded.

B2.4. Definition Let $n \in \mathcal{O}$. Therefore the ordinal coded by *n* is v_n where

- (base case) $v_0 = 0$;
- (successor case) $v_{\text{code}(0,n)} = v_n + 1$; and
- (limit case) if $\operatorname{code}(1, e) \in \mathcal{O}$ with $\llbracket e \rrbracket(k) = n_k$ for $k < \omega$, then $\nu_{\operatorname{code}(1, e)} = \sup_{k < \omega} \nu_{n_k}$.
- We call an ordinal α notated iff there is an $n \in \mathcal{O}$ with $\alpha = \nu_n$.

We will use these to define $\emptyset^{(\alpha)}$ for $\alpha < \omega_1^{CK}$, but first we should show that these give precisely the recursive ordinals. The idea is that ordinal notations give an alternative characterization of ω_1^{CK} , mostly just by giving names to recursive ordinals. This has the affect of trivially showing ω_1^{CK} is countable just because there are only countably many notations for $\alpha < \omega_1^{CK}$. But of course, to do this, we need to show that this really does give an alternative characterization of ω_1^{CK} , which is the main goal of this section.

B2.5. Theorem $\omega_1^{CK} = \{v_n : n \in \mathcal{O}\}$. That is to say, all recursive ordinals are notated, and all notated ordinals are recursive.

Note that it's rather difficult to directly compare whether $n, m \in O$ are codes for the same ordinal. This is mostly just due to the fact that at limit stages, there are many possible ways to take the supremum, and showing the ordinals are cofinal in the other ordinals just via their codes isn't computable generally. In this sense, it's more appropriate to view O as having a tree-like order given by \leq_O , which can be seen as the tree of how these ordinal notations were built up.

It should be obvious that if $n \leq_{\mathcal{O}} m$ then $v_n \leq v_m$, but the converse need not happen: $\leq_{\mathcal{O}}$ is not linear. But the following tells us we can still can regard $\leq_{\mathcal{O}}$ as a tree order, and thus get that every branch is well-ordered.

− **B2**•6. Theorem $\leq_{\mathcal{O}}$ is a well-founded partial ordering. Moreover, if $n \in \mathcal{O}$, then $\leq_{\mathcal{O}} \upharpoonright$ pred $_{\leq_{\mathcal{O}}}(n)$ is linear.

Proof .:.

It should be clear that $\leq_{\mathcal{O}}$ is well-founded as it's defined by structural recursion. Alternatively, one may note that $n \leq_{\mathcal{O}} m$ as ordinal notations implies $n \leq m$ as natural numbers. That $\leq_{\mathcal{O}}$ is a partial ordering should be clear with transitivity built into the definition.

To see that $\leq_{\mathcal{O}} \upharpoonright \operatorname{pred}_{\leq_{\mathcal{O}}}(n)$ is linear, proceed by induction on $\leq_{\mathcal{O}}$ to show that $\leq_{\mathcal{O}}$ is total: every $m_0, m_1 \leq_{\mathcal{O}}$ is comparable, i.e. has $m_0 \leq_{\mathcal{O}} m_1, m_1 \leq_{\mathcal{O}} m_0$, or $m_1 = m_0$. If n = 0, then this requires $m_1 = m_0 = n = 0$. If $n = \operatorname{code}(0, m)$ and $i \in \{0, 1\}$, then $m_i \leq_{\mathcal{O}} n$ is equivalent to $m_i = n \lor m_i \leq_{\mathcal{O}} m$. Obviously $m_i = n$ gives the result, and for $m_0, m_1 \leq_{\mathcal{O}} m$, the inductive hypothesis then gives that m_0 and m_1 are $\leq_{\mathcal{O}}$ -comparable.

If n = code(1, e) where $[\![e]\!](k) = n_k$, then $m_0, m_1 \le n$ implies $m_0 = n_{k_0}$ and $m_1 = n_{k_1}$ for some $k_0 \ne k_1 < \omega$. But then by Definition B2 • 3, $k_0 < k_1$ implies $m_0 \le \omega m_1$, and similarly for $k_1 < k_0$.

In fact, we can pretty easily calculate the order type of $(\operatorname{pred}_{\leq o}(n), \leq o)$ for any $n \in \mathcal{O}$: v_n .

```
- B2•7. Result 
Let n \in \mathcal{O}. Therefore \langle \operatorname{pred}_{\leq_{\mathcal{O}}}(n), \leq_{\mathcal{O}} \rangle \cong \langle \nu_n, \in \rangle.
```

ORDINALS AND COMPUTABILITY

Proof .:.

Proceed by induction on $\leq_{\mathcal{O}}$. Write α_n for the order type of $\langle \operatorname{pred}_{\leq_{\mathcal{O}}}(n), \leq_{\mathcal{O}} \rangle$. For n = 0, this is clear. For $n = \operatorname{code}(0, m)$, $\operatorname{pred}_{\leq_{\mathcal{O}}}(n) = \operatorname{pred}_{\leq_{\mathcal{O}}}(m) \cup \{m\}$ inductively has order-type $\alpha_n = \alpha_m + 1 = \nu_m + 1 = \nu_n + 1$. For $n = \operatorname{code}(1, e)$, since $\operatorname{pred}_{\leq_{\mathcal{O}}}(n) = \bigcup_{m < \omega} \operatorname{pred}_{\leq_{\mathcal{O}}}(\llbracket e \rrbracket(m))$, it follows that $\alpha_n = \sup_{m < \omega} \alpha_{\llbracket e \rrbracket(m)} = \sup_{m < \omega} \nu_{\llbracket e \rrbracket(m)} = \nu_n$.

Hence we can also chop off at initial segments to get another notated ordinal. In particular, $\{v_n : n \in O\}$ is closed downwards.

B2.8. Corollary Let $n \in \mathcal{O}$ with $\alpha < v_n$. Therefore there is some $m \leq_{\mathcal{O}} n$ with $\alpha = v_m$.

Proof .:.

 $\langle \alpha, \in \rangle$ is an initial segment of $\langle \nu_n, \in \rangle \cong \langle \operatorname{pred}_{\leq_{\mathcal{O}}}(n), \leq_{\mathcal{O}} \rangle$. In particular, for $f : \nu_n \to \operatorname{pred}_{\leq_{\mathcal{O}}}(n)$ the isomorphism, $\nu_{f(\alpha)} = \alpha$.

This fact will be very useful as well-founded relations admit a kind of *effective* (often used to mean a generalization of *computable* or explicit) version of transfinite recursion.

§ B2 a. Effective transfinite recursion

If we are going to talk about relations and functions on ordinals through their notations, we need a better way of defining these things. The way we would normally do this in set theory is just to define them by transfinite recursion. But it's not clear that this will be *effective* in the sense that we can deal with this in a computable way through notations.

B2 a • 1. Theorem (Effective Transfinite Recursion) Let \leq be a wellfounded relation over $X = \operatorname{dom}(\leq) \cup \operatorname{ran}(\leq) \subseteq \omega$ (writing $x \prec y$ for $x \leq y \land x \neq y$). Let $f: \omega \to \omega$ be a (total) computable function. Suppose for all $e \in \omega$ and $x \in X$, $\operatorname{pred}_{\prec}(x) \subseteq \operatorname{dom}(\llbracket e \rrbracket) \to x \in \operatorname{dom}(\llbracket f(e) \rrbracket)$. Therefore, for some $e_0 \in \omega$, $\llbracket e_0 \rrbracket = \llbracket f(e_0) \rrbracket$. Moreover, $X \subseteq \operatorname{dom}(\llbracket f(e_0) \rrbracket)$.

Proof .:.

By the second recursion theorem, there is an e_0 such that $\llbracket e_0 \rrbracket = \llbracket f(e_0) \rrbracket$. Now if $X \not\subseteq \operatorname{dom}(\llbracket e_0 \rrbracket)$, then there is some \preccurlyeq -minimal $x \in X \setminus \operatorname{dom}(\llbracket e_0 \rrbracket)$ which then has $\operatorname{pred}_{\prec}(x) \subseteq \operatorname{dom}(\llbracket e_0 \rrbracket)$ and thus $x \in \operatorname{dom}(\llbracket f(e_0) \rrbracket) = \operatorname{dom}(\llbracket e_0 \rrbracket)$, a contradiction.

Note that \leq need not be computable, as in the case of $\leq_{\mathcal{O}}$. A simple example of using effective transfinite recursion is defining addition over ordinal notation: a function $+_{\mathcal{O}} : \omega \times \omega \to \omega$ that gives an ordinal notation to the sum of the ordinals that its two inputs notate.

- B2 a • 2. Example There is a computable $+_{\mathcal{O}} : \omega \times \omega \rightharpoonup \omega$ such that • $\mathcal{O} \times \mathcal{O} \subseteq \operatorname{dom}(p)$ with $p''\mathcal{O} \times \mathcal{O} \subseteq \mathcal{O}$; and • $v_{p(n,m)} = v_n + v_m$ for all $n, m \in \mathcal{O}$.

Proof .:.

By the S_n^m -theorem, there is an injective, total, computable s where for all e, e', x, y,

 $[s(e, e', x)](y) \stackrel{\circ}{=} [e](x, [e'](y)).$

So now we define $+_{\mathcal{O}}$ by effective transfinite recursion. For any $e, n, m \in \omega$, consider the function P defined by

$$P(e, n, m) \stackrel{\circ}{=} \begin{cases} n & \text{if } m = 0\\ \operatorname{code}(0, \llbracket e \rrbracket(n, m')) & \text{if } m = \operatorname{code}(0, m')\\ \operatorname{code}(1, s(e, e', n)) & \text{if } m = \operatorname{code}(1, e')\\ \operatorname{undefined} & \text{otherwise.} \end{cases}$$

As a computable function, $P = \llbracket p \rrbracket$ for some $p < \omega$ and therefore by the S_n^m -theorem, there is a (total) computable $f : \omega \to \omega$ such that for each $e \in \omega$, $\llbracket f(e) \rrbracket (n,m) = P(e,n,m)$.

 $\leq_{\mathcal{O}}$ is well-founded, and clearly pred $_{\leq_{\mathcal{O}}}(x) \in \text{dom}(\llbracket e \rrbracket)$ implies $x \in \text{dom}(\llbracket f(e) \rrbracket)$ as we only need to continually decode the things that resemble ordinal notations building up to x. Thus by Effective Transfinite Recursion (B2 a • 1), there is an $e_0 \in \omega$ with $\llbracket e_0 \rrbracket = \llbracket f(e_0) \rrbracket$. We then define $+_{\mathcal{O}} = \llbracket e_0 \rrbracket$ and write $n +_{\mathcal{O}} m$ for $+_{\mathcal{O}}(n, m)$.

We get by an easy induction on *m* that

- $n + \mathcal{O} 0 = n;$
- $n + \mathcal{O} \operatorname{code}(0, m) = \operatorname{code}(0, n + \mathcal{O} m)$; and

• $n +_{\mathcal{O}} \operatorname{code}(1, m) = \operatorname{code}(1, e')$ where m and e' compute m_k and $n +_{\mathcal{O}} m_k$ for $k < \omega$ respectively. And therefore, assuming $+_{\mathcal{O}} "\mathcal{O} \times \mathcal{O} \subseteq \mathcal{O}, +_{\mathcal{O}}$ satisfies, as desired,

- $v_{n+\varrho_0} = v_n + v_0;$
- $v_{n+ocode(0,m)} = v_{n+ocom} + 1$; and
- $v_{n+\rho \operatorname{code}(1,m)} = \sup_{k < \omega} v_{n+\rho m_k}$ where $[m](k) = m_k$.

Now we show that if $n, m \in \mathcal{O}$, $n +_{\mathcal{O}} m \in \mathcal{O}$. This is the result of induction: m = 0 clearly yields the result. $m = \operatorname{code}(0, m')$ yields $n +_{\mathcal{O}} m = \operatorname{code}(0, n +_{\mathcal{O}} m') \in \mathcal{O}$ inductively. The following claim gives the limit case.

– Claim 1 –

Let $n, m, m' \in \mathcal{O}$ with $m_0 \leq_{\mathcal{O}} m'$ and assume $n +_{\mathcal{O}} m, n +_{\mathcal{O}} m' \in \mathcal{O}$. Therefore $n +_{\mathcal{O}} m \leq_{\mathcal{O}} n +_{\mathcal{O}} m'$.

Proof .:.

Without loss of generality, we assume $m <_{\mathcal{O}} m'$, as the result is obvious if m = m'. Proceed by induction on m'. For m' = 0, this is immediate. For $m' = \text{code}(0, m^*)$, inductively, $n +_{\mathcal{O}} m^* \in \mathcal{O}$ with $m \leq_{\mathcal{O}} m'$ and therefore

 $n +_{\mathcal{O}} m \leq_{\mathcal{O}} n +_{\mathcal{O}} m^* <_{\mathcal{O}} \operatorname{code}(0, n +_{\mathcal{O}} m^*) = n +_{\mathcal{O}} m'.$

For $m' = \operatorname{code}(1, e)$ where $\llbracket e \rrbracket(k) = m'_k$ with $m'_k \leq m'_{k+1}$ for $k < \omega$. $m <_{\mathcal{O}} m'$ then requires $m \leq_{\mathcal{O}} m'_k$ for some $k < \omega$ and therefore inductively,

$$n + {}_{\mathcal{O}} m \leqslant_{\mathcal{O}} n + {}_{\mathcal{O}} m'_k \leqslant_{\mathcal{O}} n + {}_{\mathcal{O}} m'. \qquad \qquad \dashv$$

This shows the limit case of $n, m \in \mathcal{O} \to n + \mathcal{O}m \in \mathcal{O}$ as follows. If $m = \text{code}(1, e_m)$ and $n + \mathcal{O}m = \text{code}(1, e_+)$ (here e_m and e_+ compute m_k and $(n + \mathcal{O}m)_k$ for $k < \omega$ respectively), then $m_k \leq m_{k+1}$ implies by Claim 1 that

$$(n + {\mathcal{O}} m)_k = n + {\mathcal{O}} m_k \leq_{{\mathcal{O}}} n + {\mathcal{O}} m_{k+1} = (n + {\mathcal{O}} m)_{k+1}.$$

Thus $n + \mathcal{O} m \in \mathcal{O}$ and therefore $+ \mathcal{O} " \mathcal{O} \times \mathcal{O} \subseteq \mathcal{O}$.

There are a variety of useful and expected properties of $+_{\mathcal{O}}$ interacting with $\leq_{\mathcal{O}}$. For example, consider the following.

- B2 a • 3. Corollary For all $n, m \in \omega$, 1. $n, m \in \mathcal{O}$ iff $n +_{\mathcal{O}} m \in \mathcal{O}$; 2. $n, m \in \mathcal{O}$ and $m \neq 0$ implies $n <_{\mathcal{O}} n +_{\mathcal{O}} m$; 3. $n, m, m' \in \mathcal{O}$ and $m' <_{\mathcal{O}} m$ iff $n +_{\mathcal{O}} m' <_{\mathcal{O}} n +_{\mathcal{O}} m$; 4. $n, m, m' \in \mathcal{O}$ and m = m' iff $n +_{\mathcal{O}} m = n +_{\mathcal{O}} m'$. \neg

Proof .:.

We show each direction simultaneously by induction, meaning that the inductive hypothesis is that all of the above \rightarrow directions hold, or the inductive hypothesis is that all of the \leftarrow directions hold.

- 1. It suffices to show the \leftarrow direction: that $n + \mathcal{O} m \in \mathcal{O}$ implies $n, m \in \mathcal{O}$. Proceed by induction on $n + \mathcal{O} m$.
 - If $n +_{\mathcal{O}} m = n$ then $m = 0 \in \mathcal{O}$ and $n = n +_{\mathcal{O}} m \in \mathcal{O}$.
 - If $n + \mathcal{O} m = \operatorname{code}(0, k)$ for some $k \in \mathcal{O}$, then $m = \operatorname{code}(0, m')$ for some m' and thus inductively, $k = n + \mathcal{O} m'$ implies $n, m' \in \mathcal{O}$ and so $m' \in \mathcal{O}$.
 - If $n +_{\mathcal{O}} m = \operatorname{code}(1, e)$ for some $e \in \omega$ with $\llbracket e \rrbracket(k) <_{\mathcal{O}} \llbracket e \rrbracket(k+1)$ for all $k \in \omega$, then $m = \operatorname{code}(1, e')$ for some e' where then $n +_{\mathcal{O}} \llbracket e' \rrbracket(k) \in \mathcal{O}$ for each $k \in \omega$ implies $\llbracket e' \rrbracket(k) \in \mathcal{O}$ inductively.

$$[\![e]\!](k) = n +_{\mathcal{O}} [\![e']\!](k) <_{\mathcal{O}} n +_{\mathcal{O}} [\![e']\!](k+1) = [\![e]\!](k+1)$$

implies that $\llbracket e' \rrbracket(k) <_{\mathcal{O}} \llbracket e' \rrbracket(k+1)$ by (3) for all $k < \omega$. Hence $m = \operatorname{code}(1, e') \in \mathcal{O}$.

2. Suppose $n, m \in \mathcal{O}$ with $m \neq 0$. We will show $n <_{\mathcal{O}} n +_{\mathcal{O}} m$. Proceed by induction on $m \neq 0$. For m = code(0, m'), we have inductively $n \leq_{\mathcal{O}} n +_{\mathcal{O}} m' <_{\mathcal{O}} n +_{\mathcal{O}} m$. For m a limit, inductively

 $n \leq_{\mathcal{O}} n +_{\mathcal{O}} m_k <_{\mathcal{O}} n +_{\mathcal{O}} m_{k+1} \leq_{\mathcal{O}} n +_{\mathcal{O}} m$

where m = code(1, e) with $\llbracket e \rrbracket(k) = m_k$ for $k < \omega$.

3. Suppose $n, m', m \in \mathcal{O}$ with $m' <_{\mathcal{O}} m$. Proceed by induction on m to show $n +_{\mathcal{O}} m' <_{\mathcal{O}} n +_{\mathcal{O}} m$.

- For $m = \operatorname{code}(0, m^*)$, this is clear as $n + \mathcal{O} m' \leq \mathcal{O} n + \mathcal{O} m^* < \mathcal{O} n + \mathcal{O} \operatorname{code}(0, n + \mathcal{O} m^*) = n + \mathcal{O} m$.
- For m = code(1, e) where $[\![e]\!](k) = m_k$ for $k < \omega$, we have $m' <_{\mathcal{O}} m_k$ for some $k < \omega$ where then inductively, $n +_{\mathcal{O}} m' <_{\mathcal{O}} n +_{\mathcal{O}} m_k \leq_{\mathcal{O}} n +_{\mathcal{O}} m$.

For the converse, suppose $n +_{\mathcal{O}} m' <_{\mathcal{O}} n +_{\mathcal{O}} m$. Proceed by induction on $n +_{\mathcal{O}} m$. Note $m \neq 0$ as otherwise $n \leq_{\mathcal{O}} n +_{\mathcal{O}} m' \leq_{\mathcal{O}} n +_{\mathcal{O}} 0 = n$ results in $n +_{\mathcal{O}} m' = n +_{\mathcal{O}} m$. So clearly $n +_{\mathcal{O}} m \neq 0$.

- If $n +_{\mathcal{O}} m = \operatorname{code}(0, n +_{\mathcal{O}} m^*)$ then $m = \operatorname{code}(0, m^*)$ where then inductively $m' \leq_{\mathcal{O}} m^* <_{\mathcal{O}} m$.
- If $n +_{\mathcal{O}} m = \operatorname{code}(1, e)$ then $\llbracket e \rrbracket(k) = n +_{\mathcal{O}} \llbracket e' \rrbracket(k)$ where $m' = \operatorname{code}(1, e)$. So $n +_{\mathcal{O}} m' <_{\mathcal{O}} n +_{\mathcal{O}} \llbracket e' \rrbracket(k)$ for some $k < \omega$ where then inductively $m' <_{\mathcal{O}} \llbracket e' \rrbracket(k) \leq_{\mathcal{O}} m$.
- 4. It suffices to show the \leftarrow direction: that $n +_{\mathcal{O}} m = n +_{\mathcal{O}} m'$ implies m = m'. We have by (3) that $n +_{\mathcal{O}} m \leq_{\mathcal{O}} n +_{\mathcal{O}} m'$ implies $m \leq_{\mathcal{O}} m'$ and $n +_{\mathcal{O}} m' \leq_{\mathcal{O}} n +_{\mathcal{O}} m$ implies $m' \leq_{\mathcal{O}} m$ so that m' = m. \dashv

– B2a•4. Corollary

Let $s : \omega \to \omega$ be computable such that [[s(e)]] is total for all $e \in \omega$. Therefore, there is a computable function (written with the notation $e \mapsto \sum_{\mathcal{O}, k < \omega} [[s(e)]](k)$) with the following properties:

- If $\sum_{\mathcal{O}, k < \omega} \llbracket s(e) \rrbracket(k) \in \mathcal{O}$, then $\llbracket s(e) \rrbracket(k) \in \mathcal{O}$ for all $k < \omega$;
- If $0 <_{\mathcal{O}} [[s(e)]](k)$ for all $k < \omega$, then $\sum_{\mathcal{O}, k < \omega} [[s(e)]](k) \in \mathcal{O}$.

• If $\sum_{\mathcal{O},k<\omega} [s(e)](k) \in \mathcal{O}$, then $\nu_{\sum_{\mathcal{O},k<\omega} [s(e)](k)} = \sum_{k<\omega} \nu_{[s(e)](k)}$.

Proof .:.

Define by recursion the partial sums:

$$S(e, 0) = [[s(e)]](0)$$

$$S(e, n + 1) = S(e, n) + o [[s(e)]](n + 1).$$

This definition makes sense since each $[\![s(e)]\!]$ is total. By the S_n^m -theorem, there is a (total) computable $f: \omega \to \omega$ such that for each $e < \omega$, $S(e, n) = [\![f(e)]\!](n)$ for all $n < \omega$. This $[\![f(e)]\!]$ is also total since $[\![s(e)]\!]$ is. So then we just set $\sum_{\mathcal{O}, k < \omega} [\![s(e)]\!](k)$ to be code(1, f(e)), and we now appeal to the results in Corollary B2 a \cdot 3.

• Suppose $\sum_{\mathcal{O},k<\omega} \llbracket s(e) \rrbracket(k) \in \mathcal{O}$. Thus $\llbracket f(e) \rrbracket(n) <_{\mathcal{O}} \llbracket f(e) \rrbracket(n+1) = \llbracket f(e) \rrbracket(n) +_{\mathcal{O}} \llbracket s(e) \rrbracket(n+1)$ implies $0 \leq_{\mathcal{O}} \llbracket s(e) \rrbracket(n+1) \in \mathcal{O}$.

- Similarly, if $\forall k < \omega$ ($[[s(e)]](k) \in \mathcal{O}$), then each $[[f(e)]](k) \in \mathcal{O}$ by an easy induction. Moreover, if each $[[s(e)]](k) \neq 0$, then $[[f(e)]](k) <_{\mathcal{O}} [[f(e)]](k) + [[s(e)]](k + 1) = [[f(e)]](k + 1)$ implies $\sum_{\mathcal{O},k<\omega} [[s(e)]](k) \in \mathcal{O}$.
- By an easy induction, $\nu_{\llbracket f(e) \rrbracket(N)} = \sum_{k < N} \nu_{\llbracket s(e) \rrbracket(k)}$ so that

$$\nu_{\sum_{\mathcal{O},k<\omega} \llbracket s(e)(k) \rrbracket} = \sup_{N<\omega} \nu_{\llbracket f(e) \rrbracket(n)} = \sum_{k<\omega} \nu_{\llbracket s(e) \rrbracket(k)}.$$

Another example of effective transfinite recursion is in showing $\operatorname{pred}_{\leq_{\mathcal{O}}}(n)$ is $\Sigma_1^{0,\omega}$ for each $n \in \mathcal{O}$ (although $\leq_{\mathcal{O}}$ is not $\Sigma_{\alpha}^{0,\omega}$ for any α). For the following, write W_e for dom($[\![e]\!]$).

– B2 a • 5. Example

There are computable functions predprgm : $\omega \to \omega$ (for "predecessor program") and orderprgm : $\omega \to \omega$ (for "order program") where for any $n \in \mathcal{O}$,

- 1. $W_{\text{predprgm}(n)} = \{m \in \mathcal{O} : m \leq_{\mathcal{O}} n\}; \text{ and }$
- 2. $W_{\text{orderprgm}(n)} = \{ \langle m, m' \rangle \in \mathcal{O} \times \mathcal{O} : m \leq_{\mathcal{O}} m' \leq_{\mathcal{O}} n \}.$

Proof .:.

1. It suffices to find a computable p such that for all $n, e \in \omega$, $W_{p(0)} = \emptyset$, and

$$W_{p(\operatorname{code}(0,n))} = W_{p(n)} \cup \{n\}$$

$$W_{p(\operatorname{code}(1,e))} = \bigcup_{k \in W_{1}} W_{p(\llbracket e \rrbracket(k))}$$

It's not difficult to see that such a p satisfies (1). To build such a p, we use Effective Transfinite Recursion (B2 a • 1). Let e_0 be such that $\llbracket e_0 \rrbracket = \emptyset$ so that $W_{e_0} = \emptyset$. To ease up notation, we say $W_{\llbracket e \rrbracket(n)} = \emptyset$ if $n \notin W_e$.

There are computable $s: \omega \times \omega \to \omega$ (for "successor") and $\ell: \omega \times \omega \to \omega$ (for "limit") where

$$W_{s(e,n)} = W_{[[e]](n)} \cup \{n\}, \text{ and } W_{\ell(e,e')} = \bigcup_{n \in \mathbb{N}} W_{[[e]]([[e']](n))}.$$

Proof .:.

Claim 1

Define $s': \omega^3 \rightharpoonup \omega$ and $\ell': \omega^3 \rightarrow \omega$ by

$$s'(e, n, m) \stackrel{\circ}{=} \begin{cases} 0 & \text{if } m = n \\ \llbracket\llbracket e \rrbracket(n) \rrbracket(m) & \text{otherwise.} \end{cases}$$

By the S_n^m -theorem, there is a (total) computable function $s : \omega^2 \to \omega$ where $[\![s(e, n)]\!](m) \stackrel{\circ}{=} s'(e, n, m)$ for all $e, n, m \in \omega$. It should be clear that $W_{s(e,n)} = W_{e(n)} \cup \{n\}$. Similarly, we can define the function $\ell'(e, e', n) = y$ iff $y = 0 \land \exists m_0, m_1 \in \omega$ ([[[e]] ([[e']] (n))]] $(m_0) = m_1$).

As ℓ' regarded as a relation is $\Sigma_1^{0,\omega}$, ℓ' as a function is computable. By the S_n^m -theorem, there is a (total) computable $\ell: \omega^2 \to \omega$ where $[\![\ell(e, e')]\!](n) = \ell'(e, e', n)$ for all $e, e', n \in \omega$.

We can make use of s and ℓ in defining predprgm through some p. Firstly, we must find our iteration operator f from Effective Transfinite Recursion (B2 a • 1). In particular, we set

 $g(e,n) = \begin{cases} e_0 & \text{if } n = 0\\ s(e,m) & \text{if } n = \text{code}(0,m)\\ \ell(e,e') & \text{if } n = \text{code}(1,e')\\ 0 & \text{otherwise.} \end{cases}$

This g is computable and in fact total as both s and ℓ are total. By the S_n^m -theorem, there is some (total) computable $f: \omega \to \omega$ where $[\![f(e)]\!](n) = g(e, n)$ for all $e, n < \omega$. This f will be as in Effective Transfinite Recursion (B2 a • 1). The hypothesis of the theorem is immediate because every $[\![f(e)]\!]$ is total since g is total. Therefore by Effective Transfinite Recursion (B2 a • 1), there is some e_1 where $[\![e_1]\!] = [\![f(e_1)]\!]$. This $p = [\![e_1]\!]$

then satisfies

$$W_{p(0)} = W_{e_0} = \emptyset$$

$$W_{p(\operatorname{code}(0,n))} = W_{s(e_1,n)} = W_{\llbracket e_1 \rrbracket(n)} \cup \{n\} = W_{p(n)} \cup \{n\}$$

$$W_{p(\operatorname{code}(1,e))} = W_{\ell(e_1,e)} = \bigcup_{n < \omega} W_{\llbracket e_1 \rrbracket(\llbracket e \rrbracket(n))} = \bigcup_{n < \omega} W_{p(\llbracket e \rrbracket(n))},$$

as desired. Hence p = predprgm witnesses the result.

2. This is easy to derive from (1). In particular, consider the function defined as a $\Sigma_1^{0,\omega}$ -relation

h(n, m, m') = y iff $y = 0 \land m \in W_{\text{predprgm}(n)} \land m' \in W_{\text{predprgm}(n)} \land m \in W_{\text{predprgm}(m')}$.

This yields by the S_n^m -theorem a (total) computable orderprgm : $\omega \to \omega$ where [[orderprgm(n)]](m, m') = h(n, m, m') for all $n, m, m' \in \omega$. In particular, $W_{orderprgm(n)} = \{\langle m, m' \rangle \in \mathcal{O} \times \mathcal{O} : m \leq_{\mathcal{O}} m' \leq_{\mathcal{O}} n\}$ for all $n \in \mathcal{O}$.

This allows us to show that all notated ordinals are recursive.

- B2 a • 6. Corollary
$$\{v_n : n \in \mathcal{O}\} \subseteq \omega_1^{CK}.$$

Proof .:.

Let $n \in \mathcal{O}$ be arbitrary. By Result B2 • 7 and Example B2 a • 5,

 $\langle v_n, \in \rangle \cong \langle \operatorname{pred}_{\leq_{\mathcal{O}}}(n), \leq_{\mathcal{O}} \rangle \cong \langle W_{\operatorname{predprgm}(n)}, W_{\operatorname{orderprgm}(n)} \rangle$

is a $\Sigma_1^{0,\omega}$ -well-ordering. If pred $_{\leq_{\mathcal{O}}}(n)$ is finite, then clearly ν_n is recursive. So assume pred $_{\leq_{\mathcal{O}}}(n)$ is infinite, and therefore there is some (total) computable bijection $f: \omega \to W_{\text{predprgm}(n)}$. Hence if we define

 $\langle m, m' \rangle \in R$ iff $\langle f(m), f(m') \rangle \in W_{\text{orderprgm}(n)}$,

we get that R is computable since f is total with $\operatorname{im}(f) \subseteq W_{\operatorname{predprgm}(n)} = \operatorname{dom}(W_{\operatorname{orderprgm}(n)}) \cup \operatorname{ran}(W_{\operatorname{orderprgm}(n)})$. So R is computable with $\langle \omega, R \rangle \cong \langle W_{\operatorname{predprgm}(n)}, W_{\operatorname{orderprgm}(n)} \rangle \cong \langle \nu_n, \epsilon \rangle$. Hence ν_n is recursive.

§ B2 b. From ordinals to notations

Now in showing all recursive ordinals are notated, we need a way of translating from notations to well-orders, or rather from notations to (codes of) programs computing well orders. There are a few steps in doing this. First we have the following technical lemma.

- B2 b • 1. Lemma There is a computable function $d : \omega \to \omega$ where for all $e \in \omega$, 1. $d(e) \in \mathcal{O}$ iff $W_e \subseteq \mathcal{O}$; 2. $W_e \subseteq \mathcal{O}$ implies $v_n < v_{d(e)}$ for all $n \in W_e$.

Proof ...

Consider the sums as defined in Corollary B2 a • 4. First we must enumerate the elements of $W_e \cup \{0\}$, and then we sum these codes.

- Claim 1 There is an $s: \omega \to \omega$ where

- $s: \omega \to \omega$ is total, and computable;
- [s(e)] is total for each $e < \omega$; and
- [s(e)] enumerates $W_e \cup \{0\}$ for each $e < \omega$.

Proof .:.

Define by recursion g(e, 0) = 0 $g(e, n + 1) = \begin{cases} y_1 & \text{for } y = \text{code}(y_0, y_1) \text{ the least } y \leq n \text{ where } \forall m \leq n \ (y_1 > g(e, m)) \land \llbracket e \rrbracket (y_0) = y_1, \\ 0 & \text{if there is no such } y \leq n. \end{cases}$ $g \text{ is computable, total, and } n \mapsto g(e, n) \text{ enumerates } W_e \cup \{0\}. \text{ By the } S_n^m \text{-theorem, there is an } s \text{ where } \\ \llbracket s(e) \rrbracket (n) = g(e)(n) \text{ for } e, n < \omega. \text{ This } s \text{ has the desired properties.} \qquad \neg$ We then define $d(e) = \sum_{0,n < \omega} \text{code}(0, \llbracket s(e) \rrbracket (n)) \text{ and use the properties of Corollary B2 a • 4.}$ 1. Suppose $d(e) \in \mathcal{O}$. So that $\text{code}(0, \llbracket s(e)(n) \rrbracket \text{ and thus } \llbracket s(e) \rrbracket (n) = n \text{ for some } m < \omega \text{ os } d(e) \in \mathcal{O}.$ 2. Suppose $W_e \subseteq \mathcal{O}$ and thus $d(e) \in \mathcal{O}$. Let $n \in W_e$ so that $\llbracket s(e) \rrbracket (m) = n \text{ for some } m < \omega. \text{ We then have } v_n < v_n + 1 = v \llbracket s(e) \rrbracket (m) + 1 \leq \sum_{k \leq m} (v \llbracket s(e) \rrbracket (k) + 1) \leq v_d(e). \qquad \neg$

The benefit of this lemma is proving the next result, connecting well-founded relations with notations that potentially extend them. First we identify $\Sigma_1^{0,\omega}$ -relations with their codes in the following sense: R_e is the relation defined by

$$R_e(x, y)$$
 iff $\langle x, y \rangle \in \operatorname{dom}(\llbracket e \rrbracket^2)$,

where $\llbracket e \rrbracket^2 : \omega^2 \to \omega$ is the computable partial function computed by $e \in \omega$. Note that for functions with one input, we still refer to W_e for dom($\llbracket e \rrbracket^1$).

– B2b•2. Lemma

There is a computable function relcode : $\omega \rightarrow \omega$ such that

1. R_e is well-founded iff relcode $(e) \in \mathcal{O}$;

2. R_e is well-founded implies the order-type of R_e is $\leq v_{\text{relcode}(e)}$.

Proof .:.

Generally, we want to overshoot and then take an initial segment. So it's useful to know that we can do this in a computable way with the codes.

— Claim 1

There is a (total) computable $i: \omega^2 \to \omega$ such that for any $e, c, R_{i(e,c)}$ is $R_e \upharpoonright \text{pred}_{R_e}(c)$. In other words, for any $e, c, n, m \in \omega, R_{i(e,c)}(n,m)$ iff $R_e(n,m) \land R_e(n,c) \land R_e(m,c)$.

Proof .:.

Define $I: \omega^4 \to \omega$ as follows and apply the S_n^m -theorem:

 $I(e, c, n, m) = \llbracket e \rrbracket(n, m) + \llbracket e \rrbracket(n, c) + \llbracket e \rrbracket(m, c)$

Thus there is a (total) computable $i: \omega^2 \to \omega$ where $[\![i(e,c)]\!](n,m) = I(e,c,n,m)$ for all $e,c,n,m < \omega$. In particular, $\langle n,m \rangle \in R_{i(e,c)}$ iff $\langle n,m \rangle \in \operatorname{dom}([\![i(e,c)]\!])$ iff $\langle n,m \rangle, \langle n,c \rangle, \langle m,c \rangle \in \operatorname{dom}([\![e]\!])$ iff $\langle n,m \rangle, \langle n,c \rangle, \langle m,c \rangle \in \operatorname{dom}([\![e]\!])$ iff $\langle n,m \rangle, \langle n,c \rangle, \langle m,c \rangle \in \operatorname{R}_e$, as desired.

In particular, $R_{i(e,c)} = \emptyset$ if $c \notin \text{dom}(R_e) \cup \text{ran}(R_e)$. Now we order the domain and ranges of these relations.

— Claim 2

There is a (total) computable $c : \omega^2 \to \omega$ such that for any $e, e', n \in \omega$,

$$W_{c(e,e')} = \begin{cases} \emptyset & \text{if } R_{e'} = \emptyset \\ \{ \llbracket e \rrbracket (i(e',n)) : n < \omega \} & \text{otherwise.} \end{cases}$$

Proof .:.

Consider the function C(e, e', m) defined to be the least $n = \operatorname{code}(n_0, n_1, n_2, n_3)$ where $\llbracket e' \rrbracket (n_0, n_1) = n_2$ and $m = \llbracket e \rrbracket (i(e', n_3))$. If there is no such n, C(e, e', m) is undefined. Note that C(e, e', m) is defined iff $R_{e'} \neq \emptyset$ and $m \in \{\llbracket e \rrbracket (i(e', n)) : n < \omega\}$. C is computable and hence by the S_n^m -theorem, there is some (total) computable $c : \omega^2 \to \omega$ where $\llbracket c(e, e') \rrbracket (m) = C(e, e', m)$ and therefore $W_{c(e,e')} = \emptyset$ if $R_{e'} = \emptyset$ and otherwise $W_{c(e,e')} = \{\llbracket e \rrbracket (i(e', n)) : n < \omega\}$.

Now we can use the function d from Lemma B2 b • 1 in conjunction with Effective Transfinite Recursion (B2 a • 1) as follows. By the Second Recursion Theorem, there is an e_0 where for all $e' \in \omega$, $[\![e_0]\!](e') = d(c(e_0, e'))$. Write relcode = $[\![e_0]\!]$ and $c'(e) = c(e_0, e)$ so that relcode(e) = d(c'(e)). It follows that

$$W_{c'(e)} = \begin{cases} \emptyset & \text{if } R_e = \emptyset \\ \{\text{relcode}(i(e, n)) : n < \omega\} & \text{otherwise.} \end{cases}$$

This finishes our definition of relcode, and now we must confirm the associated properties.

1. Suppose R_e is well-founded. If $R_e = \emptyset$, then clearly $W_{c'(e)} = \emptyset$ so that $relcode(e) = d(c'(e)) \in \mathcal{O}$ by Lemma B2 b • 1 (1).

Now suppose relcode $(e) = d(c'(e)) \in \mathcal{O}$, implying that $W_{c'(e)} \subseteq \mathcal{O}$, and therefore relcode $(i(e, n)) \in \mathcal{O}$ for all $n < \omega$. Moreover, by Lemma B2 b • 1 (2), $v_{\text{relcode}(i(e,n))} < v_{\text{relcode}(e)}$ for all $n < \omega$. By induction on the notations according to the ordinals they code, we may assume inductively that $R_{i(e,n)}$ is well-founded for all $n < \omega$. But then R_e must be well-founded, as otherwise, any ill-founded R_e -sequence lies below some $n \in \omega$ and thus $R_{i(e,n)}$ would be ill-founded; a contradiction.

2. Suppose R_e is well-founded. By (1), relcode $(e) = d'(c(e)) \in \mathcal{O}$ so that relcode $(i(e, n)) \in \mathcal{O}$ and $\nu_{\text{relcode}(i(e,n))} < \nu_{\text{relcode}(e)}$ for all $n < \omega$. So by induction on the notations according to the ordinals they code, the ordertype of $R_{i(e,n)} = R_e \upharpoonright \text{pred}_{R_e}(n)$ is at most $\nu_{\text{relcode}(i(e,n))}$ for each $n < \omega$. Therefore, the ordertype of R_e , being $\sup_{n < \omega}$ ordertype $(R_{i(e,n)})$, is at most $\sup_{n < \omega} \nu_{\text{relcode}(i(e,n))} \le \nu_{\text{relcode}(e)}$.

It turns out that we cannot improve the inequality to an equality in (2) above. While the order type of R_e will be an initial segment of the ordinal notated by relcode(e), we won't be able to figure out how and where to cut this in a uniform, computable way. Nevertheless, this does give the final result needed to prove Theorem B2 • 5: that notated ordinals are precisely the recursive ones.

– B2b•3. Corollary

Every recursive ordinal is notated. In particular, $\omega_1^{CK} \subseteq \{v_n : n \in \mathcal{O}\}$ and so $\omega_1^{CK} = \{v_n : n \in \mathcal{O}\}.$

Proof .:.

Let $\alpha < \omega_1^{CK}$ be arbitrary. Let $\langle \operatorname{dom}(R) \cup \operatorname{ran}(R), R \rangle \cong \langle \alpha, \in \rangle$ with $R \subseteq \omega^2$ computable. Say $R = R_e$ for some $e \in \omega$. Since R_e is well-founded, $\operatorname{relcode}(e) \in \mathcal{O}$ with $\alpha \leq \nu_{\operatorname{relcode}(e)}$ by Lemma B2 b • 2. Since $\{\nu_n : n \in \mathcal{O}\}$ is closed downwards by Corollary B2 • 8, $\alpha = \nu_n$ for some $n \leq_{\mathcal{O}} \operatorname{relcode}(e)$ and therefore $\omega_1^{CK} \subseteq \{\nu_m : m \in \mathcal{O}\}$. The other containment is Corollary B2 • 6.

Section B3. The Length of the Hyperarithmetical Hierarchy

Section B4. Identifying the Hyperarithmetical Hierarchy with Δ_1^1

Section B5. Admissible Sets and Kripke–Platek Set Theory

Admissible Sets and Kripke–Platek Set Theory

Appx §B5

Appendix C. Gory Details of the Forcing Relation

Section C1. Gory Detail of the Forcing Relation and Its Definability

It's not recommended to read any of this section. The results of it are useful, but the proofs are long, technical, and uninteresting. I will repeat for emphasis: *do not read this section if you do not have to*. This is mostly for the curious and the skeptical. First we show the definability of each relation $\{\langle p, \vec{\tau} \rangle \in \mathbb{P} \times \mathbb{V}^{\mathbb{P}} : p \Vdash ``\varphi(\vec{\tau})"\}$ for φ a formula. To do this, we just straight up define a relation \Vdash^* with all the properties we'd like it to have, and then we show that it is equivalent to \Vdash . The motivation behind the definition is a result about \Vdash . First, we have the following useful fact.

- C1 · 1. Definition

Let \mathbb{P} be a preorder with $p \in \mathbb{P}$. A set $D \subseteq \mathbb{P}$ is *dense below* p iff for every $q \leq p$, there is some $r \leq q$ with $r \in D$. Equivalently, D is dense below p iff $D \cup (\mathbb{P} \setminus \mathbb{P}_{\leq p})$ is dense in \mathbb{P} (here $\mathbb{P}_{\leq p} = \{q \in \mathbb{P} : q \leq p\}$).

— C1 · 2. Result

Let $V \vDash \mathsf{ZFC}$ be a transitive model. Let $\mathbb{P} \in V$ be appropriate for forcing. Suppose G be \mathbb{P} -generic over V with $p \in G$. Therefore, $G \cap D \neq \emptyset$ for every $D \subseteq \mathbb{P}$ that is dense below p.

Proof .:.

As \mathbb{P} is appropriate for forcing, we can always extend. In particular, a set D is dense iff $D \setminus \{q \in \mathbb{P} : p \leq^{\mathbb{P}} q\}$ is dense (removing an initial segment doesn't change long-term behavior of being able to extend into the set). Hence $D \cup (\mathbb{P} \setminus \mathbb{P}_{\leq p})$ is dense implies $D \cup (\mathbb{P} \setminus \{q \in \mathbb{P} : q \leq^{\mathbb{P}} p \lor p \leq^{\mathbb{P}} q\})$ is dense. Hence G has a non-empty intersection with this. So there is some $q \in G$ in D or else not comparable to p.

- C1 · 3. Motivation

Let $V \vDash \mathsf{ZFC}$ be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. Let φ be a formula. Therefore, the following are equivalent.

- $p \Vdash \varphi$;
- $\forall p^*$
- $D = \{p^* is dense below <math>p$.

The proof of this result will follow from the rest of our work in this section.

C1 • 4. Definition Let $\mathbb{P} = \langle \mathbb{P}, \leq \rangle$ be a preorder. We define $p \Vdash^* ``\varphi(\vec{\tau})$ ", read as p *-forces $``\varphi(\vec{\tau})$ ", by structural induction on φ and \mathbb{P} -name rank of $\vec{\tau}$. • $p \Vdash^* ``\tau_1 = \tau_2$ " iff for each $\langle \sigma_1, q_1 \rangle \in \tau_1$, D_1 is dense below p; and D_2 is too for each $\langle \sigma_2, q_2 \rangle \in \tau_2$; where $D_1 = \{p^* \leq p : p^* \leq q_1 \rightarrow \exists \langle \sigma, q \rangle \in \tau_2 (p^* \leq q \land p^* \Vdash^* ``\sigma = \sigma_1")\}$ $D_2 = \{p^* \leq p : p^* \leq q_2 \rightarrow \exists \langle \sigma, q \rangle \in \tau_1 (p^* \leq q \land p^* \Vdash^* ``\sigma = \sigma_2")\}.$ • $p \Vdash^* ``\tau_1 \in \tau_2"$ iff $\{p^* \leq p : \exists \langle \sigma, q \rangle \in \tau_2 (p^* \leq q \land q \Vdash^* ``\sigma = \tau_1")\}$ is dense below p. • $p \Vdash^* ``\varphi(\vec{\tau}) \land \psi(\vec{\tau})"$ iff $p \Vdash^* ``\varphi(\vec{\tau})"$ and $p \Vdash^* ``\psi(\vec{\tau})"$. • $p \Vdash^* ``\neg\varphi(\vec{\tau})"$ iff every $p^* \leq p$ has $p \not\vdash^* ``\varphi(\sigma, \vec{\tau})"\}$ is dense below p.

So we always have $p \Vdash^* \mathscr{O} = \emptyset$ just vacuously. So to confirm whether $p \Vdash^* \mathscr{O} = \{\langle \emptyset, p \rangle\} = \{\langle \emptyset, q \rangle\}$, we need to see

GORY DETAIL OF THE FORCING RELATION AND ITS DEFINABILITY

Appx §C1

whether (plugging $\sigma_1 = \emptyset = \sigma_2$ and $q_1 = p$, $q_2 = q$ into Definition C1 • 4)

$$D_1 = \{p^* \leq p : p^* \leq p \to (p^* \leq q \land p^* \Vdash ``\emptyset = \emptyset'')\}$$

= $\{p^* \leq p : p^* \leq q \land p^* \Vdash ``\emptyset = \emptyset''\}$
= $\{p^* \leq p : p^* \leq q\}$ is dense below p , and
$$D_2 = \{p^* \leq p : p^* \leq q \to (p^* \leq p \land p^* \Vdash ``\emptyset = \emptyset'')\}$$

= $\{p^* \leq p : p^* \in \mathbb{P}\}$ is dense below p .

Here, D_1 represents where $\tau_1 = \{\langle \emptyset, p \rangle\}$ will be a subset of $\tau_2 = \{\langle \emptyset, q \rangle\}$ according to p. Similarly, D_2 represents when τ_2 will be a subset of τ_1 according to p: always. So $p \Vdash^* (\tau_1 = \tau_2)$ iff (because D_2 is clearly dense below p) $\{p^* \leq p : p^* \leq q\}$ is dense below p. Note that with a \mathbb{P} -generic filter $G, p \in G$ implies $G \cap D_1 \neq \emptyset$, which implies $q \in G$ and thus $(\tau_1)_G = \{\emptyset\} = (\tau_2)_G$.

This gives some motivation that \Vdash^* is well-defined: because we're always decreasing \mathbb{P} -name rank in the atomic formulas, eventually we go down to the \mathbb{P} -name \emptyset where equality and membership are easy to calculate. Of course, this doesn't mean \Vdash^* is easy to calculate, as it can be unclear whether certain sets are dense or not.

Inductively, we have the following result about \Vdash^* (in V).

C1 • 5. Lemma Let $\mathbb{P} = \langle \mathbb{P}, \leq \rangle$ be appropriate for forcing. Let φ be a formula and $\vec{\tau}$ \mathbb{P} -names. Therefore, the following are equivalent. 1. $p \Vdash^* (\varphi(\vec{\tau}))$; 2. $\forall r^* \in \pi$ $(r^* \parallel * (\pi(\vec{\tau})))$;

∀p* ≤ p (p* ⊩* "φ(τ̃)");
 D = {p* ≤ p : p* ⊩* "φ(τ̃)"} is dense below p, i.e. D ∪ (P \ P≤p) is dense in P.

Proof .:.

- That (2) implies (1) is immediate since $p \leq p$. That (2) implies (3) is also immediate since then $D = \mathbb{P}_{\leq p}$.
- Suppose (1) holds, working towards (2). We proceed by structural induction on φ and the P-rank of τ.
 More precisely, let φ be the <_{lex}-least formula with some τ where (1) holds but (2) fails. Then we set τ to be witnesses of least P-rank. Let p* ≤ p be such that p* ⊮ "φ(τ)".
 - If " $\varphi(\vec{\tau})$ " is of the form " $\tau_1 = \tau_2$ ", " $\tau_1 \in \tau_2$ ", or " $\exists x \ \psi(x, \vec{\tau})$ ", then (2) follows easily, since those sets being dense below p implies they are dense below p*.
 - If " $\varphi(\vec{\tau})$ " is of the form " $\neg \psi(\vec{\tau})$ ", then every $p^* \leq p$ has $p^* \not\models^*$ " $\psi(\vec{\tau})$ ". In particular, for any $p^{**} \leq p^* \leq p, p^{**} \not\models^*$ " $\psi(\vec{\tau})$ " and therefore $p^* \Vdash^*$ " $\neg \psi(\vec{\tau})$ " by definition. So (2) holds.
 - If " $\varphi(\vec{\tau})$ " is of the form " $\theta(\vec{\tau}) \wedge \psi(\vec{\tau})$ ", then $p \Vdash^* "\theta(\vec{\tau})$ " and $p \Vdash^* "\psi(\vec{\tau})$ " so inductively every $p^* \leq p$ has $p^* \Vdash^* "\theta(\vec{\tau})$ " and $p^* \Vdash^* "\psi(\vec{\tau})$ " and so the conjunction is *-forced: $p^* \Vdash^* "\varphi(\vec{\tau})$ " so (2) holds.
- Suppose (3) holds, working towards (1). We again proceed by induction on φ and $\vec{\tau}$.
 - If " $\varphi(\vec{\tau})$ " is of the form " $\tau_1 = \tau_2$ ", " $\tau_1 \in \tau_2$ ", or " $\exists x \ \psi(x, \vec{\tau})$ ", then (2) follows easily just from properties of denseness. Explicitly, if $\{p^* \leq p : D' \text{ is dense below } p^*\}$ is dense below p, then D' is dense below p: we just extend twice. To show (1), we just need to notice that we don't need to restrict the dense set definitions of Definition C1 4 to extensions of p. Then we're working with the same dense sets for all elements and thus we can apply this observation.
 - If "φ(τ̃)" is of the form "¬ψ(τ̃)", then suppose (1) fails: there is some p* ≤ p with p* I⊢* "ψ(τ̃)". By density of D, there is a p** I⊢* "¬ψ(τ̃)", meaning p** I⊬ "ψ(τ̃)". But this contradicts that (1) implies (2) since p* *-forces it and p** ≤ p*. Hence (1) holds.
 - Suppose " $\varphi(\vec{\tau})$ " is of the form " $\theta(\vec{\tau}) \wedge \psi(\vec{\tau})$ ". Since

$$D = \{ p^* \leq p : p^* \Vdash^* ``\theta(\vec{\tau}) \land \psi(\vec{\tau}) ``\}$$

is dense below p, then in particular,

$$D \subseteq D_0 = \{ p^* \leqslant p : p^* \Vdash^* ``\theta(\vec{\tau})" \}$$
GORY DETAIL OF THE FORCING RELATION AND ITS DEFINABILITY

Appx §C1

$$D \subseteq D_1 = \{ p^* \leqslant p : p^* \Vdash^* ``\psi(\vec{\tau})" \}$$

are both dense below p. Inductively, then $p \Vdash^* "\theta(\vec{\tau})"$ and $p \Vdash^* "\psi(\vec{\tau})"$, so p *-forces the conjunction: $p^* \Vdash^* "\varphi(\vec{\tau})"$, meaning so (1) holds.

This allows us to show the analogues of Definition $31 D \cdot 8$ and Theorem $31 B \cdot 6$ for *-forcing. We unfortunately need to prove these simultaneously rather than focusing on just one or the other.

C1.6. Lemma Let $V \models \mathsf{ZFC}$ be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. Let $p \in \mathbb{P}$, $\vec{\tau} \in V^{\mathbb{P}}$, and G be \mathbb{P} -generic over V. Therefore

- (a) $p \Vdash^* "\varphi(\vec{\tau})"$ implies $p \Vdash "\varphi(\vec{\tau})"$; and
- (b) if $\mathbf{V}[G] \models "\varphi(\vec{\tau}_G)$ ", then there is some $p \in G$ with $p \Vdash^* "\varphi(\vec{\tau})$ ".

Proof .:.

As before, proceed by structural induction on φ and \mathbb{P} -rank of $\vec{\tau}$ (by which I mean induction on the max of the \mathbb{P} -ranks of $\vec{\tau}$). Let G be \mathbb{P} -generic over V with $p \in G$, $\mathbb{P} = \langle \mathbb{P}, \leq^{\mathbb{P}} \rangle$. First we show the result for atomic formulas, and then we induct on φ .

(a) Suppose $p \Vdash^* ``\varphi(\vec{\tau})$ ''. We must show $V[G] \vDash ``\varphi(\vec{\tau}_G)$ ''.

• Suppose " $\varphi(\vec{\tau})$ " is " $\tau_1 = \tau_2$ ". We shall show $(\tau_1)_G \subseteq (\tau_2)_G$ in V[G], because the other containment is similar. So let $(\sigma_1)_G \in (\tau_1)_G$ be arbitrary. Therefore, there is some $q_1 \in G$ with $\langle \sigma_1, q_1 \rangle \in \tau_1$. Since $p^* \Vdash "\tau_1 = \tau_2$ ", by Definition C1 • 4, the set

$$D_1 = \{ p^* \leq p : p^* \leq q_1 \to \exists \langle \sigma, q \rangle \in \tau_2 \ (p^* \leq q \land p^* \Vdash^* "\sigma = \sigma_1") \}$$

is dense below p. Hence, as G is generic, $G \cap D_1 \neq \emptyset$. Hence there is some p^* and $\langle \sigma, q \rangle \in \tau_2$ such that

1.
$$p^* \in G;$$

2. $p^* \leq p, q_1, q;$ and
3. $p^* \Vdash^* ``\sigma = \sigma_1``.$

(1) and (2) imply $q \in G$ and thus $\sigma_G \in (\tau_2)_G$. (3) implies by the inductive hypothesis that $p \Vdash "\sigma = \sigma_1"$ (we're looking at the same formula φ but now with parameters σ and σ_1 which have maximum \mathbb{P} -name rank less than the maximum \mathbb{P} -name rank of τ_2 and τ_1) and so $\mathbf{V}[G] \models$ " $(\sigma_1)_G = \sigma_G \in (\tau_2)_G$ ". As $(\sigma_1)_G$ was arbitrary, it follows that $\mathbf{V}[G] \models "(\tau_1)_G \subseteq (\tau_2)_G$ ". The other containment follows analogously.

• Suppose " $\varphi(\vec{\tau})$ " is " $\tau_1 \in \tau_2$ ". Since $p \Vdash^* "\tau_1 \in \tau_2$ ", the set

$$D = \{ p^* \leq p : \exists \langle \sigma, q \rangle \in \tau_2 \ (p^* \leq q \land q \Vdash^* ``\sigma = \tau_1") \}$$

is dense below p. In particular, $G \cap D \neq \emptyset$ and so there is a p^* and $\langle \sigma, q \rangle \in \tau_2$ such that

1. $p^* \in G$; 2. $p^* \leq p, q$; and 3. $p^* \Vdash^* ``\sigma = \tau_1$.

(1) and (2) imply that $q \in G$ so that $V[G] \models "\sigma_G \in (\tau_2)_G$ ". (3) implies by the previous case above that $V[G] \models "\sigma_G = (\tau_1)_G$ " and thus $V[G] \models "(\tau_1)_G \in (\tau_2)_G$ ".

- (b) Suppose $V[G] \models "\varphi(\vec{\tau})$ ". We must show there is some $p \in G$ with $p \Vdash^* "\varphi(\vec{\tau})$ ".
 - Suppose $V[G] \models "(\tau_1)_G = (\tau_2)_G$ ". To see that some $p \in G$ has $p \Vdash^* "\tau_1 = \tau_2$ ", it suffices to consider the dense set of Lemma C1•5. In particular, consider the set of all $p^* \leq p$ such that $p^* \Vdash^* "\tau_1 = \tau_2$ ". This isn't exactly easy to get a handle on, so instead consider the set *D* of *p* where $p \Vdash "\tau_1 = \tau_2$ " or we have a conflict with the dense sets of Definition C1•4: either
 - (i) there is a $\langle \sigma_1, q_1 \rangle \in \tau_1$ where $p \leq q_1$, and for every $\langle \sigma_2, q_2 \rangle \in \tau_2$ and every $q_2^* \leq q_2$, if $q_2^* \Vdash^* \sigma_1 = \sigma_2^*$ then $q_2^* \perp p$; or

GORY DETAIL OF THE FORCING RELATION AND ITS DEFINABILITY

(ii) there is a $\langle \sigma_2, q_2 \rangle \in \tau_2$ where $p \leq q_2$, and for every $\langle \sigma_1, q_2 \rangle \in \tau_1$ and every $q_1^* \leq q_1$, if $q_1^* \Vdash^* \sigma_1 = \sigma_2^*$, then $q_1^* \perp p$.

The idea is that there can be no $p \in G$ that satisfies either (i) or (ii). The issue is that (the handling of these two are analogous) if p satisfies (i) along with $\langle \sigma_1, q_1 \rangle \in \tau_1$, we would have $q_1 \in G$ so that $\mathbf{V}[G] \models "(\sigma_1)_G \in (\tau_1)_G = (\tau_2)_G$ ", meaning $\mathbf{V}[G] \models "(\sigma_1)_G = (\sigma_2)_G$ " for some $\langle \sigma_2, q_2 \rangle \in \tau_2$. By the inductive hypothesis on \mathbb{P} -name rank, there is then some $q \in G$ where $q \Vdash "\sigma_1 = \sigma_2$ ". Without loss of generality (G is a filter) we can assume $q \leq q_2$ so that by (i) $q \perp p$, contradicting that $q, p \in G$ and G is a filter.

Thus if there are no $p \Vdash^* "\tau_1 = \tau_2"$ in G, then $G \cap D = \emptyset$. So it suffices to show that D is dense, yielding that $G \cap D \neq \emptyset$ and thus there is a $p \in G$ that *-forces " $\tau_1 = \tau_2$ ". So let $p \in \mathbb{P}$ be arbitrary, working towards a $p^* \in D$. Assume without loss of generality that $p \not\Vdash^* "\tau_1 = \tau_2$ ". Thus by Definition C1 • 4 (the other possibility being similar, yielding an extension witnessing to (ii)) there is a $\langle \sigma_1, q_1 \rangle \in \tau_1$ with

 $D_1 = \{ p^* \leq p : p^* \leq q_1 \to \exists \langle \sigma_2, q_2 \rangle \in \tau_2 \ (p^* \leq q \land p^* \Vdash^* \sigma_1 = \sigma_2") \}$

not dense below p. In particular, there is some $p^* \leq p$ with no $p^{**} \in D_1$, meaning for all $p^{**} \leq p^*$, $p^{**} \leq q_1 \land \forall \langle \sigma_2, q_2 \rangle \in \tau_2 \ (p^{**} \notin q_2 \lor p^{**} \not\models^* ``\sigma_1 = \sigma_2``).$

We now show that p^* satisfies (i). The above shows that in particular, $p^* \leq p^*$ has $p^* \leq q_1$. Note that if $\langle \sigma_2, q_2 \rangle \in \tau_2$, $q_2^* \leq q_2$ and $q_2^* \Vdash^* \sigma_1 = \sigma_2$, then any common extension $r \leq q_2^*$, p^* has $r = p^{**}$ contradict the above. Thus this would imply $q_2 \perp p^*$ and thus that p^* satisfies (i). Therefore D is dense.

• Suppose $V[G] \vDash (\tau_1)_G \in (\tau_2)_G$. This means there is some $\langle \sigma_2, q_2 \rangle \in \tau_2$ with $q_2 \in G$ and $V[G] \vDash$ $(\tau_1)_G = (\sigma_2)_G$. By the argument above, there is then some $p \in G$ with $p \Vdash^* (\tau_1 = \sigma_2)$. So if $p^* \leq p, q_2$, by Lemma C1 • 5, every $p^{**} \leq p^*$ has $p^{**} \leq q_2$ and $p^{**} \Vdash^* (\tau_1 = \sigma_2)$. Thus by definition, $p^* \Vdash (\tau_1 \in \tau_2)$.

So far, we have (a) and (b) for atomic formulas. The inductive steps are much easier, and we prove (a) and (b) by structural induction on φ .

- Suppose φ is " $\theta \wedge \psi$ ".
 - (a) If $p \Vdash^* "\varphi(\vec{\tau})$ " then p *-forces each conjunct. By the inductive hypothesis, p forces each conjunct, and thus the conjunction " $\varphi(\vec{\tau})$ ".
 - (b) If V[G] ⊨ "φ(τ̃_G)" then inductively there are p₁, p₂ ∈ G with p₁ ⊩* "θ(τ̃)" and p₂ ⊩* "ψ(τ̃)". As G is a filter, there is some common extension p ≤ p₁, p₂ where then p *-forces both (by Lemma C1 • 5) and thus the conjunction "φ(τ̃)".
- Suppose φ is " $\neg \psi$ ". This case is the only reason why we needed to prove (a) and (b) together.
 - (a) Suppose p ⊢* "¬ψ(τ)" but some generic G has p ∈ G with V[G] ⊨ "ψ(τ)". By the inductive hypothesis on (b), there is some q ∈ G with q ⊢* "ψ(τ)". But then a common extension p* ≤ p, q has (by Lemma C1 5) p* ⊢* "ψ(τ)", contradicting Definition C1 4.
 - (b) Suppose $V[G] \models \neg \psi(\vec{\tau})$. Consider the set D of $p \in \mathbb{P}$ that decide ψ :

$$D = \{ p \in \mathbb{P} : p \Vdash^* ``\psi(\vec{\tau})`' \lor p \Vdash^* ``\neg\psi(\vec{\tau})`' \}.$$

It should be clear that *D* is dense in **V**, since either we can extend an arbitrary *p* to a $p^* \Vdash^* \psi(\vec{\tau})^*$ or else every $p^* \leq p$ doesn't *-force " $\psi(\vec{\tau})$ ", in which case $p \leq p$ *-forces " $\neg \psi(\vec{\tau})$ ". Hence $G \cap D \neq \emptyset$ as witnessed by some $p \in G$. We obviously can't have $p \Vdash^* \psi(\vec{\tau})^*$ as this would imply inductively that $p \Vdash \psi(\vec{\tau})^*$ and thus $\mathbf{V}[G] \models \psi(\vec{\tau})^*$. Hence $p \Vdash^* \neg \psi(\vec{\tau})^*$ witnesses the result.

- Suppose φ is " $\exists x \psi$ ".
 - (a) Suppose $p \Vdash^* \exists x \ \psi(x, \vec{\tau})$ " so that $\{p^* \leq p : \exists \sigma \in V^{\mathbb{P}} \ (p^* \Vdash^* "\psi(\sigma, \vec{\tau})")\}$ is dense below p. Thus if G is generic and $p \in G$, then there is some $p^* \leq p$ with $p^* \Vdash^* "\psi(\sigma, \vec{\tau})$ " for some \mathbb{P} -name σ . By the inductive hypothesis, $V[G] \models "\psi(\sigma_G, \vec{\tau}_G)$ " and thus $V[G] \models "\exists x \ \psi(x, \vec{\tau}_G)$ ". As G was arbitrary, it follows that $p \Vdash "\exists x \ \psi(x, \vec{\tau})$ ".

GORY DETAIL OF THE FORCING RELATION AND ITS DEFINABILITY

(b) Suppose $V[G] \models \exists x \ \psi(x, \vec{\tau}_G)$ ". Let σ_G be a witness to this. Thus $V[G] \models \psi(\sigma_G, \vec{\tau}_G)$ " and so inductively, there is some $p \in G$ with $p \Vdash^* \psi(\sigma, \vec{\tau})$ " and in particular, since every $p^* \leq p$ *-forces this, by Definition C1 • 4 $p \Vdash^* \exists x \ \psi(x, \vec{\tau})$ ".

This allows us to prove the desired results about the actual forcing relation.

C1.7. Corollary Let $V \models \mathsf{ZFC}$ be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. Let $p \in \mathbb{P}$, $\vec{\tau} \in V^{\mathbb{P}}$, and φ a formula. Therefore in V, $p \Vdash^* "\varphi(\vec{\tau})"$ iff $p \Vdash "\varphi(\vec{\tau})"$.

Proof .:.

The (\rightarrow) direction follows from Lemma C1 • 6 (a). For the (\leftarrow) direction, suppose $p \Vdash ``\varphi(\vec{\tau})"$, but $p \not\Vdash ``\varphi(\vec{\tau})"$. Hence by Lemma C1 • 5, $D = \{p^* \leq p : p^* \Vdash ``\varphi(\vec{\tau})"\}$ is not dense below p, meaning there is some $p^* \leq p$ that cannot be extended into D, i.e. every $p^{**} \leq p^*$ has $p^{**} \not\Vdash ``\varphi(\vec{\tau})"$, i.e. $p^* \Vdash ``\neg\varphi(\vec{\tau})"$. But then $p^* \Vdash ``\neg\varphi(\vec{\tau})"$, contradicting that $p \geq p^* \Vdash ``\varphi(\vec{\tau})"$.

The above proof requires some philosophical assumptions: namely that we can force over V with \mathbb{P} . This is clear when V is countable by Theorem 31 D • 1,ⁱ but otherwise, one could read the conclusion of the proof above that instead p^* has no generic G with $p^* \in G$.

One consequence of the equivalence between forcing and *-forcing is the following from Lemma C1 • 6 (b).

- C1.8. Corollary Let $V \models \mathsf{ZFC}$ be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. Let G be \mathbb{P} -generic over V. Let $\vec{\tau} \in V^{\mathbb{P}}$. Therefore $V[G] \models "\varphi(\vec{\tau}_G)$ " iff there is some $p \in G$ with $p \Vdash "\varphi(\vec{\tau})$ ".

Moreover, we can finally confirm the results of Motivation 31 B • 3. The only interesting case here is (4).

C1.9. Theorem — Let $V \models \mathsf{ZFC}$ be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. For $p \in \mathbb{P}$, write $p^* \in \mathbb{P}$ for an arbitrary $p^* \leq p$ (an arbitrary point in time after p). Let φ be a formula with parameters in $V^{\mathbb{P}}$. Therefore,

- 1. $p \Vdash \varphi$ iff every $p^* \Vdash \varphi$;
- 2. $p \Vdash \neg \varphi$ iff every $p^* \not\Vdash \varphi$, i.e. you can conclude it's false iff you will never discover that it's true;
- 3. $p \Vdash "\varphi \land \psi"$ iff $p \Vdash \varphi$ and $p \Vdash \psi$;

4. if $p \Vdash \exists x \ \varphi(x)$ then there is some $p^* \leq^{\mathbb{P}} p$ and τ where $p^* \Vdash \varphi(\tau)$; and

5. if $p \Vdash \varphi$, and φ is logically equivalent to ψ , then $p \Vdash \psi$;

Proof .:.

- 1. This follows from Lemma C1 5 and Corollary C1 7: $p \Vdash \varphi$ implies $p \Vdash^* \varphi$, which implies every $p^* \Vdash^* \varphi$, which implies every $p^* \Vdash \varphi$. The converse follows similarly.
- 2. This follows from Corollary C1 7 and Definition C1 4.
- 3. This follows from Definition 31 B 2.
- 4. Suppose p ⊨ "∃x φ(x)". Let p ∈ G which is P-generic over V. Since V[G] ⊨ "∃x φ(x)", there is some π ∈ V^P where V[G] ⊨ φ(π_G) and thus some condition of G forces this: q ∈ G has q ⊨ "φ(π)". As G is a filter, there is some common extension p* ≤^P p, q which then forces "φ(π)".
- 5. If φ is logically equivalent to ψ , then any generic extension $V[G] \vDash "\varphi \leftrightarrow \psi"$ so if $p \in G$ and $V[G] \vDash \varphi$, then clearly $V[G] \vDash \psi$.

 $^{{}^{}i}\mathcal{D} = \{D \in V : V \vDash "D \text{ is dense in } \mathbb{P}"\} \subseteq V \text{ must also be countable if } V \text{ is.}$

Section C2. ZFC in the Generic Extension

Our goal is now to prove Theorem 31 D • 13. Doing this amounts mainly to finding the right \mathbb{P} -names for certain sets that a particular axiom of ZFC claims the existence of. Now ostensibly, we could just apply a result like Corollary C1 • 8 and begin by, for each axiom φ of ZFC, finding an element of the preorder that forces ZFC. This isn't exactly easy to do if, say, $\mathbb{1}^{\mathbb{P}} \Vdash$ ZFC (which will be the case if $\mathbf{V} \models$ ZFC). Mostly this is because while \Vdash is defined, that doesn't mean it's computable, as whether certain sets are dense isn't always immediate.

We now collect together the implications of how much set theory V satisfies on how much set theory V[G] satisfies. As a bit of notation, we refer to V as the *ground model* and V[G] as the *generic extension*. Also, P refers to the powerset axiom while ZF⁻ refers to ZF-P+Col, where Col is the axiom scheme of collection. Collection is strictly stronger than replacement, as there is a complicated forcing where replacement holds in the generic extension, but not collection [30].

− C2 · 1. Theorem − Let $V \models$ ZFC be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. Let G be \mathbb{P} -generic over V. Therefore,

- $\mathbf{V} \models \mathbf{ZF}^-$ implies $\mathbf{V}[G] \models \mathbf{ZF} \mathbf{P};$
- $\mathbf{V} \vDash \mathsf{ZF}$ implies $\mathbf{V}[G] \vDash \mathsf{ZF}$;
- $V \vDash$ ZFC implies $V[G] \vDash$ ZFC.

To prove these inequalities, we need to prove various closure properties of the generic extension given by appropriate names in the ground model. The existence of these names follows from the amount of set theory the ground model satisfies.

Note that I will often use "a name", "a \mathbb{P} -name", and "a name in V" all for the same thing: an element of $V^{\mathbb{P}}$ for some given element of V[G].

- C2.2. Theorem - Let V be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. Let G be \mathbb{P} -generic over V. Suppose $V \models ZF - P$. Therefore $V[G] \models ZF - P$ - Rep, where Rep is the axiom scheme of replacement.

Proof .:.

- The axioms of extensionality and foundation follow from the fact that V[G] is transitive.
- The empty set axiom follows from the fact that $\emptyset \in V$ so that $\check{\emptyset} \in V$ and thus $\check{\emptyset}_G = \emptyset \in V[G]$.
- Pairing follows easily: for $x, y \in V[G]$, let $\dot{x}, \dot{y} \in V^{\mathbb{P}}$ be two \mathbb{P} -names for x and y respectively: $\dot{x}_G = x$ and $\dot{y}_G = y$. Consider the \mathbb{P} -name in V

$$\tau = \{ \langle \dot{x}, \mathbb{1}^{\mathbb{P}} \rangle, \langle \dot{y}, \mathbb{1}^{\mathbb{P}} \rangle \}.$$

Since any filter has $\mathbb{1}^{\mathbb{P}} \in G$, $\tau_G = \{\dot{x}_G, \dot{y}_G\} = \{x, y\}$. Hence V[G] is closed under pairs, and so as a transitive set, $V[G] \models \text{Pair}$.

• Comprehension requires some work. Let φ be a formula, and $x \in V[G]$. We'd like to show $\{y \in x : V[G] \models "\varphi(x, y, \vec{w})"\} \in V[G]$ for any parameters $\vec{w} \in V[G]$.

So let \vec{w} be \mathbb{P} -names for the parameters, and \dot{x} a \mathbb{P} -name for x. For any $y \in x$ (with name $\dot{y} \in \text{dom}(\dot{x})$) such that $\mathbf{V}[G] \models "\varphi(x, y, \vec{w})$ ", there is some $p \in G$ such that $p \Vdash "\dot{y} \in \dot{x} \land \varphi(\dot{x}, \dot{y}, \vec{w})$ ". So consider the set

 $\tau = \{ \langle \sigma, p \rangle \in \operatorname{dom}(\dot{x}) \times \mathbb{P} : p \Vdash ``\sigma \in \dot{x} \land \varphi(\dot{x}, \sigma, \vec{w})" \}.$

Thus any $\sigma_G \in \tau_G$ has $\langle \sigma, p \rangle \in \tau$ with $p \in G$, and hence $\mathbf{V}[G] \models "\sigma \in x \land \varphi(x, \sigma, \vec{w})$ ". The argument given above shows that any $y \in x$ with $\mathbf{V}[G] \models "\varphi(x, y, \vec{w})$ " has some \dot{y} and $p \in G$ with $\langle \dot{y}, p \rangle \in \tau$. Hence this τ witnesses this arbitrary instance of comprehension, and thus $\mathbf{V}[G] \models \text{Comp.}$

ZFC IN THE GENERIC EXTENSION

For union, for X ∈ V[G], we need to show ∪ X ∈ V[G]. Because comprehension holds, we only need to show there is some Y ∈ V[G] with ∪ X ⊆ Y, because then we can consider in V[G] the set {y ∈ Y : ∃x ∈ X (y ∈ x)} = ∪ X.

So let \dot{X} be a name for X. Consider the set

 $\tau = \{ \langle \sigma, p \rangle : \exists \langle \sigma', p' \rangle \in \dot{X} \ (\langle \sigma, p \rangle \in \operatorname{dom}(\sigma')) \}.$

This clearly works as $x \subseteq \tau_G$ not only for $\langle \dot{x}, p \rangle \in \dot{X}$ with $p \in G$ but for all x with $\dot{x} \in \text{dom}(x)$. Hence $\bigcup \dot{X}_G = \bigcup X \subseteq \tau_G$. Thus $V[G] \models \text{Union}$.

• For infinity, just note that the name $\check{\omega} = \{\langle \check{n}, \mathbb{1}^{\mathbb{P}} \rangle : n \in \omega\} \in V^{\mathbb{P}}$ witnesses that $\omega \in V[G]$.

With the addition of powerset in the ground model, we also get powerset in the generic extension.

- C2·3. Lemma Let V be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. Let G be \mathbb{P} -generic over B. Suppose $V \models ZF$. Therefore $V[G] \models P$

Proof .:.

Let $x \in V[G]$. We need to show that $\mathcal{P}(x) \cap V[G] \in V[G]$, meaning that there is a set in V[G] that collects every subset of x in V[G]. By Theorem C2 • 2, $V[G] \models \text{Comp so it suffices to find } Y \in V[G]$ with $\mathcal{P}(X) \cap V[G] \subseteq Y$, since then we just consider $\mathcal{P}(x) \cap V[G] = \{y \in Y : y \subseteq x\} \in V[G]$.

So let \dot{x} be a name for x. If $\sigma_G \subseteq x$ in V[G], then there is some $p \in \mathbb{P}$ with $p \Vdash "\sigma_G \subseteq \dot{x}$ ". So consider $\tau = \{\langle \sigma, p \rangle \in \mathcal{P}(\operatorname{dom}(\dot{x}) \times \mathbb{P}) \times \mathbb{P} : p \Vdash "\sigma \subseteq x".\}$

This is a set in V by the powerset axiom in V. If $\langle \sigma, p \rangle \in \tau$ with $p \in G$, then $V[G] \models "\sigma_G \subseteq x$ " and thus $\tau_G \subseteq \mathcal{O}(x) \cap V[G]$. Recall Result 32 E • 2, which says that any $y \in V[G]$ with $y \subseteq x$ has a name $\dot{y} \in \mathcal{O}(\operatorname{dom}(\dot{x}) \times \mathbb{P})$. So if $V[G] \models "y \subseteq x$ ", there is some $p \in G$ with $p \Vdash "\dot{y} \subseteq \dot{x}$ ", and so $\langle \dot{y}, p \rangle \in \tau$ has $y = \dot{y}_G \in \tau_G$. Hence $\tau_G = \mathcal{O}(x) \cap V[G]$ witnesses this instance of powerset, and so $V[G] \models \mathbb{P}$.

This shows that $\mathbf{V} \models \mathsf{ZF}$ implies $\mathbf{V}[G] \vdash \mathsf{ZF} - \mathsf{Rep}$. In order to confirm replacement we need the axiom scheme of collection in the ground model. This follows from powerset and replacement, but without powerset, we might not have the axiom scheme of collection. So when we jump from $\mathbf{V} \models \mathsf{ZF} - \mathsf{P}$ to $\mathbf{V} \models \mathsf{ZF}$, we can confirm two axioms in $\mathbf{V}[G]$: P and Rep .

First we introduce the axiom scheme of collection, and then we show this follows from ZF. We introduce this axiom, because it is used in the proof that $V \models ZF$ implies $V[G] \models ZF$. Of course, we could just proof the particular instance(s) we need during the proof, but this isn't exactly instructive.

– C2•4. Definition

The axiom scheme of collection (Col) states the following: if φ is a relation on a set D, then there is a set containing φ -relatives of each $x \in D$. Symbolically, Col consists of all formula of the form

$$\forall \vec{w}, D \ (\forall x \in D \ \exists y \ \varphi(x, y, D, \vec{w}) \rightarrow \exists R \ \forall x \in D \ \exists y \in R \ \varphi(x, y, D, \vec{w})).$$

where φ is a formula.

The \vec{w} just allow parameters. Note that this clearly stronger than replacement, which requires φ to define a function over D:

$$\forall \vec{w} \forall D \; (\forall x \in D \; \exists ! y \; \varphi(x, y, \vec{w}) \to \exists R \; \forall x \in D \; \exists y \in R \; \varphi(x, y, \vec{w})).$$

<mark>── C2•5. Lemma</mark> -ZF ⊢ Col

Proof .:.

For each φ , \vec{w} , and D, consider the collection of all relatives of elements in D:

$$R' = \{ y : \exists x \in D \ \varphi(x, y, D, \vec{w}) \}.$$

*)

Note that this is potentially a proper class. But with powerset, we can consider

 $R = \{ y : \exists x \in D \ (\varphi(x, y, D, \vec{w}) \land \forall z \ (\varphi(x, z, D, \vec{w}) \to \operatorname{rank}(z) \ge \operatorname{rank}(y))) \}.$

This will be a set, because we've defined a function $f : D \to V$ where f(x) is the least rank of a y with $\varphi(x, y, D, \vec{w})$. This yields $f"D \subseteq Ord$ as a set of ordinals, and thus $R \subseteq V_{\sup f"D}$ yields that R is a set by comprehension.

The above idea (considering only the elements of least rank) has been dubbed "Scott's trick" as per Scott's Trick $(9 \text{ C} \cdot 1)$.

- C2•6. Lemma

Let V be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. Let G be \mathbb{P} -generic over V. Suppose $V \models \mathsf{ZF} - \mathsf{P} + \mathsf{Col}$. Therefore $V[G] \models \mathsf{Rep}$.

Proof .:.

Let φ be a formula with parameters in V[G], and let $D \in V[G]$. Suppose

$$\mathbf{V}[G] \models ``\forall x \in D \exists ! y \varphi(x, y, D)`'.$$

We need to find a \mathbb{P} -name for the range of φ restricted to D. Note that there is some $p_D \in \mathbb{P}$ forcing (*) (translated with parameters as \mathbb{P} -names).

Consider the formula $\psi(p, \sigma, \tau)$ stating:

 $p \in \mathbb{P} \land \left(p \Vdash ``\varphi(\sigma, \tau, \dot{D})" \lor \neg \exists \pi \ p \Vdash ``\varphi(\sigma, \pi, \dot{D})" \right).$

In **V**, for each $\langle \sigma, p \rangle \in \operatorname{dom}(\dot{D}) \times \mathbb{P}$, there is a $\tau \in V^{\mathbb{P}}$ where $\psi(p, \sigma, \tau)$ holds (τ can be anything if $\neg \exists \pi p \Vdash$ " $\psi(\sigma, \pi, \dot{D})$ "). By collection in **V**, there is a set $R \subseteq V^{\mathbb{P}}$ where each $\langle \sigma, p \rangle \in \operatorname{dom}(\dot{D}) \times \mathbb{P}$ has a $\tau \in R$. As $R \subseteq V^{\mathbb{P}}$ is a set, $\rho = R \times \mathbb{P}$ is a \mathbb{P} -name.

To see that $\rho_G \in V[G]$ satisfies our requirements, suppose $V[G] \vDash "x \in D$ ". We can take $x = \sigma_G$ for $\sigma \in \text{dom}(\dot{D})$. Since (*) holds, there is some y where $V[G] \vDash "\varphi(x, y, D)$ ". This is forced by some $p \in \mathbb{P}$: $p \Vdash$ " $\varphi(\sigma, \dot{y}, \dot{D})$ ". Hence there is a $\tau \in R$ where $p \Vdash "\varphi(\sigma, \tau, \dot{D})$ ", and thus $V[G] \vDash "\tau_G \in \rho_G \land \varphi(x, \tau_G, D)$ ", yielding the result. This shows this arbitrary instance of replacement holds in V[G], and thus $V[G] \vDash Rep$. \dashv

So we can conclude $V \models ZF$ implies $V[G] \models ZF$. The last thing to consider is choice. There are multiple versions of AC, but we will consider one that's easy to use. In particular, we're using the version that says every set is covered by an ordinal.ⁱⁱ

– C2•7. Theorem

Let V be a transitive model we can force with $\mathbb{P} \in V$ over, where \mathbb{P} is appropriate for forcing. Let G be \mathbb{P} -generic over V. Suppose $V \models \mathsf{ZFC}$. Therefore $V[G] \models \mathsf{ZFC}$.

Proof .:.

We have by the previous lemmas that $V[G] \vDash ZF$. So we only need to prove $V[G] \vDash AC$, and it suffices to show that for any $x \in V[G]$, there is an $f \in V[G]$ and an $\alpha \in Ord \cap V[G]$ where $V[G] \vDash "f : \alpha \to x$ is surjective". Let \dot{x} be a name for x. By AC in V, there is a surjection $F : \alpha \to dom(\dot{x})$ for some $\alpha \in Ord \cap V$. Thus

$$f = \{ \langle \langle \langle \xi, F(\xi) \rangle \rangle, \mathbb{1}^{\mathbb{P}} \rangle : \xi < \alpha \}$$

works. (Here, $\langle\!\langle a, b \rangle\!\rangle$ is a name for $\langle a_G, b_G \rangle$, and in particular is $\{\langle\!\langle a, \mathbb{1}^P \rangle\!\rangle, \mathbb{1}^P \rangle\!\rangle, \langle\!\langle a, \mathbb{1}^P \rangle\!\rangle, \langle b, \mathbb{1}^P \rangle\!\rangle, \mathbb{1}^P \rangle\!\rangle$.) If we consider f_G , we have that any of its elements is of the form $\langle\xi, F(\xi)_G\rangle$ where $F(\xi) \in \operatorname{dom}(\dot{x})$ and $\xi < \alpha$. And so we can regard f_G as a function from α , and it should be clear that $\mathbf{V}[G] \models ``x \subseteq \operatorname{im}(f_G)$ ''. Hence this version of AC holds in $\mathbf{V}[G]$.

ⁱⁱWe get an injection from $g: x \to \alpha$ just by setting g(y) to be the least β with $f(\beta) = y$ where $f: \alpha \to x$ is the surjection. This yields a well-order of x. Note that in ZF, we have that x can always be surjected onto various ordinals, but the reverse is equivalent to x having a well-order.

Bibliography

- [1] Carlos Antos and Victoria Gitman, Modern Class Forcing (2019). Unpublished manuscript.
- [2] James Baumgartner, Iterated forcing, Surveys in Set Theory (A. R. D. Mathias, ed.), London Mathematical Society Lecture Note Series, vol. 87, Cambridge University Press, 1983, pp. 1–59.
- [3] Andrés Caicedo and Ralf Schindler, Projective well-orderings of the reals, Archive for Mathematical Logic 45 (2006), 783–793.
- [4] James Cummings, Iterated Forcing and Elementary Embeddings, Handbook of Set Theory (Matthew Foreman and Akihiro Kanamori, eds.), Vol. 2, Springer, Dordrecht, 2010, pp. 775–883.
- [5] Donald Cohn, Measure Theory, Birkhäuser, 1980.
- [6] Harvey Diamond and Gregory Gelles, *Relations Among Some Classes of Subsets of R*, The American Mathematical Monthly 91 (1984), no. 1, 19–22.
- [7] Herbert Enderton, A Mathematical Introduction to Logic, 2nd ed., Academic Press, 2001.
- [8] Ilijas Farah, The Extender Algebra and Σ_1^2 -Absoluteness (2011). to appear in The Cabal Seminar, Volume IV.
- [9] Melvin Chris Fitting, Intuitionistic Logic Model Theory and Forcing, Studies in Logic and the Foundations of Mathematics, vol. 60, North-Holland, 1969.
- [10] Farmer S (https://math.stackexchange.com/users/910439/farmer-s), (Long) extender in its own ultrapower, Mathematics Stack Exchange. URL: https://math.stackexchange.com/q/4224925 (version: 2021-08-15).
- [11] Sy-David Friedman, Constructibility and Class Forcing, Handbook of Set Theory (Matthew Foreman and Akihiro Kanamori, eds.), Springer Science+Business Media B.V., 2010, pp. 557–604.
- [12] D. H. Fremlin, Consequences of Martin's Axiom, Cambridge Tracts in Mathematics, vol. 84, Cambridge University Press, 1984.
- [13] Serge Grigorieff, Intermediate Submodels and Generic Extensions in Set Theory, Annals of Mathematics 101 (1975), no. 3, 447–490.
- [14] Joel David Hamkins, Gap Forcing, Israel Journal of Mathematics 125 (2001), 237-252.
- [15] Leo Harrington, The contructible reals can by (almost) anything (1974). Unpublished.
- [16] Thomas Jech, The Axiom of Choice, Studies in Logic and the Foundations of Mathematics, vol. 75, North-Holland, 1973. Republished by Dover Publications, 2008.
- [17] _____, Set Theory: The Third Millennium Edition, revised and expanded, Springer Monographs in Mathematics, Springer-Verlag Berlin Heidelberg, 2003.
- [18] Alexander Kechris, Classical Descriptive Set Theory, Graduate Texts in Mathematics, vol. 156, Springer-Verlag New York, 1995.
- [19] _____, On Projective Ordinals, The Journal of Symbolic Logic 39 (1974), no. 2, 269–282.
- [20] Kenneth Kunen, Set Theory: An Introduction to Independence Proofs, 6th ed., Studies in Logic and the Foundations of Mathematics, vol. 102, North-Holland, 1995.
- [21] A. Lévy and M. Solovay, Measurable cardinals and the continuum hypothesis, Israel Journal of Mathematics 5 (1967), 234-248.
- [22] Arnold Miller, Descriptive Set Theory and Forcing: How to prove theorems about Borel sets the hard way, Lecture Notes in Logic, vol. 4, Springer-Verlag Berlin Heidelberg, 1995.
- [23] Yiannis Moschovakis, Descriptive Set Theory, 2nd ed., Mathematical Surveys and Monographs, vol. 155, American Mathematical Society, 2009.
- [24] Assaf Rinot, Chain Conditions of Products, and Weakly Compact Cardinals, The Bulletin of Symbolic Logic 20 (2014), no. 3, 293-314.
- [25] Jonas Reitz, The Ground Axiom, Ph.D., 2008.
- [26] Gerald Sacks, Higher Recursion Theory, 2nd ed., Perspectives in Logic, Cambridge University Press, 2016.
- [27] Ralf Schindler, Set Theory: Exploring Independence and Truth, Universitext, Springer International Publishing, 2014.
- [28] Terence Tao, Analysis I: Third Edition, Texts and Readings in Mathematics, vol. 37, Springer Singapore, 2016.
- [29] _____, Analysis II: Third Edition, 3rd ed., Texts and Readings in Mathematics, vol. 38, Springer Singapore, 2016.
- [30] Andrzej M. Zarach, Replacement → Collection, Gödel '96: Logical foundations of mathematics, computer science and physics—Kurt Gödel's legacy, Brno, Czech Republic, August 1996, proceedings (Petr Hájek, ed.), Lecture Notes in Logic, vol. 6, Springer-Verlag, 1996.