Set Theory: The Independence Phenomenon

Peter Koellner

November 13, 2013

Contents

1 The Standard System **7**

1.1 Language of Set Theory . . . 7
1.2 Informal Characterisation . . . 7
1.3 Axioms . . . 9
1.4 Sets versus Classes . . . 11
1.5 Basic Notions . . . 13
1.6 Choice . . . 15

2 Ordinals **17**

2.1 Transfinite Induction . . . 17
2.2 Ordinals . . . 20
2.3 Ordinal Arithmetic . . . 23
2.4 The Hierarchy . . . 24

3 Cardinals **25**

3.1 Choice . . . 25
3.2 Cardinality . . . 26
3.3 Cardinal Arithmetic . . . 29
3.4 Cofinality . . . 31

4 Trees **34**

4.1 Suslin's Hypothesis . . . 34
4.2 Aronszajn Trees . . . 36
4.3 Suslin Trees . . . 39

5 Measurability **41**

5.1 The Measure Problem . . . 41
5.2 Borel Sets . . . 42

5.3 Projective Sets 45

6 Constructibility 47

6.1 Absoluteness and Reflection 47
6.2 The constructible universe 53
6.3 The Axiom $V = L$ 59

7 Forcing 63

7.1 Basic Forcing 64
7.2 The Continuum Hypothesis. 77
7.3 Product Forcing 87

Introduction

1 Incompleteness in Arithmetic

There are many ways of introducing set theory. I shall do so in terms of the incompleteness phenomenon.

Let us start with the structure of the natural numbers: $(\mathbb{N}, 0, S, +, \times, \exp)$. We begin by writing down truths about this structure and at a certain point we arrive at the axiom system PA. This system appears to be robust in that the theorems proved by informal means can be recast as formal derivations within PA. This apparent robustness leads one to suspect that PA is in some sense "maximal", that is, that it embodies all of the truths of arithmetic. Indeed there are areas where such completeness holds. For example, by a result of Tarski, Euclidean geometry admits a complete axiomatisation. But in arithmetic the situation is different.

Theorem. (Gödel) *Assume that* PA *is consistent. Then* $\mathrm{PA} \nvdash \mathrm{Con}(\mathrm{PA})$.

There are two approaches to overcoming incompleteness in arithmetic—the first due to Turing and the second due to Gödel. Turing's approach is to shift to a richer *system*. The trouble with PA is that it does not prove Con(PA). To overcome this limitation we shift to the system $\mathrm{PA} + \mathrm{Con}(\mathrm{PA})$. The trouble is that the new system—although able to capture the truth Con(PA)—is unable to capture the truth $\mathrm{Con}\big(\mathrm{PA} + \mathrm{Con}(\mathrm{PA})\big)$. But we can fix this by extending our system to $\mathrm{PA} + \mathrm{Con}(\mathrm{PA}) + \mathrm{Con}\big(\mathrm{PA} + \mathrm{Con}(\mathrm{PA})\big)$. In this way we obtain a series of systems each one of which captures a truth missed by its predecessor. This procedure can be extended into the (countable) transfinite. Unfortunately, if one drives the iteration by means which are "implicit" in PA then the process closes off at a system which has come to be known as *Predicative Analysis* at which point the incompleteness phenomenon reappears.

Gödel's approach is to shift to a richer *structure*. The next natural structure after the natural numbers is the structure of the *powerset* of the natural numbers: $(P(\mathbb{N}), \mathbb{N}, +, \times, \in)$. As before we write down all of the "obvious" truths of this structure and we arrive at the axiom system PA^2 of *second-order* arithmetic. This system proves the consistency of Predicative Analysis and much more. But alas it too is incomplete. Fortunately, we can shift to an even richer structure by applying the powerset again. As before we arrive at the axiom system PA^3 of *third-order* arithmetic and so on.

2 Set Theory

In describing Gödel's approach to overcoming the incompleteness phenomenon in arithmetic and it's higher-order analogues we introduced the subject matter of set theory. Starting with the simplest set of all—the empty set $\varnothing$—one iterates the powerset into the (uncountable) transfinite. When we write down the "obvious" truths about this structure we arrive at the axiom system ZFC. But we arrive at more. For example, we arrive at the axiom that asserts that there is an iterate of the powerset operation—a level of the hierarchy of sets—that satisfies the axioms ZFC and, furthermore, *that the hierarchy continues beyond this*. In contrast to PA there is no clear end in sight of the "obvious" truths of set theory precisely because (in contrast to the subject matter of arithmetic) there is no clear end in sight to the subject matter of set theory.

The existence of such strong extensions of ZFC—through the extension of the hierarchy—encourages the view that although the incompleteness phenomenon will never go away we can rest assured that for any "obvious" system such as ZFC there will be an "obvious" extension that captures the truths missed by the first. Can such a view be sustained?

3 Three Fundamental Problems

It turns out that there are examples of independence which are much more intractable than the Gödel sentences. The most natural of these occur already at the beginnings of set theory. Let me mention three such fundamental problems associated with the three main areas of *cardinal arithmetic*, *combinatorics*, and *descriptive set theory*.

As soon as one takes the powerset $P(\mathbb{N})$ of the natural numbers $\mathbb{N}$ one is led to ask 'how "large" is $P(\mathbb{N})$?' This question—which we will make precise—was asked by Cantor when he created set theory. He conjectured that the size of $P(\mathbb{N})$ is the smallest possible, a conjecture known as the *continuum hypothesis* (CH). This is the fundamental problem of cardinal arithmetic. The second problem—the problem of combinatorics—is *Suslin's hypothesis* (SH). And the third problem—the problem of descriptive set theory—is what I will call the *measurability hypothesis* (MH).

It would take us to far afield to pause to give a precise formulation of each of these problems. I would, however, like to say something about MH. There is a famous theorem of Banach and Tarski according to which one can take a sphere, partition it into finitely many pieces, and rearrange those pieces via rigid motions to obtain a sphere of *twice* the size. This seems incredible. It turns out that the pieces are *very* bizarre and that it cannot be done with "simple" pieces.[1] The question is 'how complex must the pieces be?' Well, if the pieces are "very simple" ($\utilde{\Sigma}^1_1$) then it cannot be done. What then if the pieces are just "simple" ($\utilde{\Sigma}^1_2$)? Consider the following statement:

SPHERE. One **cannot** take a sphere, partition it into finitely many "simple" ($\utilde{\Sigma}^1_2$) pieces, rearrange those pieces via rigid motions and put them back together to form a sphere of twice the size.

4 Independence in Set Theory

People tried very hard to settle CH, SH, and MH. But they tried in vain. For they were working within a theory that could be formalised in ZFC and it is now known that ZFC *cannot* settle these questions. In particular, SPHERE (which is a version of MH) cannot be settled in ZFC. What is more, these limitations of ZFC also apply to the "obvious" extensions of ZFC mentioned above.

These limitations are much more dramatic than the limitation of ZFC with respect to Con(ZFC). There is something benign about the latter limitation. One reason is that if we have good reason for believing a theory T then we have good reason for believing the statement Con(T). More importantly, if we know that Con(ZFC) is a limitation then we know whether or not Con(ZFC) is true. I am not saying that all consistency statements are

[1]It should be mentioned that there are *other* paradoxical decompositions of the sphere—for example, those due to Dougherty and Foreman—that *do* involve simple pieces.

benign; for example, there is nothing benign about Con(ZFC+'There is a supercompact cardinal'). Rather, I am saying that consistency statements that *we know to be instances of incompleteness* are benign. The reason such statements are benign is that if you know that a consistency statement φ is independent of a correct theory T (which includes a minimal amount of arithmetic, such as Robinson's arithmetic Q) then you know that φ is true. (Something more general is true. Consider a statement in the language of arithmetic of the form $\varphi = \forall n\, \psi(n)$ where ψ is *bounded* in the sense that all of its quantifiers are of the form '$\exists x < n$' or $\forall x < n$ where n is a numeral. Such statements are called 'Π^0_1-statements'. Con(PA) is such a statement. The point is that if a Π^0_1-statement φ is independent of a theory which is correct about bounded statements—for example, Q—then φ must be true. Here is the proof: Suppose, for contradiction, that φ is not true. Then $\exists n\, \neg\psi(n)$. Then $\neg\psi(k)$ for some k. But then $\mathrm{Q} \vdash \neg\psi(k)$. So $\mathrm{Q} \vdash \exists n\, \neg\psi(n)$, which contradicts the assumption that φ is independent of T).

The limitations of ZFC (and its "obvious" extensions) with respect to CH, SH, and MH are *not* benign in the above sense. Let me summarise the contrast between benign and serious limitations. The statement Con(ZFC) is independent of ZFC. It exemplifies a limitation of ZFC. This is a benign limitation because the recognition that it is a limitation suffices to determine its truth-value, in fact, it must be true. that it is a limitation implies that it is true. This is the case with all Π^0_1-statement. The statement CH is independent of ZFC. This is a serious limitation because the fact that it is a limitation does not imply anything about its truth-value.

The serious limitations pose serious problems. For they put us in a situation where we know that there are statements that our system is not capturing and yet do not know whether or not these statements are true.

5 Hope: New Axioms

These notes fall naturally into three parts. In Part I we introduce the basics of set theory, concentrating on the standard system, ZFC. In Part II we introduce the techniques used for establishing independence—most notably, the constructible universe and iterated forcing—and we use these to establish the independence of our three problems and much more. In Part III we give a glimpse of new axioms—most notably large cardinal axioms and forcing axioms—and show that they have bearing on our problems.

Chapter 1

The Standard System

1.1 Language of Set Theory

Our language LST is the language of first-order logic with two additional symbols: '$=$' and '$\in$'. Atomic formulas are thus of the form '$x = y$' and '$x \in y$' and molecular formulas are built up from these in the standard fashion using the connectives '$\neg$', '$\wedge$', '$\vee$', '$\rightarrow$', '$\leftrightarrow$' and the quantifiers '$\forall$' and '$\exists$'. Our logic is that of classical first-order logic with identity. We use '$\exists x \in y \; \varphi(x)$' to abbreviate '$\exists x \, (x \in y \wedge \varphi(x))$', and '$\forall x \in y \; \varphi(x)$' to abbreviate '$\forall x \, (x \in y \rightarrow \varphi(x))$', and '$\exists! x \, \varphi(x)$' to abbreviate '$\exists x \, (\varphi(x) \wedge \forall y \, (\varphi(y) \leftrightarrow x = y))$'.

Our reason for being explicit about the language of set theory is that many of our concerns will be *metamathematical*. For example, we shall be concerned with definability and independence.

1.2 Informal Characterisation

There are two primitive operations in the foundations of set theory: the *set formation* operation and the *powerset* operation.

The *set formation* operation—which I shall abbreviate 'set of' or, more explicitly, 'set of x's' where "the variable x ranges over some given kind of objects" (Gödel 1964)—takes a given kind K of objects and forms a single object y, called a 'set', which has as members exactly those objects of kind K. The statement 'x is a member of y' is abbreviated '$x \in S$'. As an example consider the natural numbers 0, 1, 2, 3, ... Applying the operation 'set of' to this given kind we obtain the set $\{0, 1, 2, 3, \dots\}$, an object which has as members all and only the natural numbers. This object is denoted 'ω'. The 'set of' operation is such that the resulting set is *constituted* by its elements; more precisely, sets have the feature that two sets A and B are the same if and only if they have the same members. In this respect sets differ from other objects, such as concepts, which also, in some sense, collect the objects of a given kind. For compare the concepts *natural number* and *natural number less than some prime number*. These concepts apply to exactly the same things (that is, they have exactly the same things "falling under" them) and yet they are *not* the same; that is, a concept, in contrast to a set, is not determined by the objects that "fall under" it.

The second primitive notion is the operation of *powerset*. First, let us say that a set x is a *subset* of another set y—symbolised $x \subseteq y$—if every member of x is a member of y. For example, the set of even numbers is a subset of the set of natural numbers. Now the powerset operation is the operation which, given a set x, produces the set $P(x)$ of *all arbitrary* subsets of x. Here an *arbitrary* subset is to be contrasted with a *definable* subset. As examples of definable subsets of ω consider the set of even numbers and the set of odd numbers. Of course, there are only as many definable sets as there are definitions. As we shall see there are far more arbitrary subsets of ω than definable subsets.

Given the above primitives we are now in a position to give an informal characterisation of the iterative hierarchy of sets.

The powerset operation can be iterated. Thus, starting with a given set, say the set of the natural numbers ω, we can take its powerset to obtain $P(\omega)$ and then we can take the powerset of this in turn to obtain $P(P(\omega))$, and so on. In this way we generate a hierarchy ω, $P(\omega)$, $P(P(\omega))$, $\ldots$ Our starting point in this case was ω. But we could have started with any set. For example, we could have started with the set of giraffes or the set of fundamental particles. Of all possible starting points there is a simplest, namely, the set which contains nothing, that is, the *empty set*, which we shall denote '$\varnothing$'. Let us start here, at the simplest point, and construct the hierarchy of *pure sets*.

The first level of the hierarchy is $V_0 = \varnothing$. We can take the powerset of this to obtain $V_1 = P(\varnothing)$. Continuing in this way we obtain $V_2 = P(P(\varnothing))$, $V_3 = P(P(P(\varnothing)))$, $\ldots$ What next? At this point we can apply the set formation operation to obtain the set that contains x if and only if x is a member of V_n for some n. We call this level V_ω. We are then in a position to start taking powersets again. After iterating the powerset operation through all of the natural numbers (for the second) time we can apply the set formation operation to obtain a level that we will call $V_{\omega+\omega}$. We are then in a position to start taking powersets again ...

Three salient features of the above account should be noted. First, we applied an operation, the *powerset operation*, to get from one successor stage to the next. Second, we applied an operation, the *summing up operation* (a special case of the set formation operation), to pass through limit stages. Third, we ended our description with *the elusive three dots* ... The first two operations are clear but the three dots are something of a mystery since they never seem to go away. For example, suppose we have applied the summing up operation ω-many times. Even then we are not at an end. For we can apply the summing up operation yet again to get a new level—labeled '$V_{\omega\cdot\omega}$' in the above diagram—at which point we can start taking powersets again ...

1.3 Axioms

The standard system ZF consists of the following infinite list of axioms:

(1) (Extentionality) Two sets that have the same elements are identical:

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

(2) (Emptyset) There is a set $\varnothing$ that contains no elements:

$$\exists a\,\forall x\,(x \notin a).$$

(3) (Pairing) For any two sets x and y there is a set $\{x, y\}$ that contains only x and y:

$$\forall x\,\forall y\,\exists a\,\forall z\,\big(z \in a \leftrightarrow (z = x \vee z = y)\big).$$

(4) (Union) For any set x there is a set $\cup x$ consisting of the members of members of x:

$$\forall x\,\exists a\,\forall z\,\big(z \in a \leftrightarrow \exists y\,(z \in y \in x)\big).$$

(5) (Powerset) For any set x there is a set $P(x)$ consisting of all of the subsets of x (where y is a subset of x—symbolised '$y \subseteq x$'—if and only if $\forall z\,(z \in y \to z \in x)$):

$$\forall x\, \exists a\, \forall y\, \big(y \in a \leftrightarrow y \subseteq x\big).$$

(6) (Infinity) There is a set which contains $\varnothing$ and is such that if it contains y then it also contains $y \cup \{y\}$ (where $y \cup \{y\} = \cup\{y, \{y\}\}$):

$$\exists x\, \big(\varnothing \in x \wedge \forall y\, (y \in x \to y \cup \{y\} \in x)\big).$$

(7) (Foundation) There are no infinitely descending $\in$-chains:

$$\forall x\, \big(x \neq \varnothing \to \exists y \in x\, \forall z \in x\, (z \notin y)\big).$$

$(8)_\varphi$ (Replacement) For each formula φ of LST that defines (with parameters) a function F and for each set a there is a set b that contains $F(x)$ if and only if a contains x:

$$\begin{aligned} \forall p_1 \cdots \forall p_n\, \forall a\, \big[&\forall y \in a\, \exists! z\, \varphi(y, z, p_1, \ldots, p_n) \\ &\to \exists b\, \forall z\, \big(z \in b \leftrightarrow \exists y \in a\, \varphi(y, z, p_1, \ldots, p_n)\big)\big]. \end{aligned}$$

Exercise 1.1. Axioms (2)–(5) assert the existence of *a* set with certain properties. Show that in each case there is *exactly one* set with the corresponding properties. (This ensures that our notation '$\varnothing$', '$\{x, y\}$', '$\cup x$', '$P(x)$' is unambiguous.)

Exercise 1.2. Give an informal proof that the following are equivalent:

(a) There is no sequence of sets $x_1, x_2, x_3, \ldots$ such that for all natural numbers n, $x_{n+1} \in x_n$.

(b) $\forall x\, \big(x \neq \varnothing \to \exists y \in x\, \forall z \in x\, (z \notin y)\big)$.

(Thus in our statement of Foundation the formal statement and the informal statement are equivalent.) Isolate the key principle that you used in your informal proof. This principle is *not* provable in ZF.

There are two schemata that are closely related to Replacement.

(i)$_\varphi$ (Comprehension) For each formula φ of LST that defines (with parameters $p_1, \ldots, p_n$) a property P and for each set a there is a set $\{x \in a \mid \varphi(x, p_1, \ldots, p_n)\}$ that contains exactly those members of a that have P:

$$\forall p_1 \cdots \forall p_n \, \forall a \, \exists b \, \forall x \left(x \in b \leftrightarrow (x \in a \wedge \varphi(x, p_1, \ldots, p_n)) \right).$$

(ii)$_\varphi$ (Collection) For each formula φ of LST that defines a function F and for each set a there is a set b that contains $F(x)$ if a contains x:

$$\forall p_1 \cdots \forall p_n \, \forall a \left[\forall y \in a \, \exists! z \, \varphi(y, z, p_1, \ldots, p_n) \right.$$
$$\left. \to \exists b \, \forall z \left(\exists y \in a \, \varphi(y, z, p_1, \ldots, p_n) \to z \in b \right) \right].$$

Exercise 1.3. Assume ZF – Replacement. Show that Replacement is equivalent to the conjunction of Comprehension and Collection.

1.4 Sets versus Classes

In the statement of Comprehension we introduced the notation for *set abstraction*: '$\{x \in a \mid \varphi(x, p_1, \ldots, p_n)\}$' where a is a set, $p_1, \ldots, p_n$ are parameters, and φ is a formula. Extensionality implies that this expression—the *set abstract*—is unambiguous.

It is important to remember that the introduction of set abstracts is justified by Comprehension. So it is crucial to include the restriction '$\in a$' in '$\{x \in a \mid \varphi(x, p_1, \ldots p_n)\}$'. Still, it is often convenient to use the notation of *unrestricted* abstraction. For example, it is often convenient to speak as if there is an object $\{x \mid x = x\}$ that contains "all" sets. But when doing so we must remember that Comprehension does not apply. For example, we are not warranted in assuming that '$\{x \mid x = x\}$' denotes a set. Indeed this assumption contradicts Foundation. For if '$\{x \mid x = x\}$' denotes a set v then $v \in v$. Now you might think 'So much the worse for Foundation'. But Foundation is not really the issue. For if v is a set then Comprehension implies that $r = \{x \in v \mid x \notin x\}$ is a set and this leads to a contradiction since if $r \in r$ then $r \notin r$ and if $r \notin r$ then $r \in r$. In other contexts this is known as *Russell's paradox*. In our present context the lesson is that when we introduce the notation for unrestricted abstraction we cannot always assume that it denotes a set.

There are cases where unrestricted abstracts *do* denote sets. For example, if x is a set then

$$`\{z \mid \exists y\,(z = \{y\} \wedge y \in x)\}'$$

denotes a set since it denotes what is denoted by

$$`\{z \in P(x) \mid \exists y\,(z = \{y\} \wedge y \in x)\}'.$$

When the appropriate restriction is obvious, we will often omit it.

There are also cases where an unrestricted abstract *does not* denote a set and yet we will use this notation anyway. For example we will use '$\{x \mid x = x\}$'. Now there is a temptation to assume that although such an abstract does not denote a *set* it nevertheless denotes something and such a something is usually called a *proper class*. (Abstracts in general are said to denote *classes* and those which do not denote sets are said to denote *proper* classes.) We will resist this temptation but will employ the language of classes as a convenient but possibly misleading way of speaking; for example, we will write things like 'consider the proper class $\{x \mid x = x\}$' when we really mean 'consider the proper class-abstract '$\{x \mid x = x\}$'. Now, when an unrestricted abstract does not denote a set we must be careful never to allow such an abstract to flank the left hand side of '$\in$'. We can, however, allow unrestricted abstracts to flank the *right* hand side of '$\in$' since we can treat

$$`a \in \{x \mid \varphi(x)\}'$$

as a notational variant of

$$`\varphi(a)'.$$

We will also allow unrestricted abstracts to flank '$=$' and '$\subseteq$', treating

$$`\{x \mid \varphi(x)\} = \{x \mid \psi(x)\}'$$

as a notational variant of

$$`\forall x\,(\varphi(x) \leftrightarrow \psi(x))'$$

and

$$`\{x \mid \varphi(x)\} \subseteq \{x \mid \psi(x)\}'$$

as a notational variant of

$$`\forall x\,(\varphi(x) \rightarrow \psi(x))'.$$

We will also occasionally introduce constants for classes; for example, 'V' shall abbreviate '$\{x \mid x = x\}$'. And, on occasion, we will use upper case variables to range over classes. As above, such statements are to be regarded as notational variants of statements which speak only of sets. For example, the statement

'for each class C and for each set a, $C \cap a$ is a set'

is shorthand for Comprehension.

Up until this point I have been careful to observe the distinction between *using* an expression and *mentioning* it by flanking it with quotation marks. From this point on I will be less careful, often relying on context to disambiguate.

1.5 Basic Notions

A. Intersection and Difference. The intersection $\cap x$ of a set x is the set of elements which are members of every member of x:

$$\cap x = \{z \mid \forall y \in x\,(z \in y)\}.$$

Let $x \cap y = \cap\{x, y\}$ and $x_1 \cap \cdots \cap x_n = \cap\{x_1, \ldots, x_n\}$. The difference $x - y$ between two sets x and y is the set of elements in the former but not in the latter:

$$x - y = \{z \mid z \in x \wedge z \notin y\}.$$

B. Ordered sets. Sets are indifferent to order; for example, $\{x_1, x_2\} = \{x_2, x_1\}$. We should like to define an *ordered pair* (x_1, x_2) with the property that $(x_1, x_2) = (y_1, y_2)$ iff $x_1 = y_1$ and $x_2 = y_2$. To this end we set

$$(x_1, x_2) = \{\{x_1\}, \{x_1, x_2\}\}.$$

Exercise 1.4. Show that so defined the ordered pair has the desired property.

Ordered n-tuples are defined inductively by

$$(x_1, \ldots, x_n, x_{n+1}) = ((x_1, \ldots, x_n), x_{n+1}).$$

By a generalisation of the above exercise we have that $(x_1, \ldots, x_n) = (y_1, \ldots, y_n)$ iff $x_1 = y_1$ and $\cdots$ and $x_n = y_n$.

C. Products. The *product* $X \times Y$ of two classes X and Y is the class of all ordered pairs (x, y) such that $x \in X$ and $y \in Y$. If X and Y are sets then by Comprehension

$$X \times Y = \{(x, y) \in PP(X \cup Y) \mid x \in X \wedge y \in Y\}$$

is also a set. If X and Y are proper classes then the definition of the class $X \times Y$ can be obtained from the definition of X and Y. (Similar remarks apply throughout.) More generally, the product of sets $X_1, \dots, X_n$ is defined inductively by

$$X_1 \times \cdots \times X_n \times X_{n+1} = (X_1 \times \cdots \times X_n) \times X_{n+1}$$

When $X_1 = \cdots = X_n$ we set $X^n = X_1 \times \cdots \times X_n$.

D. Relations: An *n-ary relation* R on a class X is a subclass of X^n. We shall often write '$R(x_1, \dots x_n)$' instead of '$(x_1, \dots, x_n) \in R$' and when R is binary we will often write '$x \, R \, y$' instead of '$(x, y) \in R$'. For a binary relation R on a class X,

$$\begin{aligned} \text{dom}(R) &= \{x \in X \mid \exists y \, x \, R \, y\}, \\ \text{ran}(R) &= \{y \in X \mid \exists x \, x \, R \, y\}, \text{ and} \\ \text{field}(R) &= \text{dom}(R) \cup \text{ran}(R) \end{aligned}$$

are respectively the *domain*, *range*, and *field* of the relation R. The *inverse* of a binary relation R on X is:

$$R^{-1} = \{(y, x) \mid (x, y) \in R\}.$$

A binary relation R on X is *reflexive* if $\forall x \in X \, (x \, R \, x)$; *symmetric* if $\forall x, y \in X \, (x \, R \, y \leftrightarrow y \, R \, x)$; and *transitive* if $\forall x, y, z \in X \, (x \, R \, y \wedge y \, R \, z \rightarrow x \, R \, z)$. A binary relation is an *equivalence relation* if it is reflexive, symmetric, and transitive. Given an equivalence relation R on a set X and an element $x \in X$, $[x]_R = \{y \in X \mid x \, R \, y\}$ is the *R-equivalence class* of x.

Exercise 1.5. Let R be an equivalence relation on a set X. Show that the R-equivalence classes *partition* X:

(1) $\cup\{[x]_R \mid x \in X\} = X$ and

(2) $\forall x, y \in X \, (\neg(x \, R \, y) \leftrightarrow [x]_R \cap [y]_R = \varnothing)$.

Why did we assume that X is a set?

E. Functions: A *function* is a binary relation f such that each element in $\text{dom}(f)$ is related to only one element in $\text{ran}(f)$: $\forall x \in \text{dom}(f)\, \exists! y \in \text{ran}(f)\, (x\, f\, y)$. In such a situation we can, without ambiguity, use '$f(x)$' to denote the element of $\text{ran}(f)$ that f associates with x.

When we write

$$\begin{aligned} f : X &\to Y \\ x &\mapsto y \end{aligned}$$

this it is to be taken to mean that f is a function such that $\text{dom}(f) \subseteq X$, $\text{ran}(f) \subseteq Y$, and $f(x) = y$. We are thus allowing this notation to include *partial* functions from X to Y, that is, functions $f : X \to Y$ such that $\text{dom}(f) \subsetneq X$. We shall, however, adopt the convention that when we write $f : X \to Y$ without qualification it is to be understood that f is *total*, that is, that $\text{dom}(f) = X$. For set X and Y,

$$Y^X = \{f \mid f : X \to Y\}$$

is thus the set of all functions from X to Y. A function is *onto* if $\text{ran}(f) = Y$. And a function f is *one-to-one* if for all $x, y \in \text{dom}(f)$, if $x \neq y$ then $f(x) \neq f(y)$.

For $A \subseteq X$,

$$f \restriction A = \{(x, y) \in f \mid x \in A\}$$

is the *restriction of f to A* and

$$f``A = \{f(x) \mid x \in A\}.$$

is the *image of A under f*. If f is one-to-one then

$$f^{-1} = \{(y, x) \mid (x, y) \in f\}$$

is a function, called the *inverse* of f.

1.6 Choice

We our now in a position to state our last axiom of ZFC.

(9) (Choice) For every set x consisting of non-empty sets there is a (choice) function f which picks an element from each set $y \in x$:

$$\forall x \big(\forall y \in x\,(y \neq \varnothing) \to \exists f\,(f \text{ is a function}, \\ \operatorname{dom}(f) = x,\ \text{and } \forall y \in x\,(f(y) \in y)\big).$$

Exercise 1.6. Show that Choice is equivalent to:

$$\forall x\, \forall y\, \forall f : x \to y\ \exists g : \operatorname{ran}(f) \to x\ (g \subseteq f^{-1}).$$

Chapter 2

Ordinals

2.1 Transfinite Induction

Let R be a binary relation on a class X. R is *strict* if for all $x \in X$, $\neg(x\,R\,x)$. R is *linear* if for all $x, y \in X$, $x\,R\,y$ or $y\,R\,x$. For each $x \in X$,

$$\mathrm{ext}_R(x) = \{y \in X \mid y\,R\,x\}$$

is the *R-extension* of x. The *transitive closure* $\mathrm{TC}(R)$ is the intersection of all transitive relations containing R. For each $x \in R$,

$$\mathrm{pred}_R(x) = \{y \in X \mid y\,\mathrm{TC}(R)\,x\}$$

is the class of *R-predecessors* of x. A subset Y of X is *R-transitive* if for all $x \in Y$, $\mathrm{pred}_R(x) \subseteq Y$. An element $x \in Y \subseteq X$ is *R-minimal* (with respect to Y) if there is no $y \in Y$ such that $y \neq x$ and $y\,R\,x$. As a non-mathematical example consider the binary relation R of *being a parent of*. Then $\mathrm{ext}_R(x)$ is the set of parents of x, $\mathrm{TC}(R)$ is the relation of *being an ancestor of*, $\mathrm{pred}_R(x)$ is the set of ancestors of x and an element x is R-minimal just in case x does not have parents.

Definition 2.1. Let R be a binary relation on a class X. R is *well-founded* iff

(1) every non-empty set $Y \subseteq X$ has an R-minimal element (with respect to Y) and

(2) for all $x \in X$, $\mathrm{ext}_R(x)$ is a set.

Definition 2.2. Let R be a binary relation on a class X. R is *well-ordered* iff

(1) R is well-founded and

(2) R is strict and linear.

When we say '(R, X) is well-founded' we mean 'R is a well-founded relation on X'. Since the field X of R can be extracted from R we will often write 'R' instead of '(X, R)'.

Theorem 2.3. (Transfinite Induction) *Suppose R is a well-founded relation on X. If $Y \subseteq X$ is such that for all $x \in Y$*

$$\mathrm{pred}_R(x) \subseteq Y \to x \in Y$$

then $Y = X$.

Proof. Suppose for contradiction that $Y \subsetneq X$. Letting a be an R-minimal element in $X - Y$, we have that

$$\mathrm{pred}_R(a) \subseteq Y.$$

But then $a \in Y$, which is a contradiction. □

Theorem 2.4. (Definition by Transfinite Recursion) *Suppose R is a well-founded relation on X. Suppose $G : X \times V \to V$. Then there is a unique function $F : X \to V$ such that*

$$\forall x \in X \left(F(x) = G(x, F{\restriction}\mathrm{pred}_R(x))\right).$$

Proof. Uniqueness: Suppose F_1 and F_2 are two such functions. Let

$$Y = \{x \in X \mid F_1(x) = F_2(x)\}.$$

Since Y is such that for all $x \in Y$,

$$\mathrm{pred}_R(x) \subseteq Y \to x \in Y,$$

it follows, by transfinite induction, that $Y = X$, i.e. $F_1 = F_2$.

Existence: Let us say that $f : D \to V$ is *good* if

(1) $D \subseteq X$ and $\forall x \in D \, (\mathrm{pred}_R(x) \subseteq D)$ and

(2) $\forall x \in D\,\big(f(x) = G(x, F\restriction \mathrm{pred}_R(x))\big)$.

By the proof of uniqueness we have that good functions agree on their common domain; that is, if $f_1 : D_1 \to V$ and $f_2 : D_2 \to V$ are good, then

$$f_1\restriction D_1 \cap D_2 = f_2\restriction D_1 \cap D_2.$$

Now, if $f : \mathrm{pred}_R(x) \to V$ is good, then $f' : \{x\} \cup \mathrm{pred}_R(x) \to V$ is good where

$$f'(y) = \begin{cases} f(y) & \text{if } y \in \mathrm{pred}_R(x) \\ G(x, f\restriction \mathrm{pred}_R(x)) & \text{if } y = x. \end{cases}$$

So, by transfinite induction, we have that $\cup\{\mathrm{dom}(f) \mid f \text{ is good}\} = X$. Since all good functions agree on their common domain we have that $F = \cup\{f \mid f \text{ is good}\}$ is as desired. □

Exercise 2.1. Why didn't we just set $F(x) = G(F\restriction \mathrm{pred}_R(x))$?

Suppose that (X, R) is a well-ordering. For $x \in R$, let

$$\mathrm{I}^R_x = \big(\mathrm{pred}_R(x), R \cap (\mathrm{pred}_R(x) \times \mathrm{pred}_R(x))\big).$$

I^R_x and R itself are called *R-initial segments* of R.

Two well-orderings (X, R) and (Y, S) are *isomorphic*—written '$(X, R) \cong (Y, S)$' or '$R \cong S$'—if there is a function $f : X \to Y$ that is onto, one-to-one and order-preserving (i.e. $x\,R\,y \leftrightarrow f(x)\,S\,f(y)$ for all $x, y \in X$).

Theorem 2.5. (Comparability) *Suppose (X, R) and (Y, S) are well-orderings. Then either*

(1) $\mathrm{I}^R_x \cong R$ *for some* $x \in X$, *or*

(2) $R \cong \mathrm{I}^S_y$ *for some* $y \in Y$, *or*

(3) $R \cong S$.

Proof. For $A \subseteq X$, $f : A \to Y$, and $x \in X$ let

$$G(x, f) = \begin{cases} \text{the } S\text{-least element in } Y - f``\mathrm{pred}_R(x) & \text{if such exists} \\ \text{undefined} & \text{otherwise} \end{cases}$$

and let F be defined from G by transfinite recursion. By transfinite induction we have:

$$\forall x \in \mathrm{dom}(F) \left(F``\mathrm{pred}_R(x) = \mathrm{pred}_S(F(x))\right).$$

So if $a\,R\,b$ then $F(b) \notin \mathrm{pred}_S(F(a))$ and hence $F(a)\,S\,F(b)$; and if $\neg(a\,R\,b)$, then $b\,R\,a$, and so $F(b)\,S\,F(a)$, and hence $\neg(F(a)\,S\,F(b))$. Thus F is order-preserving. It is clear that F is one-to-one. There are three cases

(1) $\exists x \in X\,(\mathrm{dom}(F) = \mathrm{pred}_R(x))$

(2) $\mathrm{dom}(F) = X$ and $\exists y \in Y\,(\mathrm{ran}(F) = \mathrm{pred}_S(y))$

(3) $\mathrm{dom}(F) = X$ and $\mathrm{ran}(F) = Y$

which imply the corresponding claims in the statement of the theorem. □

2.2 Ordinals

The class W of well-orderings in V is partitioned into equivalence classes by the relation $\cong$ of isomorphism. We now wish to select a canonical representative from each of these equivalence classes; that is, we seek a class $C \subseteq W$ such that each element of C is *canonical*, the elements of C are *linearly-ordered* (under initial segment), and C is *comprehensive* in the sense that for every element of W there is an element of C to which it is isomorphic.

Let us motivate our eventual choice: In order to ensure comprehensiveness we need to ensure that whenever C contains an element R there is another element R' of C which is minimally longer than R. Given R we can obtain such an R' by picking *any* element z that is not in $X = \mathrm{dom}(R)$ and place it after all of the elements of X. But to ensure canonicity we have to take care to choose an element Z in such a way that it is canonically related to X. What is the simplest set that is not in X and is canonically related to X? Answer: X. So as the domain of our relation R' we take $X \cup \{X\}$ and as our ordering we take the ordering that coincides with R on X and coincides with $\in$ with respect to the new element. Notice that the predecessors of X are precisely its members. Furthermore, had we chosen our elements canonically at each stage then this would be true of each element. Let us call a set *transitive* if it is $\in$-transitive; in other words a set X is transitive if whenever $x \in y \in X$, $x \in X$. Thus an element of C, if it is to be canonical, must be transitive and well-ordered by $\in$.

Definition 2.6. An *ordinal* is a set that is transitive and well-ordered by $\in$. For ordinals α and β we set $\alpha < \beta$ if and only if $\alpha \in \beta$ and say that α is *less than* β .

The first four ordinals are:

$$\varnothing,\ \{\varnothing\},\ \{\varnothing,\{\varnothing\}\},\ \{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}.$$

We shall write the above ordinals as $0, 1, 2, 3$ and so on through the natural numbers. As variables and constants for ordinals we take the lower case Greek letters: α, β, γ, ... The class of ordinals is denoted by 'On'.

Lemma 2.7. *Assume* $\alpha, \beta \in \mathrm{On}$*. The following are equivalent:*

(1) $\alpha \in \beta$

(2) $\alpha \subsetneq \beta$.

Proof. (1) $\to$ (2): Suppose $\gamma \in \alpha \in \beta$. Then $\gamma \in \beta$ by transitivity. So $\alpha \subseteq \beta$. Clearly, $\alpha \neq \beta$ since otherwise we would have $\beta \in \beta$.

(2) $\to$ (1): Assume $\alpha \subsetneq \beta$. Let γ be the $\in$-least element in $\beta - \alpha$. Note $\gamma \in \mathrm{On}$. Now, $\alpha = \mathrm{pred}_{(\beta,\in)}(\gamma)$ and so $\alpha = \gamma$ by Exercise 2.2(2). Thus $\alpha \in \beta$. □

Thus the ordering $<$ on On is just $\subsetneq$ and $\leqslant$ is just $\subseteq$.

Theorem 2.8. (Linearity) *If* $\alpha, \beta \in \mathrm{On}$ *then* $\alpha \leqslant \beta$ *or* $\beta \leqslant \alpha$*.*

Proof. We have that $\alpha \cap \beta \in \mathrm{On}$. We have to show that either $\alpha \cap \beta = \alpha$ or $\alpha \cap \beta = \beta$. Assume for contradiction that $\alpha \cap \beta \subsetneq \alpha$ and $\alpha \cap \beta \subsetneq \beta$. By the previous lemma this means that $\alpha \cap \beta \in \alpha$ and $\alpha \cap \beta \in \beta$, i.e. $\alpha \cap \beta \in \alpha \cap \beta$, which is a contradiction. □

Exercise 2.2. Prove the following:

(1) If $\alpha \in \mathrm{On}$ and $\beta \in \alpha$ then $\beta \in \mathrm{On}$.

(2) If $\alpha \in \mathrm{On}$, then $\alpha = \{\beta \mid \beta < \alpha\}$.

(3) If $C \subseteq \mathrm{On}$ is a non-empty class then $\bigcap C \in \mathrm{On}$ and $\bigcap C$ is the largest ordinal less than or equal to every ordinal in C.

(4) If $C \subseteq \mathrm{On}$ is a non-empty set then $\bigcup C \in \mathrm{On}$ and $\bigcup C$ is the least ordinal greater than or equal to every ordinal in C.

(5) If $\alpha \in \mathrm{On}$ then $\alpha \cup \{\alpha\}$ is the least ordinal greater than α.

Definition 2.9. For $\alpha \in \mathrm{On}$, $\alpha + 1 = \alpha \cup \{\alpha\}$.

Theorem 2.10. (Comprehensiveness) *For every well-ordering (X, R) there is an ordinal α such that $(X, R) \cong (\alpha, \in)$.*

Proof. Fix a well-ordering (X, R).

Claim. $\forall a \in X\, \exists! \alpha \in \mathrm{On}\, (\mathrm{I}_a^R \cong (\alpha, \in))$.

Proof. Suppose not. Let $a \in X$ be the R-least counter-example. Then

$$\forall b \in X\, \big(b\, R\, a \to \exists! \beta \in \mathrm{On}\, (\mathrm{I}_b^R \cong (\beta, \in))\big).$$

By Replacement let

$$\alpha = \{\beta \mid \exists\, b\, R\, a\, (\mathrm{I}_b^R \cong (\beta, \in))\}.$$

Note that $\alpha \in \mathrm{On}$ and

$$\mathrm{I}_a^R \cong (\alpha, \in)$$

(as witnessed by the map $\pi : X \to \alpha$ sending b to the unique β such that $\mathrm{I}_b^R \cong (\beta, \in)$). It is easy to see that there is a unique such α. This contradicts our assumption that a was a counterexample. □

Now do the same thing again. By Replacement let

$$\gamma = \{\beta \mid \exists\, b \in X\, (\mathrm{I}_b^R \cong (\beta, \in))\}.$$

We have $\gamma \in \mathrm{On}$ and $(X, R) \cong (\gamma, \in)$. □

Definition 2.11. Suppose (X, R) is a well-ordering. Then the *ordertype* $\mathrm{ot}(X, R)$ of (X, R) is the unique ordinal α such that $(X, R) \cong (\alpha, \in)$.

Definition 2.12. α is a *successor ordinal* iff there exists a largest $\beta < \alpha$. α is a *limit* ordinal iff there does not exist a largest $\beta < \alpha$.

Exercise 2.3. Show that there are arbitrarily large limit ordinals.

The least limit ordinal greater than 0 is denoted 'ω'.

2.3 Ordinal Arithmetic

Definition 2.13. Addition, multiplication, and exponentiation of ordinals is defined by transfinite recursion as follows:

(1) (i) $\alpha + 0 = \alpha$
  (ii) $\alpha + (\beta + 1) = (\alpha + \beta) + 1$ for all β
  (iii) $\alpha + \lambda = \bigcup_{\beta<\lambda} \alpha + \beta$ for all limit ordinals $\lambda > 0$

(2) (i) $\alpha \cdot 0 = 0$
  (ii) $\alpha \cdot (\beta + 1) = (\alpha \cdot \beta) + a$ for all β
  (iii) $\alpha \cdot \lambda = \bigcup_{\beta<\lambda} \alpha \cdot \beta$ for all limit ordinals $\lambda > 0$

(3) (i) $\alpha^0 = 1$
  (ii) $\alpha^{\beta+1} = \alpha^\beta \cdot \alpha$ for all β
  (iii) $\alpha^\lambda = \bigcup_{\beta<\lambda} \alpha^\beta$ for all limit ordinals $\lambda > 0$

It is of use to get a more concrete picture of these operations.

Given two well-ordered sets (X, R) and (Y, S) the lexicographic ordering on $X \times Y$ is:

$$(x_1, y_1) <_{\text{lex}} (x_2, y_2) \leftrightarrow x_1 \, R \, x_2 \vee (x_1 = x_2 \wedge y_1 \, S \, y_2).$$

For ordinals α and β let

$$\mathscr{F}(\beta, \alpha) = \{f : \beta \to \alpha \mid f(\gamma) = 0 \text{ for all but finitely many } \gamma\}$$

and for $f, g \in \mathscr{F}(\beta, \alpha)$ define

$$f <^* g \leftrightarrow f \neq g \text{ and if } \gamma \in \beta \text{ is largest such that } f(\gamma) \neq g(\gamma) \text{ then } f(\gamma) < g(\gamma).$$

Exercise 2.4. Show that for all ordinals α and β,

$$\alpha + \beta = \text{ot}((\{0\} \times \alpha) \cup (\{1\} \times \beta), <_{\text{lex}})$$
$$\alpha \cdot \beta = \text{ot}(\beta \times \alpha, <_{\text{lex}})$$
$$\alpha^\beta = \text{ot}(\mathscr{F}(\beta, \alpha), <^*).$$

Exercise 2.5. Show that

(1) $1 + \omega = \omega$ and $\omega + 1 \neq \omega$

(2) $2 \cdot \omega = \omega$ and $\omega \cdot 2 = \omega + \omega$.

Exercise 2.6. Suppose $\beta > 0$. Show that for any ordinal α there is a unique γ_1 and a unique $\gamma_2 < \beta$ such that $\alpha = \beta \cdot \gamma_1 + \gamma_2$.

Theorem 2.14. (Cantor Normal Form) *Suppose $\alpha > 0$. Then α can be uniquely written in the form*

$$\omega^{\beta_1} \cdot k_1 + \cdots + \omega^{\beta_n} \cdot k_n$$

where $n > 0$, $\alpha \geqslant \beta_1 > \cdots > \beta_n$ and $k_1, \ldots, k_n$ are non-zero natural numbers.

Proof. This is proved by transfinite induction on ordinals α. The base case is trivial since $1 = \omega^0 \cdot 1$. Assume then that we have proved the theorem for all $\bar{\alpha} < \alpha$. Choose β to be greatest such that $\omega^\beta \leqslant \alpha$. By the above exercise there is a unique γ_1 and a unique $\gamma_2 < \omega^\beta$ such that $\alpha = \omega^\beta \cdot \gamma_1 + \gamma_2$. Note that γ_1 is finite since otherwise $\omega^{\beta+1} \leqslant \alpha$. If $\gamma_2 = 0$ then we are done. If $\gamma_2 > 0$ then, since $\gamma_2 < \alpha$ it has the desired form, by our induction hypothesis. Thus we have that every ordinal can be written in Cantor normal form. A simple induction establishes uniqueness. □

2.4 The Hierarchy

We are now in a position to give a more precise characterisation of the hierarchy of sets. Let

$$\begin{aligned} V_0 &= \varnothing \\ V_{\alpha+1} &= P(V_\alpha) \\ V_\lambda &= \textstyle\bigcup_{\alpha<\lambda} V_\alpha \quad \text{for } \lambda \text{ a limit ordinal.} \end{aligned}$$

Exercise 2.7. Recall that we set $V = \{x \mid x = x\}$. Show that $V = \bigcup_{\alpha \in \mathrm{On}} V_\alpha$.

Definition 2.15. The *rank* of a set x (abbreviated $\mathrm{rank}(x)$) is the least α such that $x \in V_{\alpha+1}$.

Exercise 2.8. By transfinite recursion define

$$\begin{aligned} \rho(\varnothing) &= 0 \\ \rho(x) &= \textstyle\bigcup_{y \in x}(\rho(y) + 1). \end{aligned}$$

Show that for all x,

$$\mathrm{rank}(x) = \rho(x).$$

Chapter 3

Cardinals

The axiom of choice is used at certain points in the theory of cardinals. We begin with a brief discussion of this and related axioms and then develop the theory of cardinals, prolonging appeal to Choice for as long as possible.

3.1 Choice

Theorem 3.1. (Zermelo, 1904) *Assume* AC*. Then any set X can be well-ordered.*

Proof. Let $f : P(X) - \{\varnothing\} \to X$ be a choice function. Define $h : \beta \to X$ by transfinite recursion by the condition: For each α, set

$$h(\alpha) = f(X - \operatorname{ran}(h \restriction \alpha))$$

provided $\operatorname{ran}(h \restriction \alpha) \subsetneq X$ and let β be the least ordinal α such that $\operatorname{ran}(h \restriction \alpha) = X$. Letting R be such that $a\,R\,b$ iff $h^{-1}(a) < h^{-1}(b)$ we have that R well-orders X. □

Exercise 3.1. Assume that any set X can be well-ordered. Prove AC.

Exercise 3.2. A *partially ordering* is a relation (X, R) which is strict and transitive. Let $C \subseteq X$. C is a *chain* if C is linearly ordered by R; an element $x \in X$ is an *upper bound* of C if for every $c \in C$, $c\,R\,x$ or $c = x$; an element $x \in X$ is *maximal* if there is no element $x' \in X$ such that $x\,R\,x'$. Prove that the following are equivalent:

(1) AC

(2) (Zorn's Lemma) If (X, R) is a non-empty partial order such that every chain has an upper bound then (X, R) has a maximal element.

There is a very important consequence of Choice, one that is so simple that it is easy to overlook.

Definition 3.2. (Dependent Choice (DC)) Suppose that (X, R) is a non-empty binary relation such that for every $x \in X$ there is a $y \in X$ such that $x \, R \, y$. Then there is an infinite sequence $x_0, x_1, x_2, \ldots, x_n, \ldots$ such that for all $n \in \omega$, $x_n \, R \, x_{n+1}$.

This is the principle that was implicit in the informal proof of Exercise 1.2.

Exercise 3.3. Show that AC implies DC.

A set X is *countable* if there is a function $f : X \to \omega$ that is one-to-one and onto.

Definition 3.3. (Countable Choice (AC_ω) Every countable set of non-empty sets admits a choice function.

(We will use $\mathrm{AC}_\omega(X)$ to abbreviate the statement that every countable set of non-empty subsets of X admits a choice function.)

Exercise 3.4. Show that DC implies AC_ω.

3.2 Cardinality

Definition 3.4. For sets X and Y, we write $X \approx Y$ and say that X is *equinumerous* with Y if there is a function $f : X \to Y$ that is one-to-one and onto. We write $X \preccurlyeq Y$ if there is a function $f : X \to Y$ that is one-to-one.

Exercise 3.5. Let $E = \{0, 2, 4, \ldots, 2n, \ldots\}$. Show that

(1) $E \approx \omega$

(2) $\omega \times \omega \approx \omega$.

Theorem 3.5. (Cantor-Bernstein) $X \preccurlyeq Y \wedge Y \preccurlyeq X \to X \approx Y$.

Proof. Suppose $f : X \to Y$ and $g : Y \to X$ are one-to-one. Define X_n and Y_n by induction on n as follows: $X_0 = X$, $Y_0 = Y$, $X_{n+1} = g``Y_n$, $Y_{n+1} = f``X_n$. Notice that

$$X_0 \supseteq X_1 \supseteq \cdots \quad \text{and} \quad Y_0 \supseteq Y_1 \supseteq \cdots$$

Let

$$X_\infty = \bigcap_{n<\omega} X_n \quad \text{and} \quad Y_\infty = \bigcap_{n<\omega} Y_n.$$

Define $h : X \to Y$ by

$$h(x) = \begin{cases} f(x) & \text{if } x \in X_n - X_{n+1} \text{ for some even } n \\ g^{-1}(x) & \text{if } x \in X_n - X_{n+1} \text{ for some odd } n \\ f(x) & \text{if } x \in X_\infty. \end{cases}$$

We have that

$$\begin{aligned} f\restriction X_{2n} - X_{2n+1} &\to Y_{2n+1} - Y_{2n+2}, \\ g^{-1}\restriction X_{2n+1} - X_{2n+2} &\to Y_{2n} - Y_{2n+1} \\ f\restriction X_\infty &\to Y_\infty \end{aligned}$$

are one-to-one and onto. Thus h is as desired. □

Exercise 3.6. In the above proof how do we know that $f\restriction X_\infty \to Y_\infty$ is onto? Do we necessarily have $f\restriction X_\infty = g^{-1}\restriction Y_\infty$?

Definition 3.6. An ordinal α is a *cardinal* if there is no ordinal $\beta < \alpha$ such that $\beta \approx \alpha$.

The finite cardinals are $0, 1, 2, 3, \ldots$ and the first infinite cardinal is ω. As variables and constants for cardinals we take the lower case Greek letters: $\kappa, \lambda, \ldots$ The class of cardinals is denoted 'CARD'.

Our next goal is to show that CARD is a proper class.

Theorem 3.7. (Cantor) *For any set X, $P(X) \not\approx X$.*

Proof. Suppose, for contradiction, that X is such that $P(X) \approx X$ and let $f : X \to P(X)$ be onto. Consider the set

$$D = \{x \in X \mid x \notin f(x)\}.$$

Since f is onto, $D \in \operatorname{ran}(f)$. Let $d \in X$ be such that $f(d) = D$. Now, if $d \in D$, then $d \notin f(d) = D$; and if $d \notin D = f(d)$ then $d \in D$. So we have a contradiction. □

Assuming AC one can use this theorem to show that CARD is a proper class. For AC implies that every set can be well-ordered and so we have that every set is equinumerous to a cardinal number. In particular, for any given cardinal number κ, $P(\kappa)$ is equinumerous to a cardinal number which, by Cantor's theorem, is greater than κ. One can circumvent this appeal to Choice.

Exercise 3.7. Without assuming AC show that CARD is a proper class.

Thus the following definition is legitimate in ZF.

Definition 3.8. For each cardinal κ, κ^+ is the least cardinal greater than κ.

Lemma 3.9. *If $A \subseteq$ CARD is a set, then $\bigcup A \in$ CARD.*

Proof. If A has a largest element κ the $\cup A = \kappa \in$ CARD. Assume then, for contradiction, that A does not have a largest element and $\cup A \notin$ CARD. Then there is an $\alpha \in \cup A$ such that $\alpha \approx \cup A$. Now choose $\kappa \in \cup A$ such that $\alpha < \kappa$. Since $\kappa \subseteq \cup A \approx \alpha$, $\kappa \preccurlyeq \alpha$, which is impossible. □

We can now describe the members of CARD.

Definition 3.10. By transfinite recursion on On, let

$$\begin{aligned} \aleph_0 &= \omega \\ \aleph_{\alpha+1} &= (\aleph_\alpha)^+ \\ \aleph_\lambda &= \textstyle\bigcup_{\alpha<\lambda} \aleph_\alpha \text{ for limit ordinals } \lambda. \end{aligned}$$

We will also write ω_α for $\aleph_\alpha$.

Theorem 3.11. *If $\kappa \in$ CARD then either κ is finite or $\kappa = \aleph_\alpha$ for some $\alpha \in$ On.*

Proof. Suppose $\kappa \in$ CARD and κ is not finite. Let α be least such that $\aleph_\alpha \geqslant \kappa$. Either α is a successor or α is a limit. If α is a successor, say $\alpha = \beta + 1$, then $\aleph_\beta < \kappa$ and so $\aleph_\alpha = \kappa$ as otherwise we violate the minimality of α. If α is a limit then for every $\gamma < \alpha$, $\aleph_\gamma < \kappa$ and so $\aleph_\alpha = \kappa$ as otherwise we violate the minimality of α. □

Exercise 3.8. Show that $\mathbb{R} \approx \omega^\omega \approx 2^\omega$.

Continuum Hypothesis (CH). (Cantor, 1878) *Every subset of 2^ω that cannot be mapped one-to-one into ω is equinumerous with 2^ω.*

It is important to note that we have made no appeal to Choice in this section.

3.3 Cardinal Arithmetic

A. Addition and Multiplication. Recall that ordinal addition and multiplication are such that

$$\begin{aligned}\alpha+\beta &= \mathrm{ot}((\{0\}\times\alpha)\cup(\{1\}\times\beta),<_{\mathrm{lex}})\\ \alpha\cdot\beta &= \mathrm{ot}(\beta\times\alpha,<_{\mathrm{lex}}).\end{aligned}$$

The corresponding operations on cardinals are obtained by "ignoring order".

Definition 3.12. (AC) $|X|$ is the least ordinal α such that $X\approx\alpha$.

Definition 3.13. For cardinals κ and λ,

$$\begin{aligned}\kappa+\lambda &= |(\{0\}\times\kappa)\cup(\{1\}\times\lambda)|\\ \kappa\cdot\lambda &= |\kappa\times\lambda|\end{aligned}$$

are the operations of cardinal addition and multiplication.

Definition 3.14. (AC) A set X is *finite* if $|X|<\omega$, *infinite* if $|X|\geqslant\omega$, *countable* if $|X|\leqslant\omega$, and *uncountable* if $|X|>\omega$.

Exercise 3.9. Recall that a set X is *finite* iff $|X|<\omega$. Let us say that a set X is *D-finite* if there is no $f:X\to X$ such that f is one-to-one and $\mathrm{ran}(f)\subsetneq X$. Let us say that a set X is *D^*-finite* if there exists $f:X\to X$ such that there does not exist a non-empty $Y\subsetneq X$, $f``Y\subseteq Y$.

(1) Show that X is D-finite iff X is finite. Which axioms did you need?

(2) Show that X is D*-finite iff X is finite. Which axioms did you need?

Notice that we do not need AC to draw the distinction between finite, infinite, countable and uncountable *ordinals*. Let us then continue for the time being without AC.

Clearly the operations of cardinal addition and multiplication differ on finite sets. Surprisingly they coincide on infinite sets.

Theorem 3.15. *If κ is an infinite cardinal then $\kappa\cdot\kappa=\kappa$.*

Proof. Assume that the theorem holds for all cardinals $\alpha < \kappa$. Our goal is to define an ordering of $\kappa \times \kappa$ of ordertype κ. For $\alpha, \beta, \gamma, \delta$ set

$$\begin{aligned}(\alpha,\beta) \lhd (\gamma,\delta) \text{ iff either } &\max(\alpha,\beta) < \max(\gamma,\delta)\\ &\text{or } \max(\alpha,\beta) = \max(\gamma,\delta) \wedge (\alpha,\beta) <_{\text{lex}} (\gamma,\delta).\end{aligned}$$

Each $(\alpha,\beta) \in \kappa \times \kappa$ has at most $|\max(\alpha,\beta)+1 \times \max(\alpha,\beta)+1| = \bar{\kappa}$-many $\lhd$-predecessors. By the induction hypothesis, $\bar{\kappa} < \kappa$ and so $\text{ot}(\kappa \times \kappa) \leqslant \kappa$. □

Corollary 3.16. *If κ and λ are non-zero cardinals such that one is infinite then $\kappa + \lambda = \kappa \cdot \lambda = \max(\kappa, \lambda)$.*

In short, the operations of cardinal addition and multiplication are straightforward. On finite cardinals these operations coincide with the standard operations. And in the case where one cardinal is infinite these operations collapse into the trivial operation of maximisation.

Note that in this subsection we have made no appeal to AC.

B. Exponentiation. Recall that ordinal exponentiation is such that

$$\alpha^\beta = \text{ot}(\mathscr{F}(\beta,\alpha), <^*)$$

where

$$\mathscr{F}(\beta,\alpha) = \{f : \beta \to \alpha \mid f(\gamma) = 0 \text{ for all but finitely many } \gamma\}$$

and for $f, g \in \mathscr{F}(\beta,\alpha)$

$$\begin{aligned}f <^* g \leftrightarrow f \neq g \text{ and if } &\gamma \in \beta \text{ is largest such that}\\ &f(\gamma) \neq g(\gamma) \text{ then } f(\gamma) < g(\gamma).\end{aligned}$$

The corresponding operation on cardinals is *not* obtained by "ignoring order".

Definition 3.17. For cardinals κ and λ, $\kappa^\lambda = |\kappa^\lambda|$ is the operation of cardinal exponentiation.

(In the above definition the second instance of 'κ^λ' stands for the set of functions from λ to κ. It is unfortunate that the same notation is used for this set and the corresponding cardinal. To remedy this defect some authors write '${}^{\lambda}\kappa$' for the set of functions. We will not follow this course, but instead rely on context to disambiguate.)

In contrast to the preceding section we now need to invoke AC. For in contrast to the sets $(\{0\} \times \kappa) \cup (\{1\} \times \lambda)$ and $\kappa \times \lambda$ there is no natural way to well-order κ^λ. (Indeed one cannot prove in ZF that there is a definable well-ordering of κ^λ.)

Lemma 3.18. $(\kappa^\lambda)^\mu = \kappa^{\lambda \cdot \mu}$.

Proof. We need to show that there is a bijection between $\{f : \lambda \times \mu \to \kappa\}$ and $\{f : \mu \to \kappa^\lambda\}$. For each function f in the first set let, for $\alpha < \mu$, $f_\alpha : \lambda \to \kappa$ be defined by $f_\alpha(\beta) = f(\beta, \alpha)$, and let $f^* : \mu \to \kappa^\lambda$ be defined by $f^* : \alpha \to f_\alpha$. Then the map sending f to f^* is a bijection between the two sets. □

Cardinal exponentiation is highly non-trivial. The goal is to determine the operations $\kappa \mapsto 2^\kappa$ and $(\kappa, \lambda) \mapsto \kappa^\lambda$. It is useful to examine the relationship between these two operations. There are two cases: (1) $\lambda \geqslant \kappa$ and (2) $\lambda < \kappa$. The first case is quite simple.

Lemma 3.19. *If $2 \leqslant \kappa \leqslant \lambda$ and λ is infinite, then $\kappa^\lambda = 2^\lambda$.*

Proof.

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

□

The second case is quite complicated. To state one of the fundamental results in this area we need to introduce the notion of cofinality.

3.4 Cofinality

Definition 3.20. Let α be a limit ordinal. The *cofinality* $\mathrm{cof}(\alpha)$ of α is the least ordinal β such that there is a function $f : \beta \to \alpha$ that is cofinal in the sense that for all $\bar{\alpha} < \alpha$ there is a $\bar{\beta} < \beta$ such that $f(\bar{\beta}) \geqslant \bar{\alpha}$.

Notice that $\mathrm{cof}(\alpha)$ is a limit ordinal.

Definition 3.21. A limit ordinal α is *regular* if $\mathrm{cof}(\alpha) = \alpha$ and *singular* otherwise.

Lemma 3.22. *Suppose that α is a limit ordinal. Then there is map $f : \mathrm{cof}(\alpha) \to \alpha$ that is cofinal and increasing (in the sense that if $\alpha < \beta$ then $f(\alpha) < f(\beta)$).*

Proof. Let $g : \mathrm{cof}(\alpha) \to \alpha$ be cofinal. We extract an increasing cofinal map from g by setting

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

□

Lemma 3.23. *Let α and β be limit ordinals and suppose $f : \alpha \to \beta$ is cofinal and non-decreasing (in the sense that if $\beta < \gamma < \alpha$ then $f(\beta) \leqslant f(\gamma)$). Then $\mathrm{cof}(a) = \mathrm{cof}(\beta)$.*

Proof. (1) $\mathrm{cof}(\beta) \leqslant \mathrm{cof}(\alpha)$: Let $g : \mathrm{cof}(\alpha) \to \alpha$ be cofinal. Then $g \circ f : \mathrm{cof}(\alpha) \to \beta$ is cofinal. (2) $\mathrm{cof}(\alpha) \leqslant \mathrm{cof}(\beta)$: Let $h : \mathrm{cof}(\beta) \to \beta$ be cofinal. Define $k : \mathrm{cof}(\beta) \to \alpha$ by $k(\gamma) =$ the least δ such that $\delta > k(\bar{\gamma})$ for all $\bar{\gamma} < \gamma$ and $f(\delta) > h(\gamma)$. □

Lemma 3.24. *If α is a limit ordinal then $\mathrm{cof}(\alpha)$ is a regular cardinal.*

Proof. First, $\mathrm{cof}(\alpha)$ is a limit ordinal. Second, $\mathrm{cof}(\alpha)$ is regular since $\mathrm{cof}(\mathrm{cof}(\alpha)) = \mathrm{cof}(\alpha)$. Third, $\mathrm{cof}(\alpha)$ is a cardinal since otherwise the bijection with a smaller ordinal would witness that $\mathrm{cof}(\alpha)$ is not regular. □

Up until this point we have made no appeal to Choice. But we now need it for the following basic result.

Lemma 3.25. (AC) *Suppose $|A| = \kappa^+$. Then there is no collection of sets $\{A_\alpha \mid \alpha < \kappa\}$ such that for each $\alpha < \kappa$, $|A_\alpha| \leqslant \kappa$ and $A = \bigcup_{\alpha<\kappa} A_\alpha$.*

Proof. By AC we can choose, for each α, a function $f_\alpha : A_\alpha \to \kappa$. Let $g : \kappa \times \kappa \to \kappa$ be a bijection and let

$$\begin{aligned} h : \textstyle\bigcup_{\alpha<\kappa} A_\alpha &\to \kappa \times \kappa \\ x &\mapsto (\alpha, f_\alpha(x)). \end{aligned}$$

Then $h \circ g$ witnesses that $\bigcup_{\alpha<\kappa} A_\alpha \leqslant \kappa$. □

Theorem 3.26. (AC) *For every cardinal κ, κ^+ is regular.*

Proof. Suppose, for contradiction, that κ^+ is not regular and let $f : \bar{\kappa} \to \kappa^+$ be cofinal, where $\bar{\kappa} \leqslant \kappa$. Then we would be able to write κ^+ as a union of less than or equal to κ-many sets of size less than or equal to κ, contradicting the previous lemma. □

Thus in ZFC all successor cardinals are regular. But there are many limit cardinals which are singular. For example, $\aleph_\omega$ is a singular cardinal. Are there limit cardinals that are regular?

Definition 3.27. κ is *weakly inaccessible* if $\kappa > \omega$ and κ is a regular limit cardinal. κ is a *strong limit* if for all $\lambda < \kappa$, $2^\lambda < \kappa$. κ is (*strongly*) *inaccessible* if κ is weakly inaccessible and a strong limit.

As we shall see later, one cannot prove the existence of weakly inaccessible cardinals in ZFC.

Theorem 3.28. (König) *For every infinite cardinal κ, $\kappa^{\mathrm{cof}(\kappa)} > \kappa$.*

Proof. Let $f : \mathrm{cof}(\kappa) \to \kappa$ be cofinal. Suppose, for contradiction, that $\kappa^{\mathrm{cof}(\kappa)} = \kappa$. Let $\{g_\alpha \mid \alpha < \kappa\}$ enumerate all of the functions from $\mathrm{cof}(\kappa)$ into κ. Now define

$$\begin{aligned} g &: \mathrm{cof}(\kappa) \to \kappa \\ &\beta \mapsto \text{the least element in } \kappa - \{g_\alpha(\beta) \mid \alpha < f(\beta)\}. \end{aligned}$$

Note that for no $\alpha < \kappa$ do we have $g = g_\alpha$. □

Chapter 4

Trees

4.1 Suslin's Hypothesis

A *linear ordering* is a structure $(L, \leqslant_L)$ such that $\leqslant_L$ is a relation on L such that for all $x, y, z \in L$, $x \leqslant_L x$, $x \leqslant_L y \wedge y \leqslant_L x \rightarrow x = y$, $x \leqslant_L y \leqslant_L z \rightarrow x \leqslant_L z$, and $x \leqslant_L y \vee y \leqslant_L x$. We will often identify a structure with its underlying domain, writing 'L' instead of '$(L, \leqslant_L)$'. Define $x <_L y$ to be $x \leqslant_L y \vee x \neq y$. A linear ordering $(L, \leqslant_L)$ is *dense* if whenever $x, z \in L$ and $x <_L z$ there is a $y \in L$ such that $x <_L y <_L z$. A linear ordering is *without endpoints* if for all $y \in L$ there exists $x, z \in L$ such that $x <_L y <_L z$.

Theorem 4.1. (Cantor) *Suppose $(A, \leqslant_A)$ and $(B, \leqslant_B)$ are countable dense linear orderings without endpoints. Then $(A, \leqslant_A) \cong (B, \leqslant_B)$.*

Proof. Let $A = \{a_n \mid n < \omega\}$ and $B = \{b_n \mid n < \omega\}$. We shall construct an isomorphism $f : A \to B$ step-wise by a "back and forth" argument. Suppose that we have defined $f(a_0), f^{-1}(b_0), \ldots, f(a_n), f^{-1}(b_n)$. If $f(a_{n+1})$ has not already been defined then we let it be any element in B such that so defined f is order preserving. We can always choose such an element since B is dense and without endpoints and we have only used up finitely many elements. $f^{-1}(b_{n+1})$ is defined similarly using the density of A. □

Thus $(\mathbb{Q}, \leqslant)$ is uniquely characterised (up to isomorphism) as a countable dense linear ordering without endpoints.

A linear ordering $(L, \leqslant_L)$ is *complete* if every subset $X \subseteq L$ that is bounded above has a least upper bound and every subset $X \subseteq L$ that is bounded below has a greatest lower bound. A linear ordering $(L, \leqslant_L)$ is *separable* if there is a countable subset $D \subseteq L$ that is dense in L in the sense that whenever $x, z \in L$ and $x <_L z$ there is a $y \in D$ such that $x <_L y <_L z$.

Corollary 4.2. *Suppose that $(A, \leqslant_A)$ and $(B, \leqslant_B)$ are dense linear orderings without endpoints that are separable and complete. Then $(A, \leqslant_A) \cong (B, \leqslant_B)$.*

Proof. Let $\bar{A}$ and $\bar{B}$ respectively witness the separability of A and B. Let $f : \bar{A} \to \bar{B}$ be the isomorphism from the previous theorem. This isomorphism extends to an isomorphism $f^* : A \to B$ where $f^*(x)$ is the $\leqslant_B$-least upper bound of $f``\{y \in \bar{A} \mid y \leqslant_A x\}$. □

Thus $(\mathbb{R}, \leqslant)$ is uniquely characterised (up to isomorphism) as a dense linear ordering without endpoints that is complete and separable.

Let $(L, \leqslant_L)$ be a linear ordering. For $a, b \in L$ we let $(a, b)_L = \{z \in L \mid a <_L z <_L b\}$ and call such intervals *open* and we let $[a, b]_L = \{z \in L \mid a \leqslant_L z \leqslant_L b\}$ and call such intervals *closed*.

Definition 4.3. A linear ordering $(L, \leqslant_L)$ has the *Suslin property* if every collection of disjoint non-empty open intervals is countable.

Exercise 4.1. Let $(L, \leqslant_L)$ be a linear ordering. Show that if $(L, \leqslant_L)$ is separable then it has the Suslin property.

Suslin asked whether $(\mathbb{R}, \leqslant)$ can be uniquely characterised (up to isomorphism) as a dense linear ordering without endpoints that is complete and has the Suslin property.

Suslin's Hypothesis (SH). *Suppose $(L, \leqslant_L)$ is a dense linear ordering without endpoints that is complete and has the Suslin property. Then $(L, \leqslant_L)$ is separable.*

Definition 4.4. A *Suslin line* is a counterexample Suslin's hypothesis, that is, a Suslin line is a dense linear ordering without endpoints that is complete, has the Suslin property and is not separable.

It will be helpful to reformulate Suslin lines in terms of trees.

4.2 Aronszajn Trees

Let $(T, \leqslant_T)$ be a partial ordering. For each $x \in T$, let $(\cdot, x)_T = \{y \in T \mid y <_T x\}$. A *tree* is a partial ordering $(T, \leqslant_T)$ such that for every $x \in T$, $(\cdot, x)$ is a well-ordering. For each x in a tree T, let $\mathrm{ht}_T(x) = \mathrm{ot}(\cdot, x)$ be the *height* of x, $T_\alpha = \{x \in T \mid \mathrm{ht}_T(x) = \alpha\}$ be the α^{th}*-level of* T, $T_{<\alpha} = \bigcup_{\beta<\alpha} T_\beta$ be the *tree* T *below* α, and $T^x = \{y \in T \mid x \leqslant_T y\}$ is the *tree above* x. The *height of a tree*, $\mathrm{ht}(T)$, is the least ordinal α such that $T_\alpha = \varnothing$. A *chain* of T is a subset $C \subseteq T$ such that for all $x, y \in C$, ($x \leqslant_T y$ or $y \leqslant_T x$) and $(x, y)_T \subseteq C$. A *branch* of T is a maximal chain. A *cofinal branch* of T is a branch that intersects every level of T. Two distinct points x and y of T are *incomparable* if neither $x \leqslant_T y$ nor $y \leqslant_T x$. A subset $A \subseteq T$ is an *antichain* if any two distinct points in A are incompatible. A *node* of T is a maximal set of elements N such that for all $x, y \in N$, $(\cdot, x)_T = (\cdot, y)_T$. The *path of predecessors of a node* N is $\rho(N) = \{x \in T \mid \forall y \in N\, (x <_T y)\}$. The set $\rho(x, y) = (\cdot, x)_T \cap (\cdot, y)_T$ is the *common part* of $(\cdot, x)_T$ and $(\cdot, y)_T$.

Since slender tall trees tend to have long branches one is led to ask whether a tree that is very slender and very tall must have a cofinal branch. Some restrictions must be put in place if this question is to escape a trivial answer.

Exercise 4.2. Let κ be an infinite regular cardinal. Show that if we allow $|T_\alpha| \leqslant \kappa$ for each α then there is a tree of height κ that has no cofinal branch.

For this reason we restrict our attention to κ*-trees*, that is, trees of height κ such that for all $\alpha < \kappa$, $|T_\alpha| < \kappa$.

Exercise 4.3. Let κ be an infinite singular cardinal. Show that there is a κ-tree that has no cofinal branch.

For this reason we restrict our attention to κ-trees where κ is an infinite regular cardinal.

Theorem 4.5. (König) *Every* $\aleph_0$*-tree has a cofinal branch.*

Proof. Let T be an $\aleph_0$-tree. We shall build a branch $\langle x_n \mid n < \omega\rangle$ inductively, ensuring along the way that T^{x_n} is infinite. The construction does not halt since at stage $n+1$ there must be an $x_{n+1} \in T_{n+1}$ such that $T^{x_{n+1}}$ is infinite since otherwise T_n would be the union of finitely many finite trees. □

Exercise 4.4. (Hard) Let κ be a regular cardinal. Suppose that T is a tree of height κ such that for some $\lambda < \kappa$, $|T_\alpha| < \lambda$ for all $\alpha < \kappa$. Then T has a cofinal branch. [Hint: First assume that λ is regular. Let $S^\kappa_\lambda = \{\alpha < \kappa \mid \mathrm{cof}(\alpha) = \lambda\}$. For each $\alpha \in S^\kappa_\lambda$ choose $x_\alpha \in T_\alpha$ and $y_\alpha <_T x_\alpha$ such that $T^{y_\alpha} \cap T_\alpha = \{x_\alpha\}$. Now use Fodor's Lemma (which you may have to look up) to find an unbounded subsequence of the x_α's that all lie on a chain.]

For this reason we did not restrict our attention to trees of height κ such that for all $\alpha < \kappa$, $|T_\alpha| < \lambda$ for some $\lambda < \kappa$.

Definition 4.6. Let $(T, \leqslant_T)$ be a κ-tree, where κ is an infinite regular cardinal. T is *κ-Aronszajn* if T has no cofinal branch. T is *κ-Suslin* if T has no cofinal branch and no antichain of cardinality κ.

We have already seen one result on Aronszajn trees, namely, König's result that there are no $\aleph_0$-Aronszajn trees.

Theorem 4.7. *There is an $\aleph_1$-Aronszajn tree.*

Proof. Consider $(Q, \leqslant_Q)$ where Q consists of all well-ordered subsets of $\mathbb{Q}$ with a maximal element and $\leqslant_Q$ is the initial segment ordering. This tree is of height ω_1 and does not have a cofinal branch. But it is not an Aronszajn tree since it has uncountable levels. We shall isolate a slender subtree T of Q that is Aronszajn. The key is to inductively maintain

$(*)_\alpha$ For each $\gamma < \beta < \alpha$, for each $x \in T_\gamma$, and for each $q \in \mathbb{Q}$ such that $\max(x) < q$ there is a $y \in T_\beta$ such that $x <_T y$ and $\max(y) < q$

for each $\alpha < \omega_1$.

Case 1: $\alpha = \bar{\alpha} + 1$. Let

$$T_\alpha = \{x^\frown q \mid x \in T_{\bar{\alpha}},\ q \in \mathbb{Q},\ \text{and } \max(x) < q\}.$$

Thus $|T_\alpha| = \aleph_0$ and $(*)_{\alpha+1}$ holds.

Case 2: α is a limit ordinal. Let $\langle \alpha_n \mid n < \omega\rangle$ be an increasing sequence that that is cofinal in α. Consider $x \in T_{<\alpha}$ and $q \in \mathbb{Q}$ such that $\max(x) < q$. Let $m = \min\{n \mid \alpha_n \geqslant \mathrm{ht}_T(x)\}$. By $(*)_\alpha$ we can inductively construct an increasing sequence $\langle x_k \mid k < \omega\rangle$ of elements of $T_{<\alpha}$ such that $x_0 = x$, $x_k \in T_{\alpha_{m+k}}$, $\max(x_k) < q$ and $\sup\{\max(x_k) \mid k < \omega\} = q$. Let $y_{x,q} = (\cup_{k<\omega} x_k)^\frown q$. Finally, let

$$T_\alpha = \{y_{x,q} \mid x \in T_{<\alpha},\ q \in \mathbb{Q},\ \text{and } \max(x) < q\}.$$

Thus $|T_\alpha| = \aleph_0$ and $(*)_{\alpha+1}$ holds.

The tree

$$T = \bigcup_{\alpha<\omega_1} T_\alpha$$

is an ω_1-Aronszajn tree. □

Definition 4.8. Let $(T, \leqslant_T)$ be a tree and let $(P, \leqslant_P)$ be a partial ordering. T is *P-embeddable* if there is a map $f : T \to P$ such that strictly increasing in the sense that for all $x, y \in T$, if $x <_T y$ then $f(x) <_P f(y)$.

Definition 4.9. A tree is *κ-special* if it is the union of κ-many antichains.

Exercise 4.5. Let $(T, \leqslant_T)$ be the Aronszajn tree from the above theorem. The map $\max : T \to \mathbb{Q}$ witnesses that T is $\mathbb{Q}$-embeddable. Show that this implies that T is $\aleph_0$-special and that T is not Suslin.

Exercise 4.6. Suppose that κ is an infinite regular cardinal and T is a κ-tree that splits cofinally often: For all $\alpha < \kappa$, if $x \in T_\alpha$ then there exist $\beta > \alpha$ and y_1,$y_2 \in T_\beta$ such that $y_1 \neq y_2$ and $x <_T y_1, y_2$. Show that if T has a cofinal branch then it has an antichain of size κ.

Exercise 4.7. (Hard) Assume κ is an infinite regular cardinal such that $\kappa^{<\kappa} = \kappa$ (where $\kappa^{<\kappa} = \sup\{\kappa^\lambda \mid \lambda < \kappa\}$). Show that there is a κ-special κ^+-Aronszajn tree. [Hint: Generalise the proof of the above theorem. In place of $\mathbb{Q}$ use $\mathbb{Q}_\kappa = \{f \in {}^\omega\kappa \mid |\{n < \omega \mid f(n) \neq 0\}| < \omega\}$ ordered lexicographically. Build a subtree $(T, \leqslant_T)$ of $(Q_\kappa, \leqslant_{Q_\kappa})$ where Q_κ consists of all well-ordered subsets of $\mathbb{Q}_\kappa$ with a maximal element and $\leqslant_{Q_\kappa}$ is the initial segment ordering. Case 1 is as before but Case 2 breaks into two subcases. In the subcase where $\mathrm{cof}(\alpha) = \kappa$ proceed as before and in the subcase where $\mathrm{cof}(\alpha) = \lambda < \kappa$ note that since $|T_{<\alpha}|^\lambda \leqslant \kappa^\lambda = \kappa$ one is justified in extending each branch in $T_{<\alpha}$ to a branch in T_α.]

Thus if CH holds there is a $\aleph_1$-special $\aleph_2$-Aronszajn tree.

Definition 4.10. A tree of height κ is *normal* if the following conditions hold:

(1) $|T_0| = 1$

(2) For all limit ordinals $\alpha < \kappa$, if $x, y \in T_\alpha$ and $(\cdot, x)_T = (\cdot, y)_T$ then $x = y$

(3) For all $\alpha < \beta < \kappa$, if $x \in T_\alpha$ then there is a $y \in T_\beta$ such that $x <_T y$

(4) For all $\alpha < \kappa$, if $x \in T_\alpha$ then there exist y_1,$y_2 \in T_{\alpha+1}$ such that $y_1 \neq y_2$ and $x <_T y_1, y_2$.

Exercise 4.8. Show that the tree T constructed in Theorem 4.7 is normal.

4.3 Suslin Trees

Lemma 4.11. *If there is an $\aleph_1$-Suslin tree then there is a normal $\aleph_1$-Suslin tree.*

Proof. Let $(T, \leqslant_T)$ be an $\aleph_1$-Suslin tree. To ensure (1) and (2) in the definition of normality we add a new point between $\rho(N)$ and N where N is a node at the base or a node at a limit level. Let $(T_2, \leqslant_{T_2})$ be the resulting tree. To ensure (3) note that a point $x \in T_2$ that cannot be extended arbitrarily high must be such that T_2^x is countable. So we prune the tree, eliminating such branches. The result $T_3 = \{x \in T_2 \mid |T_2^x| > \omega\}$ is such that for any $x \in T_3$, T_3^x meets each level T_α for $\alpha < \omega_1$. To ensure (4) note that for each $x \in T_3$ there are arbitrarily high points above x that have two immediate successors since. For if there is a level α after which all points above x lie on the same chain we would have a cofinal branch since $|T^x|$ is uncountable. Thus $T_4 = \{y \in T_3 \mid \exists y_1, y_2 \in T_{\mathrm{ht}_T(x)+1}\,(y_1 \neq y_2 \wedge y <_T y_1, y_2)\}$ is a normal $\aleph_1$-Suslin tree. □

The same proof shows that if there is a κ-Suslin tree then there is a normal κ-Suslin tree and that if there is a κ-Aronszajn tree then there is a normal κ-Aronszajn. Thus we can, without loss of generality, assume that our trees are normal.

Theorem 4.12. *There is a Suslin line iff there is a Suslin tree.*

Proof. Suppose that $(L, \leqslant_L)$ is a Suslin line. We shall construct an ω_1-Suslin tree consisting of closed intervals of L ordered by reverse inclusion. Of course, we cannot allow all closed intervals since our tree must have size ω_1. We shall construct our tree inductively in ω_1-many steps, adding a new node at each step. In the first step we choose any closed interval $I_0 = [a_0, b_0]$. For the α^{th}-step suppose that for all $\beta < \alpha$ we have defined $I_\beta = [a_\beta, b_\beta]$. Since L is not separable the set of endpoints $C = \{a_\beta \mid \beta < \alpha\} \cup \{b_\beta \mid \beta < \alpha\}$ is not dense in L and so there is a closed interval $[a_\alpha, b_\alpha]$ that does not contain a point from C. We let $[a_\alpha, b_\alpha]$ be any such interval. This construction proceeds through all countable ordinals. Thus we let $T = \{I_\alpha \mid \alpha < \omega_1\}$ and take the ordering $\leqslant_T$ to be reverse inclusion $\supseteq$. Since for any $\beta < \alpha$, $[a_\alpha, b_\alpha]$ is either contained inside $[a_\beta, b_\beta]$ or disjoint from it $(\cdot, I_\alpha)$ is a well-ordering and hence T is a tree. Every antichain of T is countable since L is Suslin. Thus T is a Suslin tree.

Suppose that $(T, \leqslant_T)$ is a Suslin tree. We may assume that T is normal. We may also assume that each $x \in T$ has infinitely many immediate successors since if it does not then we can work with $T\restriction\{\alpha < \omega_1 \mid \alpha \text{ is a limit ordinal}\}$. For each node N in T let $<_N$ be an ordering of N that is isomorphic to $(\mathbb{Q}, <)$. Let $(L, \leqslant_L)$ be such that L consists of branches in T and $\leqslant_L$ is the lexicographic order induced by $\{\leqslant_N \mid N \text{ is a node of } T\}$: For all $x, y \in L$, $x \leqslant_L y$ iff either $x \leqslant_T y$ or $x \nleqslant_T y$ and $y \nleqslant_T x$ and $x \leqslant_{N_{\rho(x,y)}} y$. L is clearly linear and dense. And since disjoint open intervals in L contain incomparable branches in T, L has the Suslin Property. It remains to see that L is not separable: Suppose D is a countable set of branches. Choose $\alpha > \sup\{\mathrm{ht}_T(x) \mid x \in D\}$ and let $x \in T$ be such that $\mathrm{ht}_T(x) > \alpha$. Then the branches of T^x form an open set in L that does not contain a member of D. □

Chapter 5

Measurability

5.1 The Measure Problem

An old problem of geometry is to determine the volume of a region. For our purposes it will be convenient to restrict ourselves to the interval $[0,1]$ and the n-dimensional cube $[0,1]^n$. Let $\Gamma \subseteq P([0,1]^n)$ be a collection of sets the volumes of which we would like to measure. In the very least Γ should contain $\varnothing$, $[0,1]^n$, and be closed under countable unions. Naturally, the function $\mu : \Gamma \to [0,1]$ giving our measure of volume should satisfy:

(1) $\mu(\varnothing) = 0$ and $\mu([0,1]^n) = 1$

(2) if $X \subseteq Y$ then $\mu(X) \leqslant \mu(Y)$

(3) if $\{X_n \mid n < \omega\} \in \Gamma^\omega$ consists of pairwise disjoint sets then $\mu(\bigcup_{n<\omega} X_n) = \sum_{n<\omega} \mu(X_n)$

(4) if X is congruent to Y (that is, if X can be transformed into Y using translations, rotations, and reflections) then $\mu(X) = \mu(Y)$.

Furthermore, we would like Γ to be as broad as possible. Ideally, Γ should consist of *all* subsets of $[0,1]^n$. Unfortunately, this is not possible.

Exercise 5.1. Assume AC. Show that there is no function $\mu : P([0,1]) \to [0,1]$ satisfying (1)–(4). [Hint: Let $X \subseteq [0,1]$ be maximal such that for any pair of distinct $x, y \in X$, $x - y$ is not rational. For each $q \in \mathbb{Q}$, consider $X_q = \{x + q \mid x \in X\}$.]

One might try to remedy the situation by taking the drastic measure of restricting (3) to *finite* collections of pairwise disjoint sets. But for $n \geqslant 3$ even this will not help.

Theorem 5.1. (Banach and Tarski) *Assume* AC*. Let U and V be arbitrary open sets in $[0,1]^n$ where $n \geqslant 3$. Then there is an $m < \omega$ and subsets $X_1, \dots, X_m$ and $Y_1, \dots, Y_m$ of $\mathbb{R}^n$ such that*

(1) *$X_1, \dots, X_m$ are disjoint and $U = X_1 \cup \dots \cup X_m$,*

(2) *$Y_1, \dots, Y_m$ are disjoint and $V = Y_1 \cup \dots \cup Y_m$, and*

(3) *X_i is congruent to Y_i for each $i \leqslant m$.*

In particular, one can take a tiny sphere in $[0,1]^3$, cut it into finitely many pieces and rearrange those pieces to form a huge sphere in $[0,1]^3$.

The lesson of these examples is that we must rest content with measuring the volume of sets in a proper subcollection $\Gamma \subsetneq P([0,1]^n)$. Instead of specifying this subcollection in advance we will follow the course of Lebesgue by taking Γ to consist of the *measurable* sets (which we shall define) and then try to determine the extent of Γ. But before defining the notion of a measurable set it will be useful to introduce the Borel sets.

5.2 Borel Sets

A *topology* on a set X is a set $\mathscr{T} \subseteq P(X)$ such that $\varnothing, X \in \mathscr{T}$, $U \cap V \in \mathscr{T}$ whenever $U, V \in \mathscr{T}$, and $\bigcup_{\alpha \in I} U_\alpha \in \mathscr{T}$ whenever $\{U_\alpha\}_{a \in I} \subseteq \mathscr{T}$. The structure $(X, \mathscr{T})$ is called a *topological space* and we will often refer to X as a topological space. The sets in $\mathscr{T}$ are called *open* and a set A is called *closed* if $X - A$ is open. Sets which are both open and closed are called *clopen*. One way to present a topology is by specifying a subset $\mathscr{B}$ of $P(X)$ and letting $\mathscr{T}$ be the intersection of all topologies containing $\mathscr{B}$. In such a situation $\mathscr{B}$ is referred to as the collection of *basic open* sets and $\mathscr{T}$ is said to be the topology *generated* from $\mathscr{B}$.

Let $(X_1, \mathscr{T}_1)$ and $(X_2, \mathscr{T}_2)$ be topological spaces. A function $f : X_1 \to X_2$ is *continuous* if for each $A \in \mathscr{T}_2$, $f^{-1}``A \in \mathscr{T}_1$. A function $f : X_1 \to X_2$ is a *homeomorphism* if f is one-to-one and onto and both f and f^{-1} are continuous.

For our purposes it is convenient to identify the reals with ω^ω. As the topology on ω^ω we take the topology generated by

$$\mathscr{B} = \{N_s \mid s \in \omega^{<\omega}\},$$

where for $s \in \omega^{<\omega}$, $N_s = \{x \in \omega^\omega \mid s \subseteq x\}$.

Exercise 5.2. Suppose $s, t \in \omega^{<\omega}$. Then

(1) $N_s \cap N_t$ is either $\varnothing$, N_s or N_t,

(2) $N_s - N_t$ is a disjoint union of basic open sets,

(3) N_s is closed, and

(4) every open set is a disjoint union of basic open sets.

As the topology on $(\omega^\omega)^n$ we take the topology generated by

$$\mathscr{B}_n = \{N_{s_1} \times \cdots \times N_{s_n} \mid s_1, \ldots, s_n \in \omega^{<\omega}\}.$$

Exercise 5.3. Show that ω^ω and $(\omega^\omega)^n$ are homeomorphic.

This is one of the main advantages of working with ω^ω. It allows us to treat ω^ω and $(\omega^\omega)^n$ as a single case. For notational simplicity we shall work with ω^ω. Everything we say extends to $(\omega^\omega)^n$.

Let $\mathscr{T}$ be the above topology on ω^ω. We define the collections $\utilde{\Sigma}^0_\alpha$, $\utilde{\Pi}^0_\alpha$, and $\utilde{\Delta}^0_\alpha$ by induction on $\alpha < \omega_1$:

$$\begin{aligned}
&\utilde{\Sigma}^0_1 = \mathscr{T}\\
&\utilde{\Pi}^0_1 = \{A \in P(\omega^\omega) \mid \omega^\omega - A \in \mathscr{T}\}\\
&\utilde{\Sigma}^0_\alpha = \{A \in P(\omega^\omega) \mid A = \textstyle\bigcup_{n<\omega} B_n \text{ where } B_n \in \bigcup_{0<\beta<\alpha} \utilde{\Pi}^0_\beta\}\\
&\utilde{\Pi}^0_\alpha = \{A \in P(\omega^\omega) \mid A = \textstyle\bigcap_{n<\omega} B_n \text{ where } B_n \in \bigcup_{0<\beta<\alpha} \utilde{\Sigma}^0_\beta\}\\
&\utilde{\Delta}^0_\alpha = \utilde{\Sigma}^0_\alpha \cap \utilde{\Pi}^0_\alpha.
\end{aligned}$$

This hierarchy is known as the *Borel hierarchy* and $B = \bigcup_{\alpha<\omega_1} \utilde{\Sigma}^0_\alpha$ is the collection of *Borel sets*. (In order for this hierarchy to be non-trivial one must assume a certain amount of choice. For example, if one does not assume *any* form of choice then it is consistent that $\utilde{\Delta}^0_4 = P(\omega^\omega)$.)

To see that the Borel hierarchy is indeed a hierarchy we require the notion of a universal set. A *pointclass* is a collection $\Gamma \subseteq P(\omega^\omega)$. We say that a pointclass Γ has a *universal set* if there is a $U \subseteq \omega^\omega \times \omega^\omega$ such that $U \in \Gamma$ and for every $A \in \Gamma$ then there is an $a \in \omega^\omega$ such that $A = U_a$, where $U_a = \{x \in \omega^\omega \mid (x, a) \in U\}$.

Exercise 5.4. Show that $\utilde{\Sigma}^0_1$ has a universal set. Show that if $\utilde{\Sigma}^0_\alpha$ has a universal set then $\utilde{\Sigma}^0_{\alpha+1}$ has a universal set.

Corollary 5.2. *For each $\alpha < \omega_1$, there is a set that is $\utilde{\Sigma}^0_\alpha$ but not $\utilde{\Pi}^0_\alpha$.*

Proof. Let $U \subseteq \omega^\omega \times \omega^\omega$ be a universal set for $\utilde{\Sigma}^0_\alpha$. If $\utilde{\Sigma}^0_\alpha \subseteq \utilde{\Pi}^0_\alpha$, then

$$D = \{x \in \omega^\omega \mid (x, x) \notin U\}$$

is $\utilde{\Sigma}^0_\alpha$ and so there is a $d \in \omega^\omega$ such that

$$\{x \in \omega^\omega \mid (x, x) \notin U\} = U_d,$$

which is impossible. □

Thus, for each α, $\utilde{\Delta}^0_\alpha \subsetneq \Sigma^0_\alpha, \Pi^0_\alpha \subsetneq \utilde{\Delta}^0_{\alpha+1}$.

Exercise 5.5. ($\mathrm{AC}_\omega(\omega^\omega)$) Show that B is the smallest σ-algebra containing the open sets. (A *σ-algebra* of sets on X is a non-empty subset $A \subseteq P(X)$ that is closed under countable unions and complements.)

We now return to the measure problem.

Definition 5.3. Suppose $\Gamma \subseteq \omega^\omega$ is closed under finite intersections and contains $\varnothing$ and ω^ω. A *measure* on Γ is a function $\mu : \Gamma \to [0, 1]$ such that

(1) $\mu(\varnothing) = 0$ and $\mu(\omega^\omega) = 1$

(2) if $X \subseteq Y$ then $\mu(X) \leqslant \mu(Y)$

(3) if $\{X_n \mid n < \omega\} \in \Gamma^\omega$ consists of pairwise disjoint sets then $\mu(\bigcup_{n<\omega} X_n) = \sum_{n<\omega} \mu(X_n)$.

We shall define a measure on open and closed sets and then use this measure to first define the class of *Lebesgue measurable* sets and then define a measure on the Lebesgue measurable sets.

We start with the basic open sets, setting $\mu(N_s) = \prod_{i<\mathrm{lh}(s)} \frac{1}{2^{s(i)+1}}$. Then for an open set U we let $A = \{s \in \omega^{<\omega} \mid N_s \subseteq U \wedge \forall t \subsetneq s\, (N_t \nsubseteq U)\}$ and set $\mu(U) = \sum_{s\in A} \mu(N_s)$. And for a closed set C we set $\mu(C) = 1 - \mu(\omega^\omega - C)$. It is easy to see that μ is a measure on $\utilde{\Sigma}^0_1 \cup \utilde{\Pi}^0_1$. Now for an arbitrary set $A \in P(\omega^\omega)$ let

$$\mu^* = \inf\{\mu(U) \mid A \subseteq U \wedge U \in \utilde{\Sigma}^0_1\}$$

be the *outer measure* of A and let

$$\mu_* = \sup\{\mu(C) \mid C \subseteq A \wedge C \in \underset{\sim}{\Pi}^0_1\}$$

be the *inner measure* of A. We say that A is *Lebesgue measurable* if $\mu^*(A) = \mu_*(A)$ and we set

$$\Gamma = \{A \mid A \text{ is Lebesgue measurable}\}$$

and define a measure on the Lebesgue measurable sets thus:

$$\begin{aligned} \mu : \Gamma &\to [0,1] \\ A &\mapsto \mu^*(A). \end{aligned}$$

It is straightforward to show (by standard ϵ-δ arguments) that Γ is a σ-algebra and so we have:

Theorem 5.4. *Borel sets are Lebesgue measurable.*

5.3 Projective Sets

The projective hierarchy extends the Borel hierarchy. For $X \subseteq \omega^\omega \times \omega^\omega$ let $\exists^{\mathbb{R}} X = \{x \in \omega^\omega \mid \exists y \in \omega^\omega \, (x,y) \in X\}$ and for $\Gamma \subseteq P(\omega^\omega \times \omega^\omega)$ let $\exists^{\mathbb{R}}\Gamma = \{\exists^{\mathbb{R}} X \mid X \in \Gamma\}$. By induction on $n < \omega$ let:

$$\begin{aligned} \underset{\sim}{\Sigma}^1_1 &= \exists^{\mathbb{R}} \underset{\sim}{\Pi}^0_1 \\ \underset{\sim}{\Pi}^1_n &= \{X \in \omega^\omega \mid \omega^\omega - X \in \underset{\sim}{\Sigma}^1_n\} \\ \underset{\sim}{\Sigma}^1_{n+1} &= \exists^{\mathbb{R}} \underset{\sim}{\Pi}^1_n \\ \underset{\sim}{\Delta}^1_n &= \underset{\sim}{\Sigma}^1_n \cap \underset{\sim}{\Pi}^1_n. \end{aligned}$$

This hierarchy is the *projective hierarchy* and a set X is said to be *projective* if $X \in \underset{\sim}{\Sigma}^1_n$ for some n. A useful operation in this connection is *Suslin's operation* $\mathscr{A}$: For $\langle C_s \mid s \in \omega^{<\omega}\rangle$ such that each $C_s \subseteq \omega^\omega$, $\mathscr{A}(\langle C_s \mid s \in \omega^{<\omega}\rangle) = \{x \in \omega^\omega \mid \exists f \in \omega^\omega \, \forall n \, (x \in C_{f \restriction n})\}$.

Lemma 5.5. *If $X \subseteq \omega^\omega$ is $\underset{\sim}{\Sigma}^1_1$ then there is a sequence $\langle C_s \mid s \in \omega^{<\omega}\rangle$ of clopen sets such that $X = \mathscr{A}(\langle C_s \mid s \in \omega^{<\omega}\rangle)$.*

Proof. We first show that if $A \in \underset{\sim}{\Sigma}^1_1$ then there is a tree T on $(\omega \times \omega)^{<\omega}$ such that $A = \text{proj}[T]$, where $[T] = \{x \in \omega^\omega \mid \forall n < \omega\,(x \restriction n \in T)\}$ is the set of infinite branches through T and $\text{proj}[T] = \{x \in \omega^\omega \mid \exists y \in \omega^\omega\,(x, y) \in [T]\}$: We have $A = \exists^{\mathbb{R}} C$ where $C \subseteq (\omega^\omega \times \omega^\omega)$ is closed. Let $\pi : \omega^\omega \times \omega^\omega \to (\omega \times \omega)^\omega$ be the natural bijection. Let $C^* = \pi``C$ and let $T_{C^*} \subseteq (\omega \times \omega)^\omega$ be such that $C^* = [T_{C^*}]$. So $A = \text{proj}[T_{C^*}]$.

For each $s \in \omega^{<\omega}$, let $C_s = \{x \mid (x \restriction \text{dom}(s), s) \in T_{C^*}\}$. Note that these sets are clopen and $A = \mathscr{A}(\langle C_s \mid s \in \omega^{<\omega}\rangle)$. □

Exercise 5.6. The family of Lebesgue measurable sets is closed under Suslin's operation.

As a corollary we have:

Theorem 5.6. (Luzin) *$\underset{\sim}{\Sigma}^1_1$ sets are Lebesgue measurable.*

Is Γ broader than $\underset{\sim}{\Sigma}^1_1$?

Measurability Hypothesis (MH). *All $\underset{\sim}{\Sigma}^1_2$ sets are Lebesgue measurable.*

Chapter 6

Constructibility

In this chapter we will show that ZF cannot refute AC and that ZFC cannot refute CH, ¬SH, or ¬MH. This will be achieved by showing that each of these statements is satisfied in a certain class size *inner model* of ZF. This inner model L was introduced by Gödel and is known as the *constructible universe*.

Our strategy will be to first establish Con(ZF + $V = L$) and then, working in ZF + $V = L$ derive AC, CH, ¬SH, and ¬MH. The most natural way to execute the first step is to construct a model of ZF + $V = L$. Unfortunately, if we wish to work within ZF then the construction of such a model is prohibited by Gödel's second incompleteness theorem. As a first pass we will sidestep this difficulty by working in the stronger theory $\mathrm{ZF}^+ =$ ZF + "κ is an inaccessible cardinal". The result will be $\mathrm{ZF}^+ \vdash (\mathrm{Con(ZF)} \rightarrow \mathrm{Con(ZFC + CH + \neg SH + \neg MH)})$. Since from a foundational point of view it is of interest to prove relative consistency results in weak theories we will use the method of relativisation to sharpen the above result to $\mathrm{PA} \vdash (\mathrm{Con(ZF)} \rightarrow \mathrm{Con(ZFC + CH + \neg SH + \neg MH)})$.

6.1 Absoluteness and Reflection

A. Syntax. The language $\mathscr{L}$ of set theory has as basic symbols the parentheses '(' and ')', the binary relation symbols '=' and '$\in$', the connectives '$\neg$' and '$\vee$', the quantifier '$\exists$' and countably many variables 'x_i', where i ranges over ω. (The other symbols '$\wedge$', '$\rightarrow$', '$\leftrightarrow$', and '$\forall$' are treated here as defined symbols.) Another language which is useful is the extension $\mathscr{L}_M$ of

$\mathscr{L}$ obtained by adding a constant symbol $\dot{a}$ for each a in the class M.

Now we would like to talk of the syntax of set theory in set theory. We do this by inductively associating a set $\ulcorner\varphi\urcorner$ with each formula φ of $\mathscr{L}_M$. ($\ulcorner\varphi\urcorner$ is called the *Gödel set* of φ.) For the basic symbols let

$$\begin{array}{lll} \ulcorner(\urcorner = 0 & \ulcorner\in\urcorner = 3 & \ulcorner\exists\urcorner = 6 \\ \ulcorner)\urcorner = 1 & \ulcorner\neg\urcorner = 4 & \ulcorner x_i\urcorner = (7,i) \text{ for } i \in \omega \\ \ulcorner=\urcorner = 2 & \ulcorner\vee\urcorner = 5 & \ulcorner\dot{a}\urcorner = (8,a) \text{ for } a \in M \end{array}$$

and by induction on formula complexity let

$$\begin{aligned} \ulcorner x_i = x_j\urcorner &= (\ulcorner=\urcorner, \ulcorner x_i\urcorner, \ulcorner x_j\urcorner) \\ \ulcorner x_i \in x_j\urcorner &= (\ulcorner\in\urcorner, \ulcorner x_i\urcorner, \ulcorner x_j\urcorner) \\ \ulcorner\neg\varphi\urcorner &= (\ulcorner\neg\urcorner, \ulcorner\varphi\urcorner) \\ \ulcorner(\varphi \vee \psi)\urcorner &= (\ulcorner(\urcorner, \ulcorner\vee\urcorner, \ulcorner\varphi\urcorner, \ulcorner\psi\urcorner, \ulcorner)\urcorner) \\ \ulcorner(\exists x_i\, \varphi)\urcorner &= (\ulcorner(\urcorner, \ulcorner\exists\urcorner, \ulcorner x_i\urcorner, \ulcorner\varphi\urcorner, \ulcorner)\urcorner). \end{aligned}$$

For example, $\ulcorner\exists x_0\,(x_0 \in x_1)\urcorner = \big(0, 6, (7,0), \big(3, (7,0), (7,1)\big), 1\big)$.

Later on we will continue this development, formalising various syntactic notions in set theory.

It will be important to keep track of the complexity of our definitions and for this reason we introduce the *Levy hierarchy* of formulas. An $\mathscr{L}$-formula is $\Delta_0 = \Sigma_0 = \Pi_0$ if all of its quantifiers are bounded; Σ_{n+1} if it is of the form $\exists x_1 \cdots \exists x_k\, \varphi$ where φ is Π_n; and Π_{n+1} if it is of the form $\forall x_1 \cdots \forall x_k\, \varphi$ where φ is Σ_n.

Exercise 6.1. Show that each of the following notions has a Δ_0 definition:

$\mathrm{empty}(x) : x = \varnothing$	$\mathrm{ord}(x) : x$ is an ordinal
$\mathrm{up}(x) : x$ is an unordered pair	$\mathrm{succ}(x) : x$ is a successor ordinal
$\mathrm{op}(x) : x$ is an ordered pair	$\lim(x) : x$ is a limit ordinal
$\mathrm{rel}(x) : x$ is a relation	$\mathrm{omega}(x) : x = \omega$
$\mathrm{func}(x) : x$ is a function	$\mathrm{union}(x, y) : x = \cup y$.
$\mathrm{trans}(x) : x$ is transitive	

B. Semantics. An *$\mathscr{L}$-model* is a pair (M, E) where $M \neq \varnothing$ and $E \subseteq M \times M$. (The intention is that E is to interpret '$\in$'.) An *assignment* for a model

(M, E) and a formula is a function $s : \mathrm{VAR} \to M$, where $\mathrm{VAR} = \{(7, i) \mid i \in \omega\}$ is the set of variables. The notion of a formula φ being satisfied in M under an assignment s is defined inductively as follows:

$$
\begin{array}{lll}
(M, E) \models (x_i \in x_n)[s] & \text{iff} & s(x_i) \in s(x_j) \\
(M, E) \models (x_i = x_n)[s] & \text{iff} & s(x_i) = s(x_j) \\
(M, E) \models (\neg\psi)[s] & \text{iff} & (M, E) \not\models \psi[s] \\
(M, E) \models (\psi_1 \vee \psi_2)[s] & \text{iff} & (M, E) \models \psi_1[s] \text{ or } (M, E) \models \psi_2[s] \\
(M, E) \models (\exists x_i \psi)[s] & \text{iff} & \exists s' \forall x \in \mathrm{VAR}(x \neq x_i \to (s(x) = s'(x))) \\
 & & \text{ and } (M, E) \models \psi[s'].
\end{array}
$$

Later on we will give a more precise definition with an eye toward complexity.

Let us adopt the following two conventions. First, whenever we write '$\varphi(x_1, \ldots, x_n)$' it is to be understood that $x_1, \ldots, x_n$ are the only free variables in φ. Second, for $\bar{a} \in M^n$, a sequence of set parameters, we will write '$(M, E) \models \varphi[\bar{a}]$' as shorthand for '$(M, E) \models \varphi[s]$' where $s : \mathrm{VAR} \to M$ is such that $(s(x_1), \ldots, s(x_n)) = \bar{a}$.

Let (M, E) be an $\mathscr{L}$-model. A relation $X \subseteq M^n$ is (M, E)-*definable* if there is a formula $\varphi(x_1, \ldots, x_n)$ such that $X = \{\bar{a} \in M^n \mid (M, E) \models \varphi[\bar{a}]\}$. A relation $X \subseteq M^n$ is (M, E)-*definable with parameters* if there is a formula $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and a parameter sequence $\bar{b} \in M^m$ such that $X = \{\bar{a} \in M^n \mid (M, E) \models \varphi[\bar{a}, \bar{b}]\}$.

We are mainly interested in $\mathscr{L}$-models of the form (M, E) where M is transitive and $E = \in\restriction(M \times M)$. Whenever M is transitive we will use 'M' to abbreviate '$(M, \in\restriction(M \times M))$'.

Exercise 6.2. Let M and N be transitive $\mathscr{L}$-models such that $M \subseteq N$. Let $\varphi(x_1, \ldots, x_n)$ be a formula of $\mathscr{L}$ and let $\bar{a} \in M^n$ be a parameter sequence. Show

(a) If $\varphi \in \Delta_0$ then $M \models \varphi[\bar{a}]$ iff $N \models \varphi[\bar{a}]$.

(b) If $\varphi \in \Sigma_1$ then $M \models \varphi[\bar{a}]$ implies $N \models \varphi[\bar{a}]$ but not conversely.

(c) If $\varphi \in \Pi_1$ then $N \models \varphi[\bar{a}]$ implies $M \models \varphi[\bar{a}]$ but not conversely.

Suppose that M is a nonempty transitive set and we would like to show that $M \models$ Emptyset. This might seem trivial since M contains $\varnothing$. But there is a subtlety here. We have to ensure that M thinks that the object $a = \varnothing$ has the property of being the set that does not have any elements. Let us be more precise. We would like to show that $M \models \exists x_0\,(\text{empty}(x_0))$. It would suffice to show that $M \models (\text{empty}(x_0))[a]$. Now we know that $V \models (\text{empty}(x_0))[a]$. Since M is transitive, $M \subseteq V$ and 'empty(x_0)' is Δ_0, Exercise 6.2 implies that $M \models (\text{empty}(x_0))[a]$. In short, $M \models$ Emptyset because M contains the emptyset *and* since the property of being the emptyset is absolute M recognizes that the *object* $\varnothing$ has the *property* of being the emptyset.

The following exercise is an elaboration on this theme. By appealing to absoluteness it reduces the task of verifying that a transitive model satisfies a certain axiom to checking that the model satisfies an appropriate closure condition. Suppose that M is transitive. Let $\text{Def}^M = \{X \subseteq M \mid X \text{ is } M\text{-definable with parameters}\}$.

Exercise 6.3. Let M be a transitive set.

(1) $M \models \text{Extensionality} + \text{Foundation}$

(2) $M \models$ Emptyset iff $\varnothing \in M$

(3) $M \models$ Pairing iff $\forall x, y \in M\,(\{x, y\} \in M)$

(4) $M \models$ Union iff $\forall x \in M\,(\cup x \in M)$

(5) $M \models$ Powerset iff $\forall x \in M\,(P(x) \cap M \in M)$

(6) $M \models$ Infinity whenever $\omega \in M$, but not conversely

(7) $M \models$ Comprehension iff $\forall C \in \text{Def}^M\, \forall a \in M\,(C \cap a \in M)$

(8) $M \models$ Collection iff $\forall F \in \text{Def}^M\, \forall a \in M$ (if $F : M \to M$ then $F``a \in M$).

C. Reflection. Let $\mathscr{M} = (M, E)$ and $\mathscr{N} = (N, F)$ be $\mathscr{L}$-models such that $M \subseteq N$. We say that $\mathscr{M}$ is a *substructure* of $\mathscr{N}$ and write $\mathscr{M} \subseteq \mathscr{N}$ if $E = F \cap (M \times M)$. We say that $\mathscr{M}$ is an *elementary substructure* of $\mathscr{N}$ and write $\mathscr{M} \prec \mathscr{N}$ if for all $\varphi(x_1, \ldots, x_n) \in \mathscr{L}$ and for all $\bar{a} \in M^n$,

$$\mathscr{M} \models \varphi[\bar{a}] \text{ iff } \mathscr{N} \models \varphi[\bar{a}].$$

Exercise 6.4. (Tarski-Vaught Test) Let $\mathscr{M} = (M, E)$ and $\mathscr{N} = (N, F)$ be $\mathscr{L}$-models such that $\mathscr{M} \subseteq \mathscr{N}$. Show that the following are equivalent:

(1) $\mathscr{M} \prec \mathscr{N}$

(2) for all $\varphi(x, y_1, \ldots, y_n) \in \mathscr{L}$ and for all $\bar{b} \in M^n$,

$$\begin{aligned}&\text{if there is an } a \in N \text{ such that } \mathscr{N} \models \varphi[a, \bar{b}]\\&\text{then there is an } a \in M \text{ such that } \mathscr{N} \models \varphi[a, \bar{b}].\end{aligned}$$

Theorem 6.1. (Reflection) *Suppose κ is a regular uncountable cardinal. Let $(\mathscr{M}_\alpha \mid \alpha < \kappa)$ be such that*

(1) *$\mathscr{M}_\alpha = (M_\alpha, E_\alpha)$ is an $\mathscr{L}$-model for each $\alpha < \kappa$*

(2) *$\mathscr{M}_\alpha \subseteq \mathscr{M}_\beta$ for every $\alpha \leqslant \beta < \kappa$*

(3) *$|M_\alpha| < \kappa$ for all $\alpha < \kappa$*

(4) *$M_\lambda = \bigcup_{\alpha<\lambda} M_\alpha$ for all limit ordinals $\lambda < \kappa$.*

Let $M = \bigcup_{\alpha<\kappa} M_\alpha$ and $E = \bigcup_{\alpha<\kappa} E_\alpha$ and let $\mathscr{M} = (M, E)$. Then

$$C = \{\alpha < \kappa \mid \mathscr{M}_\alpha \prec \mathscr{M}\}$$

is club in κ.

Proof. C is closed: Let $\lambda < \kappa$ be a limit ordinal such that $C \cap \lambda$ is unbounded in λ. We have to show that $\mathscr{M}_\lambda \prec \mathscr{M}$. Consider $\varphi(x, y_1 \ldots, y_n) \in \mathscr{L}$ and $\bar{b} \in M^n$ such that

$$\exists a \in M \, (\mathscr{M} \models \varphi[a, \bar{b}]).$$

Since $M_\lambda = \bigcup_{\alpha<\lambda} M_\alpha$, there exists $\alpha \in C \cap \lambda$ such that $\bar{b} \in (M_\alpha)^n$. But $\mathscr{M}_\alpha \prec \mathscr{M}$. So, by the Tarski-Vaught test,

$$\exists a \in M_\alpha \, (\mathscr{M} \models \varphi[a, \bar{b}]).$$

C is unbounded: For each $\varphi(x, y_1, \ldots y_n) \in \mathscr{L}$ let

$$\begin{aligned}f_\varphi &: \kappa \to \kappa\\ &\alpha \mapsto \text{the least } \beta < \kappa \text{ such that for all } \bar{b} \in {M_\alpha}^n\\ &\quad \text{if } \exists a \in M \, (\mathscr{M} \models \varphi[a, \bar{b}])\\ &\quad \text{then } \exists a \in M_\beta \, (\mathscr{M} \models \varphi[a, \bar{b}])\end{aligned}$$

This function is well defined since if for a given $\alpha < \kappa$ there is no such $f_\varphi(\alpha)$ then there would be a cofinal map from $|M_\alpha{}^n|$ into κ which is impossible since $|M_\alpha{}^n| < \kappa$ and κ is regular. Now let

$$\begin{aligned} g : \kappa &\to \kappa \\ \alpha &\mapsto \sup_{\varphi \in \mathscr{L}} f_\varphi(\alpha). \end{aligned}$$

Again, this function is well defined since κ is regular. Now if α is such that $g``\alpha \subseteq \alpha$ then, by the Tarski-Vaught test, $\mathscr{M}_\alpha \prec \mathscr{M}$. So it remains to see that the set of such α is club in κ. It is easy to see that the set of such α is closed. To see that it is unbounded for $\alpha < \kappa$, let $\alpha_0 = \alpha$, $\alpha_{n+1} = g(\alpha_n)$ and $\alpha^* = \sup_{n<\omega} \alpha_n$, and note that $\alpha^* < \kappa$ is such that $g``\alpha^* \subseteq \alpha^*$. □

D. Inaccessibles. Recall that κ is inaccessible if κ is an uncountable regular cardinal such that $2^\alpha < \kappa$ for all $\alpha < \kappa$.

Lemma 6.2. *Suppose κ is inaccessible. Then $|V_\alpha| < \kappa$ for all $\alpha < \kappa$. Hence $|V_\kappa| = \kappa$.*

Proof. The lemma is proved by induction on $\alpha < \kappa$.

Case 1: $\alpha = \bar{\alpha}+1$. Then $|V_\alpha| = 2^{|V_{\bar{\alpha}}|}$ and so $|V_\alpha| < \kappa$ since by our induction hypothesis $|V_{\bar{\alpha}}| < \kappa$.

Case 2: $\lim(\alpha)$. Consider

$$\begin{aligned} f : \alpha &\to \kappa \\ \beta &\mapsto |V_\beta|. \end{aligned}$$

Since κ is regular, f is not cofinal. So let $\lambda < \kappa$ be such that $f : \alpha \to \lambda$. Then $|V_\alpha| \leqslant |\alpha| \cdot |\lambda| < \kappa$. □

Theorem 6.3. *If κ is inaccessible then $V_\kappa \models$ ZFC.*

Proof. Notice that $\varnothing \in V_1$, $x, y \in V_\alpha \to \{x, y\} \in V_{\alpha+1}$, $x \in V_\alpha \to \cup x \in V_\alpha$, $x \in V_\alpha \to P(x) \in V_{\alpha+1}$, $\forall C \in \mathrm{Def}^{V_\kappa} \forall a \in V_\kappa\, (C \cap a \in V_\kappa)$ (since $C \cap a \in V_{\kappa+1} \cap P(a) \subseteq V_\kappa$), and $\omega \in V_{\omega+1}$. So, by Exercise 6.3, $\textsc{lev}_\kappa$ satisfies Extensionality, Foundation, Emptyset, Pairing, Powerset, Union, Infinity and Comprehension. To see that V_κ satisfies Collection, fix $F \in \mathrm{Def}^{V_\kappa}$ a function

$F : V_\kappa \to V_\kappa$ and fix $a \in V_\kappa$. $F``a$ must be contained in some V_α for $\alpha < \kappa$ since otherwise, letting $g : |a| \to a$ be a bijection, the function

$$\begin{aligned} h &: \alpha \to \kappa \\ \alpha &\mapsto \operatorname{rank}(F(g^{-1}(\alpha))) \end{aligned}$$

would be cofinal in κ, which is impossible since $|a| < \kappa$ and κ is regular. To see that V_κ satisfies AC note that any well-ordering W of $a \in V_\kappa$ is such that $W \in P(P(P(a))) \subseteq V_\kappa$. □

There is an interesting fact regarding Exercise 6.3 (5). First we need the following result.

Exercise 6.5. Suppose C is a class and E is a well-founded relation on C such that for all x and y in C if $\mathrm{ext}_E(x) = \mathrm{ext}_E(y)$ then $x = y$. Define, by transfinite recursion with respect to E, the map

$$\pi(x) = \{\pi(y) \mid y\, E\, x\}.$$

Show that $\pi : (C, E) \to (M, \in)$ is an isomorphism and that M is transitive. (The model $(M, \in)$ is known as the *Mostowski collapse* of (C, E).)

Theorem 6.4. (ZFC^+) *There is a countable transitive model M such that $M \models \mathrm{ZFC}$.*

Proof. As in the proof of reflection, let $X \prec V_\kappa$ be countable. Let

$$\pi : (X, \in) \to (M, \in)$$

be the Mostowski collapse. □

6.2 The constructible universe

A. Definition. Recall that for M transitive,

$$\mathrm{Def}^M = \{X \subseteq M \mid X \text{ is definable over } M \text{ with parameters from } M\}.$$

The hierarchy of constructible sets is defined by transfinite recursion:

$$\begin{aligned} L_0 &= \varnothing \\ L_{\alpha+1} &= \mathrm{Def}^{L_\alpha} \\ L_\lambda &= \textstyle\bigcup_{\alpha<\lambda} L_\alpha \text{ for } \lambda \text{ a limit ordinal.} \end{aligned}$$

The constructible universe is the class $L = \bigcup_{\alpha \in \mathrm{On}} L_\alpha$.

Exercise 6.6. Prove the following:

(1) For all α, L_α is transitive.

(2) For all $\alpha \leqslant \beta$, $L_\alpha \subseteq L_\beta$.

(3) For all $\alpha < \beta$, $L_\alpha \in L_\beta$.

(4) For all α, $\alpha = L_\alpha \cap \mathrm{On}$.

(5) For all $\alpha \geqslant \omega$, $|L_\alpha| = |\alpha|$.

B. ZF.

Theorem 6.5. (ZFC^+) $L_\kappa \models \mathrm{ZF}$.

Proof. Notice that $\varnothing \in L_1$, $x, y \in L_\alpha \to \{x, y\} \in L_{\alpha+1}$, $x \in L_\alpha \to \cup x \in L_\alpha$ and $\omega \in L_{\omega+1}$. So, by Exercise 6.3, L_κ satisfies Extensionality, Foundation, Emptyset, Pairing, Union, and Infinity.

Powerset: Fix $x \in L_\kappa$ and let $\alpha < \kappa$ be such that $x \in L_\alpha$. Consider

$$\begin{aligned} f : P(x) \cap L_\kappa &\to \kappa \\ y &\mapsto \text{the least } \alpha \text{ such that } y \in L_\alpha. \end{aligned}$$

Since $|P(x) \cap L_\kappa| \leqslant |P(x)| < \kappa$ and κ is regular, f is not cofinal. Letting $\beta < \kappa$ be such that $f : P(x) \cap L_\kappa \to \beta$, we have $P(x) \cap L_\kappa \subseteq L_\beta$ and so $P(x) \cap L_\kappa \in L_{\beta+1}$.

Comprehension: Fix $C \in \mathrm{Def}^{L_\kappa}$ and $a \in L_\kappa$. Let

$$C = \{x \in L_\kappa \mid L_\kappa \models \varphi[x, a, \bar{p}]\}.$$

By the reflection theorem, there is an $\alpha < \kappa$ such that

$$L_\alpha \prec L_\kappa \text{ and } a, \bar{p} \in L_\alpha.$$

So $C \cap a \in \mathrm{Def}^{L_\alpha} = L_{\alpha+1}$.

Collection: Fix $F \in \mathrm{Def}^{L_\kappa}$ a function $F : L_\kappa \to L_\kappa$ and fix $a \in L_\kappa$. $F``a$ must be contained in some L_α for $\alpha < \kappa$ as otherwise we could cofinalise κ. □

C. Encoding syntax. The sentence $V{=}L$ asserts $\forall x\, \exists \alpha\, (x \in L_\alpha)$. Our next goal is to show that $L_\kappa \models V{=}L$. This might seem trivial since, after all, for each $\alpha < \kappa$, $L_\alpha \in L_\kappa$ and $L_\kappa = \bigcup_{\alpha<\kappa} L_\alpha$. The trouble is that we must rule out the possibility that $(L_\alpha)^{L_\kappa} \subsetneq L_\alpha$ to ensure $(L_\alpha)^{L_\kappa} \in L_\kappa$. To do this we must show that the notion of being constructible is sufficiently absolute across transitive models. For this reason we revisit our definition of "$x \in L_\alpha$", showing that within the context of ZF the notion is Σ_1.

A *term* in the language $\mathscr{L}_M$ is a variable or a constant. Let $\text{TERM}(x, M, V_\omega)$ abbreviate

$$\exists i \in \omega\, (x = \ulcorner x_i \urcorner) \vee \exists a \in M\, (x = \ulcorner \dot{a} \urcorner).$$

An *atomic* formula in the language of $\mathscr{L}_M$ is a formula of the form '$y = z$' or '$y \in z$' where y and z are terms. Let $\text{ATOM}(x, M, V_\omega)$ abbreviate

$$\exists y, z \in \text{TERM}(x, M, V_\omega)\, (x = \ulcorner y = z \urcorner \vee x = \ulcorner y \in z \urcorner).$$

A *formula sequence* for a formula φ in $\mathscr{L}_M$ is a sequence such that (i) every element of the sequence is either an atomic formula or a negation, disjunction, or existential quantification of previous elements of the sequence and (ii) φ is the final element of the sequence. Let $\text{FSEQ}(F, \ulcorner \varphi \urcorner, M, V_\omega)$ abbreviate

$$\begin{aligned} \exists n \in \omega \big(&\text{func}(F) \wedge \text{dom}(F) = n + 1 \wedge \\ &\forall m \in \text{dom}(F) \big(\text{ATOM}(F(m), M, V_\omega) \vee \\ &\exists k, l < m (F(m) = \ulcorner \neg F(k) \urcorner \vee \\ &\qquad F(m) = \ulcorner F(k) \vee F(l) \urcorner \vee \\ &\qquad F(m) = \ulcorner \exists x_i F(k) \urcorner \text{ some } i \in \omega)\big) \\ &\wedge F(n) = \ulcorner \varphi \urcorner \big). \end{aligned}$$

We shall need to keep track of the number of variables that occur in formulas of a formula sequence. Let $\text{VLEN}(F, V_\omega)$ abbreviate

$$\begin{aligned} 1 + \sup\{n < \omega \mid \exists i < \omega (&\text{ATOM}(F(i), M, V_\omega) \wedge \\ &F(i)(1) = \ulcorner x_n \urcorner \vee F(i)(2) = \ulcorner x_n \urcorner)\}. \end{aligned}$$

A *satisfaction sequence* for a formula φ in $\mathscr{L}_M$ is a sequence that runs alongside a formula sequence for φ and is such that (i) each entry S is a set of sequences satisfying the corresponding formula $\varphi(S)$ in the formula sequence and (*ii*) the construction of each S from previous S_1 and S_2 mirrors the construction of $\varphi(S)$ from $\varphi(S_1)$ and $\varphi(S_2)$ (in the case where φ is not atomic). Let $\textsc{sseq}(S, F, \ulcorner\varphi\urcorner, M, V_\omega)$ abbreviate the statement

$$\begin{aligned}\exists n, r \in \omega\Big(&\textsc{fseq}(F, \ulcorner\varphi\urcorner, M, V_\omega) \wedge \mathrm{dom}(F) = n+1 \wedge F(n) = \ulcorner\varphi\urcorner \wedge\\ &\textsc{vlen}(F, V_\omega) = r \wedge \mathrm{func}(S) \wedge \mathrm{dom}(S) = n+1 \wedge\\ &\forall i, j < m\ \forall a, b \in M\ \forall m \in \mathrm{dom}(S)\\ &\quad\big((\textsc{atom}(F(m), M, V_\omega) \to \Psi(i, j, a, b)) \wedge\\ &\quad\ \exists k, l < m(F(m) = \ulcorner\neg F(k)\urcorner \to S(m) = M^r - S(k) \wedge\\ &\qquad\qquad F(m) = \ulcorner F(k) \vee F(l)\urcorner \to S(m) = S(k) \cup S(l) \wedge\\ &\qquad\qquad F(m) = \ulcorner\exists x_i\varphi\urcorner \text{ for some } i \in \omega \to\\ &\qquad\qquad\quad S(m) = \{s \in M^r \mid \exists\bar{s} \in S(k)\ \forall i' < n+1\\ &\qquad\qquad\qquad\qquad (i' \neq i \to \bar{s}(i') = s(i'))\}\big)\\ &\wedge F(n) = \ulcorner\varphi\urcorner\Big)\end{aligned}$$

where $\Psi(i, j, a, b)$ abbreviates the statement

$$\begin{aligned}&F(m) = \ulcorner x_i \in x_j\urcorner \to S(m) = \{s \in M^r \mid s(i) \in s(j)\} \wedge\\ &F(m) = \ulcorner x_i \in a\urcorner \to S(m) = \{s \in M^r \mid s(i) \in a\} \wedge\\ &F(m) = \ulcorner a \in x_j\urcorner \to S(m) = \{s \in M^r \mid a \in s(j)\} \wedge\\ &F(m) = \ulcorner a \in b\urcorner \to S(m) = \begin{cases} M^r & \text{if } a \in b\\ \varnothing & \text{otherwise}\end{cases} \wedge\\ &F(m) = \ulcorner x_i = x_j\urcorner \to S(m) = \{s \in M^r \mid s(i) = s(j)\} \wedge\\ &F(m) = \ulcorner x_i = a\urcorner \to S(m) = \{s \in M^r \mid s(i) = a\} \wedge\\ &F(m) = \ulcorner a = x_j\urcorner \to S(m) = \{s \in M^r \mid a = s(j)\} \wedge\\ &F(m) = \ulcorner a = b\urcorner \to S(m) = \begin{cases} M^r & \text{if } a = b\\ \varnothing & \text{otherwise.}\end{cases}\end{aligned}$$

Notice that $\textsc{term}(x, M, V_\omega)$, $\textsc{atom}(x, M, V_w)$, $\textsc{fseq}(F, \ulcorner\varphi\urcorner, M, V_\omega)$, $\textsc{vlen}(F, V_\omega)$, and $\textsc{sseq}(S, F, \ulcorner\varphi\urcorner, M, V_\omega)$ are Δ_0.

We are now in a position to define the notion of an assignment s satisfying a formula φ in an $\mathscr{L}$-model (M, E). Let $\textsc{sat}(M, \ulcorner\varphi\urcorner, s)$ abbreviate

$$\begin{aligned}\exists S, F, x, n\big(&x = V_\omega \wedge n \in \omega \wedge \textsc{sseq}(S, F, \ulcorner\varphi\urcorner, M, x) \wedge\\ &\qquad F(n) = \ulcorner\varphi\urcorner \wedge s \in S(n)\big).\end{aligned}$$

Notice that $\textsc{sat}(M, \ulcorner\varphi\urcorner, s)$ is Σ_1.

Exercise 6.7. Let $\textsc{fvar}(\ulcorner\varphi\urcorner, V_\omega)$ abbreviate the notion informally expressed by "the set of free variables in $\ulcorner\varphi\urcorner$". Show that "$x = \textsc{fvar}(\ulcorner\varphi\urcorner, V_\omega)$" is Δ_0.

We can not define definability with parameters. Let $\textsc{def}(x, M)$ abbreviate

$$\begin{aligned}&\big(\exists z = V_\omega \forall y \in x \exists \ulcorner\varphi\urcorner \in V_\omega \exists i \in \omega\big(\textsc{fvar}(\ulcorner\varphi\urcorner, V_\omega) = \{\ulcorner x_i \urcorner\} \wedge \\ &\qquad y = \{z \in M \mid \exists s \in M^{<\omega}\textsc{sat}(M, \ulcorner\varphi\urcorner, s) \wedge z = s(i)\}\big) \wedge \\ &\big(\forall \ulcorner\varphi\urcorner \in V_\omega \exists i \in \omega\big(\textsc{fvar}(\ulcorner\varphi\urcorner, V_\omega) = \{\ulcorner x_i \urcorner\} \wedge \\ &\qquad \{z \in M \mid \exists s \in M^{<\omega}\textsc{sat}(M, \ulcorner\varphi\urcorner, s) \wedge z = s(i)\} \in x\big).\end{aligned}$$

This formula is easily seen to be equivalent (in ZF) to a Σ_1-formula.

Let $\textsc{lev}(x, \alpha)$ abbreviate

$$\begin{aligned}\mathrm{ord}(\alpha) \wedge \exists f\big(&\mathrm{func}(f) \wedge \mathrm{dom}(f) = \alpha + 1 \wedge \\ &f(0) = \varnothing \wedge \\ &\forall \lambda \in \mathrm{dom}(f)\big(\lim(\lambda) \to f(\lambda) = \textstyle\bigcup_{\beta<\lambda} f(\beta)\big) \wedge \\ &\forall \bar\gamma \forall \gamma \in \mathrm{dom}(f)\big(\gamma = \bar\gamma + 1 \to \textsc{def}(f(\gamma), f(\bar\gamma)\big) \wedge \\ &x = f(\alpha)\big).\end{aligned}$$

Let $\textsc{constr}(x)$ abbreviate

$$\exists y, \alpha(\textsc{lev}(y, \alpha) \wedge x \in y).$$

Let $V{=}L$ abbreviate

$$\forall x(\textsc{constr}(x)).$$

Notice that $\textsc{def}(x, M)$ and $\textsc{lev}(x, \alpha)$ are Σ_1 and that the statement $V{=}L$ is Π_2.

D. Absoluteness and Constructibility.

Exercise 6.8. Assume that $\lambda > \omega$ is a limit ordinal. Show that for all α such that $\omega < \alpha < \lambda$

$$L_\lambda \models \textsc{def}(L_{\alpha+1}, L_\alpha).$$

Theorem 6.6. *If M is transitive and $M \models V{=}L$ then $M = L_\lambda$ for some limit ordinal λ.*

Proof. Since $M \models V = L$ and α first appears in $L_{\alpha+1}$, $\mathrm{On} \cap M$ must be a limit ordinal. Let $\lambda = \mathrm{On} \cap M$. Consider $x \in M$. Since $M \models V = L$ there is a b and an α such that

$$M \models \textsc{lev}(b, \alpha) \wedge x \in b.$$

The point is that since $\textsc{lev}(b, \alpha)$ is Σ_1, M is correct in its calculation of L_α. (For, by Σ_1-upward absoluteness, we have

$$V \models \textsc{lev}(b, \alpha)$$

and so $b = L_\alpha$.) This implies (i) $x \in L_\alpha$ and hence $M \subseteq L_\lambda$ and (ii) $L_\alpha \in M$ and hence $L_\lambda \subseteq M$. □

Theorem 6.7. *If λ is a limit ordinal, then $L_\lambda \models V = L$.*

Proof. Fix $x \in L_\lambda$. Let $\alpha < \lambda$ be such that $x \in L_\alpha$. To prove the theorem it suffices to show that

$$L_\lambda \models \textsc{lev}(L_\alpha, \alpha)$$

and to show this it suffices to show that L_λ contains a witness $f = \langle L_\beta \mid \beta \leqslant \alpha \rangle$ to the existential in the statement $\textsc{lev}(L_\alpha, \alpha)$. (Look at the definition of $\textsc{lev}(L_\alpha, \alpha)$. Once we have the witness for the existential we just need to know that L_λ is correct about the statement following '$\exists f$' in the statement $\textsc{lev}(L_\alpha, \alpha)$. The only non-trivial part of this statement is $\textsc{def}(f(\gamma), f(\bar{\gamma}))$. But in Exercise 6.8 we saw that L_λ is indeed correct in its calculation of this.) So it suffices to prove the following:

Lemma 6.8. *If λ is a limit ordinal, the for all $\alpha < \lambda$,*

$$\langle L_\beta \mid \beta < \alpha \rangle \in L_\lambda.$$

In fact, $\langle L_\beta \mid \beta < \lambda \rangle \in \mathrm{Def}^{L_\lambda}$.

Proof. The lemma is proved by induction on limit ordinals λ. The base case $\lambda = \omega$ is trivial. So assume that the lemma is true for all limit ordinals up to and including $\bar{\lambda}$ and consider $\lambda = \bar{\lambda} + \omega$. We claim that $\langle L_\beta \mid \beta < \bar{\lambda} \rangle = \{(\alpha, x) \in L_{\bar{\lambda}} \mid L_{\bar{\lambda}} \models \textsc{lev}(x, \alpha)\}$. To see this note that by induction $L_{\bar{\lambda}}$ contains the relevant existential witnesses, namely, $\langle L_\beta \mid \beta \leqslant \alpha \rangle$ for each $\alpha < \bar{\lambda}$. Thus $\langle L_\beta \mid \beta < \bar{\lambda} \rangle \in L_\lambda$. Now by induction on $n < \omega$, we have $\langle L_\beta \mid \beta < \bar{\lambda} + n \rangle \in L_\lambda$. □

This completes the proof of the theorem. □

Corollary 6.9. *$L_\kappa \models \mathrm{ZF} + V = L$.*

6.3 The Axiom $V = L$

A. Choice.

Exercise 6.9. Show that there is a function

$$F : \mathrm{On} \to V$$
$$\alpha \mapsto <_\alpha$$

such that for all $\alpha \geqslant \omega$,

(1) $<_\alpha$ is a well-ordering of L_α,

(2) if $\alpha < \beta$ then $<_\alpha \subseteq <_\beta$ and for all $x \in L_\alpha$ and $y \in L_\beta - L_\alpha$, $x <_\beta y$, and

(3) there is a Σ_1-formula $\varphi(x_0, x_1, x_2)$ such that for all limit ordinals $\lambda \geqslant \omega$

$$L_\lambda \models \varphi[\alpha, x, V_\omega] \text{ iff } \mathrm{ord}(\alpha) \wedge x = <_\alpha .$$

As a consequence we have:

Theorem 6.10. $\mathrm{ZF} + V = L \vdash \mathrm{AC}$.

B. Continuum Hypothesis.

Lemma 6.11. (Condensation) *Assume* $\mathrm{ZF} + V = L$. *Suppose* $X \prec L_\lambda$ *where* λ *is a limit ordinal. Then*

$$(X, \in \restriction (X \times X)) \cong L_{\bar{\lambda}}$$

for some limit ordinal $\bar{\lambda}$.

Proof. Since $(X, \in \restriction (X \times X))$ is wellfounded and extensional we can take its Mostowski collapse, letting

$$\pi : X \to M$$

be the isomorphism where M is transitive. Since $L_\lambda \models V = L$, $X \models V = L$ and hence $M \models V = L$ which implies, by Theorem 6.6, that $M = L_{\bar{\lambda}}$ for some limit ordinal $\bar{\lambda}$. □

Theorem 6.12. $\mathrm{ZF} + V = L \vdash \mathrm{GCH}$.

Proof. Let κ be an infinite cardinal. We aim to show that $2^\kappa \leqslant \kappa^+$. Now in Exercise 6.6 we saw that for all $\alpha \geqslant \omega$, $|L_\alpha| = |\alpha|$. In particular, $|L_{\kappa^+}| = \kappa^+$. So it suffices to show that $(P(\kappa))^L \subseteq L_{\kappa^+}$.

Let $A \subseteq \kappa$ and let λ be a limit ordinal such that $A \in L_\lambda$. Now let $X \prec L_\lambda$ be such that $A \in X$, $\kappa \subseteq X$, and $|X| = \kappa$, and let

$$\pi : X \to L_{\bar{\lambda}}$$

be an isomorphism. Notice that $\bar{\lambda} < \kappa^+$ and $\pi(A) = A$ since for all $\alpha < \kappa$, $\pi(\alpha) = \alpha$. ($\pi(A) = \{\pi(\alpha) \mid \alpha \in A \cap X\} = \{\alpha \mid \alpha \in A\} = A$.) Thus, each $A \subseteq \kappa$ appears in some $L_{\bar{\lambda}}$ where $\bar{\lambda} < \kappa^+$. □

Thus, each real in L appears in some L_α where $\alpha < \omega_1$. In is natural to ask whether there are stretches where no new reals appear. More precisely, for $\alpha, \beta < \omega_1$, let us say that $[\alpha, \alpha + \beta]$ is a *β-gap* if $P(\omega) \cap (L_{\alpha+\beta} - L_\alpha) = \varnothing$. Our question, then, is whether there are β-gaps.

Exercise 6.10. Show that for each $\beta < \omega_1$ there is a β-gap.

C. Fine Structure. In the proof of GCH we showed that every subset $A \subseteq \kappa$ appears in some $L_{\alpha+1}$ of cardinality κ. In particular, if $A \in L_{\alpha+1} - L_\alpha$ then $|L_\alpha| = \kappa$. This is a fact about L, that is, L contains a function $f : \kappa \to L_\alpha$ that is onto. It is interest to ask where the first such function appears. It turns out that it appears as soon as possible, namely in $L_{\alpha+1}$. One can ask a finer question. Each such $A \in L_{\alpha+1}$ is definable with parameters over L_α. Suppose that A is definable with parameters over L_α via a Σ_n-formula. Then one can ask where the first collapsing function appears with respect to the order of definability. It turn out that it appears as soon as possible even in this finer sense, that is, it too is definable with parameters over L_α via a Σ_n-formula. These remarkable results are due to Jensen, who initiated and essentially completed the fine structural analysis of L. This analysis warrants a separate course. But to give the reader a hint of this beautiful subject we pause to prove a representative case.

Theorem 6.13. *Assume* ZF + $V = L$. *Let κ be an infinite cardinal and let $A \subseteq \kappa$. Let λ be a limit ordinal such that $A \in L_{\lambda+1} - L_\lambda$. Then there exists $f : \kappa \to L_\lambda$ such that f is onto and $f \in L_{\lambda+1}$.*

Proof. By Exercise 6.9 the wellordering $<_\lambda$ of L_λ is Σ_1 definable (say, via ψ) over L_λ. For $\varphi(x, y_1, \ldots, y_n) \in \mathscr{L}$ and $\bar{b} \in (L_\lambda)^n$ let

$$f_\varphi(\bar{b}) = \begin{cases} <_\lambda \text{-least } a \text{ such that } L_\lambda \models \varphi[a, \bar{b}] & \text{if such exists} \\ \varnothing & \text{otherwise.} \end{cases}$$

Fix a formula $\varphi \in \mathscr{L}$ such that for some $\bar{b} \in L_\lambda$

$$A = \{x \in L_\lambda \mid L_\lambda \models \varphi[a, \bar{b}]\}.$$

Let p_A be the $<_\lambda$-least such $\bar{b}$. Now let Γ contain all subformulas of ψ, φ, and $V = L$ and let

$$H = \text{the closure of } \kappa \cup \{p_A\} \text{ under } f_\varphi \text{ for } \varphi \in \Gamma.$$

Let $\pi : H \to M$ be the collapsing map. Since $M \models V = L$, $M = L_{\bar{\lambda}}$ for some limit ordinal $\bar{\lambda}$.

Now notice that $\pi \restriction \kappa$ is the identity map. So, for $\alpha < \kappa$,

$$\begin{aligned} L_\lambda \models \varphi[\alpha, p_A] &\leftrightarrow H \models \varphi[\alpha, p_A] \\ &\leftrightarrow L_{\bar{\lambda}} \models \varphi[\pi(\alpha), \pi(p_A)] \\ &\leftrightarrow L_{\bar{\lambda}} \models \varphi[\alpha, \pi(p_A)]. \end{aligned}$$

Thus A is definable over $L_{\bar{\lambda}}$. So $L_\lambda = L_{\bar{\lambda}}$!

Now, what about H? Is it a *proper* subset of L_λ that managed to get rearranged into all of L_λ when we took its collapse?

Well, p_A is definable over L_λ as the $<_\lambda$-least $\bar{b}$ such that A is definable over L_λ via φ and $\bar{b}$. Now $\pi(A) = A$, $\pi(\ulcorner\varphi\urcorner) = \ulcorner\varphi\urcorner$ and so $\pi(p_A) = \varphi_A$. But there is an L_λ-definable map from $\kappa \cup \{p_A\}$ onto H. So, by the elementarity of π, $H = L_\lambda$! □

D. Suslin's Hypothesis. Suppose κ is an infinite regular cardinal. $C \subseteq \kappa$ is *club* (closed, unbounded) in κ if C is unbounded and for all $\alpha < \kappa$ if $\sup(C \cap \alpha) = \alpha$ then $\alpha \in C$. $S \subseteq \kappa$ is *stationary* in κ if $S \cap C \neq \varnothing$ for every C club in κ.

Exercise 6.11. Show that if $\langle C_\alpha \mid \alpha < \beta \rangle$ is such that C_α is club in κ and $\beta < \kappa$ then $\bigcap_{\alpha<\beta} C_\alpha$ is club in κ.

Exercise 6.12. (Fodor's lemma) Suppose κ is an infinite regular cardinal, S is stationary in κ, and $f : S \to \kappa$ is *regressive* in the sense that for all $a \in S$, $f(\alpha) < \alpha$. Show that there exists a β such that

$$\{\alpha \in S \mid f(\alpha) = \beta\}$$

is stationary.

Definition 6.14. Let κ be an infinite regular cardinal. $\Diamond_\kappa$ is the statement asserting the existence of a sequence $\langle C_\alpha \mid \alpha < \kappa \rangle$ such that

(a) $C_\alpha \subseteq \alpha$ for all $\alpha < \kappa$ and

(b) for any $X \subseteq \kappa$, $\{\alpha \mid X \cap \alpha = C_\alpha\}$ is stationary for all $\alpha < \kappa$.

Think of $\Diamond_\kappa$ as asserting the existence of a "guessing sequence" for subsets of κ—a *single* sequence that can "stationarily guess" *any* $X \subseteq \kappa$.

We shall primarily be interested in $\Diamond = \Diamond_{\omega_1}$. Note that, like CH, $\Diamond$ is a statement about $P(\omega_1)$. In fact, it can be viewed as a strong version of CH.

Exercise 6.13. Show that $\Diamond$ implies CH.

Theorem 6.15. (Jensen) *Assume* ZF + $V = L$. *Then* $\Diamond$ *holds.*

Theorem 6.16. (Jensen) *Assume* $\Diamond$. *Then there is a Suslin tree.*

Corollary 6.17. ZF + $V = L \vdash \neg$SH.

D. Removing the Inaccessible

Definition 6.18. (Relativization to L) For φ atomic let $\varphi^L = \varphi$. Inductively define $(\varphi \vee \psi)^L = \varphi^L \vee \psi^L$, $(\neg\varphi)^L = \neg(\varphi^L)$, $(\exists v\, \varphi(v))^L = \exists v\, \textsc{constr}(v) \wedge (\varphi(v))^L$.

The proofs show:

Theorem 6.19. ZF $\vdash \varphi^L$ *for each* φ *of* ZF + $V = L$.

E. Measurability Hypothesis.

Chapter 7

Forcing

Having used the method of "inner models" to show that ZFC cannot refute CH, ¬SH, or ¬MH we now aim to prove the companion result that ZFC cannot refute ¬CH, SH, or MH. For this a new method is needed—the method of "outer models".

To see the need for a new method suppose that we wish to establish the above independence results using the method of inner models. This method produced a model $L \subseteq V$ such that for any $n < \omega$ we could prove in ZFC that L satisfies ZFC$_n$. Clearly L will not serve our present purposes since (as we showed in the last chapter) L satisfies $V = L$ and hence CH, ¬SH, and ¬MH. As a next attempt one might try to show that V, or some model N such that $L \subsetneq N \subsetneq V$, is a model of $V \neq L$. But this too is impossible since to do so would be to refute $V = L$ in ZFC.

The independence technique we seek is one that works in ZFC and since in ZFC we cannot rule out the possibility that $V = L$ this technique must work in L. To ensure that we meet this constraint let us for the time being adopt ZFC + $V = L$ as our background theory. To make matters simpler let us in addition assume that κ is an inaccessible cardinal. (It is easy to see that if κ is inaccessible then κ is inaccessible in L and so our additional assumption is not in conflict with our assumption that $V = L$.) One strategy is to find a *countable* model of $V \neq L$. We start by letting $X \prec L_\kappa$ be countable and then collapsing to obtain $L_\lambda \cong X$. Of course, L_λ is not our desired model since it satisfies ZFC + $V = L$. But there is hope that we might extend this model by adding a real x to get an outer model $L_\lambda[x]$ that satisfies ZFC + ¬CONSTR(x). At first this might strike one as hopeless. For x *is* constructible and it would appear that $L_\lambda[x]$ must be able to recognise this fact. The point is that $L_\lambda[x]$ need not recognise that x is constructible since the notion of being constructible is Σ_1 and there is no guarantee that $L_\lambda[x]$ contains the existential witness to this fact. So, contrary to appearances, there is hope of producing such an outer model. But a difficulty remains. For we want to ensure that $L_\lambda[x]$ satisfies ZFC and for this purpose not any x will do. For example, if x happens to code a function f from ω onto L_λ then $L_\lambda[x]$ will think that there is an onto map from λ onto the entire universe, a dramatic violation of ZFC. So we have to ensure that x does not have any such problematic properties. Our strategy will be to focus on x such that L_λ knows enough about x that if there is a problem with $L_\lambda[x]$ satisfying ZFC then there will also be a problem with L_λ satisfying ZFC. We also wish to arrange that L_λ knows enough about $L_\lambda[x]$ to notice that it satisfies the negation of the statement we wish to show to be independent. The trick is to do all of this without providing L_λ with enough knowledge to actually calculate x!

In general we will not confine ourselves to L. Instead we will work in ZFC. We will start with a countable transitive model M of ZFC, called the *ground model*. This model will contain a partial order $\mathbb{P} = (P, \leqslant)$ that will determine a *generic extension* $M[G]$. Our initial concerns will be:

(A) (Existence) To show that $M[G]$ exists

(B) (Truth) To show that truth in $M[G]$ can be suitably approximated in M (which in turn will enable us)

(C) (Preservation) To show that $M[G]$ satisfies ZFC.

Our next concern will be to show that for suitably chosen $\mathbb{P}$ the generic extension $M[G]$ satisfies the statement the independence of which we wish to establish.

7.1 Basic Forcing

A. The generic extension $M[G]$. Let M be a transitive model of ZFC and let $\mathbb{P} = (P, \leqslant)$ be a partial ordering in M. Throughout we shall assume that our partial orderings $\mathbb{P}$ have a maximal element $1_\mathbb{P}$ and we shall abuse notation by writing $\mathbb{P}$ for P. Neither the assumption nor the abuse is necessary but each renders the presentation smoother.

Definition 7.1. Let M be a transitive model of ZFC and let $\mathbb{P} = (P, \leqslant)$ be a partial ordering in M.

(1) $D \subseteq \mathbb{P}$ is *dense in* $\mathbb{P}$ if $\forall p \in \mathbb{P}\, \exists q \leqslant p\, (q \in D)$.

(2) $G \subseteq \mathbb{P}$ is *M-generic* if

  (a) $\forall p \in G\, \forall q \geqslant p\, (q \in G)$,

  (b) $\forall p, q \in G\, \exists r \in G\, (r \leqslant p, q)$, and

  (c) $\forall D \in M\, (D \text{ is dense} \rightarrow G \cap D \neq \varnothing)$.

This is one of the central definitions in the theory of forcing. To get a feel for the notion of genericity it is worthwhile to establish some equivalent formulations.

Definition 7.2. Let M be a transitive model of ZFC and let $\mathbb{P} = (P, \leqslant)$ be a partial ordering in M. For $p, q \in \mathbb{P}$ we write $p \perp q$ if there does not exist an $r \leqslant p, q$ and we write $p \parallel q$ if there does exist an $r \leqslant p, q$. In the former case we say that p and q are *incompatible* and in the latter case we say that p and q are *compatible*. Further, we say that

(1) $D \subseteq \mathbb{P}$ is *dense below* p if $\forall q \leqslant p \in \mathbb{P}\, \exists r \leqslant q\, (r \in D)$

(2) $X \subseteq \mathbb{P}$ is *open* if $\forall p \in X\, \forall q \leqslant p\, (q \in X)$,

(3) $X \subseteq \mathbb{P}$ is *predense* if $\{q \mid \exists p \in X\, (q \leqslant p)\}$ is dense (open),

(4) $A \subseteq \mathbb{P}$ is an *antichain* if $\forall p, q \in A\, (p \neq q \rightarrow p \perp q)$, and

(5) $A \subseteq \mathbb{P}$ is a *maximal antichain* if A is an antichain and $\forall p \in \mathbb{P}\, \exists q \in A\, (p \parallel q)$.

Exercise 7.1. Let M be a transitive model of ZFC and let $\mathbb{P} = (P, \leqslant)$ be a partial ordering in M. Show that we get an equivalent definition of M-genericity in each of the following changes to Definition 7.1:

(1) Replace $2(c)$ with '$G \neq \varnothing$ and $\forall p \in G\, \forall D \in M\, (D \text{ is dense below } p \rightarrow G \cap D \neq \varnothing)$'.

(2) Replace 'dense' in $2(c)$ by 'dense open'.

(3) Replace 'dense' in $2(c)$ by 'predense'.

(4) Replace ‘dense’ in $2(c)$ by ‘maximal antichain’.

(5) Replace $2(b)$ with the weaker condition ‘$\forall p, q \in G\,(p \parallel q)$’.

The first thing that we need to ensure is that M-generics exist. In the case where M is a *countable* transitive model this is almost immediate. (In this chapter we will focus on countable transitive models. But later, when we consider forcing axioms, it will be of interest to meet more than countably many dense sets and so we will with cases where the models are not assumed to be countable. Indeed we will be talking of V-generics.)

Theorem 7.3. (Existence) *Let M be a countable transitive model of* ZFC *and let $\mathbb{P} \in M$ be a partial order. Then there is an M-generic $G \subseteq \mathbb{P}$.*

Proof. Let $\langle D_i \mid i < \omega \rangle$ enumerate the sets in M that are dense in $\mathbb{P}$. Pick $p_0 \in D_0$. Since D_1 is dense in $\mathbb{P}$ there is a $p_1 \leqslant p_0$ such that $p_1 \in D_1$. Continuing in this way we define a sequence of conditions

$$p_0 \geqslant p_1 \geqslant p_2 \geqslant \cdots$$

such that $p_n \in D_n$ for each $n < \omega$. Thus,

$$G = \{p \in \mathbb{P} \mid \exists n < \omega\,(p \geqslant p_n)\}$$

is an M-generic. □

The next thing that we need to ensure is that if $G \subseteq \mathbb{P}$ is M-generic then it is non-trivial in the sense that it is not an element of M. This will be true for any non-trivial partial ordering. Let us say that a partial ordering $\mathbb{P}$ is *splitting* if for every $p \in \mathbb{P}$ there exist $q, r \leqslant p$ such that $q \perp r$.

Exercise 7.2. Suppose that M is a transitive model of ZFC and $\mathbb{P} \in M$ a partial order that is splitting. Show that if $F \in M$ is a filter over $\mathbb{P}$ then $\{p \in \mathbb{P} \mid p \notin F\}$ is dense in $\mathbb{P}$. In particular, if $G \subseteq \mathbb{P}$ is M-generic, then $G \notin M$.

Let us now turn to the model $M[G]$. To define $M[G]$ and prove that it satisfies some of the basic axioms of ZFC it will suffice to assume only that $G \subseteq \mathbb{P}$ is a filter (which we will always assume is non-empty (and hence contains $1_{\mathbb{P}}$)). To ensure that the remaining axioms of ZFC hold in $M[G]$ we will have to assume, in addition, that G is M-generic.

Suppose that M is a transitive model of ZFC, $\mathbb{P} \in M$ is a partial order, and $G \subseteq \mathbb{P}$ is a filter. The model $M[G]$ is obtained by interpreting *names* in M via the filter G. The $\mathbb{P}$-*names* are defined by transfinite recursion:

$$\begin{aligned} \text{Name}_0 &= \varnothing \\ \text{Name}_{\alpha+1} &= P(\text{Name}_\alpha \times \mathbb{P}) \\ \text{Name}_\lambda &= \textstyle\bigcup_{\alpha<\lambda} \text{Name}_\alpha, \text{ if } \lambda \text{ is a limit} \\ \text{Name} &= \textstyle\bigcup_{\alpha<\text{On}} \text{Name}_\alpha. \end{aligned}$$

For $\tau \in \text{Name}$, let n-rank(τ) be the least α such that $\tau \in \text{Name}_\alpha$. The interpretation τ^G of a name τ under G is defined by transfinite recursion:

$$\tau^G = \{\sigma^G \mid \langle \sigma, p\rangle \in \tau \wedge p \in G\}.$$

The generic extension $M[G]$ of M is defined by setting

$$M[G] = \{\tau^G \mid \tau \in \text{Name}^M\}.$$

With each element $x \in M$ we associate a name that is recursively defined as follows

$$\check{x} = \{\langle \check{y}, 1_{\mathbb{P}}\rangle \wedge y \in x\}$$

and we let

$$\dot{G} = \{\langle \check{p}, p\rangle \mid p \in \mathbb{P}\}.$$

Exercise 7.3. Suppose M is a transitive model of ZFC, $\mathbb{P} \in M$ is a partial order, and $G \subseteq \mathbb{P}$ is a filter. Then

(1) $\check{x}^G = x$, for each $x \in M$,

(2) $\dot{G}^G = G$, and

(3) $\text{rank}(\sigma^G) \leqslant \text{n-rank}(\sigma)$, for each $\sigma \in \text{Name}^M$.

Corollary 7.4. *Suppose M is a transitive model of* ZFC*, $\mathbb{P} \in M$ is a partial order, and $G \subseteq \mathbb{P}$ is a filter. Then $M \subseteq M[G]$, $G \in M[G]$, $M[G]$ is transitive, and* $(\text{On})^M = (\text{On})^{M[G]}$.

Proof. It is clear from the exercise that $M \subseteq M[G]$ and $G \in M[G]$. To see that $M[G]$ is transitive consider $x \in M[G]$. Let $\tau \in \text{Name}^M$ be such that $\tau^G = x$. Consider $y \in \tau^G$. Since $\tau^G = \{\sigma^G \mid \langle \sigma, p\rangle \in \tau \wedge p \in G\}$ there exists $(\sigma, p) \in \tau$ such that $y = \sigma^G$. But since M is transitive, $\sigma \in M$ and so $y \in M[G]$. Finally, since M and $M[G]$ are transitive, $(\mathrm{On})^M = \mathrm{On} \cap M$ and $(\mathrm{On})^{M[G]} = \mathrm{On} \cap M[G]$; since $M \subseteq M[G]$,$\mathrm{On} \cap M \leqslant \mathrm{On} \cap M[G]$, and by (3) of the exercise, $\mathrm{On} \cap M[G] \leqslant \mathrm{On} \cap M$. Thus, $(\mathrm{On})^M = (\mathrm{On})^{M[G]}$. □

Theorem 7.5. *Suppose M is a transitive model of* ZFC*, $\mathbb{P} \in M$ is a partial order, and $G \subseteq \mathbb{P}$ is a filter. Then $M[G]$ satisfies Extensionality, Foundation, Emptyset, Infinity, Pairing and Union.*

Proof. It is immediate that $M[G]$ satisfies Extensionality and Foundation since $M[G]$ is transitive. In addition, since $M[G]$ is transitive and $\varnothing, \omega \in M[G]$ (as $(\mathrm{On})^M = (\mathrm{On})^{M[G]}$ and this is true in M), $M[G]$ satisfies Emptyset and Infinite. It remains to verify Pairing and Union.

Pairing. For ${\sigma_1}^G, {\sigma_2}^G \in M[G]$, letting $\tau = \{\langle \sigma_1, 1_{\mathbb{P}} \rangle, \langle \sigma_2, 1_{\mathbb{P}} \rangle\}$, we have $\tau^G = \{{\sigma_1}^G, {\sigma_2}^G\}$. Moreover, the name τ is in M since M satisfies Pairing.

Union. For $\sigma^G \in M[G]$, letting

$$\tau = \{\langle \rho, r \rangle \mid \exists \rho' \, \exists p, q \in \mathbb{P} \, (\langle \rho, p \rangle \in \rho' \wedge \langle \rho', q \rangle \in \sigma) \wedge r \leqslant p, q\},$$

we have $\tau^G = \cup \sigma^G$. Moreover, the name τ is in M since M satisfies Union. □

Note: The assumption that 'M satisfies ZFC' can be replaced by 'M satisfies ZF' throughout this subsection. In the background theory we have made use of a fragment of AC in only one place, namely, in Theorem 7.3 we made use of DC. But in the cases of interest we will be able to dispense with this use since the partial ordering under consideration will have a definable well-ordering.

B. The forcing relation. To establish that the remaining axioms hold in $M[G]$ we will have to do additional work. To appreciate the problem imagine that we are trying to show that $M[G]$ satisfies Comprehension. Let $\sigma^G, {\sigma_0}^G, \ldots, {\sigma_n}^G \in M[G]$ and $\varphi(x, x_0, \ldots, x_n) \in \mathscr{L}$. We aim to show that the set $A = \{x \in \sigma^G \mid M[G] \models \varphi[x, {\sigma_0}^G, \ldots, {\sigma_n}^G]\}$ is in $M[G]$. This will involve showing that there is a name τ in M such that $\tau^G = A$. Now, in the verification of the axioms thus far we always used the axiom in M to build a name which when interpreted certified this instance of the axiom in $M[G]$. For example, given ${\sigma_1}^G, {\sigma_2}^G \in M[G]$ we used Pairing in M to construct a name for $\{{\sigma_1}^G, {\sigma_2}^G\}$. Following this pattern, we would like to be able to use Comprehension in M to build a name for A. We can certainly get our hands on the ingredients $\sigma, \sigma_0, \ldots, \sigma_n$ in M. But in order to design our name so that when interpreted it will give us A we need to have some grip on when $M[G]$ satisfies $\varphi[x, \sigma_0, \ldots, \sigma_n]$. But of course we cannot do this for *any* filter $G \subseteq \mathbb{P}$. For the filter could code up random information that M cannot see. We shall thus have to restrict our attention to certain, well-behaved filters. The goal is to find a collection of filters such that with regard to this collection *truth in* $M[G]$ can be approximated by *truth in* M. It turns out that M-genericity is the key ingredient.

Now, there is no hope that M will be able to determine whether $M[G]$ satisfies an arbitrary statement with parameters—say, $\varphi({\sigma_1}^G, \ldots, {\sigma_n}^G)$—for the simple reason that M cannot see $M[G]$. Note, however, that if we fix a condition $p \in \mathbb{P}$ and restrict our attention to generics G such that $p \in G$ then M will be able to see a fragment of $M[G]$ (since $M[G]$ is completely determined by the names (which M can see) and the generic (which, by our assumption, M can partially see)). Thus, although we cannot reason about a particular generic extension $M[G]$ in M we can reason about fragments of such models and there is hope that we might be able to say something about *truth in all models* $M[G]$ *where* G *is such that* $p \in G$. More precisely, for a given condition $p \in \mathbb{P}$ and a given sentence $\varphi({\sigma_1}^G, \ldots, {\sigma_n}^G)$ there is hope that within M we might define the relation $p \Vdash\!\!= \varphi(\sigma_1, \ldots, \sigma_n)$ that holds iff for all M-generics G such that $p \in G$, $M[G] \models \varphi({\sigma_1}^G, \ldots, {\sigma_n}^G)$. According to this relation, p settles the truth of $\varphi({\sigma_1}^G, \ldots, {\sigma_n}^G)$ *regardless* of G.

The trick is to capture this notion (which involves quantification over objects that are external to M) inside M. The situation is analogous to that involved in the completeness theorem. The completeness theorem shows that the semantic notion of truth in all models of a theory T is equivalent to the syntactic notion of provability in T. We seek a theorem that shows that the semantic notion of truth in all models $M[G]$ such that $p \in G$ is equivalent to a notion (which we shall denote $\Vdash$) that is definable in M.

We shall define $\Vdash$ in M by a careful induction, bearing in mind along the way our aim to capture the semantic notion $\Vdash\!\!=$. Three key features of $\Vdash$ that we wish to secure are:

(1) (Persistence) If $p \Vdash \varphi$ and $q \leqslant p$ then $q \Vdash \varphi$.

(2) (Consistency) If $p \Vdash \varphi$ then $p \nVdash \neg\varphi$.

(3) (Completeness) If $p \nVdash \varphi$ then $\exists q \leqslant p\, (q \Vdash \neg\varphi)$.

We will then show that for every M-generic G the following are equivalent:

(A) $M[G] \models \varphi({\sigma_1}^G, \dots, {\sigma_n}^G)$.

(B) $\exists p \in G\, (p \Vdash \varphi)$.

This will suffice to show that $\Vdash$ and $\Vvdash$ coincide: For the first direction assume that $p \Vdash \varphi$. Assume for contradiction that $p \not\Vvdash \varphi$. Let G be an M-generic such that $p \in G$ and $M[G] \models \neg\varphi$. Then, by the above equivalence, $\exists q \in G\, (q \Vdash \neg\varphi)$. We may assume (by Persistence and the fact that G is a filter) that $q \leqslant p$. But then $q \Vdash \varphi$ and $q \Vdash \neg\varphi$ which is a contradiction (by Consistency). For the other direction assume that $p \Vvdash \varphi$. Assume, for contradiction, that $p \nVdash \varphi$. Then (by Completeness) $\exists q \leqslant p\, (q \Vdash \neg\varphi)$. Let G be an M-generic such that $q \in G$. Then, by the above equivalence, $M[G] \models \neg\varphi$, which is a contradiction.

Let us now motivate the particulars of the definition of the forcing relation. The induction is an induction on both rank and formula complexity. The clauses involving $\tau_1 = \tau_2$ and $\tau_1 \in \tau_2$ are by induction on rank. Assume that we have defined both $p \Vdash \tau_1 \in \tau_2$ and $p \Vdash \tau_1 \in \tau_2$ for terms up to a given rank and we have established at this level of complexity that $\Vdash$ and $\Vvdash$ coincide. We then wish to define these relations for terms of the next higher rank and establish the coincidence at that level.

Keeping in mind our desire to capture $\Vvdash$ we want $p \Vdash \tau_1 \in \tau_2$ to hold iff every generic G through p hits a point $q \leqslant p$ that *ensures* ${\tau_1}^G$ to be equal to a member τ^G of ${\tau_2}^G$. By induction q will ensure that ${\tau_1}^G$ is equal to a member τ^G of ${\tau_2}^G$ iff $q \Vdash \tau = \tau_1$. Thus, we want $p \Vdash \tau_1 \in \tau_2$ to hold iff

$$\{q \mid \exists\langle\tau, r\rangle \in \tau_2\, (q \leqslant r \wedge q \Vdash \tau_1 = \tau)\}$$

is dense below p. Granting our assumption that $q \Vdash \tau_1 = \tau$ has already been defined in M, the above set will be definable in M and so we will have defined $q \Vdash \tau_1 \in \tau_2$ in M.

Similarly, we want $p \Vdash \tau_1 = \tau_2$ to hold iff for any potential member $\langle\sigma_1, q_1\rangle$ of τ_1 that gets into ${\tau_1}^G$ via a generic G through p, G hits a point $r \leqslant p$ that ensures ${\sigma_1}^G$ to be equal a member ${\sigma_2}^G$ of ${\tau_2}^G$ and likewise for any potential member $\langle\sigma_2, q_2\rangle$ of τ_2. By induction $r \leqslant p$ will ensure that ${\sigma_1}^G$ is equal to a member ${\sigma_2}^G$ of ${\tau_2}^G$ iff $r \Vdash \sigma_1 = \sigma_2$. Thus, we want $p \Vdash \tau_1 = \tau_2$ to hold iff (i) for all $\langle\sigma_1, q_1\rangle \in \tau_1$ the set

$$D_{\sigma_1,q_1} = \{r \mid r \leqslant q_1 \to \exists\langle\sigma_2, q_2\rangle \in \tau_2\, (r \leqslant q_2 \wedge r \Vdash \sigma_1 = \sigma_2)\}$$

is dense below p and (ii) for all $\langle \sigma_2, q_2 \rangle \in \tau_2$ the set

$$D_{\sigma_2,q_2} = \{r \mid r \leqslant q_2 \to \exists \langle \sigma_1, q_1 \rangle \in \tau_1 \, (r \leqslant q_1 \wedge r \Vdash \sigma_1 = \sigma_2)\}$$

is dense below p.

Though technical, there is nothing surprising in these clauses. The clauses for conjunction and existential quantification are even more straightforward. The only clause that is surprising is the clause for negation. For the motivation of this clause it is best to examine the proof.

Definition 7.6. Let M be a transitive model of ZF and let $\mathbb{P} \in M$ be a partial order. For $p \in \mathbb{P}$ and for $\varphi(\sigma_1, \ldots \sigma_n)$ in the forcing language, the relation $\pi \Vdash \varphi(\sigma_1, \ldots, \sigma_n)$ is defined by transfinite recursion as follows:

(1) $p \Vdash \tau_1 = \tau_2$ iff (i) for all $\langle \sigma_1, q_1 \rangle \in \tau_1$, the set

$$D_{\sigma_1,q_1} = \{r \mid r \leqslant q_1 \to \exists \langle \sigma_2, q_2 \rangle \in \tau_2 \, (r \leqslant q_2 \wedge r \Vdash \sigma_1 = \sigma_2)\}$$

is dense below p and (ii) for all $\langle \sigma_2, q_2 \rangle \in \tau_2$, the set

$$D_{\sigma_2,q_2} = \{r \mid r \leqslant q_2 \to \exists \langle \sigma_1, q_2 \rangle \in \tau_1 \, (r \leqslant q_1 \wedge r \Vdash \sigma_1 = \sigma_2)\}$$

is dense below p.

(2) $p \Vdash \tau_1 \in \tau_2$ iff

$$\{q \mid \exists \langle \tau, r \rangle \in \tau_2 \, (q \leqslant r \wedge q \Vdash \tau_1 = \tau)\}$$

is dense below p.

(3) $p \Vdash \varphi \wedge \psi$ iff $p \Vdash \varphi$ and $p \Vdash \psi$.

(4) $p \Vdash \neg\varphi$ iff $\forall q \leqslant p \, (q \nVdash \varphi)$.

(5) $p \Vdash \exists x \, \varphi(x, \sigma_1, \ldots, \sigma_n)$ iff $\{q \mid \exists \tau \, (q \Vdash \varphi(\tau, \sigma_1, \ldots, \sigma_n))\}$ is dense below p.

Exercise 7.4. Prove the following:

(i) If $p \Vdash \varphi$ and $q \leqslant p$ then $q \Vdash \varphi$.

(ii) If $\{q \mid q \Vdash \varphi\}$ is dense below p then $p \Vdash \varphi$.

Corollary 7.7. *If* $p \nVdash \varphi$ *then* $\exists q \leqslant p\,(q \Vdash \neg\varphi)$.

Proof. Suppose $p \nVdash \varphi$. By the exercise $\{q \mid q \Vdash \varphi\}$ is *not* dense below p, i.e. $\exists \bar{p} \leqslant p \forall q \leqslant \bar{p}\,(q \nVdash \varphi)$. So, by the definition of $p \Vdash \neg\varphi$, we have that $\bar{p} \Vdash \neg\varphi$. □

Thus we have

(1) (Persistence) If $p \Vdash \varphi$ and $q \leqslant p$ then $q \Vdash \varphi$.

(2) (Consistency) If $p \Vdash \varphi$ then $p \nVdash \neg\varphi$.

(3) (Completeness) If $p \nVdash \varphi$ then $\exists q \leqslant p\,(q \Vdash \neg\varphi)$.

Corollary 7.8. $\{p \mid p \Vdash \varphi$ or $p \Vdash \neg\varphi\}$ *is dense.*

Proof. Fix p. If $\pi \Vdash \varphi$ then we are done. So assume $p \nVdash \varphi$. Then, by the previous corollary, there exists $q \leqslant p$ such that $q \Vdash \neg\varphi$. □

Exercise 7.5. Show that $p \Vdash \neg\neg\varphi$ iff $p \Vdash \varphi$.

Theorem 7.9. (Fundamental Forcing Theorem) *Suppose M is a transitive model of* ZF, $\mathbb{P} \in M$ *is a partial order, and* $G \subseteq \mathbb{P}$ *is an M-generic filter. Then*:

(1) (Truth) *For all for all* $\sigma_1, \ldots, \sigma_n \in \mathrm{Name}^M$, $M[G] \models \varphi[{\sigma_1}^G, \ldots, {\sigma_n}^G]$ *iff* $\exists p \in G\ p \Vdash \varphi(\sigma_1, \ldots, \sigma_n)$.

(2) (Definability) *The relation* $(p \Vdash \varphi(\sigma_1, \ldots, \sigma_n))^M$ *is uniformly definable over M.*

Proof. The second part of the theorem is immediate from the definition. The proof of the first part is by induction. The case $\tau_1 = \tau_2$ is proved first, in isolation, by induction on rank. The case $\tau_1 \in \tau_2$ is then proved on the basis of the previous case, again by induction on rank. The remaining cases are proved by induction on formula complexity.

Case $\tau_1 = \tau_2$. ($\rightarrow$) Assume ${\tau_1}^G = {\tau_2}^G$. It will suffice to show that $\{p \in \mathbb{P} \mid p \Vdash \tau_1 = \tau_2\}$ is dense. Notice that $p \nVdash \tau_1 = \tau_2$ is equivalent to the statement that (i) $\exists\langle\sigma_1, q_1\rangle \in \tau_1$ such that D_{σ_1,q_1} is not dense below p or (ii) $\exists\langle\sigma_2, q_2\rangle \in \tau_2$ such that D_{σ_1,q_1} is not dense below p. Now, to say that D_{σ_1,q_1} is not dense below p is to say that

$$\exists \bar{p} \leqslant p \forall r \leqslant \bar{p}\left(r \leqslant q_1 \wedge \forall\langle\sigma_2, q_2\rangle \in \tau_2\, \neg(r \leqslant q_2 \wedge r \Vdash \sigma_1 = \sigma_2)\right)$$

and likewise for D_{σ_2,q_2}. It follows that the set

$$\begin{aligned} D = \big\{p \in \mathbb{P} \mid p \Vdash \tau_1 = \tau_2 \\ &\vee \exists\langle\sigma_1, q_1\rangle \in \tau_1 \big[\forall r \leqslant p \,\big(r \leqslant q_1 \wedge \forall\langle\sigma_2, q_2\rangle \in \tau_2 \\ &\qquad \neg(r \leqslant q_2 \wedge r \Vdash \sigma_1 = \sigma_2)\big)\big] \\ &\vee \exists\langle\sigma_2, q_2\rangle \in \tau_2 \big[\forall r \leqslant p \,\big(r \leqslant q_2 \wedge \forall\langle\sigma_1, q_1\rangle \in \tau_1 \\ &\qquad \neg(r \leqslant q_2 \wedge r \Vdash \sigma_1 = \sigma_2)\big)\big]\big\} \end{aligned}$$

is dense. So there is a $p \in G \cap D$. We claim that $p \in D$ in virtue of the first disjunct, that is, $p \Vdash \tau_1 = \tau_2$. For suppose that $p \in D$ in virtue of the second disjunct, at p. (The case of the third disjunct is similar.) Then there is a $\langle\sigma_1, q_1\rangle \in \tau_1$ such that

$$\forall r \leqslant p\big(r \leqslant q_1 \wedge \forall\langle q_2, \sigma_2\rangle \in \tau_2 \, \neg(r \leqslant q_2 \wedge r \Vdash \sigma_1 = \sigma_2)\big).$$

Since $p \in G$, $q_1 \in G$ and so ${\sigma_1}^G \in {\tau_1}^G = {\tau_2}^G$. Let $\langle\sigma_2, q_2\rangle \in \tau_2$ witness ${\sigma_1}^G \in {\tau_2}^G$, that is, be such that $\langle\sigma_2, q_2\rangle \in \tau_2$ and $q_2 \in G$ and ${\sigma_1}^G = {\sigma_2}^G$. Now, in this case, we have been assuming, as our induction hypothesis, that the theorem holds for all terms or rank less than $\max(\text{rank}(\tau_1), \text{rank}(\tau_2))$. This is true of σ_1 and σ_2 (since they are in τ_1 and τ_2, respectively. So, since ${\sigma_1}^G = {\sigma_2}^G$, by the induction hypothesis, there exists $q \in G$ such that $q \Vdash \sigma_1 = \sigma_2$. Since $p, q_2, q \in G$, there exists an $r \in G$ such that $r \leqslant p, q_2, q$ and so we have found (for our initial $\langle\sigma_1, q_1\rangle \in \tau_1$) an $r \leqslant p, q_1$ such that $\exists\langle\sigma_2, q_2\rangle$ (namely, our chosen such) such that $r \leqslant q_2$ and $r \Vdash \sigma_1 = \sigma_2$ (by Persistence), which is a contradiction.

($\leftarrow$) Assume $p \in G$ is such that $p \Vdash \tau_1 = \tau_2$. We aim to show that ${\tau_1}^G \subseteq {\tau_2}^G$. (The converse is similar.) Fix $a \in {\tau_1}^G$. Let $\langle\sigma_1, q_1\rangle$ be such that $q_1 \in G$ and $a = {\sigma_1}^G$. Let $\bar{p} \leqslant p, q_1$ be such that $r \in G$. So $\bar{p} \Vdash \tau_1 = \tau_2$. Thus

$$\{r \mid r \leqslant q_1 \rightarrow \exists\langle\sigma_2, q_2\rangle \in \tau_2 \, (r \leqslant q_2 \wedge r \Vdash \sigma_1 = \sigma_2)\}$$

is dense below $\bar{p}$. Now the antecedent is vacuous since we are considering $r \leqslant \bar{p} \leqslant q_1$. So G hits this set at a point r and there exists $\langle\sigma_2, q_2\rangle \in \tau_2$ such that

$$r \leqslant q_2 \quad\text{and}\quad r \Vdash \sigma_1 = \sigma_2.$$

Since $q_2 \in G$, ${\sigma_2}^G = {\sigma_1}^G$. Since $r \in G$, ${\sigma_1}^G = {\sigma_2}^G$, by the induction hypothesis. Thus $a = {\sigma_1}^G \in {\tau_1}^G$.

Case $\tau_1 \in \tau_2$. ($\to$) Assume ${\tau_1}^G \in {\tau_2}^G$. Let $\langle \tau, r\rangle \in \tau_2$ be such that $r \in G$ and ${\tau_1}^G = \tau^G$. By the induction hypothesis, there exists $q \in G$ such that $q \Vdash \tau_1 = \tau$. Letting $p \leqslant r, q$ we have $p \Vdash \tau_1 \in \tau_2$.

($\leftarrow$) Assume $p \in G$ is such that $p \Vdash \tau_1 \in \tau_2$, that is, such that

$$\{q \mid \exists \langle \tau, r\rangle \in \tau_2\, (q \leqslant r \wedge q \Vdash \tau_1 = \tau)\}$$

is dense below p. Since G is generic it hits this set at a point, say q. So there exists $(\tau, r) \in \tau_2$ such that $q \leqslant r$ and $q \Vdash \tau_1 = \tau$. We have $r \in G$ and so $\tau^G \in {\tau_2}^G$ and (by the induction hypothesis) ${\tau_1}^G = \tau^G$. Thus ${\tau_1}^G \in {\tau_2}^G$.

Case $\varphi \wedge \psi$. ($\to$) Assume $M[G] \models \varphi \wedge \psi$. So $M[G] \models \varphi$ and $M[G] \models \psi$. By our induction hypothesis, there exist $p_1, p_2 \in G$ such that

$$p_1 \Vdash \varphi \quad \text{and} \quad p_2 \Vdash \psi.$$

Letting $p \leqslant p_1, p_2$ be such that $p \in G$, we have $p \Vdash \varphi \wedge \psi$.

($\leftarrow$) Assume $p \in G$ is such that $p \Vdash \varphi \wedge \psi$. So $p \Vdash \varphi$ and $p \Vdash \psi$ and hence, by our induction hypothesis, $M[G] \models \varphi$ and $M[G] \models \psi$ and so $M[G] \models \varphi \wedge \psi$.

Case $\neg\varphi$. ($\to$) Assume $M[G] \models \neg\varphi$. The set

$$D = \{p \mid p \Vdash \varphi \quad \text{or} \quad p \Vdash \neg\varphi\}$$

is dense. So there exists $p \in G \cap D$. Now we can't have $p \Vdash \varphi$ since, by our induction hypothesis, this would imply $M[G] \models \varphi$. So $p \Vdash \neg\varphi$.

($\leftarrow$) Assume $p \in G$ is such that $p \Vdash \neg\varphi$. Assume, for contradiction, that $M[G] \models \varphi$. By the induction hypothesis, there exists $q \in G$ such that $q \Vdash \varphi$. Letting r be such that $r \leqslant p, q$, we have $r \Vdash \neg\varphi$ and $r \Vdash \varphi$, which is a contradiction.

Case $\exists x\, \varphi(x, \sigma_1, \ldots, \sigma_n)$. ($\to$) Assume $M[G] \models \exists x\, \varphi(x, {\sigma_1}^G, \ldots, {\sigma_n}^G)$. Let $\tau \in \mathrm{Name}^M$ be such that $M[G] \models \varphi(\tau^G, {\sigma_1}^G, \ldots, {\sigma_n}^G)$. Then, by our induction hypothesis, there exists $p \in G$ such that $p \Vdash \varphi(\tau, \sigma_1, \ldots, \sigma_n)$. Trivially,

$$\{q \mid \exists \tau\, (q \Vdash \varphi(\tau, \sigma_1, \ldots, \sigma_n))\}$$

is dense below p. Thus, $p \Vdash \exists x\, \varphi(x, \sigma_1, \ldots, \sigma_n)$.

($\leftarrow$) Assume $p \in G$ is such that $p \Vdash \exists x\, \varphi(x, \sigma_1, \ldots, \sigma_n)$, i.e.,

$$\{q \mid \exists \tau\, (q \Vdash \varphi(\tau, \sigma_1, \ldots, \sigma_n))\}$$

is dense below p. By our induction hypothesis $M[G] \models \varphi(\tau^G, {\sigma_1}^G, \ldots, {\sigma_n}^G)$, and thus we have $M[G] \models \exists x\, \varphi(x, {\sigma_1}^G, \ldots, {\sigma_n}^G)$. □

C. Preservation of ZFC.

Theorem 7.10. *Suppose that M is a transitive model of* ZF*, $\mathbb{P} \in M$ is a partial order, and $G \subseteq \mathbb{P}$ is M-generic. Then $M[G]$ is a model of* ZF*. If, in addition, M satisfies* AC*, then $M[G]$ satisfies* AC*.*

Proof. We have already shown that $M[G]$ satisfies Extensionality, Foundation, Emptyset, Pairing, Union and Infinity. It remains to verify Comprehension, Powerset, Collection and Choice.

Comprehension. Suppose $\sigma^G, {\sigma_0}^G, \ldots, {\sigma_n}^G \in M[G]$ and $\varphi(x, x_0, \ldots, x_n) \in \mathscr{L}$. We aim to show

$$A = \{x \in \sigma^G \mid M[G] \models \varphi[x, {\sigma_0}^G, \ldots, {\sigma_0}^G]\}$$

is in $M[G]$. Notice that by the Fundamental Forcing Theorem

$$A = \{\rho^G \mid \rho \in \mathrm{dom}(\sigma) \text{ and } \exists p \in G\,(p \Vdash \rho \in \sigma \wedge \varphi(\rho, \sigma_0, \ldots, \sigma_n))\}.$$

Thus, letting

$$\tau = \{\langle \rho, p \rangle \in \mathrm{dom}(\sigma) \times \mathbb{P} \mid p \Vdash \rho \in \sigma \wedge \varphi(\rho, \sigma_0, \ldots, \sigma_n)\}$$

we have $A = \tau^G$. To see this, first note that τ is indeed in M (using Definability and Comprehension). Second,

$$\tau^G = \{\rho^G \mid \rho \in \mathrm{dom}(\sigma) \text{ and } \exists p \in G\,(p \Vdash \rho \in \sigma \wedge \varphi(\rho, \sigma_0, \ldots, \sigma_n))\}$$

which is just A.

Powerset. Suppose $\sigma^G \in M[G]$. We wish to show $(P(\sigma^G))^{M[G]}$ and for this it suffices (by Comprehension) to find $\tau^G \in M[G]$ such that $(P(\sigma^G))^{M[G]} \subseteq \tau^G$.

First, we show that for each $\rho^G \subseteq \sigma^G$ there is a "bounded" name $\bar{\rho}$ such that $\bar{\rho}^G = \rho^G$. Let

$$\bar{\rho} = \{\langle \pi, p \rangle \in \mathrm{dom}(\sigma) \times \mathbb{P} \mid p \Vdash \pi \in \rho\}.$$

It is easy to see (using the Fundamental Forcing Theorem) that $\bar{\rho}^G = \rho^G$.

We can now collect together all of the bounded names: Let

$$\tau = \{\langle \rho, 1_{\mathbb{P}} \rangle \mid \rho \in \mathrm{Name}^M \wedge \rho \subseteq \mathrm{dom}(\sigma) \times \mathbb{P}\}.$$

Using Powerset in M we have that $\tau \in \mathrm{Name}^M$. Moreover, each bounded name appears in $\mathrm{dom}(\tau)$. Thus, $(P(\sigma^G))^{M[G]} \subseteq \tau^G$.

Collection. Suppose $\sigma^G, {\sigma_0}^G, \ldots, {\sigma_n}^G \in M[G]$ and let $\varphi(x, y, x_0, \ldots, x_n) \in \mathscr{L}$ be such that

$$M[G] \models \forall x \in \sigma^G \, \exists y \, \varphi[x, y, {\sigma_0}^G, \ldots, {\sigma_n}^G].$$

We seek τ^G such that

$$M[G] \models \forall x \in \sigma^G \, \exists y \in \tau^G \, \varphi[x, y, {\sigma_0}^G, \ldots, {\sigma_n}^G].$$

Let

$$\begin{aligned} f : \mathrm{dom}(\sigma) \times \mathbb{P} &\to M \\ \langle \rho, p \rangle &\mapsto \begin{cases} \mu\alpha \, \exists \rho' \in \mathrm{Name}^M_\alpha \, (p \Vdash \varphi(\rho, \rho', \sigma_0, \ldots, \sigma_n)) & \text{if such exists} \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

(Here we have used the notation '$\mu\alpha$' for 'the least α such that'.) By Definability and Collection in M, the range of f is bounded. Let β be such that $\mathrm{ran}(f) \subseteq V^M_\beta$. Let

$$\tau = \{\langle \rho', 1_{\mathbb{P}} \rangle \mid \rho' \in \mathrm{Name}^M_\beta\}.$$

Then $\tau \in M$. It is easy to see (by the Fundamental Forcing Theorem) that

$$M[G] \models \forall x \in \sigma^G \, \exists y \in \tau^G \, \varphi[x, y, {\sigma_0}^G, \ldots, {\sigma_n}^G].$$

Choice. Suppose that $M \models \mathrm{AC}$. In models of ZF, AC is equivalent to the statement that for all x there exists an ordinal α and a surjective map $f : \alpha \to y$ such that $x \subseteq y$. Consider $x = \sigma^G$ in $M[G]$. In M, let α and f be such that $f : \alpha \to \mathrm{dom}(\sigma)$ is surjective. Let

$$\begin{aligned} i_G : \mathrm{Name}^M &\to M[G] \\ \sigma &\mapsto \sigma^G \end{aligned}$$

be the interpretation map. Since for each $x \in M$, $i_G \restriction x \in M[G]$, we have that $i_G \circ f \in M[G]$. Now

$$i_G \circ f : \alpha \to y$$

is surjective and $\sigma^G \subseteq y$. Thus, $M[G] \models \mathrm{AC}$. □

7.2 The Continuum Hypothesis.

We are now in a position to complete the proof of the independence of CH from ZFC. Let us first establish the independence of $V = L$. Let M be a countable transitive model of ZFC, let $\mathbb{P}$ be any partial ordering in M that is splitting and let $G \subseteq M$ be M-generic. Since $\mathbb{P}$ is splitting, by Exercise 7.2, we have that $G \notin M$. Thus $M \subsetneq M[G]$. But $L^{M[G]} = L^M \subseteq M$ and so $G \in M[G]$, yet $G \notin L^{M[G]}$. Thus, $M[G] \models \mathrm{ZFC} + V \neq L$.

The above argument starts with the assumption that there is a countable transitive model of ZFC and this assumption is a good deal stronger than the assumption that ZFC is consistent. We now refine the proof, in a way that will prove to be a template for what follows.

Theorem 7.11. *Assume that* ZFC *is consistent. Then* $\mathrm{ZFC} + V \neq L$ *is consistent.*

Proof. Assume that ZFC is consistent. By the Completeness Theorem, let N be a model of ZFC. There are two key points. The first key point is that for each finite fragment F of ZFC, ZFC proves that there is a countable transitive model of F. To see this one uses reflection to get a level V_λ which satisfies F (simply reflect the conjunction of the statements in F) and then one takes the transitive collapse of a countable elementary substructure of V_λ. The second key point is that the proof that $M[G]$ satisfies ZFC is "local" in the sense that to show that $M[G]$ satisfies a finite fragment F of ZFC we only had to assume that M satisfies a (somewhat larger) finite fragment F' of ZFC. The same applies when we show that $M[G]$ satisfies $F + V \neq L$.

Let F be a finite fragment of ZFC and let F' be a finite fragment of ZFC such that assuming M is a countable transitive model of F' we can show that $M[G]$ is a countable transitive model of $F + V \neq L$. Working in N (which satisfies ZFC but need not be transitive) let (by the first point) M be a countable transitive model of F'. (Note: M might not really be transitive. But it doesn't matter.) By the second point we have (in N) that $M[G]$ is a countable transitive model of $F + V \neq L$. Since F was an arbitrary finite fragment we have shown that all finite fragments of $\mathrm{ZFC} + V \neq L$ are consistent. Thus, $\mathrm{ZFC} + V \neq L$ is consistent. □

We have not been explicit about the background theory in which the above theorem is proved. It certainly suffices to assume ZFC to run the model theoretic proof. But, in fact, much less suffices. In fact, it can be implemented in a conservative extension of PRA. The conservative extension is called WKL_0. We will not pause to describe the details of this theory. Suffice it to say that it is a subsystem of second-order arithmetic, it suffices to prove the completeness and compactness theorems, and, moreover, in a precise sense it is the weakest system in which these last theorems (which are essential for the proof) are provable. We give a sketch.

Theorem 7.12. (WKL_0) *Assume that* ZFC *is consistent. Then* $\mathrm{ZFC}+V\neq L$ *is consistent.*

Proof. Let L_T be the theory in the language of set theory extended by an additional constant $\dot{M}$. Let T be the theory in L_T that consists of the axioms of ZFC along with the infinite collection of axioms asserting "$\dot{M} \models \varphi$" for each theorem φ of ZFC. Work in WKL_0. Assume that ZFC is consistent. By reflection and compactness it follows that T is consistent. But now, by the above argument, it follows that the theory T' obtained by altering T so that the distinctive axioms are "$\dot{M} \models \varphi + V\neq L$" is consistent. So, $\mathrm{ZFC} + V\neq L$ is consistent. $\square$

In what follows we will always start with the stronger assumption that M is a countable transitive model of ZFC. But the above will serve as a template (passed over in silence) for reducing the assumption to Con(ZFC) and getting an outright relative consistency result (in WKL_0).

B. The Failure of CH. We now seek a partial ordering which adds so many reals that $M[G]$ does not satisfy CH. Let us start by adding a single real. Working inside M, consider the partial ordering $\mathrm{Add}(\omega,1) = (\{p : \omega \to 2 \mid |\mathrm{dom}(p)| < \omega\}, \subseteq)$. (Notice that we are altering our notation slightly since in writing '$p : \omega \to 2$' we are allowing partial functions.) Let $G \subseteq \mathrm{Add}(\omega,1)$ be M-generic. Since $\mathrm{Add}(\omega,1)$ is splitting, $G \notin M$. Let us examine G. Strictly speaking G is a set of conditions p. But it is naturally regarded as a function $x : \omega \to 2$ by setting $x(n) = i$ iff $\exists p \in G\,(p(n) = i)$. Notice that $M[G] = M[x]$. For this reason we will identify G with the corresponding function x. Since $D_n = \{p \mid n \in \mathrm{dom}(p)\}$ is dense in $\mathrm{Add}(\omega,1)$ for each $n < \omega$, G is a total function from ω to 2, in other words G is a real. The generic real G has some remarkable features that justify calling it 'generic'. For example, for each s be an arbitrary sequence of 0's and 1's, the generic real G contains s infinitely often! It has all finite binary information. There is nothing particularly special about G. It is quite unlike the binary expansion of π.

We would like now to add many reals. Suppose that M is a countable transitive model of ZFC and we would like a generic extension $M[G]$ in which CH fails. For definiteness suppose we would like $M[G]$ to satisfy ZFC and $2^\omega = \omega_2$. The most obvious way to do this is to select in M a partial ordering the conditions of which approximate an ${\omega_2}^M$-sequence of reals and then shoot a generic G through this partial ordering. But there are many such partial orderings not all of which will work. The trouble is that we have to make sure that when we add ${\omega_2}^M$-many new reals we do not inadvertently add a map which collapses ${\omega_2}^M$. The point is that we are adding ${\omega_2}^M$-many reals, yet we wish to ensure that $M[G] \models 2^\omega = \omega_2$. So we need to ensure that ${\omega_2}^M = {\omega_2}^{M[G]}$. It turns out that the natural generalisation of the partial ordering that we used above is one that works. Working inside M consider the partial ordering $\mathrm{Add}(\omega, \omega_2) = (\{p : \omega_2 \times \omega \to 2 \mid |\mathrm{dom}(p)| < \omega\}, \supseteq)$. Let $G \subseteq \mathrm{Add}(\omega, \omega_2)$ be M-generic. Since for each $\alpha \in \omega_2$ and $n \in \omega$, $D_{\alpha,n} = \{p \mid \langle \alpha, n\rangle \in \mathrm{dom}(p)\}$ is dense in $\mathrm{Add}(\omega, \omega_2)$, G is a function from $\omega_2 \times \omega$ to 2. (Again we are making the identification between G and the corresponding function). We can extract an ω_2-sequence of reals from G_α by letting, for $\alpha < \omega_2$, $G_\alpha(n) = i$ iff $G(\alpha, n) = i$. Since $D_{\alpha,\beta} = \{p \mid \exists n < \omega\, (p(\alpha, n) \neq p(\beta, n))\}$ is dense in $\mathrm{Add}(\omega, \omega_2)$ for each $\alpha, \beta < \omega_2$ such that $\alpha \neq \beta$, the reals G_α and G_β are distinct. Thus forcing with $\mathrm{Add}(\omega, \omega_2)$ has the effect of adding at least ${\omega_2}^M$ many reals. As noted above, this would suffice to show that $M[G] \models \neg\mathrm{CH}$ if only we knew that ${\omega_2}^{M[G]} = {\omega_2}^M$. To ensure this last point we will have to be careful about counting names for reals.

It will be useful at the outset to work in a more general setting since in addition to adding subsets of ω we would like to add subsets of κ for any infinite cardinal.

Definition 7.13. For $I, J \in V$, and $\nu \in$ CARD, let

$$\mathrm{Fn}_\nu(I, J) = (\{p : I \to J \mid |\mathrm{dom}(p)| < \nu\}, \supseteq).$$

For $\kappa \in$ CARD and $\lambda \in \mathrm{On}$, let

$$\mathrm{Add}(\kappa, \lambda) = \mathrm{Fn}_\kappa(\lambda \times \kappa, 2)$$

Definition 7.14. Let κ be an uncountable cardinal. A partial ordering $\mathbb{P}$ has the *κ-chain condition* iff every antichain in $\mathbb{P}$ has size strictly less than κ. (When we say that a partial ordering has the *countable chain condition* (*c.c.c.*) we mean that it has the ω_1-chain condition.)

The importance of the notion of being κ-c.c. is that it provides M with "$<\kappa$-good approximations" of functions $f : X \to Y$ in $M[G]$:

Lemma 7.15. *Let M be a transitive model of* ZFC *such that* $(\mathbb{P}$ *is* κ*-c.c.*$)^M$ *Let $G \subseteq \mathbb{P}$ be M-generic. Suppose $X, Y \in M$ and $f : X \to Y$ is an element of $M[G]$. Then in M there is an $F : X \to P(Y)$ such that for all $x \in M$*

(1) $f(x) \in F(x)$ *and*

(2) $(|F(x)| < \kappa)^M$.

Proof. Let $\dot{f}$ be such that $\dot{f}^G = f$. Let $p \in G$ be such that

$$p \Vdash \dot{f} : \check{X} \to \check{Y}.$$

Work in M. Using AC let, for each $x \in X$, $A(x)$ be a maximal set of pairwise incompatible conditions $q \leqslant p$ such that

$$\forall q \in A(x) \, \exists y \in Y \; q \Vdash \dot{f}(\check{x}) = \check{y}.$$

Let $F(x) = \{y \in Y \mid \exists p \in A(x) \, p \Vdash \dot{f}(\check{x}) = \check{y}\}$. These are the "possible values" of $f(x)$. Since $\mathbb{P}$ is κ-c.c., $|A(x)| < \kappa$ and so $|F(x)| < \kappa$. It remains to prove (1).

Suppose $f(x) = y$. Let $q' \in G$ be such that $p \Vdash \dot{f}(\check{x}) = \check{y}$. We may assume that $q' \leqslant p$. By the maximality of $A(x)$ there must exist $q \in A(x)$ such that $q \parallel q'$. Letting $r \leqslant q, q'$ we have that $r \Vdash \dot{f}(\check{x}) = \check{y}$. But $q \Vdash \dot{f}(\check{x}) = \check{z}$ for some $z \in Y$. Thus $z = y$ and hence $f(x) = y \in F(x)$. □

Definition 7.16. Let M be a transitive model of ZF and let $\mathbb{P} \in M$. We say that $\mathbb{P}$ *preserves cardinals* if for all M-generic $G \subseteq \mathbb{P}$, for all κ

$$(\kappa \in \text{CARD})^M \leftrightarrow (\kappa \in \text{CARD})^{M[G]}$$

and we say that $\mathbb{P}$ *preserves cofinalities* if for all M-generic $G \subseteq \mathbb{P}$, for all α

$$(\text{cof}(\alpha))^M = (\text{cof}(\alpha))^{M[G]}$$

and we say that $\mathbb{P}$ *preserves regularity* if for all M-generic $G \subseteq \mathbb{P}$, for all κ

$$(\kappa \in \text{REG})^M \leftrightarrow (\kappa \in \text{REG})^{M[G]}.$$

Likewise, we will say that $\mathbb{P}$ *preserves cardinals* $\leqslant \lambda$ when the above condition holds for all $\kappa \leqslant \lambda$, we say that $\mathbb{P}$ *preserves cofinalities* $\leqslant \lambda$ when the above condition holds for all α such that $(\text{cof}(\alpha) \leqslant \lambda)^M$, etc.

Exercise 7.6. Let M be a transitive model of ZF and let $\mathbb{P} \in M$. Show that the following are equivalent

(1) $\mathbb{P}$ preserves cofinalities $\leqslant \lambda$

(2) $\mathbb{P}$ preserves regularity $\leqslant \lambda$.

Show that a similar equivalence holds when '$\leqslant$' is replaced by '$\geqslant$'. Show that assuming $M \models$ ZFC if $\mathbb{P}$ preserves cofinalities $\leqslant \lambda$ then $\mathbb{P}$ preserves cardinals $\leqslant \lambda$ (and likewise for $\geqslant \lambda$). Thus, to show that $\mathbb{P}$ preserves cardinals and cofinalities it suffices to show that $\mathbb{P}$ preserves regularity.

Theorem 7.17. *Let M be a transitive model of* ZFC *and let $\mathbb{P} \in M$ be such that* ($\mathbb{P}$ *is κ-c.c. and κ is regular*)^M. *Then $\mathbb{P}$ preserves cardinals and cofinalities $\geqslant \kappa$.*

Proof. By the exercise it suffices to show that $\mathbb{P}$ preserves regularity $\geqslant \kappa$. Consider $\lambda \geqslant \kappa$ such that $\lambda \in \text{REG}^M$. Assume, for contradiction, that $\lambda \notin \text{REG}^{M[G]}$ and let $f \in M[G]$ witness this, say $f : \bar{\lambda} \to \lambda$ is cofinal where $\bar{\lambda} < \lambda$. Since ($\mathbb{P}$ is κ-c.c.)^M, M has a $< \kappa$-approximation F (as in Lemma 7.15). Since $\lambda \in \text{REG}^M$, F is bounded in λ. But since for all $\alpha < \bar{\lambda}$, $f(\alpha) \in F(\alpha)$ and f is cofinal in λ, F must be cofinal in λ. □

Corollary 7.18. *Let M be a transitive model of* ZFC *and let $\mathbb{P} \in M$ be such that* ($\mathbb{P}$ *is c.c.c.*)^M. *Then $\mathbb{P}$ preserves all cardinals and cofinalities.*

Thus it remains to show that $\text{Add}(\omega, \omega_2)$ is c.c.c. Toward this end we introduce the notion of a Δ-system.

Definition 7.19. Let C be a collection of finite sets. C is a Δ-*system* if there exists R (called the *root* of C) such that

$$\forall X, Y \in C\, (X \neq Y \to X \cap Y = R).$$

Lemma 7.20. (Δ-system lemma) *Suppose that C is an uncountable collection of finite sets. Then there exists an uncountable $\bar{C}$-system $\bar{C} \subseteq C$.*

Proof. Since C is uncountable there are uncountable many members of the same cardinality. So we may assume that all of the members of C have cardinality n. The lemma is proved by induction on $n < \omega$. The base case $n = 1$ is trivial since we can take $R = \varnothing$. Assume then that the lemma is true for n and let C be an uncountable collection of finite sets of cardinality $n+1$. There are two cases:

Case 1: There is an element a that is contained in uncountably many elements of C. Let

$$D = \{X - \{a\} \mid X \in C \wedge a \in X\}.$$

By induction D has an uncountable Δ-system $\bar{D}$. So $\bar{C} = \{X \cup \{a\} \mid X \in \bar{D}\}$ is an uncountable Δ-system for C.

Case 2: There is no element a that is contained in uncountably many elements of C. Since each element a belongs to at most countably many elements of C we can construct a sequence $\bar{C} = \{X_\alpha \mid \alpha < \omega_1\}$ inductively as follows: Suppose we have constructed $\{X_\xi \mid \zeta < \alpha\}$ where $\alpha < \omega_1$. There are only countably many elements $a \in \bigcup_{\xi<\alpha} X_\zeta$, each of which is absent from a tail of elements of C. So there is an element X of C that is disjoint from each X_ζ for $\xi < \alpha$. Let X_α be such an element. $\bar{C}$ is an uncountable Δ-system for C with root $\varnothing$. □

Lemma 7.21. *For all κ, $\mathrm{Add}(\omega, \kappa)$ is c.c.c.*

Proof. Assume, for contradiction, that

$$A = \{p_\alpha \mid \alpha < \omega_1\}$$

is an antichain in $\mathrm{Add}(\omega, \kappa)$. Let $C = \{\mathrm{dom}(p_\alpha) \mid \alpha \in A\}$. Let $\bar{C}$ be an uncountable Δ-system for C with root R. Let

$$B = \{p_\alpha \mid \mathrm{dom}(p_\alpha) \in \bar{C}\}.$$

Since there are only finitely many functions $r : R \to 2$ there is an uncountable set $\bar{B} \subseteq B$ consisting of conditions p_α such that $p_\alpha \restriction R = p$ for some fixed p. But then any two elements in C are compatible, which is a contradiction. □

Theorem 7.22. *Let M be a countable transitive model of* ZFC*. Then there is a generic extension $M[G]$ that satisfies* $\mathrm{ZFC} + \neg\mathrm{CH}$.

Proof. Let $G \subseteq (\mathrm{Add}(\omega, \omega_2))^M$ be M-generic. There are at least ${\omega_2}^M$ many reals in $M[G]$. But since $(\mathrm{Add}(\omega, \omega_2)$ is c.c.c$)^M$, ${\omega_2}^M = {\omega_2}^{M[G]}$. Thus $M[G] \models \mathrm{ZFC} + \neg\mathrm{CH}$. □

We now wish to show that if we assume, in addition that $M \models$ GCH, then, $M[G] \models 2^\omega = \omega_2$. To do this we must be careful about counting names.

Names can be redundant. For example, if τ is a name that contains $\langle \check{x}, 1_\mathbb{P} \rangle$ then whenever $G \subseteq \mathbb{P}$ is M-generic we shall have $x \in \tau^G$ and so if τ also contains $\langle \check{x}, p \rangle$ then this element is redundant as far as ensuring that $x \in \tau^G$ is concerned. Similarly, if τ contains $\langle \check{x}, p \rangle$ for an antichain A of p then whenever $G \subseteq \mathbb{P}$ is M-generic we shall have $x \in \tau^G$ and so if τ also contains $\langle \check{x}, q \rangle$ where $q \notin A$ then this element is redundant as far as ensuring that $x \in \tau^G$ is concerned.

In the proof that $M[G] \models$ Powerset we showed that for each $\rho^G \subseteq \sigma^G$ that there is a "bounded" name $\bar{\rho}^G$ such that $\bar{\rho}^G = \rho^G$. We took

$$\bar{\rho} = \{ \langle \pi, p \rangle \in \mathrm{dom}(\sigma) \times \mathbb{P} \mid p \Vdash \pi \in \rho \}.$$

Now we show that there is an even better name, one with as few redundancies as possible.

Definition 7.23. Suppose $\sigma \in$ Name with respect to a partial order $\mathbb{P}$. An *antichain name* (with respect to σ) is a name ρ is a subset of $\bar{\rho}$ of $\mathrm{dom}(\sigma) \times \mathbb{P}$ with the feature that for each $\pi \in \mathrm{dom}(\sigma)$, $\{p \mid \langle \pi, p \rangle \in \bar{\rho})\}$ is an antichain.

Lemma 7.24. *Suppose M be a countable transitive model of* ZFC*, $\mathbb{P} \in M$ is a partial order, $G \subseteq \mathbb{P}$ is M-generic, and $\sigma^G \in M[G]$. Then, for each $\rho^G \subseteq \sigma^G$, there is an antichain name $\bar{\rho}$ such that $\bar{\rho}^G = \rho^G$.*

Proof. For each $\pi \in \mathrm{dom}(\sigma)$, let

$$B_\pi = \{ p \in \mathbb{P} \mid p \Vdash \pi \in \rho \}.$$

Notice that this set is a section of the name used to show $M[G] \models$ Powerset. Now, use AC to select a subset, letting, for each $\pi \in \mathrm{dom}(\sigma)$,

$$A_\pi \subseteq B_\pi$$

be a maximal antichain in B_π (that is, a maximal set of incompatible conditions in B_π). Finally, let

$$\bar{\rho} = \bigcup \{ \{\pi\} \times A_\pi \mid \pi \in \mathrm{dom}(\sigma) \}.$$

It is clear that $\bar{\rho}^G \subseteq \rho^G$. [If $a \in \bar{\rho}^G$ then $a = \pi^G$ for some $\pi \in \mathrm{dom}(\sigma)$ and $p \in G \cap A_\pi$ and for such p, $p \Vdash \pi \in \rho$ and so $\pi^G \in \rho^G$.] To see that $\rho^G \subseteq \bar{\rho}^G$, assume $a \in \rho^G$. So $a = \pi^G$ for some $\langle \pi, p \rangle \in \rho$ where $p \in G$. It follows that G must hit B_π. But then, since A_π is maximal, G must hit A_π and this point, say p, ensures that $\pi^G \in \bar{\rho}^G$. □

In particular, every subset ρ^G of an ordinal α in $M[G]$ has a name of the form

$$\bigcup\{\{\check{\beta}\} \times A_\beta \mid \check{\beta} \in \mathrm{dom}(\check{\alpha})\}$$

where A_β is an antichain.

Lemma 7.25. *Let M be a transitive model of* ZFC *and suppose* ($\mathbb{P}$ *is κ-c.c.*)^M. *Let $G \subseteq \mathbb{P}$ be M-generic and let α be a cardinal of $M[G]$. Then in $M[G]$ there are at most*

$$|(|\mathbb{P}|^{<\kappa})^\alpha|^M$$

subsets of α.

Proof. Since $\mathbb{P}$ is κ-c.c. there are at most $(|\mathbb{P}|^{<\kappa})^M$ antichains of $\mathbb{P}$ in M. Thus there are at most $|(|\mathbb{P}|^{<\kappa})^\alpha|^M$ many antichain names for subsets of α. □

Corollary 7.26. *Let M be a countable transitive model of* ZF. *Let $G \subseteq (\mathrm{Add}(\omega, \omega_2))^M$ be M-generic. Then $M[G] \models 2^\omega = \omega_2$.*

B. Collapsing Cardinals. We now show that one can use forcing to show that if ZFC is consistent then ZFC + CH is consistent, thereby establishing the independence of CH from ZFC.

Definition 7.27. For $\kappa \in \textsc{card}$ infinite and $X \in V$, let

$$\mathrm{Col}(\kappa, X) = \mathrm{Fn}_\kappa(\kappa, X).$$

Definition 7.28. For κ an uncountable cardinal, a partial ordering $\mathbb{P}$ is *κ-closed* ($<\kappa$-closed) if whenever $\alpha \leqslant \kappa$ ($\alpha < \kappa$) and

$$p_0 \geqslant p_1 \geqslant \cdots \geqslant p_\beta \geqslant \cdots$$

for $\beta < \alpha$ is a descending sequence of conditions, there is a condition $p \in \mathbb{P}$ such $p \leqslant p_\beta$ for all $\beta < \alpha$.

Lemma 7.29. *Suppose $\kappa \in \textsc{reg}$, $X \in V$, and $\lambda \in \mathrm{On}$ where $\lambda \geqslant \kappa$. Then $\mathrm{Col}(\kappa, X)$ is κ-closed and $\mathrm{Col}(\kappa, \lambda)$ is κ-closed and $(\lambda^{<\kappa})^+$-c.c.*

Proof. This is straightforward. □

The κ-c.c. ensured that for $X, Y \in M$, M contained "$<\kappa$-good approximations" F of functions $f : X \to Y$ in $M[G]$. This ensures that cardinals and cofinalities $\geqslant \kappa$ were preserved. The condition of κ-closure ensures that $M[G]$ contains no new γ-sequences of M, for any $\gamma < \kappa$. This ensures that cardinals and cofinalities $le\kappa$ are preserved.

Lemma 7.30. *Let M be a transitive model of* ZFC *and suppose* ($\mathbb{P}$ *is κ-closed*)^M. *Let $G \subseteq \mathbb{P}$ be M-generic. Then ${}^{\gamma}M \cap M[G] \subseteq M[G]$.*

Proof. Let $f : \gamma \to M$, $f \in M[G]$. Then for some ξ,

$$f : \gamma \to V_\xi^M.$$

Let $f = \dot{f}^G$ and let $p \in G$ be such that

$$p \Vdash \dot{f} : \check{\gamma} \to \check{V}_\xi^M.$$

Inductively define

$$\begin{aligned} p_0 &= p \\ p_{\alpha+1} &= \text{some } q \leqslant p_\alpha \left(\exists x \in V_\xi^M (q \Vdash \dot{f}(\check{\alpha}) = \check{x})\right) \\ p_\lambda &= \text{some lower bound of } \{p_\alpha \mid \alpha < \lambda\}, \text{ for } \lambda \in \text{Lim}. \end{aligned}$$

Let $q \leqslant p_\alpha$, all $\alpha < \lambda$. Then

$$q \Vdash \dot{f} = \check{h}$$

where

$$h(\alpha) = \text{ unique } x \text{ such that } p_{\alpha+1} \Vdash \dot{f}(\check{a}) = \check{x}.$$

Thus, $\forall r \leqslant p \, \exists q \leqslant r \, \exists h \in M$ such that

$$q \Vdash \dot{f} = \check{h}.$$

So, some such q is in G, which completes the proof. □

Corollary 7.31. *Let M be a transitive model of* ZFC *and suppose* ($\mathbb{P}$ *is κ-closed*)^M. *Then $\mathbb{P}$ preserves cardinals and cofinalities $\leqslant \kappa$.*

Theorem 7.32. *Assume* Con(ZFC). *Then* Con(ZFC + CH).

Proof. Let M be a countable transitive model of ZFC, let $G \subseteq (\text{Col}(\omega_1, \mathbb{R}))^M$ be M-generic, and let $f_G = \bigcup G$. So $f_G \in M[G]$ is a map from $(\omega_1)^M$ onto $\mathbb{R}^G$. Since ($\mathbb{P}$ is ω_1-closed)^M, $(\omega_1)^M = (\omega_1)^{M[G]}$ and $\mathbb{R}^M = \mathbb{R}^{M[G]}$. Thus, $M[G] \models \text{CH}$. □

Theorem 7.33. *Assume* Con(ZFC)*. Then* Con(ZFC + ${\omega_1}^L$ is countable)*.*

Proof. Let M be a countable transitive model of ZFC, let $G \subseteq (\mathrm{Col}(\omega, \omega_1))^M$ be M-generic, and let $f_G = \bigcup G$. So $f_G \in M[G]$ is a map from ω onto $(\omega_1)^M$ and thus $M[G]$ thinks $(\omega_1)^M$ is countable. But $({\omega_1}^L)^{M[G]} \leqslant (\omega_1)^M$. □

There are many variations on this, for example:

Corollary 7.34. *Assume* Con(ZFC)*. Then* Con(ZFC + ${\omega_1}^L$ is countable + $\omega_1 = {\omega_2}^L$)*.*

Proof. The same as the previous proof except starting with M a model of ZFC + $V = L$ (and so $M = L_\lambda$ for some countable λ). □

Definition 7.35. For $\kappa \in$ CARD infinite and $\lambda \in$ On, let

$$\mathrm{Col}(\kappa, {<}\lambda)$$

be the partial order consisting of partial functions $p : \kappa \times \lambda \to \lambda$ such that $|p| < \kappa$ and for all $(\alpha, \beta) \in \mathrm{dom}(p)$, $p(\alpha, \beta) < \beta$.

Lemma 7.36. *For $\kappa \in$ REG and $\lambda \in$ On, $\mathrm{Col}(\kappa, {<}\lambda)$ is κ-closed.*

Theorem 7.37. *If $\kappa \in$ CARD and $\lambda > \kappa$ is strongly inaccessible, then $\mathrm{Col}(\kappa, {<}\lambda)$ is λ-c.c.*

Proof. Let A be a maximal antichain in $\mathrm{Col}(\kappa, {<}\lambda)$. Let $\eta > \lambda$ be such that V_η satisfies some sufficiently large fragment of ZFC. Build an chain of elementary substructures $\langle X_\alpha \mid \alpha \leqslant \kappa \rangle$ $(V_\eta, A, \kappa, \lambda)$ such that $A \in X_0$ and for all $\alpha < \beta \leqslant \kappa$

(1) $X_\alpha \prec X_\beta \prec (V_\eta, A, \kappa, \lambda)$,

(2) $|X_\alpha| < \lambda$,

(3) ${}^{<\kappa}X_\alpha \subseteq X_{\alpha+1}$,

(4) $X_\alpha \cap \lambda \in \lambda$, and

(5) $X_\alpha = \bigcup_{\xi<\alpha} X_\xi$ for $\alpha \in$ Lim.

This is done by letting $X_{\alpha+1}$ be the Skolem closure of ${}^{<\kappa}X_\alpha \cup X_\alpha \cup (X_\alpha \cap \lambda)$ inside $(V_\eta, A, \kappa, \lambda)$. Notice

$$X_\kappa \cap \lambda \in \lambda \text{ and } {}^{<\kappa}X_\kappa \subseteq X_\kappa.$$

This is the key point. The point is that for $p \in A$, $\mathrm{ran}(p \restriction X_\kappa) \subseteq X_\kappa$ (since if $(\alpha, \beta) \in \mathrm{dom}(p)$ then $\beta \in X_\kappa$ and so $p(\alpha, \beta) < \beta \in X_\kappa$) and so $p \restriction X_\kappa \in X_\kappa$ (since ${}^{<\kappa}X_\kappa \subseteq X_\kappa$).

□

Theorem 7.38. Con(ZFC + there is a strongly inaccessible) *iff* Con(ZFC + ${\omega_1}^L$ is strongly inaccessible).

C. More General Results.

Exercise 7.7. $\mathrm{Fn}_\nu(I, J)$ is $(|J|^{<\nu})^+$-c.c.

Thus for $\kappa \in \text{REG}$ and $\lambda \in \mathrm{On}$, $\mathrm{Add}(\kappa, \lambda)$ is $(\kappa^{<\kappa})^+$-c.c. So if $\kappa^{<\kappa} = \kappa$ then $\mathrm{Add}(\kappa, \lambda)$ is κ-c.c. and so preserves cofinalities and cardinals $\geqslant \kappa$. Since $\mathrm{Add}(\kappa, \lambda)$ is $<\kappa$-closed this means that, under the assumption that $\kappa^{<\kappa} = \kappa$, $\mathrm{Add}(\kappa, \lambda)$ preserves all cardinals and cofinalities.

Exercise 7.8. Let M be a countable transitive model of ZFC + GCH. Let $\kappa \in \text{REG}^M$ and let $\lambda \in \mathrm{On}^M$ be such that $(\mathrm{cof}(\lambda) > \kappa)^M$. Let $G \subseteq \mathrm{Add}(\kappa, \lambda)^M$ be M-generic. Show that $M[G] \models 2^\kappa = \lambda$.

Exercise 7.9. Let M be a countable transitive model of ZFC. Find a forcing notion $\mathbb{P} \in M$ such that if $G \subseteq \mathbb{P}$ is M-generic then $M[G] \models \Diamond$. (It follows that $M[G] \models \neg$SH.)

7.3 Product Forcing

Suppose that we wish to use forcing to obtain a model in which $2^\omega = \omega_5$ and $2^{\omega_5} = \omega_7$. The natural way to do this is to start with a countable transitive model M that satisfies GCH and then force over M with $\mathrm{Add}(\omega, \omega_5)^M$ to obtain $M[G]$ and then force over $M[G]$ with $\mathrm{Add}(\omega_5, \omega_7)^{M[G]}$ to obtain $M[G][H]$. It is easy to see that this works. Now suppose that we wish to render $2^\omega = \omega_5$ and $2^{\omega_2} = \omega_7$. The natural way to do this is to force over M (as above) with $\mathrm{Add}(\omega, \omega_5)^M$ to obtain $M[G]$ and then force over $M[G]$

with $\mathrm{Add}(\omega_2, \omega_7)^{M[G]}$ to obtain $M[G][H]$. This doesn't work. The trouble is that $M[G]$ does not satisfy $(\omega_2)^{<\omega_2} = \omega_2$—we upset this fact with our first forcing—and so there is no guarantee that $\mathrm{Add}(\omega_2, \omega_7)$ has the ω_2-c.c. In fact, it collapses ${\omega_5}^{M[G]}$ to ω_2.

Exercise 7.10. Show that $M[G][H] \models 2^\omega = \omega_2$.

One way around this difficulty is to work backwards: First force over M with $\mathrm{Add}(\omega_2, \omega_7)^M$ to obtain $M[G]$ and then force over $M[G]$ with $\mathrm{Add}(\omega, \omega_5)^M$ to obtain $M[G][H]$.

Exercise 7.11. Show that this works.

It is easy to see that the backward approach works for finitely many steps; more precisely, if M is a countable transitive model of ZFC + GCH and $\kappa_1, \ldots, \kappa_n \in \textsc{reg}^M$ and $\lambda_1 \leqslant \ldots \leqslant \lambda_n \in \textsc{card}^M$ are such that $(\mathrm{cof}(\lambda_i) > \kappa_i)^M$ for all $i \leqslant n$, then letting $G_n \subseteq \mathrm{Add}(\kappa_n, \lambda_n)$ be M-generic, $G_{n_1} \subseteq \mathrm{Add}(\kappa_{n-1}, \lambda_{l-1})$ be $M[G_n]$-generic, $\ldots$, and $G_{n-1} \subseteq \mathrm{Add}(\kappa_1, \lambda_1)$ be $M[G_n]$ $\cdots$ $[G_2]$-generic, we obtain a model $M[G_n] \cdots [G_1]$ in which $2^{\kappa_1} = \lambda_1 \wedge \cdots \wedge 2^{\kappa_n} = \lambda_n$. But when we deal with an infinite sequence of regular cardinals κ_i troubles arise since it is unclear of what to do at limit stages. For this reason we pursue another approach.

A. Two steps.

Definition 7.39. For partial orders $\mathbb{P}$ and $\mathbb{Q}$ let

$$\mathbb{P} \times \mathbb{Q} = (\{(p, q) \mid p \in \mathbb{P} \wedge q \in \mathbb{Q}\}, \leqslant_{\mathbb{P}\times\mathbb{Q}})$$

where $(\bar{p}, \bar{q}) \leqslant_{\mathbb{P}\times\mathbb{Q}} (p, q)$ iff $\bar{p} \leqslant_{\mathbb{P}} p$ and $\bar{q} \leqslant_{\mathbb{Q}} q$.

We will suppress the subscript on '$\leqslant_{\mathbb{P}\times\mathbb{Q}}$' when it is clear from the context which ordering is intended.

Lemma 7.40. (Product Lemma) *Let M be a countable transitive model and let $\mathbb{P}$ and $\mathbb{Q}$ be partial orders of M. The following are equivalent:*

(1) *$F \subseteq \mathbb{P} \times \mathbb{Q}$ is M-generic.*

(2) *$F = G \times H$ where $G \subseteq \mathbb{P}$ is M-generic and $H \subseteq \mathbb{Q}$ is $M[G]$-generic.*

Proof. $(1 \to 2)$. Assume that $F \subseteq \mathbb{P} \times \mathbb{Q}$ is M-generic. Let

$$G = \{p \in \mathbb{P} \mid \exists q \in \mathbb{Q}\, (p,q) \in F\}$$
$$H = \{q \in \mathbb{Q} \mid \exists p \in \mathbb{P}\, (p,q) \in F\}.$$

Clearly, $F \subseteq G \times H$.

Claim 1. $F = G \times H$

Proof. Fix $(p,q) \in G \times H$. Let $q^* \in \mathbb{Q}$ and $p^* \in \mathbb{P}$ be such that $(p,q^*) \in F$ and $(p^*,q) \in F$. Choose $(\bar{p},\bar{q}) \in F$ such that $(\bar{p},\bar{q}) \leqslant (p,q^*),(p^*,q)$. Since $(\bar{p},\bar{q}) \leqslant (p,q)$, $(p,q) \in F$. □

It is easy to see that G and H are filters.

Claim 2. *$G \subseteq \mathbb{P}$ is M-generic.*

Proof. Let $D \in M$ be a dense in $\mathbb{P}$. It follows that $D \times \mathbb{Q}$ is dense in $\mathbb{P} \times \mathbb{Q}$ and so there is a point $(p,q) \in F \cap (D \times \mathbb{Q})$. Thus, $p \in G \cap D$. □

Claim 3. *$H \subseteq \mathbb{Q}$ is $M[G]$-generic.*

Proof. Let $D \in M[G]$ be dense in $\mathbb{Q}$. Let $\dot{D}$ be a $\mathbb{P}$-name such that $\dot{D}^G = D$. Let $p^* \in G$ be such that

$$p^* \Vdash \dot{D} \text{ is dense.}$$

Consider $(p^*, 1_{\mathbb{Q}})$. We claim that the set $\{(p,q) \in \mathbb{P} \times \mathbb{Q} \mid p \Vdash \check{q} \in \dot{D}\}$ is dense below $(p^*,1_{\mathbb{Q}})$: To see this fix $(p',q') \leqslant (p^*,1_{\mathbb{Q}})$ such that

$$p' \Vdash \dot{D} \text{ is dense.}$$

So $p' \Vdash \exists q \leqslant q'\, (\check{q} \in \dot{D})$. So there is a $p \leqslant p'$ and a q such that $p \Vdash \check{q} \in \dot{D}$.

Let $(p,q) \in F$ be in this set. So

$$p \Vdash \check{q} \in \dot{D}.$$

Since $p \in G$, $q^G \in H \cap D$. □

$(2 \to 1)$. Assume that $F = G \times H$ is such that $G \subseteq \mathbb{P}$ is M-generic and $H \subseteq \mathbb{Q}$ is $M[G]$-generic. It is easy to see that F is a filter. So it remains to show that F is M-generic. Let $D \subseteq \mathbb{P} \times \mathbb{Q}$ be dense open. Consider

$$D_G = \{q \in \mathbb{Q} \mid \exists p \in G\, (p,q) \in D\}.$$

It suffices to show that D_G is dense in $\mathbb{Q}$ since then there exists $q \in G \cap D_G$ such that there exists $p \in G\,(p,q) \in D$ and so $(p,q) \in F \cap D_G$. Fix $q^* \in \mathbb{Q}$. Given any $p \in \mathbb{P}$ there is (by the density of D) a point

$$(\bar{p}, \bar{q}) \leqslant (p^*, q^*)$$

and so $\bar{p} \leqslant p$. Hence

$$D = \{\bar{p} \in \mathbb{P} \mid \exists \bar{q}\,(\bar{p}, \bar{q}) \in D \wedge \bar{q} \leqslant q^*\}$$

is dense in $\mathbb{P}$. So there exists $\bar{p} \in G$ that hits this set. Thus $\bar{q} \in D_G$. □

Corollary 7.41. *Suppose $G \subseteq \mathbb{P}$ is M-generic and $H \subseteq \mathbb{Q}$ is $M[G]$-generic. Then G is $M[H]$-generic.*

Proof. By the lemma we have that $G \times H \subseteq \mathbb{P} \times \mathbb{Q}$ is M-generic. So $H \times G \subseteq \mathbb{Q} \times \mathbb{P}$ is M-generic. Thus, by the lemma, $H \subseteq \mathbb{Q}$ is M-generic and $G \subseteq \mathbb{P}$ is $M[H]$-generic. □

So $M[G \times H] = M[G][H] = M[H][G]$.

Exercise 7.12. Let $G \times H \subseteq \mathbb{P} \times \mathbb{Q}$ be M-generic. Show that if $X \in M[G] \cap M[H]$ and $X \subseteq M$ then $X \in M[G]$.

Exercise 7.13. Show that $\bigcap\{M[G] \mid G \subseteq \mathbb{P} \text{ is } M\text{-generic}\} = M$.

B. More Steps. Let $(\mathbb{P}_i \mid i \in I)$ be a sequence of partial orderings. For $p \in \prod_{i \in I} \mathbb{P}_i$, let $\text{supp}(p) = \{i \mid p(i) \neq 1_{\mathbb{P}_i}\}$ be the *support* of p. The *κ-support* product of $(\mathbb{P}_i \mid i \in I)$ is

$$(\{p \in \textstyle\prod_{i \in I} \mathbb{P}_i \mid |\text{supp}(p)| < \kappa\}, \leqslant)$$

where $p \leqslant q$ iff $\forall i \in I\,(p(i) \leqslant_{\mathbb{P}_i} q(i))$.

Suppose that M is a countable transitive model of ZFC and suppose ($\mathbb{Q}$ is κ-closed)^M. In Lemma 7.30 we showed that if $\mathbb{Q}$ is κ-closed and if $H \subseteq \mathbb{Q}$ is M-generic then $M[H]$ does not contain any new functions $f : \kappa \to M$. Suppose now that ($\mathbb{P}$ is κ^+-c.c.)^M and we first force with $\mathbb{P}$ to obtain $M[G]$ and then with $\mathbb{Q}$ to obtain $M[G][H]$. We should like to show that in the step from $M[G]$ to $M[H]$ no new functions $f : \kappa \to M$ are added. But we cannot simply appeal to Lemma 7.30 since we have no guarantee that ($\mathbb{Q}$ is κ-closed)$^{M[G]}$.

Lemma 7.42. (Easton) *Suppose M is a transitive model of* ZFC *and* ($\mathbb{P}$ *is κ^+-c.c. and $\mathbb{Q}$ is κ-closed*)^M. *Suppose $G \times H \subseteq \mathbb{P} \times \mathbb{Q}$ is M-generic. If $f : \kappa \to M$ is such that $f \in M[G][H]$, then $f \in M[G]$.*

Proof. Consider $f : \kappa \to M$ such that $f \in M[G][H]$. Let $\dot{f}$ be such that $\dot{f}^{G\times H} = f$. For $\alpha < \kappa$ let

$$D_\alpha = \{q \in \mathbb{Q} \mid \forall p \in \mathbb{P}\, \exists \bar{p} \leqslant p\, \exists x \in M\, ((\bar{p}, q) \Vdash \dot{f}(\check{\alpha}) = \check{x})\}.$$

Claim. *D_α is dense in $\mathbb{Q}$.*

Proof. Fix $q' \in \mathbb{Q}$. We seek $q \leqslant q'$ such that $q \in D_\alpha$. We will do this by a "possible values" argument that makes use of both the fact that $\mathbb{P}$ is κ^+-c.c. and that $\mathbb{Q}$ is κ-closed. By induction in M we will construct a sequence

$$\{(p_\gamma, q_\gamma) \mid \gamma < \beta\}$$

such that

(1) $\beta < \kappa^+$

(2) $(p_\gamma, q_\gamma) \Vdash \dot{f}(\check{a}_\gamma) = \check{x}_\gamma$ for some $\check{x}_\gamma \in M$

(3) $\{p_\gamma \mid \gamma < \beta\}$ is a maximal antichain in $\mathbb{P}$

(4) $q^{'} \geqslant q_0 \geqslant q_1 \geqslant \cdots \geqslant q_\gamma \geqslant \cdots$ for $\gamma < \beta$.

(This will suffice as then by the κ-closure of $\mathbb{Q}$ we can take $q \leqslant q_\gamma$ for all $\gamma < \beta$. Since $\{p_\gamma \mid \gamma < \beta\}$ is a maximal antichain, any point p can be extended to a point $\bar{p} \leqslant p_\gamma$ for some $\gamma < \beta$ and so we have $(\bar{p}, q) \Vdash \dot{f}(\check{\alpha}) = \check{x}_\gamma$.)

Now for the construction. For the base case let $p_0 \in \mathbb{P}$, $q_0 \leqslant q'$ and $\check{x}_0 \in M$ be such that

$$(p_0, q_0) \Vdash \dot{f}(\check{\alpha}) = \check{x}_0.$$

For the inductive step assume that we have defined $(p_{\bar{\gamma}}, q_{\bar{\gamma}})$ for $\bar{\gamma} < \gamma$ such that (1)–(4) hold. If $\{p_{\bar{\gamma}} \mid \bar{\gamma} < \gamma\}$ is a maximal antichain then we are done. If not then let $p_\gamma \in \mathbb{P}$, q_γ and $\check{x}_\gamma \in M$ be such that for all $\bar{\gamma} < \gamma$, $p_\gamma \perp p_{\bar{\gamma}}$, $q_\gamma \leqslant q_{\bar{g}}$, and $(p_\gamma, q_\gamma) \Vdash \dot{f}(\check{\alpha}) = \check{x}_\gamma$. ∎

Since $\mathbb{Q}$ is κ-closed

$$D = \bigcap_{\alpha<\beta} D_\alpha$$

is dense in $\mathbb{Q}$. Fix $q^* \in H \cap D$. Working in M, for each $\alpha < \kappa$ let E_α be the dense set given by q^* and D_α, that is, let

$$E_\alpha = \{\bar{p} \in \mathbb{P} \mid \exists x \in M\, (\bar{p}, q) \Vdash \dot{f}(\check{\alpha}) = \check{x}\}.$$

In $M[G]$ we can compute f as follows:

$$f(\alpha) = x \text{ iff } \exists p \in E_\alpha \cap G\, ((p, q) \Vdash \dot{f}(\check{\alpha}) = \check{x})\}.$$

□

In particular $(P(\kappa))^{M[G][H]} \subseteq (P(\kappa))^{M[G]}$.

Corollary 7.43. *Let M be a transitive model of* ZFC. *Suppose that $\mathbb{P} \in M$ is a partial order such that for all $\kappa \in \text{REG}^M$, $\mathbb{P} \cong \mathbb{P}^- \times \mathbb{P}^+$ where ($\mathbb{P}^-$ is κ^+-c.c. and $\mathbb{P}^+$ is κ-closed)^M. Then $\mathbb{P}$ preserves cardinals and cofinalities.*

Proof. By Exercise 7.6 it suffices to show that $\mathbb{P}$ preserves regularity. Let $\kappa \in \text{REG}^M$. Assume, for contradiction, that $\kappa \notin \text{REG}^{M[G]}$ and let $f : \text{cof}(\kappa) \to \kappa$ be a cofinal map. Since $\text{cof}(\kappa) \in \text{REG}^M$, we have $\mathbb{P} \cong \mathbb{P}^- \times \mathbb{P}^+$ where ($\mathbb{P}^-$ is $\text{cof}(\kappa)^+$-c.c. and $\mathbb{P}^+$ is $\text{cof}(\kappa)$-closed)^M. Let $G \subseteq \mathbb{P}$ be M-generic. We have $G = G^- \times G^+$. Since $\mathbb{P}^+$ is $\text{cof}(\kappa)$-closed, $f \in M[G]$ by Easton's Lemma. Since $\mathbb{P}^-$ is $\text{cof}(\kappa)^+$-c.c., M contains an $F : \text{cof}(\kappa) \to \kappa$ such that $\text{ran}(f) \subseteq \text{ran}(F)$, contradicting $\kappa \in \text{REG}^M$. □

Definition 7.44. Let M be a transitive model of ZFC. Let $I \subseteq \text{REG}^M$ be a set in M. An *Easton function* for M is an $F : I \to M$ such that for all $\kappa, \lambda \in I$

(1) $\kappa < \lambda \to F(\kappa) \leqslant F(\lambda)$

(2) $\text{cof}(F(\kappa)) > \kappa$.

Definition 7.45. Let F be an Easton function for M. The *Easton forcing* associated with F is the partial ordering

$$\begin{aligned}\mathbb{P}_F = \{p \mid\ & \text{dom}(p) = \text{dom}(F) \wedge \forall \kappa \in \text{dom}(p) \\ & (p(\kappa) \in \text{Add}(\kappa, F(\kappa))\omega \wedge |\text{supp}(p{\restriction}\kappa)| < \kappa)\}.\end{aligned}$$

Notice that the support condition $|\text{supp}(p{\restriction}\kappa)| < \kappa$ is trivial when κ is a successor cardinal. So the support condition is only doing work at weakly inaccessible cardinals.

Theorem 7.46. (Easton) *Let M be a transitive model of* ZFC + GCH, *let F be an Easton function for M, and let $G \subseteq \mathbb{P}_F$ be M-generic. Then $\mathbb{P}_F$ preserves all cardinals and cofinalities and*

$$M[G] \models \forall\kappa \in \mathrm{dom}(F)\,(2^\kappa = F(\kappa)).$$

Proof. For the first part it suffices (by Exercise 7.6) to show that $\mathbb{P}_F$ preserves regulars. Fix $\kappa \in \textsc{reg}^M$. Factor $\mathbb{P}_F$ as

$$\mathbb{P}_F \cong \mathbb{P}^{\leqslant\kappa} \times \mathbb{P}^{>\kappa}$$

where

$$\begin{aligned}\mathbb{P}^{\leqslant\kappa} &= \{p \restriction (\kappa+1) \mid p \in \mathbb{P}_F\}\\ \mathbb{P}^{>\kappa} &= \{p \restriction (\mathrm{dom}(F) - (\kappa+1)) \mid p \in \mathbb{P}_F\}\end{aligned}$$

Claim 1. $\mathbb{P}_F$ *preserves cardinals and cofinalities.*

Proof. By Corollary 7.43 it suffices to show that in M

$$\begin{aligned}&\mathbb{P}^{\leqslant\kappa} \text{ is } \kappa^+\text{-c.c. and}\\ &\mathbb{P}^{>\kappa} \text{ is } \kappa \text{ closed.}\end{aligned}$$

For the first, assume, for contradiction, that $\{p_\alpha \mid \alpha < \kappa^+\}$ is an antichain. Note that for all $\alpha < \kappa^+$, $|\mathrm{supp}(p_\alpha)| < \kappa$ (using the Easton support condition in the case where κ is a limit). So there are $\kappa^{<\kappa} = \kappa$ many possibilities for $\mathrm{supp}(p_\alpha)$. We may therefore assume that $\mathrm{supp}(p_\alpha) = S$ for all $\alpha < \kappa^+$. So each p_α has the form

$$\begin{aligned}p_\alpha : S &\to \textstyle\bigcup_{\bar\kappa \in S} \mathrm{Add}(\bar\kappa, F(\bar\kappa))\\ \bar\kappa &\mapsto \mathrm{Add}(\bar\kappa, F(\bar\kappa))\end{aligned}$$

which we can identify with $p_\alpha \in \mathrm{Add}(\kappa, F(\kappa))$. But $\mathrm{Add}(\kappa, F(\kappa))$ has the κ^+-c.c. in M, which is a contradiction.

For the second, note that for $p \in \mathbb{P}^{>\kappa}$ and $\lambda \in \mathrm{dom}(p)$, $p(\lambda) \in \mathrm{Add}(\lambda, F(\lambda))$ where $\mathrm{Add}(\lambda, F(\lambda))$ is κ-closed. □

We clearly have, for $\kappa \in \mathrm{dom}(F)$,

$$(2^\kappa \geqslant F(\kappa))^{M[G]}.$$

As before, the proof of the converse involves counting antichain names.

CLAIM 2. *For all* $\kappa \in \mathrm{dom}(F)$, $(2^\kappa \leqslant F(\kappa))^{M[G]}$.

Proof. Write $G = G^{\leqslant\kappa} \times \mathrm{G}^{>\kappa}$ as in the Product Lemma. Since $\mathbb{P}^{\leqslant\kappa}$ is κ^+-c.c.

$$(2^\kappa)^{M[G^{\leqslant\kappa}]} \leqslant (|\mathbb{P}^{\leqslant\kappa}|^{\kappa\cdot\kappa})^M.$$

Working in M we have $|\mathbb{P}^{\leqslant\kappa}| \leqslant F(\kappa)$ and so

$$(2^\kappa)^{M[G^{\leqslant\kappa}]} \leqslant (|\mathbb{P}^{\leqslant\kappa}|)^{\kappa\cdot\kappa} \leqslant F(\kappa)^\kappa = F(\kappa)$$

by GCH and $\mathrm{cof}(F(\kappa)) > \kappa$. (For the equality we have used the fact that $F(\kappa)^\kappa = \bigcup_{\alpha<F(\kappa)} \alpha^\kappa$ and $|\alpha^\kappa| \leqslant 2^{|\alpha|\cdot\kappa} = (|\alpha|\cdot\kappa)^+ \leqslant F(\kappa)$.) □

This completes the proof of the theorem. □