# FORCING SUMMER SCHOOL
# LECTURE NOTES

SPENCER UNGER

**Abstract.** These are the lecture notes for the forcing class in the 2014 UCLA Logic Summer School. They are a revision of the 2013 Summer School notes. The most substantial differences are a new proof of the forcing theorems in section 10.3 and a final section on the consistency of the failure of the axiom of choice. Both write-ups are by Sherwood Hachtman; the former is based on the author's notes from a course given by Richard Laver.

These notes would not be what they are without a previous set of lectures by Justin Palumbo. The introduction to the forcing language and write-ups of basic results in sections 10.1 and 10.2 are his, and many of the other lectures follow his notes.

### The Axioms of ZFC, Zermelo-Fraenkel Set Theory with Choice

- *Extensionality*: Two sets are equal if and only if they have the same elements.
- *Pairing*: If $a$ and $b$ are sets, then so is the pair $\{a, b\}$.
- *Comprehension Scheme*: For any definable property $\phi(u)$ and set $z$, the collection of $x \in z$ such that $\phi(x)$ holds, is a set.
- *Union*: If $\{A_i\}_{i \in I}$ is a set, then so is its union, $\bigcup_{i \in I} A_i$.
- *Power Set*: If $X$ is a set, then so is $\mathcal{P}(X)$, the collection of subsets of $X$.
- *Infinity*: There is an infinite set.
- *Replacement Scheme*: For any definable property $\phi(u, v)$, if $\phi$ defines a function on a set $a$, then the pointwise image of $a$ by $\phi$ is a set.
- *Foundation*: The membership relation, $\in$, is well-founded; i.e., every nonempty set contains a $\in$-minimal element.
- *Choice*: If $\{A_i\}_{i \in I}$ is a collection of nonempty sets, then there exists a choice function $f$ with domain $I$, so that $f(i) \in A_i$ for all $i \in I$.

Foundation is equivalent to the statement that every set belongs to some $V_\alpha$; Choice is equivalent to the statement that every set can be well-ordered (Zermelo's Theorem). ZFC without the Axiom of Choice is called ZF.

**§1. The Continuum Problem.** The most fundamental notion in set theory is that of well-foundedness.

Definition 1.1. A binary relation $R$ on a set $A$ is **well-founded** if every nonempty subset $B \subseteq A$ has a minimal element, that is, an element $c$ such that for all $b \in B$, $b \mathrel{R} c$ fails.

This document was last modified on September 6, 2014.

Definition 1.2. A linear order $<$ on a set $W$ is a **well-ordering** if it is well-founded.

Remark 1.3. We collect some remarks:

- In a well-order if $c$ is a minimal element, then $c \leq b$ for all $b$. So 'minimal' in this case means 'least'.
- Every finite linear order is a well-order.
- The set of natural numbers is a well-ordered set, but the set of integers is not.
- The Axiom of Choice is equivalent to the statement 'Every set can be well-ordered'.

We will now characterize all well-orderings in terms of ordinals. Here are a few definitions.

Definition 1.4. A set $z$ is **transitive** if for all $y \in z$ and $x \in y$, $x \in z$.

Definition 1.5. A set $\alpha$ is an **ordinal** if it's transitive and well-ordered by $\in$.

Proposition 1.6. *We have the following easy facts:*

1. $\varnothing$ *is an ordinal.*
2. *If* $\alpha$ *is a ordinal, then the least ordinal greater than* $\alpha$ *is* $\alpha \cup \{\alpha\}$. *We call this ordinal* $\alpha + 1$.
3. *If* $\{\alpha_i \mid i \in I\}$ *is a collection of ordinals, then* $\bigcup_{i \in I} \alpha_i$ *is an ordinal.*
4. *We write* On *for the class of ordinals.* On *is well-ordered by* $\in$.

Ordinals from the bottom up:

- $0 = \varnothing$
- $1 = \{\varnothing\}$
- $2 = \{\varnothing, \{\varnothing\}\}$
- $3 = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$

$\vdots$

- $\omega = \{0, 1, 2, 3...\}$
- $\omega + 1 = \{0, 1, 2, 3, ..., \omega\}$

The next proposition captures why ordinals are interesting.

Proposition 1.7. *Every well-ordering is isomorphic to a unique ordinal: For every well-ordering* $(W, <)$ *there are an ordinal* $\alpha$ *and a bijection* $f : W \to \alpha$ *such that* $a < b$ *if and only if* $f(a) < f(b)$.

Before we move on to talking about cardinals we record some terminology about ordinals.

Definition 1.8. Let $\alpha$ be an ordinal.

- $\alpha$ is a **successor ordinal** if $\alpha = \beta + 1$ for some ordinal $\beta$.
- $\alpha$ is a **limit ordinal** if there is an infinite increasing sequence of ordinals $\langle \alpha_i \mid i < \lambda \rangle$ such that $\alpha = \bigcup_{i < \lambda} \alpha_i$.

Cardinals are special ordinals. The Axiom of Choice makes two possible definitions of cardinal equivalent.

Definition 1.9. An ordinal $\alpha$ is a **cardinal** if there is no surjection from an ordinal less than $\alpha$ onto $\alpha$.

Clearly each $n \in \omega$ and $\omega$ itself are cardinals. We define $|A|$ to be the least ordinal $\alpha$ such that there is a bijection from $A$ to $\alpha$. $|A|$ is called the cardinality of $A$. It is not hard to see that $|A|$ is a cardinal. Note that $|A| = |B|$ if and only if there is a bijection from $A$ to $B$. The following theorem makes it easier to prove that two sets have the same cardinality.

Theorem 1.10 (Cantor-Schroder-Bernstein). *If there are an injection from $A$ to $B$ and an injection from $B$ to $A$, then there is a bijection from $A$ to $B$.*

Proof. Without loss of generality we can take $A$ and $B$ to be disjoint, since we can replace $A$ by $\{0\} \times A$ and $B$ by $\{1\} \times B$. We let $f : A \to B$ and $g : B \to A$ be injections. We construct a bijection $h : A \to B$. Let $a \in A$ and define the set

$$S_a = \{\dots f^{-1}(g^{-1}(a)), g^{-1}(a), a, f(a), g(f(a)), \dots\}.$$

Let $b \in B$ and define the set

$$S_b = \{\dots g^{-1}(f^{-1}(b)), f^{-1}(b), b, g(b), f(g(b)) \dots\}.$$

Some note is due on these definitions. At some point we may be unable to take the inverse image. Suppose that $c \in A \cup B$ and $S_c$ stops moving left because we cannot take an inverse image, if the left-most element of $S_c$ is in $A$, then we call it $A$-terminating, otherwise we call it $B$-terminating.

Observe that if $c_1, c_2 \in A \cup B$ and $c_1 \in S_{c_2}$, then $S_{c_1} = S_{c_2}$.

Define $h$ as follows. Let $a \in A$. If $S_a$ is $A$-terminating or does not terminate, then define $h(a) = f(a)$. If $S_a$ is $B$-terminating, then $a$ is in the image of $g$, so define $h(a) = g^{-1}(a)$.

Clearly this defines a map from $A$ to $B$, we just need to check that it is a bijection. First we check that it is onto. Let $b \in B$. If $S_b$ is $A$-terminating or doesn't terminate, then $b$ is in the image of $f$ and $S_{f^{-1}(b)} = S_b$ is $A$-terminating or doesn't terminate, so we defined $h(f^{-1}(b)) = f(f^{-1}(b)) = b$ as required. If $S_b$ is $B$-terminating, then $S_{g(b)}$ is also $B$-terminating, so we defined $h(g(b)) = g^{-1}(g(b)) = b$. It follows that $h$ is onto.

Let $a_1, a_2 \in A$ and suppose that $h(a_1) = h(a_2)$. We will show that $a_1 = a_2$. If $S_{a_1}$ and $S_{a_2}$ are either

1. both $A$-terminating or non-terminating; or
2. both $B$-terminating,

then $a_1 = a_2$ follows from the injectivity of $f$ or $g$.

Suppose for a contradiction that $S_{a_1}$ is $A$-terminating or nonterminating and $S_{a_2}$ is $B$-terminating. Then by the definition of $h$, $f(a_1) = h(a_1) = h(a_2) = g^{-1}(a_2)$. It follows that $S_{a_1} = S_{a_2}$ which is a contradiction. $\dashv$

We are now ready to introduce cardinal arithmetic.

Definition 1.11. Let $\kappa$ and $\lambda$ be cardinals.

- $\kappa + \lambda$ is the cardinality of $\{0\} \times \kappa \cup \{1\} \times \lambda$.
- $\kappa \cdot \lambda$ is the cardinality of $\kappa \times \lambda$.
- $\kappa^\lambda$ is the cardinality of the set ${}^\lambda\kappa = \{f \mid f : \lambda \to \kappa\}$.

If $\kappa$ and $\lambda$ are infinite, then $\kappa + \lambda = \kappa \cdot \lambda = \max \kappa, \lambda$. Exponentiation turns out to be much more interesting. For any cardinal $\kappa$, $|\mathcal{P}(\kappa)| = 2^\kappa$.

THEOREM 1.12 (Cantor). *For any cardinal $\kappa$, $2^\kappa > \kappa$.*

PROOF. Suppose that there is a surjection $H$ from $\kappa$ onto $2^\kappa$. Consider the function $f : \kappa \to 2$ given by $f(\alpha) = 0$ if and only if $H(\alpha)(\alpha) = 1$. (Recall that $H(\alpha)$ is a function from $\kappa$ to 2.)

We claim that $f$ is not in the range of $H$, a contradiction. Let $\alpha < \kappa$, then $f$ is different from $H(\alpha)$, since $f(\alpha) = 0$ if and only if $H(\alpha)(\alpha) = 1$. $\dashv$

By Cantor's theorem we see that for any cardinal $\kappa$ there is a strictly larger cardinal. We write $\kappa^+$ for the least cardinal greater than $\kappa$. Moreover the union of a collection of cardinals is a cardinal. The above facts allow us to use the ordinals to enumerate all of the cardinals.

1. $\aleph_0 = \omega$,
2. $\aleph_{\alpha+1} = \aleph_\alpha^+$ and
3. $\aleph_\gamma = \bigcup_{\alpha<\gamma} \aleph_\alpha$ for $\gamma$ a limit ordinal.

We often write $\omega_\alpha$ in place of $\aleph_\alpha$. They are the same object, but we think of $\omega_\alpha$ in the context of ordinals and $\aleph_\alpha$ in the context of cardinals.

The Continuum Hypothesis (CH) states that $2^{\aleph_0} = \aleph_1$. From Cantor's theorem we know that $2^{\aleph_0} > \aleph_0$. CH is the assertion that the continuum $2^{\aleph_0}$ is the least cardinal greater than $\aleph_0$. The goal of the course is to prove that the axioms of ZFC cannot prove or disprove CH. To do this we will construct a model of ZFC where $2^{\aleph_0} = \aleph_1$ and a different model of ZFC where $2^{\aleph_0} = \aleph_2$. Don't worry if you don't know what this means, it will all be explained by the end of the course.

We now investigate a ZFC restriction on cardinal exponentiation.

DEFINITION 1.13. Let $\alpha$ be a limit ordinal. The **cofinality** of $\alpha$, $\mathrm{cf}(\alpha)$ is the least $\lambda \leq \alpha$ such that there is an increasing sequence $\langle \alpha_i \mid i < \lambda \rangle$ of ordinals less than $\alpha$ such that $\sup_{i<\lambda} \alpha_i = \alpha$.

A sequence $\langle \alpha_i \mid i < \lambda \rangle$ of ordinals in $\alpha$ is said to be **cofinal in** $\alpha$ if $\sup_{i<\lambda} \alpha_i = \alpha$. Thus $\mathrm{cf}(\alpha)$ is the shortest length of an increasing sequence cofinal in $\alpha$.

DEFINITION 1.14. A cardinal $\kappa$ is **regular** if $\mathrm{cf}(\kappa) = \kappa$ and is **singular** otherwise.

PROPOSITION 1.15. $\mathrm{cf}(\alpha)$ *is a regular cardinal.*

THEOREM 1.16. *For any cardinal $\kappa$, $\kappa < \kappa^{\mathrm{cf}(\kappa)}$.*

PROOF. Set $\lambda = \mathrm{cf}(\kappa)$ and suppose that there is a surjection $H$ from $\kappa$ onto $\kappa^\lambda$. Fix an increasing sequence $\langle \alpha_i \mid i < \lambda \rangle$ which is increasing and cofinal in $\kappa$.

We define a function $f$ which is not in the range of $H$, a contradiction. Let $f(i)$ be the least member of $\kappa \setminus \{H(\alpha)(i) \mid \alpha \leq \alpha_i\}$. Let $\alpha < \kappa$ and choose $i < \lambda$ such that $\alpha_i > \alpha$. It follows that $f(i) \neq H(\alpha)(i)$, so $f \neq H(\alpha)$. $\dashv$

Since $(2^\omega)^\omega = 2^{\omega \cdot \omega} = 2^\omega$, it follows that $\mathrm{cf}(2^\omega) > \omega$. In particular $2^\omega \neq \omega_\omega$.

**§2. Cardinal characteristics.** In this section we introduce some combinatorial notions. We start with a few examples of *cardinal characteristics of the continuum*.

Let $f, g \in {}^\omega\omega$. Recall that ${}^\omega\omega$ is the collection of functions from $\omega$ to $\omega$. We define the notion of *eventual domination*, which is a weakening of the pointwise ordering. Let $f <^* g$ if and only if there is an $N < \omega$ such that for all $n \geq N$, $f(n) < g(n)$.

Clearly we can find an upperbound in this ordering for any finite collection of functions $\{f_0, \dots f_k\}$. For $n < \omega$ we define

$$f(n) = \max\{f_0(n), \dots f_k(n)\} + 1.$$

So in fact $f$ is larger than each $f_i$ on every coordinate. What happens if we allow our collection of functions to be countable, say $\{f_i \mid i < \omega\}$. Is it still possible to find a function $f$ such that for all $i$, $f_i <^* f$?

The answer is Yes! To do this we use a diagonal argument. We know that on each coordinate we can only beat finitely many of the $f_i$. So we make sure that after the first $n$ coordinates, we always beat the $n^{th}$ function.

For $n < \omega$, we define

$$f(n) = \max_{i \leq n}(f_i(n)) + 1.$$

It is straightforward to check that this works. Now we ask if it is possible to continue, that is, to increase the size of our collection of functions to $\omega_1$. Given $\{f_\alpha \mid \alpha < \omega_1\}$, can we find a single function $f$ which eventually dominates each $f_\alpha$?

The answer to this question is sensitive to the Set Theory beyond the Axioms of ZFC. For instance, if CH holds, then the answer is no, since all of ${}^\omega\omega$ can be enumerated in $\omega_1$ steps. However, we will see that it is possible that the answer is yes if we assume Martin's Axiom.

We give a definition that captures the essence of this question.

DEFINITION 2.1. Let $\mathfrak{b}$ be the least cardinal such that there exists a family of functions $\mathcal{F}$ with $|\mathcal{F}| = \mathfrak{b}$ such that no $f : \omega \to \omega$ eventually dominates all members of $\mathcal{F}$. Such a family is called an **unbounded family**.

$\mathfrak{b}$ is a cardinal characteristic of the continuum. We can phrase our observations as a theorem about $\mathfrak{b}$.

THEOREM 2.2. $\omega < \mathfrak{b} \leq 2^{\aleph_0}$.

Our question about families of size $\omega_1$ can now be rephrased as 'Is $\mathfrak{b} > \omega_1$?'.

We now introduce another cardinal characteristic $\mathfrak{a}$.

DEFINITION 2.3. Let $A, B$ be subsets of $\omega$. We say that $A$ and $B$ are **almost disjoint** if $A \cap B$ is finite. A family $\mathcal{F}$ of infinite pairwise almost disjoint subsets of $\omega$ is **maximally almost disjoint** (MAD) if for any infinite subset $B$ of $\omega$, there is an $A \in \mathcal{F}$ such that $A \cap B$ is infinite.

An easy example of a MAD family is to take $\mathcal{F} = \{A, B\}$ where $A$ is the set of odd natural numbers and $B$ is the set of even natural numbers. In fact any partition of $\omega$ into finitely many pieces is a MAD family.

The following proposition is left as an exercise.

Proposition 2.4. *There is a MAD family of size* $2^{\aleph_0}$.

The following lemma is a part of what makes MAD families interesting.

Lemma 2.5. *There are no countably infinite MAD families.*

Proof. Let $\mathcal{F} = \{A_n \mid n < \omega\}$ be a countable family of pairwise almost disjoint sets. We will construct $B = \{b_n \mid n < \omega\}$ a subset of $\omega$ enumerated in increasing order. We ensure that $b_{n+1}$ does not belong to any of $A_0, \ldots A_n$. This ensures that $B \cap A_n$ is bounded by $b_{n+1}$ (hence it is finite). To do this let $b_0$ be any member of $A_0$ and assuming that we have defined $b_n$ for some $n$, let $b_{n+1}$ be the least member of $A_{n+1} \setminus (A_0 \cup \cdots \cup A_n)$ greater than $b_n$. This is possible since the set in question is infinite by the almost disjointness of $\mathcal{F}$. ⊣

The definition of $\mathfrak{a}$ captures our questions about the possible sizes of MAD families.

Definition 2.6. Let $\mathfrak{a}$ be the least infinite cardinal such that there is a MAD family of size $\mathfrak{a}$.

So we have proved:

Theorem 2.7. $\omega < \mathfrak{a} \leq 2^{\aleph_0}$

It turns out that $\mathfrak{a}$ and $\mathfrak{b}$ are related.

Theorem 2.8 (Solomon, 1977). $\mathfrak{b} \leq \mathfrak{a}$.

Proof. It is enough to show that any almost disjoint family of size less than $\mathfrak{b}$ is not maximal. Let $\mathcal{F} = \{A_\alpha \mid \alpha < \kappa\}$ where $\kappa < \mathfrak{b}$ be an almost disjoint family. We may assume that the collection $\{A_n \mid n < \omega\}$ are pairwise disjoint.

We seek to define a useful collection of functions from $\omega$ to $\omega$. Let $\omega \leq \alpha < \kappa$ and for $n < \omega$ define $f_\alpha(n)$ to be the least $m$ such that the $m^{th}$ member of $A_n$ is larger than all elements in $A_n \cap A_\alpha$. This defines $\{f_\alpha \mid \omega \leq \alpha < \kappa\}$ and since $\kappa < \mathfrak{b}$ there is a function $f$ which eventually dominates each $f_\alpha$.

Now we define $b_n$ to be the $f(n)^{th}$ member of $A_n$. Clearly $B = \{b_n \mid n < \omega\}$ is infinite and almost disjoint from each $A_n$, since it contains exactly one member from each $A_n$.

It remains to show that $B$ is almost disjoint from each $A_\alpha$ for $\omega \leq \alpha < \kappa$. Fix $\alpha$ and let $N$ be such that for all $n \geq N$, $f(n) > f_\alpha(n)$. For each $n \geq N$ we have that the $f(n)^{th}$ member of $A_n$ is greater than all members of $A_n \cap A_\alpha$, since $f(n) > f_\alpha(n)$. In particular $b_n$, which is the $f(n)^{th}$ member of $A_n$ is not in $A_\alpha$ for all $n \geq N$. So $B$ works. ⊣

**§3. Martin's Axiom.** We need some definitions in order to formulate Martin's Axiom.

DEFINITION 3.1. A **partially ordered set** (**poset**) is a pair $(\mathbb{P}, \leq)$ where $\leq$ is a binary relation on $\mathbb{P}$ such that $\leq$ is

1. reflexive; for all $p \in \mathbb{P}$, $p \leq p$,
2. transitive; for all $p, q, r \in \mathbb{P}$, if $p \leq q$ and $q \leq r$, then $p \leq r$, and
3. antisymmetric; for all $p, q \in \mathbb{P}$, if $p \leq q$ and $q \leq p$, then $p = q$.

We also require that our posets have a unique maximal element $\mathbb{1}_{\mathbb{P}}$, i.e. for all $p \in \mathbb{P}$, $p \leq \mathbb{1}_{\mathbb{P}}$.

For simplicity, we will always refer to 'the poset $\mathbb{P}$' instead of the poset $(\mathbb{P}, \leq)$. Elements of $\mathbb{P}$ are often called **conditions** and when $p \leq q$ we say that $p$ is an extension (or strengthening) of $q$. Posets are everywhere and we will see many examples throughout the course.

As a running example we will consider the set $\mathbb{P} = \{p \mid p : n \to 2\}$ ordered by $p_1 \leq p_2$ if and only if $p_1 \supseteq p_2$. It is not hard to check that this is a poset.

DEFINITION 3.2. Let $\mathbb{P}$ be a poset and $p, q \in \mathbb{P}$.

1. $p$ and $q$ are **comparable** if $p \leq q$ or $q \leq p$.
2. $p$ and $q$ are **compatible** if there is an $r \in \mathbb{P}$ such that $r \leq p, q$.

*Incomparable* and *incompatible* mean 'not comparable' and 'not compatible' respectively.

DEFINITION 3.3. Let $\mathbb{P}$ be a poset and $A \subseteq \mathbb{P}$. $A$ is an **antichain** if any two elements of $A$ are incompatible.

Note that for a fixed $n < \omega$ the collection $\{p \mid \mathrm{dom}(p) = n\}$ is an antichain in our example poset.

DEFINITION 3.4. Let $\mathbb{P}$ be a poset. $\mathbb{P}$ has the **countable chain condition** (is ccc) if every antichain of $\mathbb{P}$ is countable.

Our example poset is ccc for trivial reasons; the whole poset is countable.

DEFINITION 3.5. Let $\mathbb{P}$ be a poset. A subset $D \subseteq \mathbb{P}$ is **dense** if for all $p \in \mathbb{P}$ there is $q \in D$ such that $q \leq p$.

In our running example both of the following sets are dense for any $n < \omega$: $\{p \in \mathbb{P} \mid \mathrm{dom}(p) > n\}$ and $\{p \in \mathbb{P} \mid \mathrm{dom}(p) \text{ is even}\}$. How are these different?

DEFINITION 3.6. Let $\mathbb{P}$ be a poset. A subset $D \subseteq \mathbb{P}$ is **open** if for all $p \in D$ and for all $q \leq p$, $q \in D$.

The first of the two sets above is open and the second is not.

DEFINITION 3.7. A subset $G \subseteq \mathbb{P}$ is a **filter** if

1. for all $p \in G$ and $q \geq p$, $q \in G$, and
2. for all $p, q \in G$ there is $r \in G$ with $r \leq p, q$.

If $\mathcal{D}$ is a collection of dense subsets of $\mathbb{P}$, then we say that $G$ is $\mathcal{D}$-generic if for every $D \in \mathcal{D}$, $D \cap G \neq \varnothing$.

We are now ready to formulate Martin's Axiom.

Definition 3.8. MA($\kappa$) is the assertion that for every ccc poset $\mathbb{P}$ and collection of $\kappa$-many dense sets $\mathcal{D}$, there is a $\mathcal{D}$-generic filter over $\mathbb{P}$.

MA is the assertion that MA($\kappa$) holds for all $\kappa < 2^{\aleph_0}$. Roughly speaking MA asserts that if an object has a reasonable collection of approximations, then it exists.

Proposition 3.9. MA($\omega$) *holds even if we drop the ccc requirement.*

Proof. Let $\mathcal{D} = \{D_n \mid n < \omega\}$ be a collection of dense subsets of a poset $\mathbb{P}$. We construct a decreasing sequence $\langle p_n \mid n < \omega\rangle$ such that $p_n \in D_n$ for all $n$. Let $p_0 \in D_0$. Suppose we have constructed $p_n$ for some $n < \omega$. We choose $p_{n+1} \in D_{n+1}$ with $p_{n+1} \leq p_n$ by density.

We define $G = \{p \in \mathbb{P} \mid p \geq p_n \text{ for some } n < \omega\}$. It is not hard to see that $G$ is a $\mathcal{D}$-generic filter over $\mathbb{P}$. $\dashv$

Proposition 3.10. *If* MA($\kappa$) *holds, then* $\kappa < 2^{\aleph_0}$. *In particular* MA($2^{\aleph_0}$) *fails.*

Proof. Suppose that MA($\kappa$) holds. It is enough to show that given a collection $\{f_\alpha \mid \alpha < \kappa\}$ of functions from $\omega$ to 2, there is a function $g$ which is not equal to any $f_\alpha$.

Let $\mathbb{P}$ be as in our running example. We claim that for each $\alpha < \kappa$, the set $E_\alpha = \{p \mid \text{for some } n \in \text{dom}(p)\ f_\alpha(n) \neq p(n)\}$ is dense. Given a $p \in \mathbb{P}$ choose an $n \in \omega \setminus \text{dom}(p)$ and consider the condition $p \cup \{\langle n, f_\alpha(n) +_2 1\rangle\}$, which is in $E_\alpha$.

We also need $D_n = \{p \mid \text{dom}(p) > n\}$ which is dense as we discussed. We let $\mathcal{D} = \{D_n \mid n < \omega\} \cup \{E_\alpha \mid \alpha < \kappa\}$ and apply MA($\kappa$) to obtain $G$.

We claim that $g = \bigcup G$ is a function. For if $\langle n, y\rangle$ and $\langle n, y'\rangle$ are both in $g$, then there are $p, p' \in G$ so that $p(n) = y$ and $p'(n) = y'$. Since $G$ is a filter, there is some $q \leq p, p'$; so $y = q(n) = y'$, as needed. Since $G \cap D_n \neq \varnothing$ for each $n < \omega$, we have that $g$ has domain $\omega$; and since $G \cap E_\alpha \neq \varnothing$ for each $\alpha < \kappa$, we have $g \neq f_\alpha$. $\dashv$

Proposition 3.11. MA($\aleph_1$) *fails if we remove the ccc requirement.*

Proof. Let $\mathbb{P} = \{p \mid p : n \to \omega_1 \text{ for some } n < \omega\}$ ordered by $p_1 \leq p_2$ if and only if $p_1 \supseteq p_2$. We define $E_\alpha = \{p \mid \alpha \in \text{ran}(p)\}$ and $D_n = \{p \mid n \in \text{dom}(p)\}$. It is not hard to see that these sets are dense.

Let $G$ be generic for all of our dense sets. We have arranged that $g = \bigcup G$ is a surjection from $\omega$ onto $\omega_1$. Such a function cannot exist. $\dashv$

The way we have formulated MA, CH implies that MA holds for trivial reasons. It is consistent with ZFC that MA holds with the continuum large, but this result is beyond the scope of the course. The reason that we introduce MA is that its statement and applications allow us to get acquainted with core machinery of forcing (posets, filters, etc.) without being burdened by the metamathematical complications of forcing proper (which we will tackle separately soon enough).

**§4. Applications of MA to cardinal characteristics.** We continue our applications of MA by showing how MA influences cardinal characteristics of the continuum. We can view these applications as extensions of the diagonalization arguments we used to show that $\mathfrak{b}$ and $\mathfrak{a}$ are uncountable.

We will prove the following theorem.

THEOREM 4.1. MA *implies* $\mathfrak{b} = 2^{\aleph_0}$.

Using Solomon's Theorem we have,

COROLLARY 4.2. MA *implies* $\mathfrak{a} = 2^{\aleph_0}$.

Given a collection of functions of size less than continuum we need to build a ccc poset which approximates a function $f$ which dominates all of the functions in our collection. In order to satisfy the ccc requirement our approximations will be finite.

PROOF. We define a poset $\mathbb{P}$ to be the collection of pairs $(p, A)$ where $p \in {}^{<\omega}\omega$ and $A$ is a finite subset of ${}^{\omega}\omega$. For the ordering we set $(p, A) \leq (q, B)$ if and only if $p \supseteq q$, $A \supseteq B$ and for all $f \in B$ and all $n \in \mathrm{dom}(p) \setminus \mathrm{dom}(q)$, $p(n) > f(n)$. (You should check that $\leq$ is transitive.) The $p$-part of the condition is growing the function from $\omega$ to $\omega$ and the $A$-part is a collection of functions which we promise to dominate when we extend the $p$-part. The poset $\mathbb{P}$ is called the **dominating poset**.

We claim that $\mathbb{P}$ is ccc. It is enough to show that every set of conditions of size $\omega_1$ contains two pairwise compatible conditions. Let $\langle (p_\alpha, A_\alpha) \mid \alpha < \omega_1 \rangle$ be a sequence of conditions in $\mathbb{P}$. By the pigeonhole principle there is an unbounded set $I \subseteq \omega_1$ such that for all $\alpha, \beta \in I$, $p_\alpha = p_\beta$.

Let $\alpha, \beta \in I$ and define $p = p_\alpha = p_\beta$. We claim that $(p, A_\alpha \cup A_\beta)$ is a lower bound for both $(p_\alpha, A_\alpha)$ and $(p_\beta, A_\beta)$. This is clear, since the third condition for extension is vacuous. So we have actually shown that that given a sequence of $\omega_1$-many conditions in $\mathbb{P}$ there is a subsequence of $\omega_1$-many conditions which are pairwise compatible. This property is called the **$\omega_1$-Knaster property**.

We will apply MA to this poset. Let $\mathcal{F} = \{f_\alpha \mid \alpha < \kappa\}$ be a collection of functions from $\omega$ to $\omega$ where $\kappa$ is some cardinal less than $2^{\aleph_0}$. Now we need a collection of dense sets to which we will apply MA. First, we have for each $n < \omega$, the collection $D_n = \{(p, A) \mid n \in \mathrm{dom}(p)\}$. Given a condition $(p, A)$ we can just extend the $p$ to have $n$ in the domain ensuring that we choose a value larger than the maximum of the finitely many functions in $A$ on each coordinate we add. Call the extension $q$. It is clear that $(q, A) \leq (p, A)$ and $(q, A) \in D_n$. So $D_n$ is dense.

For each $\alpha < \kappa$ we define $E_\alpha = \{(p, A) \mid f_\alpha \in A\}$. Clearly this is dense, since given a condition $(q, B)$, $(q, B \cup \{f_\alpha\}) \leq (q, B)$ and is a member of $E_\alpha$.

From here the proof is easy. By MA we can choose $G$ a $\mathcal{D}$-generic filter where $\mathcal{D} = \{D_n \mid n < \omega\} \cup \{E_\alpha \mid \alpha < \kappa\}$. Let $f = \bigcup \{p \mid (p, A) \in G \text{ for some } A\}$. By the usual argument, $f \in {}^{\omega}\omega$. To see that $f$ eventually dominates each $f_\alpha$, let $\alpha < \kappa$ and choose a condition $(p, A) \in G \cap E_\alpha$. Let $N = \mathrm{dom}(p)$. We claim that for all $n \geq N$, $f(n) > f_\alpha(n)$. Fix such an $n$ and choose a condition $(q, B) \in G \cap D_n$ with $(q, B) \leq (p, A)$. By the definition of extension $q(n) > f_\alpha(n)$, but $q(n) = f(n)$ so we are done. $\dashv$

We sketch another very similar application of MA and leave some of the details as exercises.

THEOREM 4.3. *Assume* MA($\kappa$) *and let* $\mathcal{A}$ *and* $\mathcal{C}$ *be collections of size* $\leq \kappa$ *of subsets of* $\omega$ *such that for every* $y \in \mathcal{C}$ *and every finite* $F \subseteq \mathcal{A}$ *the set* $y \setminus \bigcup F$ *is infinite. There is a single subset* $Z \subseteq \omega$ *such that* $X \cap Z$ *is finite for all* $X \in \mathcal{A}$ *and* $Y \cap Z$ *is infinite for* $Y \in \mathcal{C}$.

The proof is very similar to the previous so we will define the poset and leave the rest as an exercise. Let $\mathbb{P}$ be the collection of pairs $(s, F)$ where $s \in [\omega]^{<\omega}$ and $F \subseteq \mathcal{A}$ is finite. Let $(s_0, F_0) \leq (s_1, F_1)$ if and only if $s_0 \supseteq s_1$, $F_0 \supseteq F_1$ and for all $n \in s_0 \setminus s_1$, $n \notin \bigcup F_1$.

Most of the proof is as before. Here is a helpful hint: Show that for each $n < \omega$ and $Y \in \mathcal{C}$, the set $E^n_Y = \{(s, F) \mid \text{there is } m \geq n \text{ such that } m \in s \cap Y\}$ is dense.

COROLLARY 4.4. MA *implies* $\mathfrak{a} = 2^{\aleph_0}$

Apply the previous theorem with $\mathcal{C} = \{\omega\}$.

COROLLARY 4.5. *Suppose that* MA($\kappa$) *holds. If* $\mathcal{B}$ *is an almost disjoint family of size* $\kappa$ *and* $\mathcal{A} \subseteq \mathcal{B}$, *then there is a* $Z$ *which has infinite intersection with each member of* $\mathcal{B} \setminus \mathcal{A}$ *and finite intersection with each member of* $\mathcal{A}$.

Just apply the theorem with $\mathcal{A}$ as itself and $\mathcal{C} = \mathcal{B} \setminus \mathcal{A}$. Note that the set $Z$ codes the set $\mathcal{A}$ in that if we are given $Z$ we can define $\mathcal{A} = \{A \in \mathcal{B} \mid Z \cap A \text{ is finite}\}$. This gives us the following fact.

THEOREM 4.6. MA *implies for all infinite* $\kappa < 2^{\aleph_0}$, $2^\kappa = 2^{\aleph_0}$.

PROOF. Let $\mathcal{B}$ be an almost disjoint family of size $\kappa$. It is enough to show that $|\mathcal{P}(\mathcal{B})| = 2^{\aleph_0}$.

Define $\Gamma : \mathcal{P}(\omega) \to \mathcal{P}(\mathcal{B})$ by $\Gamma(Z) = \{A \in \mathcal{B} \mid A \cap Z \text{ is finite}\}$. $\Gamma$ is surjective by the previous corollary. ⊣

COROLLARY 4.7. MA *implies* $2^{\aleph_0}$ *is regular.*

PROOF. Suppose $\mathrm{cf}(2^{\aleph_0}) = \kappa < 2^{\aleph_0}$. Then we have

$$(2^{\aleph_0})^\kappa = (2^\kappa)^\kappa = 2^\kappa = 2^{\aleph_0}$$

which violates König's Lemma, a contradiction. ⊣

**§5. Applications of MA to Lebesgue measure.** Another application of MA is to Lebesgue measure. To begin we recall some facts about Lebesgue measure. Lebesgue measure assigns a size to certain sets of real numbers. We begin by trying to extend the notion of the length of an interval. We first define a notion of **outer measure** on all sets of real numbers. Given $A \subseteq \mathbb{R}$ we define

$$\mu^*(A) = \inf \left\{ \sum_{n<\omega} (b_n - a_n) \,\middle|\, A \subseteq \bigcup_{n<\omega} (a_n, b_n) \right\}.$$

We list some properties of this outer measure. These properties will be true of the full Lebesgue measure as well.

PROPOSITION 5.1. *$\mu^*$ has the following properties:*

1. $\mu^*(\varnothing) = 0$.
2. *For all $E \subseteq F$, $\mu^*(E) \subseteq \mu^*(F)$.*
3. *For all $\{E_n | n < \omega\}$, $\mu^*(\bigcup_{n<\omega} E_n) \leq \sum_{n<\omega} \mu^*(E_n)$.*

PROOF. The first item is clear. For the second, notice that any open cover of $F$ is also an open cover of $E$. The main point is the third item. Let $\epsilon > 0$. By the definition of $\mu^*$ for each $n$ we can choose an open set $U_n$ such that $\mu^*(U_n) \leq \mu^*(E_n) + \epsilon \cdot 2^{-n-1}$.

Note that $\bigcup_{n<\omega} U_n$ is an open set covering $E = \bigcup_{n<\omega} E_n$. So we have

$$\mu^*(E) \leq \sum_{n<\omega} \mu^*(U_n) \leq \sum_{n<\omega} \left(\mu^*(E_n) + \epsilon \cdot 2^{-n-1}\right) = \sum_{n<\omega} \mu^*(E_n) + \epsilon.$$

Since $\epsilon$ was arbitrary we have the result. ⊣

It is not hard to see that $\mu^*$ returns the length of an interval, that is $\mu^*(a, b) = b - a$. Further, recall that an open subset of the real line $U$ can be written uniquely as the union of countably many disjoint open intervals. (To do this let $I_x$ be the union of all open intervals contained in $U$ with $x$ as a member. If $I_x \neq I_y$, then $I_x \cap I_y = \varnothing$. So $\bigcup_{x \in U} I_x = U$ is a disjoint union of open intervals and hence there can only be countably many intervals involved.) So if we write $U = \bigcup_{n<\omega} (a_n, b_n)$ where the intervals are pairwise disjoint, then it is clear that we have $\mu^*(U) = \sum_{n<\omega} (b_n - a_n)$.

It turns out that the outer measure $\mu^*$ is poorly behaved on arbitrary sets. To define the full Lebesgue measure we want to restrict ourselves to certain nice sets. For this, we introduce the **Borel sets**. The collection of Borel sets $\mathcal{B}$ is the smallest set which contains the open sets and is closed under countable unions and complements. (A set closed under countable unions and complements is called a $\sigma$-algebra.)

As an aside, note the Borel sets can obtained by defining

$$\mathcal{B} = \bigcap \{\mathcal{A} \mid \mathcal{A} \text{ is a } \sigma\text{-algebra containing the open sets}\}.$$

But we can also give a more concrete description of these sets as follows. Define $\mathbf{\Sigma}^0_1$ the be the collection of open sets. Then put, for $\alpha < \omega_1$,

$$\mathbf{\Pi}^0_\alpha = \{X \mid X \text{ is a complement of some } Y \in \mathbf{\Sigma}^0_\alpha\},$$
$$\mathbf{\Sigma}^0_\alpha = \{X \mid X \text{ is a countable union of sets in } \textstyle\bigcup_{\xi<\alpha} \mathbf{\Pi}^0_\alpha\}.$$

It can then be shown that $\mathcal{B} = \bigcup_{\alpha<\omega_1} \mathbf{\Sigma}^0_\alpha$.

We are now ready to define what it means to be Lebesgue measurable.

Definition 5.2. A set $A \subseteq \mathbb{R}$ is **Lebesgue measurable** if there is a Borel set $B$ such that $\mu^*(A \triangle B) = 0$. In this case the **Lebesgue measure of** $A$ is $\mu(A) = \mu^*(A)$. We call the collection of Lebesgue measurable sets $\mathcal{L}$.

We catalog some properties of Lebesgue measure.

Proposition 5.3. *$\mathcal{L}$ is the smallest $\sigma$-algebra containing the Borel sets and the sets of outer measure zero.*

Proposition 5.4. *$\mathcal{B} \neq \mathcal{L}$.*

Theorem 5.5. *$\mathcal{L}$ and $\mu$ have the following properties:*

1. *(Monotonicity) If $A, B \in \mathcal{L}$ and $A \subseteq B$, then $\mu(A) \leq \mu(B)$.*
2. *(Translation invariance) If $A \in \mathcal{L}$ and $t \in \mathbb{R}$, then $t + A = \{t + x \mid x \in A\} \in \mathcal{L}$ and $\mu(A) = \mu(t + A)$.*
3. *(Countable additivity) If $\{A_n \mid n < \omega\} \subseteq \mathcal{L}$ is a collection of pairwise disjoint sets, then $\mu(\bigcup_{n<\omega} A_n) = \sum_{n<\omega} \mu(A_n)$.*

Theorem 5.6 (AC). *There is $A \subseteq \mathbb{R}$ with $A \notin \mathcal{L}$.*

Proof. Define an equivalence relation on $\mathbb{R}$ by $x \sim y$ if and only if $|x - y|$ is rational. Note each equivalence class is countable. Let $F$ be a choice function for the equivalence classes; we can further assume $F([x]_\sim) \in [0, 1]$ for each $x$.

Let $A$ be the range of $F$. We claim $A$ is not Lebesgue measurable. Notice $\mathbb{R} = \bigcup_{q\in\mathbb{Q}}(A + q)$, so by countable additivity, we have $0 < \mu(A)$. But then $\mu(\bigcup_{n\in\omega} A + \frac{1}{n+1})$ is a subset of $[0, 2]$ with infinite measure, a contradiction. $\dashv$

Our application of MA will be to sets of measure zero and will generalize the following fact which is an easy consequence of countable sub-additivity.

Proposition 5.7. *The union of countably many measure zero sets has measure zero.*

For ease of notation we let $\mathcal{C}$ be the collection of finite unions of open intervals with rational endpoints. Note that $\mathcal{C}$ is countable. We will show that open sets can be approximated closely in measure by members of $\mathcal{C}$.

Proposition 5.8. *Let $U$ be an open set with $0 < \mu(U) < \infty$. For every $\epsilon > 0$ there is a member $Y \in \mathcal{C}$ such that $Y \subseteq U$ and $\mu(U \setminus Y) < \epsilon$.*

Proof. Let $\epsilon > 0$ and assume that $\mu(U)$ is some positive real number $m$. Write $U = \bigcup_{n<\omega}(a_n, b_n)$ where the collection $\{(a_n, b_n) \mid n < \omega\}$ is pairwise disjoint. We choose $N < \omega$ such that $\sum_{n\geq N}(b_n - a_n) < \frac{\epsilon}{2}$. For each $n < N$ we choose rational numbers $q_n, r_n$ such that $a_n < q_n < r_n < b_n$ and

$$\mu((a_n, b_n) \setminus (q_n, r_n)) = |b_n - r_n| + |q_n - a_n| < \frac{\epsilon}{2} \cdot 2^{-n-1}$$

We set $Y = \bigcup_{n<N}(q_n, r_n) \in \mathcal{C}$. An easy calculation shows that this works. $\dashv$

We are ready for our application of MA to Lebesgue measure.

Theorem 5.9. MA($\kappa$) *implies the union of $\kappa$-many measure zero sets is measure zero.*

Proof. Let $\epsilon > 0$. Define a poset $\mathbb{P}$ to be the collection of open $p \in \mathcal{L}$ such that $\mu(p) < \epsilon$ and set $p_0 \leq p_1$ if and only if $p_0 \supseteq p_1$. As usual we need to show that $\mathbb{P}$ is ccc.

Towards showing that $\mathbb{P}$ is ccc, we let $\{p_\alpha \mid \alpha < \omega_1\}$ be a collection of conditions from $\mathbb{P}$. For each $\alpha$ we know that $\mu(p_\alpha) < \epsilon$, so there is an $n_\alpha < \omega$ such that $\mu(p_\alpha) < \epsilon - \frac{1}{n_\alpha}$. By the pigeonhole principal we may assume that there is an $n$ such that $n = n_\alpha$ for all $\alpha < \omega_1$.

Now for each $\alpha$ we choose $Y_\alpha \in \mathcal{C}$ such that $Y_\alpha \subseteq p_\alpha$ and $\mu(p_\alpha \setminus Y_\alpha) < \frac{1}{2n}$. Since $\mathcal{C}$ is countable we may assume that there is a $Y \in \mathcal{C}$ such that $Y = Y_\alpha$ for all $\alpha < \omega_1$. Now let $\alpha < \beta < \omega_1$, we have

$$\mu(p_\alpha \cup p_\beta) \leq \mu(p_\alpha \setminus Y) + \mu(p_\beta \setminus Y) + \mu(Y) < \frac{1}{2n} + \frac{1}{2n} + \epsilon - \frac{1}{n} = \epsilon.$$

So $p_\alpha$ and $p_\beta$ are compatible.

We use this poset to prove the theorem. Let $\{A_\alpha \mid \alpha < \kappa\}$ be a collection of measure zero sets. We want to show that the measure of the union is zero. Let $\epsilon > 0$ and $\mathbb{P}$ be defined as above. We claim that $E_\alpha = \{p \in \mathbb{P} \mid A_\alpha \subseteq p\}$ is dense for each $\alpha < \kappa$. Let $q \in \mathbb{P}$. Since $\mu(A_\alpha) = 0$ we can find an open set $r$ such that $A_\alpha \subseteq r$ and $\mu(r) < \epsilon - \mu(q)$. Clearly $p = q \cup r \in E_\alpha$. So $E_\alpha$ is dense.

Now we apply MA to $\mathbb{P}$ and the collection of $\{E_\alpha \mid \alpha < \kappa\}$ to obtain $G$. We claim that $U = \bigcup G$ is an open set containing the union of the $A_\alpha$ and $\mu(U) \leq \epsilon$. Clearly $U$ is open since it is the union of open sets. Clearly it contains the union of the $A_\alpha$, since $G$ meets each $E_\alpha$. It remains to show that $\mu(U) \leq \epsilon$.

We claim that if $\{p_n \mid n < \omega\}$ is a subset of $G$, then $\mu(\bigcup_{n<\omega} p_n) \leq \epsilon$. Note that since each $p_n \in G$, $p_0 \cup \cdots \cup p_n \in G$. Hence $\mu(p_0 \cup \cdots \cup p_n) < \epsilon$. If we define $q_n = p_n \setminus (p_0 \cup \cdots \cup p_{n-1})$, then we have $\mu(q_0 \cup \cdots \cup q_n) = \mu(p_0 \cup \cdots \cup p_n) < \epsilon$. So we have

$$\mu\left(\bigcup_{n<\omega} p_n\right) = \mu\left(\bigcup_{n<\omega} q_n\right) = \sum_{n<\omega} \mu(q_n) \leq \epsilon$$

since each partial sum is less than $\epsilon$. This finishes the claim.

To finish the proof it is enough to show that there is a countable subset $B \subseteq G$ such that $\bigcup B = U$. Suppose that $x \in U$. Then $x \in p$ for some $p \in G$. So we can find $q_x \in \mathcal{C}$ such that $x \in q_x \subseteq p$. Since $G$ is a filter $q_x \in G$. So $G = \bigcup_{x \in U} q_x$. But $\mathcal{C}$ is countable so $B = \{q_x \mid x \in U\}$ is as required. $\dashv$

**§6. Applications of MA to ultrafilters.** Ultrafilters are an important concept in modern set theory. We introduce ultrafilters in some generality and then give an application of MA to ultrafilters on $\omega$.

DEFINITION 6.1. Let $X$ be a set. A collection $\mathcal{F} \subseteq \mathcal{P}(X)$ is a **filter on** $X$ if all of the following properties hold:

1. $X \in \mathcal{F}$ and $\varnothing \notin \mathcal{F}$.
2. If $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$
3. If $A \subseteq B$ and $A \in \mathcal{F}$, then $B \in \mathcal{F}$.

DEFINITION 6.2. A filter $\mathcal{F}$ on $X$ is **principal** if there is a set $X_0 \subseteq X$ such that $\mathcal{F} = \{A \subseteq X \mid X_0 \subseteq A\}$. Otherwise $\mathcal{F}$ is **nonprincipal**.

As an example we can always define a filter on a cardinal $\kappa$ by setting $\mathcal{F} = \{A \subseteq \kappa \mid \kappa \setminus A \text{ is bounded in } \kappa\}$. One way to think about a filter is to think of members of the filter as 'large'. If $\kappa = \omega$, then the filter that we just defined is called the **Fréchet filter**.

We want to know when a collection of sets can be extended to a filter. The following definition gives a sufficient condition.

DEFINITION 6.3. A collection of sets $\mathcal{A} \subseteq \mathcal{P}(X)$ has the **finite intersection property** if for all $A_0, \dots A_n$ from $\mathcal{A}$, $\bigcap_{i \leq n} A_i$ is nonempty.

PROPOSITION 6.4. *If $\mathcal{A} \subseteq \mathcal{P}(X)$ has the finite intersection property then there is a filter on $X$ containing $\mathcal{A}$.*

PROOF. Suppose $\mathcal{A}$ has the finite intersection property; define $\mathcal{F} = \{B \subseteq X \mid B \supseteq A_0 \cap \dots \cap A_n \text{ for some } A_1, \dots A_n \in \mathcal{A}\}$. It's easy to check $\mathcal{F}$ is a filter. ⊣

The filter in the above proof is called the filter **generated by** $\mathcal{A}$.

DEFINITION 6.5. A filter $\mathcal{F}$ on $X$ is an **ultrafilter** if for every $A \subseteq X$ either $A \in \mathcal{F}$ or $X \setminus A \in \mathcal{F}$. A filter $\mathcal{F}$ on $X$ is **maximal** if there is no filter $\mathcal{F}'$ on $X$ which properly contains $\mathcal{F}$.

PROPOSITION 6.6. *A filter is maximal if and only if it is an ultrafilter.*

PROOF. If $\mathcal{F}$ is an ultrafilter, then any set $A \notin \mathcal{F}$ must have $X \setminus A \in \mathcal{F}$. It follows that any proper extension of $\mathcal{F}$ cannot have the finite intersection property, so cannot be a filter. This shows $\mathcal{F}$ is maximal.

Conversely, if $\mathcal{F}$ is maximal, let $A \subseteq X$; one of $\mathcal{F} \cup \{A\}$ and $\mathcal{F} \cup \{X \setminus A\}$ has the finite intersection property. (Why?) One of these extensions can be further extended to a filter. Since $\mathcal{F}$ is maximal, we must have either $A$ or its complement in $\mathcal{F}$. ⊣

We have an easy example of an ultrafilter: for any $x \in X$, let $\mathcal{F}$ be the principal filter generated by $x$. Indeed, all principal ultrafilters are of this form. The following gives a more interesting class of ultrafilters.

PROPOSITION 6.7. *Every filter can be extended to an ultrafilter.*

PROOF. Given a filter $\mathcal{F}$, consider the partial order of filters containing $\mathcal{F}$ ordered by inclusion. It's easy to check that the union of any chain of filters is a filter, so by Zorn's lemma, there is a maximal such filter, and by the previous proposition, this is an ultrafilter. ⊣

We saw the Fréchet filter was $\mathcal{F} = \{A \subseteq \omega \mid \omega \setminus A \text{ is finite}\}$. Now $\mathcal{F}$ can be extended to an ultrafilter $\mathcal{U}$. $\mathcal{U}$ is nonprincipal since it contains the complement of every singleton.

Our application of MA to ultrafilters will be to construct a special kind of ultrafilter called a Ramsey ultrafilter. To motivate the definition we recall the following theorem.

THEOREM 6.8 (Ramsey). *For every $\chi : [\omega]^2 \to 2$, there is an infinite set $B$ such that $\chi$ is constant on $[B]^2$.*

PROOF. We construct three sequences $A_i, \epsilon_i, a_i$ such that $A_{i+1} \subseteq A_i$, $a_i < a_{i+1}$ and $\epsilon_i \in 2$ for all $i < \omega$. Let $a_0 = 0$ and $A_0 = \omega$. For the induction step, suppose that we have defined $A_n, a_n$ for some $n < \omega$. We choose $\epsilon_n \in 2$ such that $A_{n+1} = \{k \in A_n \setminus (a_n + 1) \mid \chi(a_n, k) = \epsilon_n\}$ is infinite and let $a_{n+1}$ be the least member of $A_{n+1}$. This completes the construction.

Let $I \subseteq \omega$ be infinite and $\epsilon \in 2$ such that for all $i \in I$, $\epsilon_i = \epsilon$. We set $B = \{a_i \mid i \in I\}$ and claim that $\chi$ is constant on $[B]^2$. Suppose $a_i < a_j$ are in $B$. Then $a_j \in A_{i+1}$ and so $\chi(a_i, a_j) = \epsilon_i = \epsilon$ as required. $\dashv$

The set $B$ we constructed is often called **monochromatic**. Here is a sample application of Ramsey's theorem.

THEOREM 6.9 (Bolzano-Weierstrass). *Every sequence of real numbers has a monotone subsequence.*

PROOF. Let $\langle a_n \mid n < \omega \rangle$ be a sequence of real numbers. We define a coloring $\chi : [\omega]^2 \to 2$ by $\chi(m, n) = 0$ if $a_m \leq a_n$ and $\chi(m, n) = 1$ otherwise. (Whenever we define a coloring we think of the domain as pairs $(m, n)$ with $m < n$.)

By Ramsey's theorem there is an infinite $B \subseteq \omega$ such that $B$ is monochromatic for $\chi$. It's easy to see that $\langle a_n \mid n \in B \rangle$ is monotone. $\dashv$

DEFINITION 6.10. An ultrafilter $\mathcal{U}$ on $\omega$ is **Ramsey** if for every coloring $\chi : [\omega]^2 \to 2$, there is $B \in \mathcal{U}$ such that $\chi$ is constant on $[B]^2$.

Note that a Ramsey ultrafilter must be nonprincipal. For let $n < \omega$ and define $\chi$ as follows. If $k > n$, then we set $\chi(n, k) = 0$ and for all other pairs $l < k$, we set $\chi(l, k) = 1$. Clearly $n$ cannot take part in any monochromatic set for $\chi$.

THEOREM 6.11. MA *implies there is a Ramsey ultrafilter.*

PROOF. There are $2^\omega$ possible colorings and we want to construct an ultrafilter with a monochromatic set for each coloring. We enumerate all of the colorings $\langle \chi_\alpha \mid \alpha < 2^\omega \rangle$ and construct a tower $T = \{A_\alpha \mid \alpha < 2^\omega\}$ such that $A_\alpha$ is monochromatic for $\chi_\alpha$. Recall that a tower of subsets of $\omega$ has the property that for all $\alpha < \beta$, $A_\beta \subseteq^* A_\alpha$.

Suppose that we have constructed $A_\alpha$ for each $\alpha < \beta$. Since MA implies $\mathfrak{t} = 2^\omega$, we can find $A \subseteq^* A_\alpha$ for all $\alpha < \beta$. By Ramsey's theorem we can find an infinite subset $A_\beta$ of $A$ which is monochromatic for $\chi_\beta$. This completes the construction.

To complete the proof we notice that our Tower $T$ has the finite intersection property! Hence $T$ can be extended to an ultrafilter $\mathcal{U}$ which is clearly Ramsey. $\dashv$

**§7. Applications of** MA **to trees.** In this section we define the notion of a tree and show that MA resolves Suslin's problem.

DEFINITION 7.1.
1. An ordering $(T, <_T)$ is a **tree** if it is wellfounded, transitive and irreflexive and for all $t \in T$ the set $\{x \in T \mid x <_T t\}$ is linearly ordered by $<_T$.
2. For $t \in T$, the **height** of $t$, denoted $\mathrm{ht}(t)$, is the order type of $\{x \in T \mid x <_T t\}$ under $<_T$.
3. For an ordinal $\alpha$, the $\alpha^{th}$ **level** of $T$, denoted $\mathrm{Lev}_\alpha(T)$, is the collection of nodes with height $\alpha$.
4. The **height** of $T$, denoted $\mathrm{ht}(T)$, is the least ordinal $\alpha$ such that $\mathrm{Lev}_\alpha(T) = \varnothing$.
5. A **branch** $b$ is a subset of $T$ which is linearly ordered by $<_T$.
6. If $\mathrm{ht}(T)$ is a regular cardinal, then we call a branch **cofinal** if the ordertype of $(b, <_T)$ is equal to $\mathrm{ht}(T)$.
7. A set $A \subseteq T$ is called an **antichain** if for all $s, t \in A$, neither $s <_T t$ nor $t <_T s$ holds.

THEOREM 7.2 (König Infinity Lemma). *Every finitely branching infinite tree has an infinite branch.*

PROOF. Here $T$ is **finitely branching** if for all $t \in T$, the set of immediate successors $\{s \in T \mid t <_T s$ and for no $u \in T$ do we have $t <_T u <_T s\}$ is finite.

Fix such a tree $T$. We can assume that $T$ has a minimum element, $t_0$. Since $T$ is infinite, $t_0$ has infinitely many nodes above it. Since $T$ is finitely branching, $t_0$ has an immediate successor $t_1$ which has infinitely many nodes above it. Continue in the same way to construct $\{t_n \mid n < \omega\}$, an infinite branch. ⊣

DEFINITION 7.3. Let $\kappa$ be a regular cardinal. A tree $T$ is a $\kappa$**-tree** if $\mathrm{ht}(T) = \kappa$ and for all $\alpha < \kappa$, $|\mathrm{Lev}_\alpha(T)| < \kappa$.

Note that an $\aleph_0$-tree satisfies the hypotheses of König Infinity Lemma. So we can restate König Infinity Lemma as 'Every $\aleph_0$-tree has a cofinal branch.'

THEOREM 7.4 (Aronszajn). *There is an $\aleph_1$-tree with no cofinal branch.*

REMARK 7.5. A tree as in the previous theorem is called an Aronszajn tree. The tree that Aronszajn constructed is **special** in the sense that there is a function $f$ from the tree to $\omega$ such that $f(x) \neq f(y)$ whenever $x <_T y$.

We can now make a general definition.

DEFINITION 7.6. A regular cardinal $\kappa$ has the **tree property** if every $\kappa$-tree has a cofinal branch.

We digress into a related question about linear orderings.

DEFINITION 7.7. Let $(L, <)$ be a linear ordering.
1. $(L, <)$ is **dense** if for all $a, b \in L$ there is $c \in L$ such that $a < c < b$.
2. $D \subseteq L$ is **dense in** $L$ if for all $a < b \in L$, there is $d \in D$ with $a < d < b$.
3. $(L, <)$ is **unbounded** if it has no greatest or least element.
4. $(L, <)$ is **complete** if every every nonempty bounded subset has a least upper bound.

Theorem 7.8 (Cantor). *Any two countable unbounded dense linear orders are isomorphic.*

The reals can be seen as the completion of the rationals with the usual ordering. One way to make this concrete is through the use of Dedekind cuts. It follows that every complete dense unbounded linear order with a countable dense subset is isomorphic to the real line $(\mathbb{R}, <)$.

The real line has another nice topological property.

Proposition 7.9. *Every collection of disjoint open subintervals of* $\mathbb{R}$ *is countable.*

So we canonize this with a definition.

Definition 7.10. A linear order $(L, <)$ has the **countable chain condition** if every collection of disjoint open subintervals is countable.

By the remarks above, the real line $(\mathbb{R}, <)$ is characterized up to isomorphism by the fact that it is a complete dense unbounded linear order with a countable dense subset. It is natural to ask whether the same holds if we replace "a countable dense subset" with "the countable chain condition".

Question 7.11 (Suslin's Problem). Suppose $(L, <)$ is a complete dense unbounded linear order that satisfies the countable chain condition. Is $(L, <)$ isomorphic to the real line?

This question cannot be resolved by ZFC, but it is resolved by MA and that is the main focus of this section. Suslin's problem can be rephrased using the following definition.

Definition 7.12. A **Suslin line** is a dense linearly ordered set that satisfies the countable chain condition but is not separable.

So Suslin's problem asks, 'Is there a Suslin line?' We return to the topic of trees by characterizing Suslin's problem in terms of trees.

Definition 7.13. A **Suslin tree** is an $\aleph_1$-tree in which all branches and antichains are countable.

Theorem 7.14. *There is a Suslin line if and only if there is a Suslin tree.*

Proof. Let $L$ be a Suslin line. We will construct a Suslin tree. The tree is a certain collection of nonempty closed subintervals of $L$ ordered by reverse inclusion. We construct intervals $I_\alpha$ for $\alpha < \omega_1$ by recursion and set $T = \{I_\alpha \mid \alpha < \omega_1\}$. Let $I_0 = [a_0, b_0]$ be arbitrary. Suppose that for some $\beta < \omega_1$ we have constructed $I_\alpha$ for all $\alpha < \beta$. We seek to define $I_\beta$. Look at the collection of endpoints of intervals so far. The set is countable and since $L$ is not separable it cannot be dense. Choose a nonempy closed interval $I_\beta$ which does not contain any of the endpoints of intervals so far. This completes the construction.

It remains to see that $(T, \supset)$ is a Suslin tree. First we show that it is a tree. Given $\alpha < \beta < \omega_1$, by the choice of $I_\beta$ we have either $I_\beta \subseteq I_\alpha$ or $I_\beta$ disjoint from $I_\alpha$. (Otherwise $I_\beta$ contains an endpoint from $I_\alpha$.) So $T$ is wellfounded and the predecessors of a point are linearly ordered. Lastly we show that all branches and antichains in $T$ are countable. It follows that $T$ has height $\omega_1$, so we will be done. If $I, J \in T$ are incomparable, then $I \cap J = \varnothing$. So every antichain in $T$ is a collection of pairwise disjoint intervals in $L$ which must be countable, since $L$ is Suslin. Next suppose that $\langle [x_\alpha, y_\alpha] \mid \alpha < \omega_1 \rangle$ is a cofinal branch. Note that $\langle x_\alpha \mid \alpha < \omega_1 \rangle$ is an increasing sequence of elements of $L$. It follows that $\langle (x_\alpha, x_{\alpha+1}) \mid \alpha < \omega_1 \rangle$ is an uncountable pairwise disjoint sequence of open intervals, which again contradicts that $L$ has the countable chain condition.

For the reverse direction, let $(T, <_T)$ be a Suslin tree. We will need some cosmetic improvements to the tree. Specifically, we assume

1. $(T, <_T)$ has a unique $<_T$-least element, that is, $|\operatorname{Lev}_0(T)| = 1$.
2. If $\alpha < \beta < \operatorname{ht}(T)$ and $x \in \operatorname{Lev}_\alpha(T)$, there is $y \in \operatorname{Lev}_\beta(T)$ such that $x <_T y$.
3. For each $x \in T$, the set of immediate $<_T$-successors of $x$ is infinite.
4. For limit $\lambda < \operatorname{ht}(T)$, if $x, y \in \operatorname{Lev}_\lambda(T)$ have the same predecessors, then $x = y$.

An $\omega_1$-tree satisfying the above conditions is called **normal**. We leave it as an exercise to show that if there is a Suslin tree, then there is a normal Suslin tree.

The domain $L$ of our Suslin line will be the set of maximal branches of $T$. To order these branches, choose, for each $x \in T$, some enumeration $\langle a^x_q \mid q \in \mathbb{Q} \rangle$ of the immediate $<_T$-successors of $x$; this may be done by condition (3) and the fact that levels of $T$ are countable. Given maximal branches $s, t \in L$, there is a unique $<_T$-largest $x \in s \cap t$, and unique $q(s), q(t) \in \mathbb{Q}$ such that $a^x_{q(s)} \in s$ and $a^x_{q(t)} \in t$. Set $s < t$ if $q(s) < q(t)$ in the usual order on $\mathbb{Q}$.

It's easy to check using normality that $(L, <)$ is a dense unbounded linear order. To see that it's ccc, notice that given any open interval $I$ of $(L, <)$, there is some $x \in T$ so that $I_x = \{s \in L \mid x \in s\}$ is contained in $I$, and if $I_1, I_2$ are disjoint, then the corresponding $x_1, x_2$ are incomparable. Finally, $(L, <)$ is not separable: given any countable subset $D$ of $L$, let $\alpha$ be the supremum of heights of branches in $D$, and take $x \in T$ with $\operatorname{ht}(x) > \alpha$. Then $I_x$ cannot contain any $s \in D$, that is, $D$ is not dense. ⊣

Theorem 7.15. $\mathrm{MA}(\aleph_1)$ *implies that there are no Suslin trees.*

Proof. Suppose for a contradiction that there is a Suslin tree $T$. By one of the homework exercises there is a Suslin subtree $T'$ with the following property: For every $\alpha < \beta < \omega_1$ and every $x \in \operatorname{Lev}_\alpha(T)$, there is a $y \in \operatorname{Lev}_\beta(T)$ such that $x < y$.

We make $T'$ into a poset $\mathbb{P}$ by reversing the order. The Suslinity of $T'$ implies that $\mathbb{P}$ is ccc. The extra condition satisfied by elements of $T'$ implies that for each $\beta < \omega_1$, $D_\alpha = \{t \in \mathbb{P} \mid \operatorname{ht}(t) > \beta\}$ is dense in $\mathbb{P}$.

By $\mathrm{MA}(\aleph_1)$ there is a $\{D_\alpha \mid \alpha < \omega_1\}$-generic filter $G$. It is not hard to see that $G$ is a cofinal branch through $T'$. This is a contradiction as $T'$ has no cofinal branches. ⊣

In particular, $\mathrm{MA}(\aleph_1)$ implies there are no Suslin lines, so answers Suslin's original question in the affirmative. In fact more is true under $\mathrm{MA}(\aleph_1)$.

Theorem 7.16. $\mathrm{MA}(\aleph_1)$ *implies that all Aronszajn trees are special.*

Note that every Suslin tree is an Aronszajn tree, but no Suslin tree is special (Exercise).

**§8. First Order Logic.** In this section we take a brief detour into first order logic. The idea for the section is to provide just enough background in first order logic to provide an understanding of forcing and independence results. We will touch briefly on both proof theory and model theory. Both of these topics deserve their own class.

The goal of the class is to prove that CH is independent of ZFC. This means that neither CH nor its negation are *provable* from the axioms of ZFC. Here are some questions that we will answer in this section:

1. What is a proof?
2. How does one prove that a statement has no proof?

We approach first order logic from the point of view of the mathematical structures that we already know. Here are some examples:

1. $(\aleph_{18}, <)$
2. $([\omega]^{<\omega}, \subseteq)$
3. $(\mathbb{Z}/7\mathbb{Z}, +_7)$
4. $\langle \mathbb{R}, +, \cdot, 0, 1\rangle$

We want to extract some common features from all of these structures. The first thing is that all have an underlying set, $\aleph_{18}, [\omega]^{<\omega}, \mathbb{Z}/7\mathbb{Z}, \mathbb{R}$. The second thing is that they all have some functions, relations or distinguished elements. Distinguished elements are called **constants**. Moreover, each function or relation has an *arity.* We formalize this with a definition.

Definition 8.1. A structure $\mathcal{M}$ is a quadruple $(M, \mathcal{C}, \mathcal{F}, \mathcal{R})$ where

1. $M$ is a set,
2. $\mathcal{C}$ is a collection of elements of $M$,
3. $\mathcal{F}$ is a collection of functions $f$ so that each $f$ has domain $M^n$ for some $n \geq 1$ and range $M$, and
4. $\mathcal{R}$ is a collection of sets $R$ so that each $R$ is a subset of $M^n$ for some $n \geq 1$.

This definition covers all of the examples above, but is a bit cumbersome in practice. We want some general way to organize structures by their type. How many constants? How many operations of a given arity? And so on. To do this we introduce the notion of a *signature.*

Definition 8.2. A **signature** $\tau$ is a quadruple $(\mathcal{C}, \mathcal{F}, \mathcal{R}, a)$ where $\mathcal{C}, \mathcal{F}, \mathcal{R}$ are pairwise disjoint and $a$ is a function from $\mathcal{F} \cup \mathcal{R}$ to $\mathbb{N} \setminus \{0\}$. The elements of $\mathcal{C} \cup \mathcal{F} \cup \mathcal{R}$ are the **non-logical symbols**.

Here we think of $a$ as assigning the arity of the function or relation. If $P$ is a function or relation symbol, then $a(p) = n$ means that $P$ is $n$-ary. Here are some examples.

1. The signature for an ordering is $\tau_< = (\varnothing, \varnothing, \{<\}, (<\mapsto 2))$. This is a bit much so usually we write $\tau_< = (<)$, since the arity of $<$ is implicit.
2. The signature for a group is $\tau_{\text{group}} = (\{1\}, \{\cdot\}, \varnothing, (\cdot \mapsto 2))$. Again we abuse notation here: Since it is easy to distinguish between function and constant symbols, we just write $\tau_{\text{group}} = (\cdot, 1)$.
3. The signature for a ring with 1 is $\tau_{\text{ring}} = (+, \cdot, 0, 1)$ (with by now standard abuse of notation).

Now we want to know when a structure has a given signature $\tau$.

Definition 8.3. A structure $\mathcal{M}$ is a **$\tau$-structure** if there is a function $i$ which takes

1. each constant symbol from $\tau$ to a member $i(c) \in M$,
2. each $n$-ary relation symbol $R$ to a subset $i(R) \subseteq M^n$ and
3. each $n$-ary function symbol $f$ to a function $i(f) : M^n \to M$.

We think of members of the signature as formal symbols and the map $i$ is the interpretation that we give to the symbols. Up to renaming the symbols each structure is a $\tau$-structure for a single signature $\tau$. Instead of writing $i(-)$ all the time, we will write $f^{\mathcal{M}}$ for the interpretation of the function symbol $f$ in the $\tau$-structure $\mathcal{M}$.

We gather some definitions.

Definition 8.4. Let $\tau$ be a signature and $\mathcal{M}, \mathcal{N}$ be $\tau$-structures.

1. $\mathcal{M}$ is a **substructure** of $\mathcal{N}$ if $M \subseteq N$ and for all $c, R, f$ from $\tau$, $c^{\mathcal{M}} = c^{\mathcal{N}}$, $R^{\mathcal{M}} = R^{\mathcal{N}} \cap M^n$ where $n = a(R)$ and $f^{\mathcal{M}} = f^{\mathcal{N}} \restriction M^k$ where $k = a(f)$.
2. A map $H : M \to N$ is a **$\tau$-homomorphism** if $H``M$ together with the natural structure is a substructure of $\mathcal{N}$.
3. A map $H : M \to N$ is an **isomorphism** if $H$ is a bijection and $H$ and $H^{-1}$ are $\tau$-homomorphisms.

Note in particular that in (1), $\mathcal{M}$ must be closed under the function $f^{\mathcal{N}}$ to be a substructure.

If you are familiar with group theory, you will see that 'substructure' in the signature $\tau_{\text{group}}$ (as we have formulated it) does not coincide with 'subgroup'. In particular $(\mathbb{N}, +, 0)$ is a substructure of $(\mathbb{Z}, +, 0)$, but it is not a subgroup. The notion of homomorphism and isomorphism are the same as those from group theory.

We now move on to talking about languages, formulas and sentences. Again we compile some large definitions.

Definition 8.5. Let $\tau$ be a signature.

1. A **word** in FOL($\tau$) is a finite concatenation of logical symbols,
$$\neg \quad \wedge \quad \vee \quad \to \quad \forall \quad \exists \quad =$$
punctuation symbols,
$$, \quad ( \quad )$$
and variables
$$v_0 \quad v_1 \quad v_2 \quad \dots$$
as well as symbols coming from the signature $\tau$: constant symbols $c \in \mathcal{C}$, function symbols $f \in \mathcal{F}$, and relation symbols $R \in \mathcal{R}$.
2. A **term** in FOL($\tau$) (a $\tau$-term) is a word formed by the following recursive rules: each constant symbol is a term; each variable is a term; and if $t_1, \dots t_n$ are terms, then $f(t_1, \dots, t_n)$ is a term when $a(f) = n$.

DEFINITION 8.6. Let $\tau$ be a signature and $\mathcal{M}$ be a $\tau$-structure. Suppose that $t$ is a $\tau$-term using variables $v_1, \dots v_n$. We define a function $t^{\mathcal{M}} : M^n \to M$ by recursion. Let $\vec{a} \in M^n$.

1. If $t = c$ where $c$ is a constant symbol, then $t^{\mathcal{M}}(\vec{a}) = c^{\mathcal{M}}$.
2. If $t = v_i$, then $t^{\mathcal{M}}(\vec{a}) = a_i$.
3. If $t = f(t_1, \dots t_n)$, then $t^{\mathcal{M}}(\vec{a}) = f(t_1^{\mathcal{M}}(\vec{a}), \dots t_n^{\mathcal{M}}(\vec{a}))$.

DEFINITION 8.7. A **formula** in FOL($\tau$) is built recursively from $\tau$-terms as follows:

1. If $t_1, t_2$ are terms, then $t_1 = t_2$ is a formula.
2. If $t_1, \dots t_n$ are terms, then $R^{\mathcal{M}}(t_1, \dots t_n)$ is a formula.
3. if $\phi$ and $\psi$ are formulas, then $\neg\phi, \phi \wedge \psi, \phi \vee \psi, \phi \to \psi, \forall v\phi$ and $\exists v\phi$ are formulas.

The formulas defined in clauses (1) and (2) are called **atomic.**

Suppose that $\exists v\psi$ occurs in the recursive construction of a formula $\phi$. We say that the **scope** of this occurrence of $\exists v$ is $\psi$. Similarly for $\forall v$. An occurrence of a variable $v$ is said to be **bound** if it occurs in the scope of an occurrence of some quantifier.

If an occurrence of a variable is not bound then it is called **free**. When we write a formula $\phi$ we typically make it explicit that there are free variables by writing $\phi(\vec{v})$. A formula with no free variables is called a **sentence**. In a given structure, a formula with $n$ free variables is interpreted like a relation on the structure. It is true for some $n$-tuples of elements and false for others.

DEFINITION 8.8. Let $\mathcal{M}$ be a $\tau$ structure and $\phi(\vec{v})$ be a formula with $n$ free variables. For $\vec{a} = (a_1, \dots a_n)$ we define a relation $\mathcal{M} \vDash \phi(\vec{a})$ by recursion on the construction of the formula.

1. If $\phi$ is $t_1 = t_2$, then $\mathcal{M} \vDash \phi(\vec{a})$ if and only if $t_1^{\mathcal{M}}(\vec{a}) = t_2^{\mathcal{M}}(\vec{a})$.
2. If $\phi$ is $R(t_1, \dots t_n)$, then $\mathcal{M} \vDash \phi(\vec{a})$ if and only if $R^{\mathcal{M}}(t_1^{\mathcal{M}}(\vec{a}), \dots t_n^{\mathcal{M}}(\vec{a}))$.
3. If $\phi$ is $\neg\psi$, then $\mathcal{M} \vDash \phi(\vec{a})$ if and only if $\mathcal{M} \nvDash \psi(\vec{a})$.
4. If $\phi$ is $\psi_1 \wedge \psi_2$, then $\mathcal{M} \vDash \phi(\vec{a})$ if and only if $\mathcal{M} \vDash \psi_1(\vec{a})$ and $\mathcal{M} \vDash \psi_2(\vec{a})$.
5. If $\phi$ is $\psi_1 \vee \psi_2$, then $\mathcal{M} \vDash \phi(\vec{a})$ if and only if $\mathcal{M} \vDash \psi_1(\vec{a})$ or $\mathcal{M} \vDash \psi_2(\vec{a})$.
6. If $\phi$ is $\psi_1 \to \psi_2$, then $\mathcal{M} \vDash \phi(\vec{a})$ if and only if $\mathcal{M} \nvDash \psi_1(\vec{a})$ or $\mathcal{M} \vDash \psi_2(\vec{a})$.
7. If $\phi$ is $\forall u\psi(\vec{v}, u)$, then $\mathcal{M} \vDash \psi(\vec{a})$ if and only if for all $b \in M$, $\mathcal{M} \vDash \psi(\vec{a}, b)$.
8. If $\phi$ is $\exists u\psi(\vec{v}, u)$, then $\mathcal{M} \vDash \psi(\vec{a})$ if and only if there exists $b \in M$ such that $\mathcal{M} \vDash \psi(\vec{a}, b)$.

We read $\mathcal{M} \vDash \phi(\vec{a})$ as '$\mathcal{M}$ models (satisfies, thinks) $\phi(\vec{a})$' or '$\phi$ holds in $\mathcal{M}$ about $\vec{a}$'.

Here is a relatively simple example of the satisfaction relation:

$$(\aleph_{18}, <) \vDash \forall\beta\exists\alpha\ \beta < \alpha$$

DEFINITION 8.9. Let $\tau$ be a signature and $\mathcal{M}, \mathcal{N}$ be $\tau$-structures.

1. $\mathcal{M}$ is an **elementary substructure of** $\mathcal{N}$ (written $\mathcal{M} \prec \mathcal{N}$) if $M \subseteq N$ and for all formulas $\phi(\vec{v})$ and $\vec{a} \in M^n$, $\mathcal{M} \vDash \phi(\vec{a})$ if and only if $\mathcal{N} \vDash \phi(\vec{a})$.
2. A map $H : M \to N$ is an **elementary embedding** if for all formulas $\phi(\vec{v})$ and all $\vec{a} \in M^n$, $\mathcal{M} \vDash \phi(\vec{a})$ if and only if $\mathcal{N} \vDash \phi(H(a_1), \dots H(a_n))$.

Elementary substructures and elementary embeddings are key points of study in model theory and also in set theory.

DEFINITION 8.10. A **theory** $T$ is a collection of $\tau$-sentences.

For example the group axioms are a theory in the signature of groups.

DEFINITION 8.11. A structure $\mathcal{M}$ **satisfies** a theory $T$ if $\mathcal{M} \models \phi$ for every $\phi \in T$.

Next we say a word or two about proofs. There is a whole field of study here, but we will only deal with it briefly. We are ready to answer the question 'What is a proof?'. To do so we forget about structures altogether and focus on formulas in a fixed signature $\tau$.

Proofs are required to follow certain *rules of inference*. Examples of rules of inference are things like *modus ponens*:

$$\text{Given } \phi \text{ and } \phi \to \psi, \text{ infer } \psi.$$

In a proof we are also allowed to use *logical axioms*. An example of a logical axiom is $\neg\neg\phi \to \phi$. This is the logical axiom that we use when we do a proof by contradiction.[2]

DEFINITION 8.12. Let $T$ be a theory and $\phi$ be a sentence. A **proof of** $\phi$ **from** $T$ is a finite sequence of formulas $\phi_1, \dots \phi_n$ such that $\phi_n = \phi$ and for each $i \leq n$, $\phi_i$ is either a member of $T$, a logical axiom, or can be obtained from some of the $\phi_j$ for $j < i$ by a rule of inference.

In this case we say that $T$ **proves** $\phi$ and write $T \vdash \phi$. Now we want to connect proofs with structures. The connection is through *soundness and completeness*. We write $T \models \phi$ if every structure which satisfies $T$ also satisfies $\phi$.

THEOREM 8.13 (Soundness). *If $T \vdash \phi$, then $T \models \phi$.*

DEFINITION 8.14. A theory $T$ is **consistent** if there is no formula $\phi$ such that $T \vdash \phi \wedge \neg\phi$.

THEOREM 8.15 (Completeness). *Every consistent theory $T$ has a model of size at most* $\max\{|\tau|, \aleph_0\}$

COROLLARY 8.16. *If $T \models \phi$, then $T \vdash \phi$.*

So now we are ready to answer the question of how one proves that a statement like CH cannot be proven nor disproven from the axioms. To show that there is no proof of CH or its negation, we simply have to show that there are two models of set theory, one in which CH holds and one in which CH fails!

REMARK 8.17. An example of the idea of independence that people have heard of comes from geometry. In particular, Euclid's parallel postulate is independent of the other four postulates. The proof involves showing that there are so-called *non-Euclidean geometries*; these are essentially models of the first four postulates in which the parallel postulate fails.

[2]For a complete list of rules of inference and logical axioms, we refer the reader to Kunen's book *The Foundations of Mathematics.*

**§9. Models of Set theory.** Armed with the model theoretic tools of the previous section, we can begin a systematic study of models of set theory. The signature for set theory is that of a single binary relation, $\tau_{\text{sets}} = (\in)$. So a model in this signature is just $(M, E)$, where $M$ is a set and $E$ is a binary relation on $M$. (Of course, the binary relation does not need to have any relation to the true membership relation $\in$.)

We saw in the last section how to build formulas in the signature $\tau_{\text{sets}}$. It's worth noting that each axiom of ZFC can be written as such a formula. For example, we can formalize the axiom of foundation as

$$\forall x(\exists y(y \in x) \to \exists z(z \in x \wedge \forall y(y \in x \to \neg(y \in z)))),$$

and for each formula $\phi(u, v_1, \dots, v_n)$, we have an instance of the axiom scheme of comprehension,

$$\forall a_1 \dots \forall a_n \forall x \exists z \forall y(y \in z \longleftrightarrow (y \in x \wedge \phi(y, a_1, \dots, a_n))).$$

Set theory is extremely powerful, since from the axioms of ZFC we can formalize classical mathematics in its entirety. That this can be done with only the single primitive notion of set membership is our whole subject's *raison d'etre*.

It will be easier for us to work with $\tau_{\text{sets}}$-structures $\mathcal{M}$ whose interpretation $\in^{\mathcal{M}}$ agrees with the true membership relation; that is, models of the form $(M, \in)$. It will also be important that our models are *transitive*. Recall a set $z$ is transitive if for every $y \in z$, $y \subseteq z$. We will say a model of set theory $(M, \in)$ is transitive if $M$ is.

Transitive models are important because they reflect basic facts about the universe of sets. For the following definition, we regard the formulas $(\exists x \in y)\phi$ and $(\forall x \in y)\phi$ as abbreviations for the formulas $\exists x(x \in y \wedge \phi)$ and $\forall x(x \in y \to \phi)$, respectively.

DEFINITION 9.1. A formula $\phi$ in the language of set theory is a **$\Delta_0$-formula** if

1. $\phi$ has no quantifiers, or
2. $\phi$ is of the form $\psi_0 \wedge \psi_1$, $\psi_0 \vee \psi_1$, $\psi_0 \to \psi_1$, $\neg\psi_0$ or $\psi_0 \leftrightarrow \psi_1$ for some $\Delta_0$-formulas $\psi_0, \psi_1$, or
3. $\phi$ is $(\exists x \in y)\psi$ or $(\forall x \in y)\psi$ where $\psi$ is a $\Delta_0$-formula.

PROPOSITION 9.2. *If $(M, \in)$ is a transitive model and $\phi$ is a $\Delta_0$-formula, then for all $\vec{x} \in M^n$, $(M, \in) \vDash \phi$ if and only if $\phi$ holds.*

To save ourselves from writing $(M, \in) \vDash \phi$, we will write $\phi^M$ instead.

PROOF. We go by induction on the complexity of the $\Delta_0$ formula. Clearly if $\phi$ is atomic, then we have $\phi$ if and only if $\phi^M$. Also if the conclusion holds for $\psi_0$ and $\psi_1$, then clearly it holds for all of the formulas listed in item (2). It remains to show the conclusion for $\phi$ of the form $(\exists x \in y)\psi(x)$ where the conclusion holds for $\psi$. Suppose $\phi^M$ holds. Then there is an $x \in M \cap y$ such that $\psi(x)^M$. So $\psi(x)$ holds and therefore so does $(\exists x \in y)\psi(x)$. Finally suppose that $\phi$ holds. Then there is $x \in y$ such that $\psi(x)$ holds. Since $y \in M$ and $M$ is transitive, the witness $x$ is in $M$. Moreover $\psi(x)^M$. Therefore $\phi^M$ holds. $\dashv$

If $M$ is a transitive model, $\phi$ is any formula and $\phi$ if and only if $\phi^M$, then we say that $\phi$ is **absolute for** $M$.

It is reasonable to ask what can be expressed by $\Delta_0$-formulas.

PROPOSITION 9.3. *The following expressions can be written as $\Delta_0$-formulas.*

1. $x = \varnothing$, *$x$ is a singleton, $x$ is an ordered pair,* $x = \{y, z\}$, $x = (y, z)$, $x \subseteq y$, *$x$ is transitive, $x$ is an ordinal, $x$ is a limit ordinal, $x$ is a natural number,* $x = \omega$.
2. $z = x \times y$, $z = x \setminus y$, $z = x \cap y$, $z = \bigcup x$, $z = \operatorname{ran}(x)$, $y = \operatorname{dom}(x)$.
3. *$R$ is a relation, $f$ is a function,* $y = f(x)$, $g = f \restriction x$.

PROOF. Exercise. ⊣

We recall some transitive models we've seen before. First, the $V$-hierarchy.

$$\begin{aligned} V_0 &= \varnothing \\ V_{\alpha+1} &= \mathcal{P}(V_\alpha) \\ V_\gamma &= \bigcup_{\alpha<\gamma} V_\alpha \text{ for } \gamma \text{ limit.} \end{aligned}$$

$V$ is then defined as the union $\bigcup_{\alpha \in \mathrm{On}} V_\alpha$. Note the sets $V_\alpha$ are transitive and increasing. Foundation asserts that every set belongs to $V$. We can thus define, for all sets $x$, $\operatorname{rk}(x)$ to be the least $\alpha$ so that $x \in V_{\alpha+1}$. Note then that $x \in y$ implies $\operatorname{rk}(x) < \operatorname{rk}(y)$.

Recall next that for an infinite cardinal $\kappa$, $H_\kappa$ is the collection of sets whose transitive closure has size less than $\kappa$. Note that $V_\omega = H_\omega$. We state a result about $H_\kappa$ for $\kappa$ regular.

THEOREM 9.4. *If $\kappa$ is regular and uncountable, then $H_\kappa$ is a transitive model of all of the axioms of* ZFC *except the power set axiom.*

We will also state and prove a theorem about the $V$ hierarchy.

THEOREM 9.5 (The Reflection theorem). *Let $\phi(x_1, \dots x_n)$ be a formula. For every set $M_0$ there are*

1. *an $M$ such that $M_0 \subseteq M$, $|M| \leq |M_0| \cdot \aleph_0$ and for all $\vec{a} \in M^n$, $\phi^M(\vec{a})$ if and only if $\phi(\vec{a})$ and*
2. *an ordinal $\alpha$ such that for all $\vec{a} \in (V_\alpha)^n$, $\phi^{V_\alpha}(\vec{a})$ if and only if $\phi(\vec{a})$.*

PROOF. Let $\phi_1, \dots \phi_n$ be an enumeration of all subformulas of $\phi$. We can assume that $\forall$ does not appear in any of the $\phi_j$, since $\forall$ can be replaced with $\neg\exists\neg$. Let $M_0$ be given.

We define by induction an increasing sequence of sets $M_i$ for $i < \omega$. Suppose that $M_i$ has be defined for some $i < \omega$. We choose $M_{i+1}$ with the following property for all $j \leq n$ and all tuples $\vec{a}$ from $M_i$:

$$\text{If } \exists x \phi_j(x, \vec{a}), \text{ then there is } b \in M_{i+1} \text{ such that } \phi_j(b, \vec{a}).$$

We use the axiom of choice to choose witnesses to these existential formulas from among the witnesses of minimal rank. It is clear that for all $i$, $|M_{i+1}| \leq |M_i| \cdot \aleph_0$. Let $M = \bigcup_{i<\omega} M_i$. Now we prove that $M$ reflects $\phi$ by induction on the complexity of formulas appearing in $\phi_1, \dots \phi_n$. The atomic formula, conjunction, disjunction, negation and implication cases are straightforward. The existential quantifier step follows from our construction of the $M_i$. Given a tuple $\vec{a}$ from $M$ and a formula $\phi_j$ for which $\exists x \phi_j(x, \vec{a})$ holds, all of the tuple's elements appear in some $M_i$ and therefore there is a witness to $\exists x \phi_j(x, \vec{a})$ in $M_{i+1}$.

The proof of the second part of the theorem is an easy modification of the first part. Instead of choosing specific witnesses to formulas, we simply inductively choose ordinals $\alpha_i$ such that $V_{\alpha_{i+1}}$ contains witness to existential formulas with parameters from $V_{\alpha_i}$. ⊣

Finally we want a solid connection between transitive and nice enough non-transitive models.

Definition 9.6. A model $(P, E)$ is

1. **well-founded** if the relation $E$ is well-founded.
2. **extensional** if for all $x, y \in P$, $\{z \in P \mid z\, E\, x\} = \{z \in P \mid z\, E\, y\}$ implies that $x = y$.

Theorem 9.7 (The Mostowski Collapse Theorem). *Every well-founded, extensional model $(P, E)$ is isomorphic to a transitive model $(M, \in)$. Moreover the set $M$ and the isomorphism are unique.*

The model $(M, \in)$ is called the **Mostowski collapse** of $(P, E)$.

Proof. Let $(P, E)$ be a well-founded, extensional model. We define a map $\pi$ on $P$ by induction on $E$. Induction on $E$ makes sense since $E$ is well-founded. Suppose that for some $x$ we have defined $\pi$ on the set $\{y \in P \mid y\, E\, x\}$. We define $\pi(x) = \{\pi(y) \mid y\, E\, x\}$. Let $M$ be the range of $\pi$.

Clearly $M$ is transitive and $\pi$ is surjective. We show that $\pi$ is one-to-one. Suppose that $z \in M$ is of minimal rank such that there are $x, y \in P$ such that $x \neq y$ and $z = \pi(x) = \pi(y)$. Since $E$ is extensional, there is $w$ such that without loss of generality $w\, E\, x$ and not $w\, E\, y$. Since $\pi(w) \in \pi(y)$, there is a $u\, E\, y$ such that $\pi(u) = \pi(w)$. This contradicts the minimality of the choice of $z$, since $\pi(u) = \pi(w) \in z$ and $u \neq w$.

To see that $M$ and $\pi$ are unique it is enough to show that if $M_1, M_2$ are transitive, then any isomorphism from $M_1$ to $M_2$ must be the identity map. This is enough since if we had $\pi_i : P \to M_i$ for $i = 1, 2$, then $\pi_2 \pi_1^{-1}$ would be an isomorphism from $M_1$ to $M_2$. Now an easy $\in$-induction shows that any isomorphism between transitive sets $M_1$ and $M_2$ must be the identity. ⊣

This allows us to prove the following theorem which is needed to fully explain consistency results.

Theorem 9.8. *For any axioms $\phi_1, \ldots \phi_n$ of* ZFC*, there is a countable transitive model $M$ such that $M \vDash \phi_1, \ldots \phi_n$.*

This is an easy application of both the reflection and Mostowski Collapse theorems.

**§10. Forcing.** This section was not written by the author of these notes: The introduction to forcing in sections 10.1 and 10.2 was written up by Justin Palumbo; the proof of the forcing theorems in section 10.3 was written up by Sherwood Hachtman.

**10.1. The Generic Extension $M[G]$.**

DEFINITION 10.1. Let $M$ be a countable transitive model of ZFC. Let $\mathbb{P}$ be a poset with $\mathbb{P} \in M$. A filter $G$ is **$\mathbb{P}$-generic over** $M$ (or just $\mathbb{P}$-generic when $M$ is understood from context, as will usually be the case) if for every set $D \in M$ which is dense in $\mathbb{P}$ we have that $G \cap D \neq \varnothing$.

LEMMA 10.2. *Let $M$ be a countable transitive model of ZFC with $\mathbb{P} \in M$. Then there is a $\mathbb{P}$-generic filter $G$. In fact, for any $p \in \mathbb{P}$ there is a $\mathbb{P}$-generic filter $G$ which contains $p$.*

PROOF. Since $M$ is countable, getting a $\mathbb{P}$-generic filter $G$ is the same as finding a $\mathcal{D}$-generic filter $G$ where

$$\mathcal{D} = \{D \in M : D \text{ is dense}\}.$$

Since MA($\omega$) always holds such a filter exists. If we want to ensure that $p \in \mathbb{G}$ we use the same proof as that of MA($\omega$), starting our construction at $p$. ⊣

Let us give a few motivating words.

Suppose we wanted to construct a model of CH, and we had given to us a countable transitive $M$, a model of ZFC. Now $M$ satisfies ZFC, so within $M$ one may define the partial order $\mathbb{P}$ consisting of all countable approximations to a function $f : \omega_1 \to \mathcal{P}(\omega)$. Of course $M$ is countable, so the things that $M$ believes are $\omega_1$ and $\mathcal{P}(\omega)$ are not actually the real objects. But for each $X \in \mathcal{P}(\omega)^M$ the set $D_X = \{p \in \mathcal{P} : X \in \operatorname{ran}(p)\}$ is dense, as is the set $E_\alpha = \{p \in \mathbb{P} : \alpha \in \operatorname{dom}(p)\}$ for each $\alpha < \omega_1^M$. So a $\mathbb{P}$-generic filter $G$ will intersect each of those sets, and will by the usual arguments yield a surjection $g : \omega_1^M \to \mathcal{P}(\omega)^M$. Thankfully, by the previous lemma, such a $G$ exists. Unfortunately there is no reason to believe $G$ is in $M$, and it is difficult to see how we would go about adding it. This is what we now learn: how to force a generic object which we can adjoin to $M$ without doing too much damage to its universe.

Given any poset $\mathbb{P}$ in $M$, and a $\mathbb{P}$-generic filter $G$, the method of forcing will give us a way of creating a new countable transitive model $M[G]$ satisfying ZFC that extends $M$ and contains $G$. Now just getting such a model is not enough. For in the example above the surjection $g : \omega_1 \to \mathcal{P}(\omega)$ defined from $G$ was a mapping between the objects in $M$. But a priori it may well be that the model $M[G]$ has a different version of $\omega_1$ and a different version of $\mathcal{P}(\omega)$ and so the CH still would not be satisfied. It turns out that in this (and many other cases) the forcing machinery will work out in our favor, and these things will not be disturbed.

It is worth pointing out that when $\mathbb{P} \in M$ then the notion of being a partial order, or being dense in $\mathbb{P}$ are absolute (written out the formulas just involve bounded quantifiers over $\mathbb{P}$). So if $D \in M$ then $M \vDash$ "$D$ is dense" exactly when $D$ really is dense. Thus the countable set $\{D \in M : D \text{ is dense}\}$ is exactly the same collection defined in $M$ to be the collection of *all* dense subsets of $\mathbb{P}$. Unless

$\mathbb{P}$ is something silly this will not actually be all the dense subsets, since $M$ will be missing some. Let us isolate a class of not-silly posets.

DEFINITION 10.3. A poset $\mathbb{P}$ is **separative** if (1) for every $p$ there is a $q$ which properly extends $p$ (i.e. $q < p$) and (2) whenever $p \not\leq q$ then there is an $r \leq p$ with $q \perp r$.

DEFINITION 10.4. A poset $\mathbb{P}$ is **non-atomic** if for any $p \in \mathbb{P}$ there exist $q, r \leq p$ which are incompatible.

Essentially every example of a poset that we have used thus far is separative. Notice that every separative poset is non-atomic.

PROPOSITION 10.5. *Suppose $\mathbb{P}$ is non-atomic and $\mathbb{P} \in M$. Let $G$ be $\mathbb{P}$-generic. Then $G \notin M$.*

PROOF. Assume $G \in M$ and consider the set $D = \mathbb{P} \setminus G$. Then $D$ belongs to $M$. Let us see that $D$ is dense. Let $p \in \mathbb{P}$ be arbitrary. Since $\mathbb{P}$ is non-atomic there are $q, r \leq p$ which are incompatible. Since $G$ is a filter, at most one of them can belong to $G$ and whichever one does not belongs to $D$.

Since $D$ is dense and $G$ is $\mathbb{P}$-generic, $G$ should intersect $D$. But that is ridiculous. ⊣

Now we will show how, given $G$ and $M$, to construct $M[G]$. Clearly the model $M$ will not know about the model $M[G]$, since $G$ can not be defined within $M$. But it will be the case that this is the only barrier. All of the tools to create $M[G]$ can assembled within $M$ itself; only a generic filter $G$ is needed to get them to run.

DEFINITION 10.6. We define the class of **$\mathbb{P}$-names** by defining for each $\alpha$ the **$\mathbb{P}$-names of name-rank** $\alpha$. (For a $\mathbb{P}$-name $\tau$ we will use $\rho(\tau)$ to denote the name-rank of $\tau$). The only $\mathbb{P}$-name of name-rank 0 is the empty set $\varnothing$. And recursively, if all the $\mathbb{P}$-names of name-rank strictly less than $\alpha$ have been defined, we say that $\tau$ is a $\mathbb{P}$-name of name-rank $\alpha$ if every $x \in \tau$ is of the form $x = \langle \sigma, p \rangle$ where $\sigma$ is a $\mathbb{P}$-name and $p \in \mathbb{P}$.

Another way of stating the definition is just to say that a set $\tau$ of ordered pairs is called a $\mathbb{P}$-name if it satisfies (recursively) the following property: every element of $\tau$ has the form $\langle \sigma, p \rangle$ where $\sigma$ is itself a $\mathbb{P}$-name and $p$ is an element of $\mathbb{P}$.

In analogy with the von Neumann hierarchy $V_\alpha$, we may define

$$\begin{aligned} V_0^{\mathbb{P}} &= \varnothing \\ V_{\alpha+1}^{\mathbb{P}} &= \mathcal{P}(V_\alpha^{\mathbb{P}} \times \mathbb{P}) \\ V_\gamma^{\mathbb{P}} &= \bigcup_{\alpha<\gamma} V_\alpha^{\mathbb{P}} \text{ for } \gamma \text{ limit.} \end{aligned}$$

Then $V_\alpha^{\mathbb{P}}$ is the set of $\mathbb{P}$-names of rank $< \alpha$. So the class of $\mathbb{P}$-names is obtained by imitating the construction of the whole universe $V$, but "tagging" all the sets at every step, by elements of $\mathbb{P}$.

It is not hard to see that the notion of being a $\mathbb{P}$-name is absolute; that is, $M \vDash$ "$\tau$ is a $\mathbb{P}$-name" exactly when $\tau$ is a $\mathbb{P}$-name. This is because the concept is defined by transfinite recursion from absolute concepts. As another piece of notation, since $\tau$ is a set of ordered pairs, it makes sense to use $\mathrm{dom}(\tau)$ as notation for all the $\sigma$ occurring in the first coordinate of an element of $\tau$.

DEFINITION 10.7. If $M$ is a countable transitive model of ZFC, then $M^{\mathbb{P}}$ denotes the collection of all the $\mathbb{P}$-names that belong to $M$.

Alone the $\mathbb{P}$-names are just words without any meaning. The people living in $M$ have the names but they do not know any way of giving them a coherent meaning. But once we have a $\mathbb{P}$-generic filter $G$ at hand, they can be given values.

DEFINITION 10.8. Let $\tau$ be a $\mathbb{P}$-name and $G$ a filter on $\mathbb{P}$. Then the **value of** $\tau$ **under** $G$, denoted $\tau[G]$, is defined recursively as the set

$$\{\sigma[G] : \langle\sigma, p\rangle \in \tau \text{ and } p \in G\}.$$

With this definition in mind, one can think of an element $\langle\sigma, p\rangle$ of a $\mathbb{P}$-name $\tau$ as saying that $\sigma[G]$ has probability $p$ of belonging to $\tau[G]$. The fact that we are calling the maximal element of our posets $\mathbb{1}$ makes this all the more suggestive, for $\mathbb{1}$ belongs to every filter $G$. So in particular, whatever $G$ is, if we have $\tau = \{\langle\varnothing, \mathbb{1}\rangle\}$ then $\tau[G] = \{\varnothing\}$. On the other hand if $\tau = \{\langle\varnothing, p\rangle\}$ for some $p$ that does not belong to $G$ then $\tau[G] = \varnothing$.

DEFINITION 10.9. If $M$ is a countable transitive model of ZFC, $\mathbb{P} \in M$, and $G$ is a filter, then $M[G] = \{\tau[G] : \tau \in M^{\mathbb{P}}\}$.

THEOREM 10.10. *If $G$ is a $\mathbb{P}$-generic filter then $M[G]$ is a countable transitive model of* ZFC *such that $M \subseteq M[G]$, $G \in M[G]$, and $M \cap \mathrm{On} = M[G] \cap \mathrm{On}$.*

Obviously $M[G]$ is countable, since the map sending a name to its interpretation is a surjection from a countable set (the names in $M$) to $M[G]$. There are a large number of things to verify in order to prove theorem (the brunt of the work being to check that $M$ satisfies each axiom of ZFC), but going through some of the verification will help us get an intuition for what exactly is going on with these $\mathbb{P}$-names.

One thing at least is not hard to see.

LEMMA 10.11. *$M[G]$ is transitive.*

PROOF. Suppose $x \in M[G]$ and $y \in x$. Then $x = \tau[G]$ for some $\tau \in M^{\mathbb{P}}$. By definition, every element of $\tau[G]$ has the form $\sigma[G]$, where $\sigma$ is a $\mathbb{P}$-name. So $y = \sigma[G]$ for some $\sigma$ with $\langle\sigma, p\rangle \in \tau$. As $M$ is transitive, $\sigma \in M$ and hence $\sigma \in M^{\mathbb{P}}$. So $y = \sigma[G] \in M[G]$. ⊣

LEMMA 10.12. *$M \subseteq M[G]$.*

PROOF. For each $x \in M$ we must devise a name $\check{x}$ so that $\check{x}[G] = x$. It turns out we can do this independently of $G$. We've already seen how to name $\varnothing$; $\check{\varnothing} = \varnothing$. The same idea works recursively for every $x$. Set $\check{x} = \{\langle\check{y}, \mathbb{1}\rangle : y \in x\}$.

Then since $\mathbb{1}$ belongs to $G$, we have by definition that $\check{x}[G] = \{\check{y}[G] : y \in x\}$ which by an inductive assumption is equal to $\{y : y \in x\} = x$. ⊣

LEMMA 10.13. *$G \in M[G]$.*

Proof. We must devise a name $\Gamma$ so that whatever $G$ is we have $\Gamma[G] = G$. Set $\Gamma = \{\langle \check{p}, p\rangle : p \in \mathbb{P}\}$. Then $\Gamma[G] = \{\check{p}[G] : p \in G\} = \{p : p \in G\} = G$. $\dashv$

Lemma 10.14. *The models $M$ and $M[G]$ have the same ordinals; that is, we have $M \cap \mathrm{On} = M[G] \cap \mathrm{On}$.*

Proof. We first show that for any $\mathbb{P}$-name $\tau$, $\mathrm{rk}(\tau[G]) \leq \rho(\tau)$. We do this by induction on $\tau$. Suppose inductively that this holds for any $\mathbb{P}$-name in the domain of $\tau$. Now each $\sigma \in \mathrm{dom}(\tau)$ clearly has $\rho(\sigma) < \rho(\tau)$. So by induction, each $\mathrm{rk}(\sigma[G]) < \rho(\tau)$. Now $\tau[G] \subseteq \{\sigma[G] : \sigma \in \mathrm{dom}(\tau)\}$. Since $\mathrm{rk}(\tau[G]) = \sup\{\mathrm{rk}(x)+1 : x \in \tau[G]\}$ and each $\mathrm{rk}(x)+1 \leq \rho(\tau)$, it must be that $\mathrm{rk}(\tau[G]) \leq \rho(\tau)$.

With that established, we show that $\mathrm{On} \cap M[G] \subseteq M \cap \mathrm{On}$ (the other inclusion is obvious). Let $\alpha \in \mathrm{On} \cap M[G]$. There is some $\tau \in M^{\mathbb{P}}$ so that $\tau[G] = \alpha$. Then $\alpha = \mathrm{rk}(\alpha) = \mathrm{rk}(\tau[G]) \leq \rho(\tau)$. Since $M$ is a model of ZFC, by absoluteness of the rank function, $\rho(\tau) \in M$. Since $M$ is transitive, $\mathrm{rk}(\tau[G])$ belongs to $M$ as well. And this is just $\alpha$. $\dashv$

Let us play around with building sets in $M[G]$ just a little bit more. Suppose for example that $\tau[G]$ and $\sigma[G]$ belong to $M[G]$, so that $\sigma, \tau \in M^{\mathbb{P}}$. Consider the name $\mathrm{up}(\sigma,\tau) = \{\langle \sigma, \mathbb{1}\rangle, \langle \tau, \mathbb{1}\rangle\}$. Then $\mathrm{up}(\sigma,\tau)[G] = \{\sigma[G], \tau[G]\}$ regardless of what $G$ we take, since $G$ always contains $\mathbb{1}$. If we define $\mathrm{op}(\sigma,\tau) = \mathrm{up}(\mathrm{up}(\sigma,\sigma), \mathrm{up}(\sigma,\tau))$ then we will always have $\mathrm{op}(\sigma,\tau)[G] = \langle \sigma[G], \tau[G]\rangle$.

At this stage, a few of the axioms of ZFC are easily verified for $M[G]$.

Lemma 10.15. *We have that $M[G]$ satisfies the axioms of extensionality, infinity, foundation, pairing, and union.*

Proof. Any transitive model satisfies extensionality, so that's done. Infinity holds since $\omega \in M \subseteq M[G]$. Foundation likewise holds, by absoluteness.

To check that $M[G]$ satisfies pairing, we must show that given $\sigma_1[G], \sigma_2[G]$ (where $\sigma_1, \sigma_2$ belong to $M^{\mathbb{P}}$) that we can find some $\tau \in M^{\mathbb{P}}$ such that $\tau[G] = \{\sigma_1[G], \sigma_2[G]\}$. What we need is precisely what $\mathrm{up}(\sigma_1, \sigma_2)$ provides.

For union, we must show given $\sigma[G] \in M[G]$ that there is a $\tau[G] \in M[G]$ such that $\bigcup \sigma[G] \subseteq \tau[G]$. Let $\tau = \{\langle \chi, \mathbb{1}\rangle : \exists \pi \in \mathrm{dom}(\sigma), \chi \in \mathrm{dom}(\pi)\}$. We claim that $\bigcup \sigma[G] \subseteq \tau[G]$. Let $x \in \bigcup \sigma[G]$. Then $x \in y$ for some $y \in \sigma[G]$. By the definition of $\sigma[G]$, $y = \pi[G]$ for some $\langle \pi, p\rangle \in \sigma$ with $p \in G$. (So $\pi \in \mathrm{dom}(\sigma)$.) Since $x \in \pi[G]$ there's $\langle \chi, p\rangle \in \pi$ with $p \in G$ such that $x = \chi[G]$. Then by definition, $\chi[G] \in \tau[G]$ as $\mathbb{1} \in G$ automatically. $\dashv$

Notice we have not used the fact that $G$ intersects dense subsets yet. Everything we've done so far could have been done just for subsets of $\mathbb{P}$ that contain $\mathbb{1}$. But such subsets can only get us so far. Let's see an example of what can go wrong if we don't require $G$ to be generic over $M$.

Let $\mathbb{P}$ be the poset of functions $p : n \times n \to 2$ for $n \in \omega$, ordered by reverse inclusion. $M$ is a countable transitive model, so $M \cap \mathrm{On}$ is a countable ordinal, say $\alpha$. Let $E$ be a well-order of $\omega$ in order-type $\alpha$. If $g : \omega \times \omega \to 2$ is the characteristic function of $E$, then we have $G = \{g \restriction n \times n : n \in \omega\}$ is a subset of $\mathbb{P}$ – indeed, it is a filter. It's not hard to see that $G$ is not generic over $M$.

Now by what we've shown already, $G \in M[G]$ and $M \cap \mathrm{On} = M[G] \cap \mathrm{On} = \alpha$. Clearly we can't have $M[G]$ a model of ZFC, though, since then we could use

$G$ to define the relation $E$ and take its transitive collapse, which is just $\alpha$. But $\alpha \notin M[G]$.

So this is one thing genericity does for us: It prevents arbitrary information about $M$ from being coded into the filter. We'll see in the next section that genericity gives a great deal more.

**10.2. The Forcing Relation.**

Definition 10.16. The **forcing language** consists of the symbols of first order logic, the binary relation symbol $\in$, and constant symbols $\tau$ for each $\tau \in M^{\mathbb{P}}$. Let $\phi(\tau_1, \ldots \tau_n)$ be formula of the forcing language, so that $\tau_1, \ldots \tau_n$ all belong to $M^{\mathbb{P}}$. Let $p \in \mathbb{P}$. We say that $p \Vdash \phi(\tau_1, \ldots \tau_n)$ (read $p$ **forces** $\varphi$) if for *every* $\mathbb{P}$-generic filter $G$ with $p \in G$ we have $M[G] \vDash \phi(\tau_1[G], \ldots \tau_n[G])$.

In order to make sense of this definition (and a few other things), let's take a breath and consider an example. Take $\mathbb{P}$ to be $\mathrm{Fn}(\omega, 2)$, the collection of finite functions whose domain is a subset of $\omega$ and which take values in $\{0, 1\}$. For each $n \in \omega$ the set $D_n = \{p \in \mathbb{P} : n \in \mathrm{dom}(p)\}$ is dense in $\mathbb{P}$, and by absoluteness belongs to $M$. Since $G$ is $\mathbb{P}$-generic, we have for each $n \in \omega$ that $D_n \cap G$ is not empty. Thus as before we can define from $G$ a function $g : \omega \to 2$ such that $g = \bigcup G$.

Once we show that $M[G]$ is a model of ZFC it will of course follow that $g \in M[G]$ since $G \in M[G]$ and $g$ is definable from $G$. But we can show this directly by devising a name $\dot{g}$ so that $\dot{g}[G] = g$. Indeed, set

$$\dot{g} = \{\langle \langle m, n\check{\rangle}, p\rangle : p(m) = n\}.$$

Then $\dot{g}[G] = \{\langle m, n\check{\rangle}[G] : p \in G \text{ and } p(m) = n\}$. Since $\langle m, n\check{\rangle}[G] = \langle m, n\rangle$, this is exactly the canonical function defined from $G$.

Let us see some examples of what $\Vdash$ means in this context. Say $p$ is the partial function with domain 3 such that $p(0) = 0, p(1) = 1, p(2) = 2$. Then, if $p \in G$ it is clear that $g(1) = 1$. In terms of forcing this is the same as saying

$$p \Vdash \dot{g}(\check{1}) = \check{1}.$$

Also notice that regardless of what $G$ contains, $g$ will always be a function from $\omega$ into 2. In other words,

$$\mathbb{1} \Vdash \dot{g} : \check{\omega} \to \check{2}.$$

The following is an important property of $\Vdash$.

Lemma 10.17. *If $p \Vdash \phi(\tau_1, \ldots, \tau_n)$ and $q \leq p$ then $q \Vdash \phi(\tau_1, \ldots, \tau_n)$.*

Proof. If $G$ is $\mathbb{P}$-generic with $q \in G$, then by definition of a filter $p \in G$. Then by definition of $p \Vdash \phi(\tau_1, \ldots, \tau_n)$, we have $M[G] \vDash \phi(\tau_1[G], \ldots, \tau_n[G])$. $\dashv$

The following theorems are the two essential tools for using forcing to prove consistency results.

Theorem 10.18 (Forcing Theorem A). *If $M[G] \vDash \phi(\tau_1[G], \ldots \tau_n[G])$ then there is a $p \in G$ such that $p \Vdash \phi(\tau_1, \ldots, \tau_n)$.*

Theorem 10.19 (Forcing Theorem B). *The relation $\Vdash$ is definable in $M$. That is, for any formula $\phi$, there's a formula $\psi$ such that for all $p \in \mathbb{P}$ and $\tau_1, \ldots \tau_n \in M^{\mathbb{P}}$ we have $M \vDash \psi(p, \tau_1, \ldots, \tau_n)$ exactly when $p \Vdash \phi(\tau_1, \ldots \tau_n)$.*

Let's take a minute to note the import of these theorems. Forcing Theorem A states that for any sentence $\varphi$ of the forcing language, one doesn't need to consult all of the filter $G$ to see that it holds in $M[G]$: It is in fact guaranteed by a single condition $p \in G$. And Forcing Theorem B states that $M$ knows when a condition $p$ guarantees $\varphi$ in this sense. So even though almost always $G \notin M$, $M$ can nonetheless "see" a lot of what's going on in the extension; and the more information the people in $M$ have about $G$ (in the sense of which conditions belong to $G$), the more statements they can accurately predict will hold in $M[G]$.

We will prove these theorems in the next section. For now we use them to finish proving Theorem 10.10. As a warm-up, we give a first example of an argument making use of Forcing Theorem A.

LEMMA 10.20. *If $p \Vdash (\exists x \in \sigma)\phi(x, \tau_1, \dots, \tau_n)$ then there is some $\pi \in \mathrm{dom}(\sigma)$ and some $q \leq p$ so that $q \Vdash \pi \in \sigma \wedge \phi(\pi, \tau_1, \dots, \tau_n)$.*

PROOF. Let $G$ be $\mathbb{P}$-generic with $p \in G$. Since $p \Vdash (\exists x \in \sigma)\phi(x, \tau_1, \dots, \tau_n)$, by definition of $\Vdash$, $M[G] \vDash (\exists x \in \sigma[G])\phi(x, \tau_1[G], \dots, \tau_n[G])$. So take $\pi[G] \in \sigma[G]$ such that we have $M[G] \vDash \phi(\pi[G], \tau_1[G], \dots, \tau_n[G])$. By definition of $\sigma[G]$ we may assume that $\pi \in \mathrm{dom}(\sigma)$. Now by Forcing Theorem A there is an $r \in G$ so that $r \Vdash \pi \in \sigma \wedge \phi(\pi, \tau_1, \dots, \tau_n)$. Since $r$ and $p$ both belong to $G$, by definition of a filter there is some $q \in G$ with $q \leq p, r$. By Lemma 10.17 we have $q \Vdash \pi \in \sigma \wedge \phi(\pi, \tau_1, \dots, \tau_n)$. ⊣

LEMMA 10.21. *$M[G]$ satisfies the Comprehension Axiom.*

PROOF. Let $\phi(x, v, y_1, \dots y_n)$ be a formula in the language of set theory, and let $\sigma[G], \tau_1[G], \dots, \tau_n[G]$ belong to $M[G]$. We must show that the set

$$X = \{a \in \sigma[G] : M[G] \vDash \phi(a, \sigma[G], \tau_1[G], \dots \tau_n[G])\}$$

belongs to $M[G]$. In other words, we must devise a name for the set. Define

$$\rho = \{\langle \pi, p\rangle : \pi \in \mathrm{dom}(\sigma),\ p \in \mathbb{P},\ p \Vdash (\pi \in \sigma \wedge \phi(\pi, \sigma, \tau_1, \dots, \tau_n))\}.$$

By Forcing Theorem B (and Comprehension applied within $M$), this set actually belongs to $M$, being defined from notions definable in $M$. So $\rho \in M^{\mathbb{P}}$. Let us check that $\rho[G] = X$. Suppose $\pi[G] \in \rho[G]$. By definition of our evaluation of names under $G$, there is some $p \in G$ such that $p \Vdash \pi \in \sigma \wedge \phi(\pi, \sigma, \tau_1, \dots, \tau_n)$. By definition of $\Vdash$ then we have that $\pi[G] \in \sigma[G]$, and $M[G] \vDash \phi(\pi[G], \sigma[G], \tau_1[G], \dots \tau_n[G])$. So indeed $\pi[G] \in X$.

Going the other way, suppose that $a \in X$. Then $a \in \sigma[G]$, and so by definition of $\sigma[G]$ there must be some $\pi$ in $\mathrm{dom}(\sigma)$ such that $a = \pi[G]$. Also, because $a \in X$, by definition of $X$ we have that $M[G] \vDash \phi(\pi[G], \sigma[G], \tau_1[G], \dots \tau_n[G])$. Applying Forcing Theorem A tells us that there is some $p \in G$ such that $p \Vdash \pi \in \sigma \wedge \phi(\pi, \sigma, \tau_1, \dots, \tau_n)$. So by definition of $\rho$, $\langle \pi, p\rangle \in \rho$. Since $p \in G$, $\pi[G] \in \rho[G]$. ⊣

Notice how in the above proof the $\rho$ we constructed does not at all depend on what $G$ actually is. This is one of the central tenets of forcing: People living in $M$ can reason out every aspect of $M[G]$ if they just imagined that some generic $G$ existed.

LEMMA 10.22. *$M[G]$ satisfies the Replacement Axiom.*

PROOF. Suppose $\phi(u, v, r, z_1, \dots z_n)$ is a fixed formula in the language of set theory, and let $\sigma[G], \tau_1[G], \dots \tau_n[G]$ be such that for every $x \in \sigma[G]$ there is a unique $y$ in $M[G]$ so that $M[G] \vDash \phi(x, y, \sigma[G], \tau_1[G], \ldots, \tau_n[G])$. We have to construct a name $\rho \in M^{\mathbb{P}}$ which witnesses replacement, i.e. so that

$$(\forall x \in \sigma[G])(\exists y \in \rho[G]) M[G] \vDash \phi(x, y, \sigma[G], \tau_1[G], \ldots, \tau_n[G]).$$

Apply Replacement within $M$ together with Forcing Theorem B to find a set $S \in M$ (with $S \subseteq M^{\mathbb{P}}$) such that

$$(\forall \pi \in \mathrm{dom}(\sigma))(\forall p \in \mathbb{P})[(\exists \mu \in M^{\mathbb{P}}(p \Vdash \phi(\pi, \mu, \tau_1, \ldots, \tau_n)) \\ \rightarrow \exists \mu \in S(p \Vdash \phi(\pi, \mu, \tau_1, \ldots, \tau_n))].$$

Actually, what we are applying here is a stronger-looking version of replacement (known as *Collection*) where we do not require the $\mu$ to be unique. In fact this is implied by replacement (and the other axioms of ZFC); this was one of the exercises in the problem sessions. So we apply it without too much guilt. Now let $\rho$ be $S \times \{\mathbb{1}\}$.

Let us see that $\rho[G]$ is as desired. We have $\rho[G] = \{\mu[G] : \mu \in S\}$. Suppose $\pi[G] \in \sigma[G]$. By hypothesis there is a $\nu[G] \in M[G]$ with

$$M[G] \vDash \phi(\pi[G], \nu[G], \sigma[G], \tau_1[G], \ldots, \tau_n[G]).$$

By Forcing Theorem A there is a $p \in G$ such that $p \Vdash \phi(\pi, \nu, \sigma, \tau_1, \ldots, \tau_n)$. So by definition of $S$ we can find $\mu$ in $S$ so that $p \Vdash \phi(\pi, \mu, \sigma, \tau_1, \ldots, \tau_n)$. Then $\mu[G] \in \rho[G]$, and since $p \in G$, applying the definition of $\Vdash$ gives

$$M[G] \vDash \phi(\pi[G], \mu[G], \sigma[G], \tau_1[G], \ldots, \tau_n[G]).$$

$\dashv$

LEMMA 10.23. *$M[G]$ satisfies the Power Set Axiom.*

PROOF. Let $\sigma[G] \in M[G]$. We must find some $\rho \in M^{\mathbb{P}}$ such that $\rho[G]$ contains all of the subsets of $\sigma[G]$ that belong to $M[G]$. Let $S = \{\tau \in M^{\mathbb{P}} : \mathrm{dom}(\tau) \subseteq \mathrm{dom}(\sigma)\}$. Notice that $S$ is actually equal to $\mathcal{P}(\mathrm{dom}(\sigma) \times \mathbb{P})$, relativized to $M$. Let $\rho = S \times \{\mathbb{1}\}$.

Let us check that $\rho$ is as desired. Let $\mu[G] \in M[G]$ with $\mu[G] \subseteq \sigma[G]$. Let

$$\tau = \{\langle \pi, p \rangle : \pi \in \mathrm{dom}(\sigma) \text{ and } p \Vdash \pi \in \mu\}.$$

Then $\tau \in S$, and so $\tau[G] \in \rho[G]$. Let us check that $\tau[G] = \mu[G]$. If $\pi[G] \in \tau[G]$, then by definition of $\tau$ there is a $p \in G$ so that $p \Vdash \pi \in \mu$ and so by definition of $\Vdash$ we have $\pi[G] \in \mu[G]$. Going the other way, if $\pi[G] \in \mu[G]$ then by Forcing Theorem A there is a $p \in G$ such that $p \Vdash \pi \in \mu$. Then $\langle \pi, p \rangle \in \tau$ and $\pi[G] \in \tau[G]$. $\dashv$

LEMMA 10.24. *$M[G]$ satisfies the Axiom of Choice.*

PROOF. It is enough to show that in $M[G]$, for every set $x$, there is some ordinal $\alpha$ and some function $f$ so that $x$ is included in the range of $f$. For then, we can define an injection $g : x \to \alpha$ by letting $g(z)$ be the least element of $f^{-1}[\{z\}]$. Such an injection easily allows us to well-order $x$.

So let $\sigma[G] \in M[G]$. Since the Axiom of Choice holds in $M$, we can well-order $\mathrm{dom}(\sigma)$, say we enumerate by $\{\pi_\gamma : \gamma < \alpha\}$. Let $\tau = \{\mathrm{op}(\check{\gamma}, \pi_\gamma) : \gamma < \alpha\} \times \{\mathbb{1}\}$. Then $\tau[G] = \{\langle \gamma, \pi_\gamma[G] \rangle : \gamma < \alpha\}$ belongs to $M[G]$, a function as desired. $\dashv$

This gives us all of the axioms of ZFC, and so Theorem 10.10 is proved.

**10.3. Proving the Forcing Theorems.** Recall the statements of the Forcing Theorems.

Theorem 10.25 (Forcing Theorem A). *$M[G] \models \varphi(\tau_1[G], \dots, \tau_n[G])$ if and only if $(\exists p \in G)$ such that $p \Vdash \varphi(\tau_1, \dots, \tau_n)$.*

Theorem 10.26 (Forcing Theorem B). *The relation $\Vdash$ is definable in $M$. That is, for a fixed $\varphi$, the class $\{\langle p, \tau_1, \dots, \tau_n\rangle \mid p \Vdash \varphi(\tau_1, \dots, \tau_n)\}$ is definable in $M$.*

We prove the forcing theorems by defining a version of the forcing relation which makes no reference to $M$-generic filters, and so makes sense in $M$. We call this relation $\Vdash^*$. Then Theorem B is automatically satisfied for $\Vdash^*$. We prove Theorem A for $\Vdash^*$, and finish by showing that $\Vdash^*$ and $\Vdash$ are very nearly equivalent. This almost equivalence is enough to give Theorems A and B.

The definition of $\Vdash^*$ is by induction on formula complexity. The definition of $\Vdash^*$ for atomic formulae, e.g. $\sigma \in \tau$, will itself be by induction on name-rank of the names $\sigma, \tau$, that is, on pairs $\{\rho(\sigma), \rho(\tau)\}$. We therefore define $\prec$ to be the lexicographical ordering on pairs of ordinals,

$$\begin{aligned}\{\alpha, \beta\} \prec \{\gamma, \delta\} \iff & \min\{\alpha, \beta\} < \min\{\gamma, \delta\} \\ & \text{or } \min\{\alpha, \beta\} = \min\{\gamma, \delta\} \text{ and } \max\{\alpha, \beta\} < \max\{\gamma, \delta\}.\end{aligned}$$

Clearly this relation on pairs is well-founded.

Suppose we are given names $\sigma, \tau \in M^{\mathbb{P}}$ so that the relations $p \Vdash^* \sigma' \in \tau'$, $p \Vdash^* \sigma' \neq \tau'$, $p \Vdash^* \neg(\sigma' \in \tau')$, and $p \Vdash^* \neg(\sigma' \neq \tau')$ have already been defined for all $p \in \mathbb{P}$ and names $\sigma', \tau'$ with $\{\rho(\sigma'), \rho(\tau')\} \prec \{\rho(\sigma), \rho(\tau)\}$. Define

$$\begin{aligned}p \Vdash^* \sigma \in \tau \iff & (\exists q \geq p)(\exists \theta) \text{ such that } \langle \theta, q\rangle \in \tau \text{ and } p \Vdash^* \neg(\theta \neq \sigma);\\ p \Vdash^* \sigma \neq \tau \iff & (\exists q \geq p)(\exists \theta) \text{ such that either}\\ & \langle \theta, q\rangle \in \sigma \text{ and } p \Vdash^* \neg(\theta \in \tau), \text{ or}\\ & \langle \theta, q\rangle \in \tau \text{ and } p \Vdash^* \neg(\theta \in \sigma);\\ p \Vdash^* \neg\varphi \iff & (\forall q \leq p) q \nVdash^* \varphi.\end{aligned}$$

Note that in each case the name-rank of $\theta$ is less than one of $\sigma, \tau$ so the inductive definition makes sense.

To finish the definition of $\Vdash^*$ we just inductively define

$$\begin{aligned}p \Vdash^* \varphi \vee \psi \iff & p \Vdash^* \varphi \text{ or } p \Vdash^* \psi;\\ p \Vdash^* \exists x\varphi \iff & \text{there is some } \tau \text{ such that } p \Vdash^* \varphi(\tau)\end{aligned}$$

and continue to use the same definition of negation given above.

We have really only defined the relation $\Vdash^*$ for sentences built up using names and the symbols $\in, \neq, \neg, \vee$, and $\exists$. Let us say such a sentence is in the **forcing$^*$ language**. Whenever we discuss the relation $\Vdash^*$, we restrict ourselves to formulas in this language. Since every formula in the full forcing language is clearly equivalent to one in the forcing$^*$ language, this will be sufficient.

Proposition 10.27. *If $p \Vdash^* \varphi$ and $r \leq p$, then $r \Vdash^* \varphi$.*

Proof. The proof is by induction on formula complexity. We prove it first for atomic forcing$^*$ formulas and their negations; this step will be by induction on pairs of name-ranks.

So suppose $r \le p$ where $p \Vdash^* \sigma \in \tau$, and inductively, that we have proved the proposition for formulas of the form $\sigma' \in \tau'$, $\sigma' \neq \tau'$, $\neg(\sigma' \in \tau')$, $\neg(\sigma' \neq \tau')$, whenever $\{\rho(\sigma'), \rho(\tau')\} \prec \{\rho(\sigma), \rho(\tau)\}$. By definition of $\Vdash^*$,

$$(\exists q \geq p)(\exists \theta)\langle \theta, q\rangle \in \tau \text{ and } p \Vdash^* \neg(\theta \neq \sigma).$$

Since $\rho(\theta) < \rho(\tau)$, we have by inductive hypothesis that $r \Vdash^* \neg(\theta \neq \sigma)$. Since $q \geq r$, we clearly have $r \Vdash^* \sigma \in \tau$ as needed.

The proof for $\sigma \neq \tau$ is similar; and the result is immediate by definition of $\Vdash^*$ for formulas built from $\neg$. The cases for $\vee$ and $\exists$ are straightforward. $\dashv$

We want to prove Theorem A for $\Vdash^*$. We make use of a non-intuitive sublemma.

Lemma 10.28 (Non-intuitive sublemma). *Suppose Theorem A holds for $\varphi$. Then for all $M$-generic $G$,*

$$(\exists q \in G)(\exists \sigma)\langle \sigma, q\rangle \in \tau \text{ and } M[G] \models \varphi$$

*if and only if*

$$(\exists p \in G)(\exists q \geq p)(\exists \sigma) \text{ such that } \langle \sigma, q\rangle \in \tau \text{ and } p \Vdash^* \varphi.$$

Proof. For the forward direction, if $q \in G$ and $\sigma$ are such that $\langle \sigma, q\rangle \in \tau$ and $M[G] \models \varphi$, then by Theorem A, let $p \in G$ be such that $p \Vdash^* \varphi$. Since $G$ is a filter and by the last proposition, we may assume $p \le q$, just what we need.

Conversely, if $p \in G$ and $q \geq p$ with $\langle \sigma, q\rangle \in \tau$ and $p \Vdash^* \varphi$, then by Theorem A, $M[G] \models \varphi$, and by upwards closure of $G$, we have $q \in G$. $\dashv$

Proof of Theorem A for $\Vdash^*$. As mentioned above, we consider only formulas $\varphi$ of the forcing* language. We prove it first for atomic formulas by induction on pairs of ranks. Assume for some $\sigma, \tau$ we have proved Theorem A for all statements of the form $\sigma' \neq \tau'$, $\sigma' \in \tau'$ and their negations, when $\{\rho(\sigma'), \rho(\tau')\} \prec \{\rho(\sigma), \rho(\tau)\}$.

$$\begin{aligned}(\exists p \in G) p \Vdash^* \sigma \in \tau &\iff (\exists p \in G)(\exists q \geq p)(\exists \theta)\langle \theta, q\rangle \in \tau \text{ and } p \Vdash^* \neg(\theta \neq \sigma)\\ &\iff (\exists q \in G)(\exists \theta)\langle \theta, q\rangle \in \tau \text{ and } M[G] \models \theta[G] = \sigma[G]\\ &\iff M[G] \models \sigma[G] \in \tau[G].\end{aligned}$$

Here the first equivalence is by the definition of $\Vdash^*$; the second is by the non-intuitive sublemma plus the inductive hypothesis; the third by definition of $\tau[G]$.

$$\begin{aligned}(\exists p \in G) p \Vdash^* \sigma \neq \tau \iff & (\exists p \in G)(\exists q \geq p)(\exists \theta) \text{ either}\\ & \langle \theta, q\rangle \in \tau \text{ and } p \Vdash^* \neg(\theta \in \sigma), \text{ or}\\ & \langle \theta, q\rangle \in \sigma \text{ and } p \Vdash^* \neg(\theta \in \tau)\\ \iff & (\exists q \in G)(\exists \theta)\langle \theta, q\rangle \in \tau \text{ and } M[G] \models \theta[G] \notin \sigma[G], \text{ or}\\ & (\exists q \in G)(\exists \theta)\langle \theta, q\rangle \in \sigma \text{ and } M[G] \models \theta[G] \notin \tau[G]\\ \iff & M[G] \models \sigma[G] \neq \tau[G].\end{aligned}$$

First equivalence by definition of $\Vdash^*$, and the second by the non-intuitive sublemma and inductive hypothesis.

Claim. $(\exists p \in G) p \Vdash^* \neg\varphi$ *if and only if it is not the case that* $(\exists p \in G) p \Vdash^* \varphi$.

Proof. It's enough to show for all $\varphi$ and $G$ that exactly one of $(\exists p \in G)p \Vdash^* \varphi$ or $(\exists p \in G)p \Vdash^* \neg\varphi$ holds. At least one holds, since using the definability of $\Vdash^*$,

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

belongs to $M$, and is dense. So $G \cap D \neq \varnothing$ is enough.

Next, at most one can hold: if $p \Vdash^* \varphi$, $q \Vdash^* \neg\varphi$ with $p, q \in G$, then let $r \in G$ with $r \leq p, q$. Then $r \Vdash^* \varphi$ and $r \Vdash^* \neg\varphi$; but this contradicts the definition of $\Vdash^*$ and negation. ⊣

So

$$\begin{aligned}(\exists p \in G)p \Vdash^* \neg\varphi &\iff \text{not } (\exists p \in G)p \Vdash^* \varphi\\ &\iff \text{not } M[G] \models \varphi\\ &\iff M[G] \models \neg\varphi.\end{aligned}$$

The first equivalence by the claim; the second by inductive hypothesis.

$$\begin{aligned}(\exists p \in G)p \Vdash^* \varphi \vee \psi &\iff (\exists p \in G)p \Vdash^* \varphi \text{ or } p \Vdash^* \psi\\ &\iff (\exists p \in G)p \Vdash^* \varphi \text{ or } (\exists p \in G)p \Vdash^* \psi\\ &\iff M[G] \models \varphi \text{ or } M[G] \models \psi\\ &\iff M[G] \models \varphi \vee \psi.\end{aligned}$$

The third equivalence by inductive hypothesis.

$$\begin{aligned}(\exists p \in G)p \Vdash^* \exists x\varphi &\iff (\exists p \in G) \text{ for some } \tau \in M^{\mathbb{P}},\ p \Vdash^* \varphi(\tau)\\ &\iff \text{for some } \tau,\ M[G] \models \varphi(\tau[G])\\ &\iff M[G] \models \exists x\varphi(x).\end{aligned}$$

This completes the proof of Theorem A for $\Vdash^*$. ⊣

Claim. *For all $\varphi$, $p \Vdash \varphi$ iff $p \Vdash^* \neg\neg\varphi$.*

Proof. Note once again that $\varphi$ is a formula in the forcing* language.

For the forward direction, assume $p \Vdash \varphi$, and $p \nVdash^* \neg\neg\varphi$. So $(\exists q \leq p)q \Vdash^* \neg\varphi$. Let $G$ be $M$-generic with $q \in G$. Note $p \in G$. So $M[G] \models \varphi$, but $M[G] \models \neg\varphi$ since $q \in G$ by Theorem A for $\Vdash^*$.

For the converse, assume $p \Vdash^* \neg\neg\varphi$, and let $G$ be $M$-generic with $p \in G$. By Theorem A for $\Vdash^*$, $M[G] \models \neg\neg\varphi$. So $M[G] \models \varphi$. ⊣

Note this proves the forcing theorems for the relation $\Vdash$. For we may let $\varphi \mapsto \varphi^*$ be some simple translation of formulas of the forcing language into equivalent ones of the forcing* language; for example, that induced by $(\sigma = \tau)^* \equiv \neg(\sigma \neq \tau)$, $(\forall x\varphi(x))^* \equiv \neg\exists x\neg(\varphi(x))^*$, $(\varphi \wedge \psi)* \equiv \neg(\neg\varphi^* \vee \neg\psi^*)$. Then $p \Vdash \varphi$ if and only if $p \Vdash \varphi^*$ if and only if $p \Vdash^* \neg\neg\varphi^*$, and this last relation is definable in $M$. Taking this to be our official definition of $\Vdash$, we have Theorem B immediately, and Theorem A follows from Theorem A for $\Vdash^*$.

**§11. The independence of CH.** As advertised we will prove the independence of CH from the axioms of ZFC. As we saw when we discussed formal proofs and model theory, it is enough to construct two models of ZFC, one in which CH holds and the other in which it fails. Of course, to do so we will need to assume the consistency of ZFC in addition to the axioms of ZFC, since otherwise there may not even be a model of ZFC to begin with; thus the independence of CH is a *relative consistency result*, in the sense that if ZFC is consistent, then so are each of the theories ZFC + CH and ZFC + ¬ CH.

In fact, we use a bit more than consistency: We will assume that there is a *transitive* model of ZFC, which a bit stronger than just the existence of a model of ZFC. This assumption can be done away with, however, by dealing with large enough finite fragments of ZFC and using the reflection theorem.

For example, suppose towards a contradiction ZFC $\vdash$ CH; then there is some finite fragment $T$ of ZFC so that $T \vdash$ CH. Using the forcing theorems, we know there is a finite theory $T'$ so that whenever $M$ is a countable transitive model of $T'$, $\mathbb{P} \in M$, and $G$ is $\mathbb{P}$-generic over $M$, then $M[G]$ satisfies $T$ (essentially, $T'$ needs to be large enough to ensure existence of the appropriate names for objects and that the needed instances of the forcing theorems hold in $M$). Furthermore (as we see later) an appropriate choice of poset $\mathbb{P}$ will ensure $M[G] \models \neg$ CH. But then by the reflection and Mostowski Collapse theorems, there is a countable transitive model $M$ of $T'$, hence also a model $M[G]$ of $T + \neg$ CH, contradicting our choice of $T$.

So in all of our forcing arguments we will just work with a transitive model of full ZFC, because we know that there is a standard way to do without it.

One more note on a common theme in forcing arguments. In general it is a bad idea to collapse $\omega_1$. What is meant by this is we do not want to pass to a generic extension $M[G]$ in which there is a function $f : \omega \to \omega_1^M$ which is surjective. From the point of view of such an extension $M[G]$, the $\omega_1$ in $M$ is a countable ordinal, and this is precisely the sort of disturbance to the universe of $M$ that we wish to avoid.

We will see two methods for arguing that $\omega_1$ is not collapsed. The key idea is to prove some property of the poset used in forcing. The first idea which we have already seen is the notion of *chain condition*. The second idea which we have not yet seen is the notion of *closure*.

**11.1. The consistency of CH.** We wish to construct a model of ZFC + CH by forcing. Given a countable transitive model $M$ of ZFC we describe a poset $\mathbb{P}$ such that whenever $G$ is $\mathbb{P}$-generic, $\omega_1^M = \omega_1^{M[G]}$ and $M[G] \vDash$ CH. The poset is easy to describe. We let $\mathbb{P} = \{p \mid p : \alpha \to 2 \text{ for some countable ordinal } \alpha\}$ ordered by extension, i.e. $p_1 \leq p_2$ if and only if $p_1 \supseteq p_2$.

To show that $\omega_1$ is preserved we develop the notion of closure of a poset.

Definition 11.1. Let $\mathbb{P}$ be a poset. $\mathbb{P}$ is **countably closed** if for every sequence of elements $\langle p_n \mid n < \omega \rangle$ of $\mathbb{P}$ such that $p_{n+1} \leq p_n$ for all $n$, there is $p \in \mathbb{P}$ such that $p \leq p_n$ for all $n$.

It is clear from the definition of $\mathbb{P}$ that it is countably closed; we just take the union of the conditions.

Lemma 11.2. *If $\mathbb{P}$ is a countably closed poset and $G$ is $\mathbb{P}$-generic over $M$, then* $\omega_1^M = \omega_1^{M[G]}$.

Proof. Assume for a contradiction that there is some $p \in \mathbb{P}$ and a $\mathbb{P}$-name $\dot{f}$ which is forced by $p$ to be a function from $\omega$ onto $\omega_1^M$. Let $n < \omega$; we claim that $D_n = \{p \in \mathbb{P} \mid p \Vdash \dot{f}(n) = \check{\alpha} \text{ for some } \alpha < \omega_1\}$ is dense in $\mathbb{P}$. Let $p \in \mathbb{P}$ and let $G$ be $\mathbb{P}$-generic with $p \in G$. In $M[G]$ there is an ordinal $\alpha < \omega_1^M$ such that $\dot{f}[G](n) = \alpha$. Choose $p' \in G$ forcing that $\dot{f}(n) = \check{\alpha}$. Since $G$ is a filter we can choose $p'' \leq p', p$. Clearly $p'' \in D_n$.

By induction build a decreasing sequence of elements of $\mathbb{P}$. Let $p_0 = p$. Given $p_n$ let $p_{n+1} \in D_n$ with $p_{n+1} \leq p_n$ and record the value $\alpha_n$ witnessing $p_{n+1} \in D_n$. Let $p_\omega \leq p_n$ for all $n$, by the countable closure of $\mathbb{P}$. Let $\alpha = \sup \alpha_n$. Let $H$ be $\mathbb{P}$-generic over $M$. Then in $M[H]$, the range of $\dot{f}[H]$ is bounded by $\alpha$; but this is a contradiction since it was supposed to be forced by $p \geq p_\omega \in H$ that $f$ was onto. $\dashv$

A similar argument shows the following.

Lemma 11.3. *If $\mathbb{P}$ is countably closed, then whenever $G$ is $\mathbb{P}$-generic over $M$,* $\mathcal{P}(\omega)^M = \mathcal{P}(\omega)^{M[G]}$.

Proof. Exercise. $\dashv$

There is a general phenomenon occurring in the previous proof. Suppose that $\dot{x}$ is a $\mathbb{P}$-name for an element of $M$. We say that a condition $p$ **decides the value of** $\dot{x}$ if it forces $\dot{x} = \check{y}$ for some $y \in M$. The collection of conditions which decide the value of such an $\dot{x}$ is *always* dense.

Next we show the following.

Lemma 11.4. *If $G$ is $\mathbb{P}$-generic where $\mathbb{P} = \{p \mid p : \alpha \to 2$ for some $\alpha < \omega_1\}$ ordered by extension, then* $M[G] \vDash \mathrm{CH}$.

Proof. Using the generic object $G$ we define a list of $\omega_1^{M[G]} = \omega_1^M$-many subsets of $\omega$. We then do a density argument to show that this list comprises all subsets of $\omega$ in $M[G]$. Work for the moment in $M[G]$. Let $g = \bigcup G$. Note that $g$ is a function from $\omega_1$ to 2. We define a collection of subsets of $\omega$, $\{x_\alpha \mid \alpha < \omega_1\}$, by $n \in x_\alpha$ if and only if $g(\omega \cdot \alpha + n) = 1$.

By the previous lemma it is enough to show that for every $x \in (\mathcal{P}(\omega))^M$, there is an $\alpha$ such that $x = x_\alpha$. For this we will do a density argument. Work in $M$ and let $x \subseteq \omega$. We claim $D_x = \{p \in \mathbb{P} \mid \text{there is } \alpha < \omega_1 \text{ such that for all } n,\ \chi_x(n) = p(\omega \cdot \alpha + n)\}$ is dense. (Here $\chi_x$ is the characteristic function of $x$.) Let $p \in \mathbb{P}$. Let $\mathrm{dom}(p) = \beta$. Let $\alpha > \beta$. It follows that for all $n < \omega$, $\omega \cdot \alpha + n \notin \mathrm{dom}(p)$. So we extend $p$ to a condition $p'$ in $D_x$ where $\alpha$ is the witness.

It follows that in $M[G]$ the map $\alpha \mapsto x_\alpha$ is a surjection from $\omega_1$ onto $\mathcal{P}(\omega)$. $\dashv$

So we have proved that $\mathrm{ZFC} + \mathrm{CH}$ is consistent.

**11.2. The consistency of $\neg$CH.** In this section we prove that there is a poset $\mathbb{P}$ such that whenever $M \models$ CH and $G$ is $\mathbb{P}$-generic, $\omega_1^M = \omega_1^{M[G]}$ and $M[G] \models 2^\omega = \omega_2$. Again the poset is easy to describe. We let $\mathbb{P} = \{p \mid$ there is $x \subseteq \omega_2$ finite such that $p : x \to 2\}$ ordered by extension.

We will show that this forcing preserves all cardinals by showing that it has the *countable chain condition.* Before showing that $\mathbb{P}$ is ccc, we show that any forcing which has the ccc preserves all cardinals.

LEMMA 11.5. *Suppose that $\mathbb{P}$ is a ccc poset. Whenever $G$ is $\mathbb{P}$-generic over $M$ and $\kappa$ is an ordinal, $M \models$ "$\kappa$ is a cardinal" if and only if $M[G] \models$ "$\kappa$ is a cardinal".*

PROOF. Let $G$ be $\mathbb{P}$-generic over $M$ and $\kappa$ be an ordinal. Notice that the reverse direction is clear. So suppose that $M \models \kappa$ is a cardinal, but $M[G] \models \kappa$ is not a cardinal. Then there is a name $\dot{f}$ such that $\dot{f}[G]$ is a surjection from some $\alpha < \kappa$ onto $\kappa$. We fix a condition $p_0 \in G$ forcing this.

For every $\beta < \alpha$, the collection $D_\beta = \{p \in \mathbb{P} \mid p$ decides $\dot{f}(\beta)\}$ is dense below $p_0$, since $\dot{f}(\beta)$ is forced by $p_0$ to be an ordinal. So if we choose $A_\beta \subseteq D_\beta$ a maximal antichain, then there is a countable set of ordinals $X_\beta$ such that whenever $p \in A_\beta$ there is an ordinal $\gamma \in X_\beta$ such that $p \Vdash \dot{f}(\beta) = \gamma$. But this means that $p_0 \Vdash \operatorname{ran}(\dot{f}) \subseteq \bigcup_{\beta<\alpha} X_\beta$ and the right hand union has size at most $\omega \cdot |\alpha| < \kappa$, contradicting that $p_0$ forces that $\dot{f}$ is onto $\kappa$. ⊣

We now recall some homework problems which will be used in showing that $\mathbb{P}$ is ccc.

Let $\kappa$ be a regular cardinal.

DEFINITION 11.6. A set $C \subseteq \kappa$ is **club** if it is unbounded in $\kappa$ and for all $\alpha < \kappa$ if $C \cap \alpha$ is unbounded in $\alpha$, then $\alpha \in C$.

LEMMA 11.7. *The collection of club subsets of $\kappa$ form a $\kappa$-complete filter.*

Recall that a filter is $\kappa$-complete if it is closed under intersections of size less than $\kappa$.

DEFINITION 11.8. A set $S \subseteq \kappa$ is **stationary** if for every club $C$ in $\kappa$, $S \cap C \neq \varnothing$.

LEMMA 11.9. *Let $S$ be a stationary set. If $F : S \to \kappa$ is a function such that $F(\alpha) < \alpha$ for all $\alpha \in S$, then there is a stationary $S' \subseteq S$ on which $F$ is constant.*

LEMMA 11.10. *If $S$ is stationary in $\kappa$, then $S$ is unbounded in $\kappa$.*

We are now ready to prove the key lemma which will be used in the proof that $\mathbb{P}$ is ccc. We prove a weak version of this lemma which is strong enough for our application. The proof we have chosen is one that generalizes to more complicated versions of the lemma.

LEMMA 11.11 (The $\Delta$-system lemma). *Let $X$ be a set of size $\omega_1$ and $\{x_\alpha \mid \alpha < \omega_1\}$ be a collection of finite subsets of $X$. There are an unbounded $I \subseteq \omega_1$ and a finite $r \subseteq X$ such that for all $\alpha, \beta \in I$, $x_\alpha \cap x_\beta = r$.*

The collection of sets $\{x_\alpha \mid \alpha \in I\}$ forms a $\Delta$-system with root $r$.

Proof. First note that it is enough to show the lemma in the case $X = \omega_1$. Since for an arbitrary $X$ of size $\omega_1$ we can use a bijection with $\omega_1$ to copy the problem. So let $\{x_\alpha \mid \alpha < \omega_1\}$ be a collection of finite subsets of $\omega_1$.

We define a function $F : \mathrm{Lim}(\omega_1) \to \omega_1$ by $F(\alpha) = \max(x_\alpha \cap \alpha)$. Since each $\alpha$ is finite, we have $F(\alpha) < \alpha$ for all limit ordinals $\alpha$. It follows that there are $S \subseteq \mathrm{Lim}(\omega_1)$ and $\delta < \omega_1$ such that for all $\alpha \in S$, $F(\alpha) = \delta$. Since there are only countably many finite subsets of $\delta$, we can choose $J \subseteq S$ unbounded and a finite $r \subseteq \delta$ such that for all $\alpha \in J$, $x_\alpha \cap \delta = r$.

Finally we construct $I$ an unbounded subset of $J$ by recursion. Suppose that we have constructed an enumeration $\gamma_\alpha$ of $I$ for all $\alpha < \beta$. The set $\bigcup_{\alpha<\beta} x_{\gamma_\alpha}$ is countable and hence bounded in $\omega_1$ by some ordinal $\eta < \omega_1$. Let $\gamma_\beta$ be the least member of $J$ greater than $\eta$.

Now we claim that $\{x_\alpha \mid \alpha \in I\}$ forms a $\Delta$-system with root $r$. Let $\alpha < \beta < \omega_1$. We will show that $x_{\gamma_\alpha} \cap x_{\gamma_\beta} = r$. By the choice of $\gamma_\beta$, $x_{\gamma_\alpha} \subseteq \gamma_\beta$. So $x_{\gamma_\alpha} \cap x_{\gamma_\beta} = x_{\gamma_\alpha} \cap x_{\gamma_\beta} \cap \gamma_\beta$. But $x_{\gamma_\beta} \cap \gamma_\beta = x_{\gamma_\beta} \cap \delta = r$. So we are done. $\dashv$

Recall the definition of $\mathbb{P}$. $\mathbb{P} = \{p \mid \text{there is a finite } x \subseteq \omega_2 \text{ such that } p : x \to 2\}$ ordered by extension.

Lemma 11.12. *$\mathbb{P}$ has the $\aleph_1$-Knaster property.*

Proof. Let $\{p_\alpha \mid \alpha < \omega_1\}$ be a sequence of conditions in $\mathbb{P}$. For each $\alpha < \omega_1$, let $x_\alpha = \mathrm{dom}(p_\alpha)$ and let $X = \bigcup_{\alpha<\omega_1} x_\alpha$. By the $\Delta$-system lemma, there are an unbounded $I \subseteq \omega_1$ and a finite set $r \subseteq X$ such that $\{x_\alpha \mid \alpha \in I\}$ forms a $\Delta$-system with root $r$.

Since there are only finitely many functions from $r$ to 2, we can assume that for all $\alpha, \beta \in I$, $p_\alpha \upharpoonright r = p_\beta \upharpoonright r$. It follows that for $\alpha, \beta \in I$, $p_\alpha \cup p_\beta$ is a condition, so we are done. $\dashv$

Lemma 11.13. *If $G$ is $\mathbb{P}$-generic over $M$, then $M[G] \vDash 2^\omega \geq \omega_2$.*

Proof. The argument is a straightforward density argument. Work in $M[G]$ and let $g = \bigcup G$. We define a collection of functions $f_\alpha : \omega \to 2$ for $\alpha < \omega_2$ by $f_\alpha(n) = 1$ if and only if $g(\omega \cdot \alpha + n) = 1$ (note that $\omega_2^M = \omega_2^{M[G]}$ since the forcing is ccc). We claim that for each pair $\alpha < \beta < \omega_2$, the set $D_{\alpha,\beta} = \{p \mid \text{there is } n \text{ such that } p(\omega \cdot \alpha + n) \neq p(\omega \cdot \beta + n)\}$ is dense. This is an argument that we have seen many times. Given a $p \in \mathbb{P}$, there is $n < \omega$ such that $\omega \cdot \alpha + n \notin \mathrm{dom}(p)$, so we are free to extend $p$ to $p' \in D_{\alpha,\beta}$. Since $G \cap D_{\alpha,\beta} \neq \varnothing$ for all $\alpha < \beta < \omega_2$, the collection $\{f_\alpha \mid \alpha < \omega_2\}$ is a set of $\omega_2$ many functions in ${}^\omega 2$. So $M[G] \vDash 2^\omega \geq \omega_2$. $\dashv$

We have shown that CH fails in $M[G]$ whenever $G$ is $\mathbb{P}$-generic over $M$. We conclude by computing the value of $2^\omega$ in the extension.

Lemma 11.14. *Let $\mathbb{P}$ be a poset and let $\dot{f}$ be a $\mathbb{P}$-name for a function from $\omega$ to 2. There is a sequence in $M$ of functions $h_n : A_n \to 2$ for $n < \omega$, where each $A_n$ is a maximal antichain in $\mathbb{P}$, such that whenever $G$ is $\mathbb{P}$-generic, $\dot{f}[G](n) = h_n(p)$ where $p$ is the unique element of $G \cap A_n$.*

PROOF. For each $n < \omega$ choose a maximal antichain $A_n$ of elements which decide the value of $\dot{f}(n)$. Choose $h_n(p)$ to be the unique element of 2 which $p$ decides to be the value of $\dot{f}(n)$. The conclusion is clear. $\dashv$

LEMMA 11.15. *If $M \vDash 2^\omega \leq \omega_2$ and $G$ is $\mathbb{P}$-generic, then $M[G] \vDash 2^\omega \leq \omega_2$.*

PROOF. Let $G$ be $\mathbb{P}$-generic. Every $f \in (2^\omega)^{M[G]}$ is coded by a sequence of functions as in the previous lemma. It is enough to count the number of such sequences of functions. To determine such a sequence of functions it is enough to choose an $\omega$ sequence of maximal antichains and an $\omega$-sequence of elements of $(2^\omega)^M$. So we have at most $({\omega_2}^\omega)^\omega \cdot (2^\omega)^\omega \leq \omega_2$ objects. $\dashv$

**§12. The Consistency of Failure of the Axiom of Choice.** In this section, use the method of forcing to produce a model of $\mathrm{ZF} + \neg\,\mathrm{AC}$. We have already seen that starting from a countable transitive model $M$ and poset $\mathbb{P} \in M$, any generic extension $M[G]$ will satisfy the Axiom of Choice, so taking the full generic extension will not get us anywhere. Instead, we look for an intermediate model $M(G)$ with $M \subseteq M(G) \subseteq M[G]$, in which all the axioms of ZF hold, and choice fails. We obtain this model by restricting to a certain class of well-behaved names.

We use the poset $\mathbb{P}$ of finite partial functions $p : \omega \times \omega \to 2$, ordered by reverse inclusion. A generic filter defines the characteristic function of a subset of $\omega \times \omega$; regarding this subset as an $\omega$-sequence of subsets of $\omega$, we have canonical names for the sections,

$$\dot{a}_n = \{\langle \check{k}, p\rangle \mid p(\langle n, k\rangle) = 1\},$$

and for the set of these sections,

$$\dot{A} = \{\langle \dot{a}_n, \mathbb{1}\rangle \mid n \in \omega\}.$$

By an easy density argument, $\dot{A}[G]$ is an infinite set of subsets of $\omega$, and is countable in $M[G]$. However, we will show that in an appropriately defined $M(G)$, there is no injection $f : \omega \to \dot{A}[G]$.

Recall a map $i : \mathbb{P} \to \mathbb{P}$ is an **automorphism** if it's bijective, and $p \leq q$ if and only if $i(p) \leq i(q)$, for all $p, q \in \mathbb{P}$; any automorphism of $\mathbb{P}$ induces a permutation of names by the inductive definition

$$i(\tau) = \{\langle i(\sigma), i(p)\rangle \mid \langle \sigma, p\rangle \in \tau\}.$$

Then $i(\check{x}) = \check{x}$ for all $x \in M$.

Note that any permutation of $\omega \times \omega$ induces an automorphism of $\mathbb{P}$ via composition. So suppose $s : \omega \to \omega$ is one-to-one and onto. We define the automorphism $i_s : \mathbb{P} \to \mathbb{P}$ by requiring

$$i_s(p)(\langle n, k\rangle) = p(\langle s^{-1}(n), k\rangle);$$

that is, we obtain $i_s(p)$ by permuting the rows of $p$ by $s$.

The idea will be to inductively define a class of names that are fixed by certain automorphisms of the form $i_s$; such names are called symmetric. The model of values of symmetric names under some generic $G$ will be a model of ZF, but there will be no symmetric name for an injection $f : \omega \to \dot{A}[G]$. The reason will be that if $\dot{f}$ is a name for such an $f$, then for each $k$ it's forced by some $p$ to take a value $f(k) = a_n$. However, if $\dot{f}$ is a *symmetric* name, then we can always change its mind using some automorphism $i_s$ with $s(n) \neq n$. Extending $p$ to $q$ forcing $\dot{f}(k) = a_{s(n)}$, we obtain $a_n = a_{s(n)}$, a contradiction to genericity.

Define the class $V^{\mathrm{HS}(\mathbb{P})}$ of **hereditarily symmetric names** by induction as follows: $\tau$ is in $V^{\mathrm{HS}(\mathbb{P})}$ if

1. there is some finite set $F \subset \omega$ so that $i_s(\tau) = \tau$ whenever $s : \omega \to \omega$ is a bijection fixing elements of $F$, and
2. every $\sigma \in \mathrm{dom}(\tau)$ is (inductively) in $V^{\mathrm{HS}(\mathbb{P})}$.

Suppose $G$ is a $\mathbb{P}$-generic filter over $M$. We then define the **symmetric extension** $M(G) = \{\tau[G] \mid \tau \in M \cap V^{\mathrm{HS}(\mathbb{P})}\}$.

We have the natural analogue of the forcing theorems for symmetric extensions.

THEOREM 12.1. *Define the symmetric forcing relation by* $p \Vdash^{\mathrm{Sym}} \phi(\tau_1, \ldots, \tau_n)$ *if and only if for every* $\mathbb{P}$*-generic* $G$ *over* $M$*, we have* $M(G) \models \phi(\tau_1[G], \ldots, \tau_n[G])$*. Then*

1. $M(G) \models \phi(\tau_1[G], \ldots, \tau_n[G])$ *if and only if there is some* $p \in G$ *so that* $p \Vdash^{\mathrm{Sym}} \phi(\tau_1, \ldots, \tau_n)$*, and*
2. *The relation* $p \Vdash^{\mathrm{Sym}} \phi(\tau_1, \ldots, \tau_n)$ *is definable in* $M$*, for each formula* $\phi$*.*

PROOF. We may define $\Vdash^{\mathrm{Sym}}$ in $M$ by induction on formula complexity. Let $p \Vdash^{\mathrm{Sym}} \sigma \in \tau$ if and only if $p \Vdash \sigma \in \tau$; similarly for $\sigma = \tau$. Next, let

$$p \Vdash^{\mathrm{Sym}} \psi_1 \vee \psi_2 \iff p \Vdash^{\mathrm{Sym}} \psi_1 \text{ or } p \Vdash^{\mathrm{Sym}} \psi_2,$$

and

$$p \Vdash^{\mathrm{Sym}} \neg\psi \iff (\forall q \leq p) q \nVdash^{\mathrm{Sym}} \psi.$$

The key step is the existential quantifier step; we put

$$p \Vdash^{\mathrm{Sym}} (\exists x)\psi(x) \iff \text{ for some } \tau \in V^{\mathrm{HS}(\mathbb{P})}, p \Vdash \psi(\tau).$$

Using the forcing theorems for $\Vdash$, it is easy to check that the theorem holds with this definition of $\Vdash^{\mathrm{Sym}}$. ⊣

From now on we'll just write $\Vdash$ for the symmetric forcing relation. These theorems allow us to show

THEOREM 12.2. *If* $M$ *is a countable transitive model of* ZFC *and* $G$ *is* $\mathbb{P}$*-generic over* $M$*, then* $M(G)$ *is a countable transitive model of the axioms of* ZF*, and* $M \subseteq M(G) \subseteq M[G]$*.*

PROOF. That $M \subset M(G) \subset M[G]$ is immediate by definition and the fact that canonical names for elements of $M$ are hereditarily symmetric. Transitivity is as before, and extensionality, infinity, and foundation are immediate. Pairing and union hold by the same argument as before, plus the verification that the names given there are symmetric.

The remaining axioms take a bit more work. We prove comprehension holds and leave the rest as an exercise. So suppose we have $\sigma, \tau_1, \ldots, \tau_n \in V^{\mathrm{HS}(\mathbb{P})}$. We need to show we have a name in $V^{\mathrm{HS}(\mathbb{P})}$ for

$$\{x \in \sigma[G] \mid M(G) \models \phi(x, \sigma[G], \tau_1[G], \ldots, \tau_n[G])\}.$$

As before, we take our name to be

$$\rho = \{\langle \pi, p \rangle \mid \pi \in \mathrm{dom}(\sigma), p \in \mathbb{P}, \text{ and } p \Vdash \pi \in \sigma \wedge \phi(\pi, \sigma, \tau_1, \ldots, \tau_n)\}.$$

We know this set works, so all we need is to show it's in $V^{\mathrm{HS}(\mathbb{P})}$. Since each $\pi \in \mathrm{dom}(\rho)$ is in $\mathrm{dom}(\sigma)$, condition 2 is satisfied. Now let $F_0$ for $\sigma$, and $F_1, \ldots, F_n$ for $\tau_1, \ldots, \tau_n$, be such that any permutation $s : \omega \to \omega$ fixing $F_i$ pointwise will have $i_s$ fixing the corresponding name. We claim $F = F_0 \cup F_1 \cup \cdots \cup F_n$ witnesses condition 1.

For suppose $s : \omega \to \omega$ is a bijection fixing $F$. By a result from the homework, we have $p \Vdash \phi(\pi, \sigma, \tau_1, \ldots, \tau_n)$ if and only if $i_s(p) \Vdash \phi(i_s(\pi), \sigma, \tau_1, \ldots, \tau_n)$; note that $i_s$ fixes each of $\sigma, \tau_1, \ldots, \tau_n$. It follows immediately that $\langle \pi, p\rangle \in \rho$ exactly when $\langle i_s(\pi), i_s(p)\rangle \in \rho$; that is, $i_s(\rho) = \rho$, and $\rho \in V^{\mathrm{HS}(\mathbb{P})}$ as needed. ⊣

Theorem 12.3. *The symmetric extension $M(G)$ is not a model of the Axiom of Choice.*

Proof. As promised, we show that $\dot{A}[G]$ is an element of $M(G)$, is infinite, and that no injection $f : \omega \to \dot{A}[G]$ can exist. First, notice that for any $n$ and any permutation $s : \omega \to \omega$, we have

$$\begin{aligned} i_s(\dot{a}_n) &= \{\langle i_s(\check{k}), i_s(p)\rangle \mid p(\langle n, k\rangle) = 1\} \\ &= \{\langle \check{k}, p\rangle \mid p(\langle s(n), k\rangle) = 1\} = \dot{a}_{s(n)}. \end{aligned}$$

It follows that whenever a permutation $s$ fixes $n$, we have $i_s(\dot{a}_n) = \dot{a}_n$, so each $\dot{a}_n$ is in $V^{\mathrm{HS}(\mathbb{P})}$. Furthermore, $i_s(\dot{A}) = \dot{A}$, since $i_s$ just permutes the elements of $\dot{A}$, and so $\dot{A} \in V^{\mathrm{HS}(\mathbb{P})}$ as well.

Clearly $\dot{A}[G]$ is an infinite set whenever $G$ is $\mathbb{P}$-generic. Suppose towards a contradiction that $f : \omega \to \dot{A}[G]$ is an injective function in $M(G)$. Then $f$ is named by some $\dot{f} \in V^{\mathrm{HS}(\mathbb{P})}$. Let $F \subset \omega$ be the finite set witnessing this, that is, whenever $s : \omega \to \omega$ is a bijection with $f(n) = n$ for all $n \in F$, we have $i_s(\dot{f}) = \dot{f}$.

Fix a condition $q$ forcing $\dot{f}$ to be an injection into $\dot{A}$, and furthermore forcing $\dot{f}(\check{i}) = \dot{a}_n$, where $n \notin F$. Note such $i, n, q$ must exist, by the forcing theorems and injectivity of $f$. Let $s : \omega \to \omega$ be a permutation fixing $F$ and moving $n$; since $p$ is finite we can further arrange that $i_s(q)$ and $q$ are compatible. We have

$$q \Vdash \dot{f}(\check{i}) = \dot{a}_n,$$

while

$$i_s(q) \Vdash \dot{f}(\check{i}) = \dot{a}_{s(n)}.$$

If we let $p \le q, i_s(q)$, we have $p \Vdash \dot{a}_n = \dot{a}_{s(n)}$. But $s(n) \neq n$, a contradiction to genericity and the fact that $D_{n,s(n)} = \{p \in \mathbb{P} \mid p(n,k) \neq p(s(n),k) \text{ for some } k\}$ is dense. ⊣