FORCING SUMMER SCHOOL
LECTURE NOTES

SPENCER UNGER

These are the lecture notes for the forcing class in the UCLA Logic Summer School of 2013. The notes would not be what they are without a previous set of notes written by Justin Palumbo. Indeed the write ups of the main forcing lemmas are his and many of the other lectures follow his notes.

1. The Continuum Problem

The concept that makes set theory stand out from other subjects is the notion of a well-ordering.

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 $b \in B$, $b \, R \, c$ fails.*

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 every $y \in z$ and every $x \in y$, $x \in z$.*

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

Some easy facts.

Proposition 1.6.
$\emptyset$ is an ordinal.
If α is a ordinal, then the least ordinal greater than α is $\alpha \cup \{\alpha\}$. We call this ordinal $\alpha + 1$.
If $\{\alpha_i \mid i \in I\}$ is a collection of ordinals, then $\bigcup_{i \in I} \alpha_i$ is an ordinal.
We write On *for the class of ordinals.* On *is well-ordered by $\in$.*

Ordinals from the bottom up:

- $0 = \emptyset$

Date: July 15th, 2013.

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

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* α *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* α *be an ordinal.*

- α *is a* successor ordinal *if* $\alpha = \beta + 1$ *for some ordinal* β*.*
- α *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* α *is a* cardinal *if there is no surjection from an ordinal less than* α *onto* α*.*

Clearly each $n \in \omega$ and ω itself are cardinals. We define $|A|$ to be the least ordinal α such that there is a bijection from A to α. $|A|$ 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. □

We are now ready to introduce cardinal arithmetic.

Definition 1.11. *Let κ and λ be cardinals.*

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

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

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

Proof. Suppose that there is a surjection H from κ onto 2^κ. 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 κ 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 $H(\alpha)(\alpha) = 1$. □

By Cantor's theorem we see that for any cardinal κ there is a strictly larger cardinal. We write κ^+ for the least cardinal greater than κ. 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 γ a limit ordinal.

We often write ω_α in place of $\aleph_\alpha$. They are the same object, but we think of ω_α 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 know investigate a ZFC restriction on cardinal exponentiation.

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

Definition 1.14. *A cardinal* κ *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* κ*,* $\kappa < \kappa^{\mathrm{cf}(\kappa)}$*.*

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

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 $f(i) > H(\alpha)(i)$, so $f \neq H(\alpha)$. $\square$

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 ω to ω. 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 ω_1. Given $\{f_\alpha \mid \alpha < \omega_1\}$ can we find a single function f which eventually dominates each f_α. 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 ω_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 ω_1 can know be rephrased as 'Is $\mathfrak{b} > \omega_1$?'.

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

Definition 2.3. *Let A, B be subsets of ω. We say that A and B are* almost disjoint *if $A \cap B$ is finite. A family $\mathcal{F}$ of pairwise almost disjoint subsets of ω is* maximally almost disjoint *(MAD) if for any infinite subset of ω B, 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 ω 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 countable 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 ω enumerated in increasing order. We ensure that b_{n+1} does not belong to any of $A_0, \dots A_n$. 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 $A_{n+1} \setminus (A_0 \cup \dots \cup A_n)$ 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 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 ω to ω. Let $\alpha < \kappa$ 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 $\{f_\alpha \mid \alpha < \kappa\}$ and since $\kappa < \mathfrak{b}$ there is a function f which eventually dominates each f_α.

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_α for $\omega \leq \alpha < \kappa$. Fix α 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_α 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 $1_{\mathbb{P}}$, *ie for all* $p \in \mathbb{P}$, $p \leq 1_{\mathbb{P}}$.

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 $\{p \mid \text{dom}(p) = n\}$ 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 $n < \omega$, $\{p \in \mathbb{P} \mid \text{dom}(p) > n\}$ and $\{p \in \mathbb{P} \mid \text{dom}(p) \text{ is even}\}$. 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* $D \in \mathcal{D}$, $D \cap G \neq \emptyset$.

We are now ready to formulate Martin's Axiom.

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

MA is the assertion that MA(κ) 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(ω) *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 $\mathbb{P}$. ☐

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

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

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_α.

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(κ) to obtain G.

Since G is a filter, $g = \bigcup G$ is a function. The density of each D_n tells us that $\text{dom}(g) = \omega$. The density of E_α tells us that $g \neq f_\alpha$. ☐

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 ω onto ω_1. Such a function cannot exist. ☐

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 it involves some of the machinery used in forcing (posets, dense sets, antichains, filters, etc).

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 should 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 $n \in \operatorname{dom}(p) \setminus \operatorname{dom}(q)$, $p(n) > f(n)$. The p-part of the condition is growing the function from ω to ω 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 ω_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_α, A_α) and (p_β, A_β). This is clear, since the third condition for extension is vacuous. So we have actually shown that that given a sequence of ω_1-many conditions in $\mathbb{P}$ there is a subsequence of ω_1-many conditions which are pairwise compatible. This property is called the ω_1-Knaster property.

We will apply MA to this poset. Let $\mathcal{F} = \{f_\alpha \mid \alpha < \kappa\}$ be a collection of functions from ω to ω where κ 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 $\{(p, A) \mid n \in \operatorname{dom}(p)\}$. 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_α.

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_α, let $\alpha < \kappa$ and choose a condition $(p, A) \in G \cap E_\alpha$. Let $N = \operatorname{dom}(p)$. 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. $\square$

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

Theorem 4.3. *Assume* $\mathrm{MA}(\kappa)$ *and let* $\mathcal{A}$ *and* $\mathcal{C}$ *be collections of size* κ *of subsets of* ω *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* $\mathrm{MA}(\kappa)$ *holds. If* $\mathcal{B}$ *is an almost disjoint family of size* κ *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 κ. It is enough to show that $|\mathcal{P}(\mathcal{B})| = 2^{\aleph_0}$.

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

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

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

Proposition 5.1. *μ^* has the following properties:*

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

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 μ^* 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

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

Since ϵ was arbitrary we have the result. □

It is not hard to see that μ^* 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 $I_x \cap I_y = \emptyset$. 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 $\mu^*(U) = \sum_{n<\omega} b_n - a_n$.

To define the full Lebesgue measure we want to restrict ourselves to certain nice sets. It turns out that the outer measure μ^* is poorly behaved on arbitrary sets. To do so 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 σ-algebra.)

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, $\mu(A) = \mu^*(A)$. *We call the collection of Lebesgue measurable sets $\mathcal{L}$.*

We catalog some properties of Lebesgue measure.

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

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

Theorem 5.5 (AC)**.** *There is $A \subseteq \mathbb{R}$ with $A \notin \mathcal{L}$.*

Theorem 5.6. *$\mathcal{L}$ and μ 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 $\mu(\bigcup_{n<\omega} A_n) = \sum_{n<\omega} \mu(A_n)$.*

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 an 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. $\square$

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

Theorem 5.9. MA(κ) *implies the union of κ-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 α 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 α 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_α and p_β 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_α 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_α 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_α, since G meets each E_α. It remains to show that $\mu(U) \leq \epsilon$.

We claim that if $\{p_n \mid n < \omega\}$ is a subset of G, then $\mu(\cup_{n<\omega} p_n) \leq \epsilon$. Note that since each $p_n \in G$, $p_0 \cup \cdots \cup p_n \in G$. Hence $\mu(p_0 \cup \ldots p_n) < \epsilon$. 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

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

since each partial sum is less than ϵ. 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. $\square$

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 ω.

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 $\emptyset \notin \mathcal{F}$.*
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 κ 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 Frechet 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 set $\mathcal{A} \subseteq \mathcal{P}(X)$ 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 intersetion property then there is a filter on X containing $\mathcal{A}$.*

Proof. Suppose $\mathcal{A}$ has the finite intersection property and define $\mathcal{F} = \{B \subseteq X \mid B \supseteq A_0 \cap \cdots \cap A_n \text{ for some } A_1, \dots A_n \in \mathcal{A}\}$. It is easy to see that $\mathcal{F}$ is a filter. □

Definition 6.5. *A filter $\mathcal{F}$ on X is an* ultrafilter *if for every $A \subseteq X$ either $A \in \mathcal{F}$ or $X \setminus A \in \mathcal{F}$.*

Definition 6.6. *A filter $\mathcal{F}$ on X is maximal if there is no filter $\mathcal{F}'$ on X which properly contains $\mathcal{F}$.*

Proposition 6.7. *A filter is maximal if and only if it is an ultrafilter.*

Proof. Exercise. □

Proposition 6.8. *Every filter can be extended to an ultrafilter.*

Proof. Exercise. □

Recall that the Frechet filter $\mathcal{F}$ is $\{A \subseteq \omega \mid \omega \setminus A \text{ is finite }\}$. 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 6.9 (Ramsey)**.** *For every $\chi : [\omega]^2 \to 2$, there is an infinite set B such that χ is constant on $[B]^2$.*

Proof. We construct three sequences A_i, ϵ_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 χ 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. □

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

Theorem 6.10 (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 χ. Suppose it is monochromatic for 0. Then for all $m < n$ from B, we have $\chi(m, n) = 0$ and hence $a_m \leq a_n$. It follows that $\langle a_n \mid n \in B\rangle$ is a monotone increasing sequence. The argument is the same if B is monochromatic for 1. □

Definition 6.11. *An ultrafilter $\mathcal{U}$ is* Ramsey *if for every coloring $\chi : [\omega]^2 \to 2$, there is $B \in \mathcal{U}$ such that χ is constant on $[A]^2$.*

Note that a Ramsey ultrafilter must be nonprincipal. Let $n < \omega$ and define χ 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 χ.

Theorem 6.12. MA *implies there is a Ramsey ultrafilter.*

Proof. There are 2^ω 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_α is monochromatic for χ_α. Recall that a tower of subsets of ω has the property that for all $\alpha < \beta$, $A_\beta \subseteq^* A_\alpha$.

Suppose that we have constructed A_α 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_β of A which is monochromatic for χ_β. 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. □

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* α*, the* α^{th} level *of* T*, denoted* $\mathrm{Lev}_\alpha(T)$*, is the collection of nodes with height* α*.*
4. *The* height *of* T*, denoted* $\mathrm{ht}(T)$*, is the least ordinal* α *such that* $\mathrm{Lev}_\alpha(T) = \emptyset$*.*
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 the* $\mathrm{ht}(T)$*.*

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

Proof. 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* κ *be a regular cardinal. A tree* T *is a* κ-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* ω *such that* $f(x) \neq f(y)$ *whenever* $x <_T y$*.*

We can now make a general definition.

Definition 7.6. *A regular cardinal* κ *has the* tree property *if every* κ*-tree has a cofinal branch.*

We digress into a related questions 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. *A set* $D \subseteq L$ *is a dense subset if for all* $a, b \in L$*, there is* $d \in D$ *such that* $a < d < b$*.*
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. 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.*

Question 7.11 (Suslin's Problem). *Let* $(L, <)$ *be 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 we can rephrase Suslin's problem as '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 tree if and only if there is a Suslin line.*

We will prove the reverse direction of this theorem since it gives an interesting method of transforming a line into a tree. The forward direction requires some cosmetic improvements to the tree, so we will leave it alone.

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_α 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_α for all $\alpha < \beta$. We seek to define I_β. 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_β 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_β we have either $I_\beta \subseteq I_\alpha$ or I_β disjoint from I_α. (Otherwise I_β contains an endpoint from I_α.) 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 ω_1, so we will be done. If $I, J \in T$ are incomparable, then $I \cap J = \emptyset$. 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. $\square$

Next we show that MA($\aleph_1$) implies that there are no Suslin trees.

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 $x \in \mathrm{Lev}_\alpha(T)$, there is a $y \in \mathrm{Lev}_\beta(T)$ 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 $\beta < \omega_1$, $D_\alpha = \{t \in \mathbb{P} \mid \mathrm{ht}(t) > \beta\}$ 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 fact more is true under $\mathrm{MA}(\aleph_1)$.

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

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 from M^n to M for some $n \geq 1$ and*
4. *$\mathcal{R}$ is a collection of subsets 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 each arity? And so on. To do this we introduce the notion of a *signature*.

Definition 8.2. *A* signature *τ 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 $\mathbb{N} \setminus \{0\}$.*

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 $\tau = (\emptyset, \emptyset, \{<\}, (< \mapsto 2))$. This is a bit much so usually we write $\tau_< = (<)$, since the arity of $<$ is implicit.
2. The signature for a ring with 1 is $\tau_{ring} = (+, \cdot, 0, 1)$. Again we abuse notation here, since it is easy to distinguish between the functions and constants (and there are not too many of each) we just write them all together.

Now we want to know when a structure has a given signature τ.

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

1. *each constant symbol from τ 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 τ-structure for a single signature τ. Instead of writing $i(-)$ all the time, we will write $f^{\mathcal{M}}$ for the interpretation of the function symbol f in the τ-structure $\mathcal{M}$.

We gather some definitions.

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

(1) *$\mathcal{M}$ is a substructure of $\mathcal{N}$ if $M \subseteq N$ and for all c, R, f from τ, $c^{\mathcal{M}} = c^{\mathcal{N}}$, $R^{\mathcal{M}} = R^{\mathcal{N}} \cap M^n$ where $n = a(R)$ and $f^{\mathcal{M}} = f^{\mathcal{N}} \upharpoonright M^k$ where $k = a(f)$.*
(2) *A map $H : M \to N$ is a τ-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 τ-homomorphisms.*

Note that if you are familiar with group theory, you will see that 'substructure' 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 τ be a signature.*

(1) *A* word *in* FOL(τ) *is a finite concatenation of logical symbols ($\neg \wedge \vee \to \forall \exists$), punctuation symbols , ( ), and variables $v_0, v_1, v_2, \dots$.*
(2) *A* term *in* FOL(τ) *(a τ-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 τ be a signature and $\mathcal{M}$ be a τ-structure. Suppose that t is a τ-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(τ) *is built recursively from τ-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 ϕ and ψ are formulas, then $\neg\phi, \phi \wedge \psi, \phi \vee \psi, \phi \to \psi, \forall v \phi$ and $\exists v \phi$ are formulas.*

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

If an occurence of a variable is not bound then it is called *free.* When we write a formula ϕ 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 τ structure and $\phi(\vec{v})$ be a formula with n free variables. For $\vec{a} = (a_1, \dots a_n)$ we define a relation $\mathcal{M} \models \phi(\vec{a})$ by recursion on the construction of the formula.*

1. *If ϕ is $t_1 = t_2$, then $\mathcal{M} \models \phi(\vec{a})$ if and only if $t_1^{\mathcal{M}}(\vec{a}) = t_2^{\mathcal{M}}(\vec{a})$.*
2. *If ϕ is $R(t_1, \dots t_n)$, then $\mathcal{M} \models \phi(\vec{a})$ if and only if $R^{\mathcal{M}}(t_1^{\mathcal{M}}(\vec{a}), \dots t_n^{\mathcal{M}}(\vec{a}))$.*
3. *If ϕ is $\neg\psi$, then $\mathcal{M} \models \phi(\vec{a})$ if and only if $\mathcal{M} \not\models \psi(\vec{a})$.*
4. *If ϕ is $\psi_1 \wedge \psi_2$, then $\mathcal{M} \models \phi(\vec{a})$ if and only if $\mathcal{M} \models \psi_1(\vec{a})$ and $\mathcal{M} \models \psi_2(\vec{a})$.*
5. *If ϕ is $\psi_1 \vee \psi_2$, then $\mathcal{M} \models \phi(\vec{a})$ if and only if $\mathcal{M} \models \psi_1(\vec{a})$ or $\mathcal{M} \models \psi_2(\vec{a})$.*
6. *If ϕ is $\forall u \psi(\vec{v}, u)$, then $\mathcal{M} \models \psi(\vec{a})$ if and only if for all $b \in M$, $\mathcal{M} \models \psi(\vec{a}, b)$.*
7. *If ϕ is $\exists u \psi(\vec{v}, u)$, then $\mathcal{M} \models \psi(\vec{a})$ if and only if there exists $b \in M$, $\mathcal{M} \models \psi(\vec{a}, b)$.*

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

Here is a relatively simple example of the satisfaction relation:

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

Next we introduce the notion of elementarity.

Definition 8.9. *Let τ be a signature and $\mathcal{M}, \mathcal{N}$ be τ-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 $\vec{a} \in M^n$, $\mathcal{M} \models \phi(\vec{a})$ if and only if $\mathcal{N} \models \phi(\vec{a})$.*
2. *A map $H : M \to N$ is an* elementary embedding *if and only if for all formulas $\phi(\vec{v})$ and all $\vec{a} \in M^n$, $\mathcal{M} \models \phi(\vec{a})$ if and only if $\mathcal{N} \models \phi(H(a_1), \dots H(a_n))$.*

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 τ-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 anwer the question 'What is a proof?'. To do so we forget about structures all together and focus on formulas in a fixed signature τ.

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

$$\phi, \phi \to \psi \Rightarrow \psi$$

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.[1]

---

[1] For a complete list of rules of inference and logical axioms, we refer the reader to Kunen's book the "The foundations of mathematics".

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

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

Theorem 8.13 (Soundness). *If $T \vdash \phi$, then $T \vDash \phi$.*

Definition 8.14. *A theory T is consistent if there is no formula ϕ 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 $T \vDash \phi$, 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.

9. Models of Set theory

We are back to set theory, but now are armed with many model theoretic tools that will make out lives easier. The signature for set theory is that of a single binary relation $\in$. Set theory is extremely powerful, since using the axioms of set theory we can formalize classical mathematics in its entirety. Since the signature for set theory has only a single binary relation, models of set theory[2] (M, E) simply consist of a set M and a binary relation E on M. The binary relation does not need to have any relation to the true membership relation $\in$. Formulas in the language of set theory are built using the methods of the previous section. Note that each of the axioms of ZFC can be written as a formula in the language of set theory.

One notion that was not important in the study of model theory in general, but becomes important in the study of models of set theory is transitivity. Recall that a set z is transitive if for every $y \in z$, $y \subseteq z$. A model of set theory is transitive if it is of the form $(M, \in)$ where M is a transitive set. Transitive models reflect basic facts about the universe of sets.

Definition 9.1. *A formula ϕ in the language of set theory is a Δ_0-formula if*

1. *ϕ has no quantifiers, or*
2. *ϕ 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 Δ_0-formulas ψ_0, ψ_1, or*
3. *ϕ is $(\exists x \in y)\psi$ or $(\forall x \in y)\psi$ where ψ is a Δ_0 formula.*

Proposition 9.2. *If $(M, \in)$ is a transitive model and ϕ is a Δ_0-formula, then for all $\vec{x} \in M^n$, $(M, \in) \models \phi$ if and only if ϕ holds.*

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

Proof. We go by induction on the complexity of the Δ_0 formula. Clearly if ϕ is atomic, then we have ϕ if and only if ϕ^M. Also if the conclusion holds for ψ_0 and ψ_1, then clearly it holds for all of the formulas listed in item (2). It remains to show the conclusion for ϕ of the form $(\exists x \in y)\psi(x)$ where the conclusion holds for ψ. Suppose ϕ^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 ϕ 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 ϕ^M holds. ☐

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

It is reasonable to ask what can be expressed by Δ_0-formulas.

Proposition 9.3. *The following expressions can be written as Δ_0-formulas.*

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

Proof. Exercise. ☐

---

[2] By 'models of set theory' we mean models some number (possibly all) of the axioms of set theory.

Recall the definitions of the V and H hierarchies.

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

For an infinite cardinal κ, H_κ is the collection of sets whose transitive closure has size less than κ. We mention a fact about H_κ.

Theorem 9.4. *If κ is regular and uncountable, then H_κ is a transitive model of all of the axioms of ZF except the powerset 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 α 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 ϕ. We can assume that $\forall$ does not appear in any of the ϕ_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 claim that M reflects ϕ 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 ϕ_j. All of the tuples elements appear in some M_i and there for 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 α_i such that $V_{\alpha_{i+1}}$ contains witness to existential formulas with parameters from V_{α_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 $x, y \in P$, $x \neq y$ implies that $\{z \in P \mid z \, E \, x\} \neq \{z \in P \mid z \, E \, y\}$*

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 π on P by induction on E. Induction on E makes sense since E is welfounded. Suppose that for some x we have defined π 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 π.

Clearly M is transitive and π is surjective. We show that π 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 wEx and not wEy. Since $\pi(w) \in \pi(y)$, there is a uEy 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 π 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 $\phi_1, \dots \phi_n$ of ZFC, there is a countable transitive model M such that $M \vDash \phi_1, \dots \phi_n$.*

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

10. Forcing

The following introduction to forcing and proofs of the basic forcing lemmas were written up by Justin Palumbo and not by the author of these notes.

10.1. $M[G]$.

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(ω) 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(ω), 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 ω_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 for us to believe that this G belongs to M. What we now learn is how to force the generic into model M without doing too much damage to the universe of M.

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 ω_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 $M \models$ “D is dense”' exactly when D really is dense. Thus the countable set $\{D \in M : \text{D is dense}\}$, 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 *(ie* $q < p$*) 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* nonatomic *if for any* $p \in \mathbb{P}$ *there exist* $q, r \leq p$ *which are incomparable.*

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

Proposition 10.5. *Say* $\mathbb{P}$ *is* nonatomic *and* $\mathbb{P} \in M$*. Let* G *be* $\mathbb{P}$*-generic. Then* $G \notin M$*.*

Proof. 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 separative there are $q, r \leq p$ which are incomparable. 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* α *the* $\mathbb{P}$*-names of name-rank* α*. (For a* $\mathbb{P}$*-name* τ *we will use* $\rho(\tau)$ *to denote the name-rank of* τ*). 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* α *have been defined, we say that* τ *is a* $\mathbb{P}$*-name of name-rank* α *if every* $x \in \tau$ *is of the form* $x = \langle \tau, p \rangle$ *where* τ *is a* $\mathbb{P}$*-name and* $p \in \mathbb{P}$*.*

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

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

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 anyway of giving them a coherent meaning. But once we are have a $\mathbb{P}$-generic filter G at hand, they can be given values.

Definition 10.8. *Let* τ *be a* $\mathbb{P}$*-name and* G *a filter on* $\mathbb{P}$*. Then the value of* τ *under* G*, denote* $\tau[G]$*, is by recursive definition the set*

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

With this definition in mind, one can think of as an element $\langle \sigma, p \rangle$ of a $\mathbb{P}$-name τ 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 $M \subseteq M[G]$ and $G \in M[G]$.*

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 $\tau \in M^{\mathbb{P}}$. Every element of $\tau[G]$ has the form $\sigma[G]$ by definition, so $y = \sigma[G]$ for some σ for which there is a pair $\langle \sigma, p\rangle \in \tau$. 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 Γ so that whatever G is we have $\Gamma[G] = G$. Set $\Gamma = \{\langle \check{p}, p\rangle : p \in \mathbb{P}\}$. Then $\Gamma[G] = \{\check{p}[G] : p \in G\} = \{p : p \in G\} = G$. ☐

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 $\mathrm{op}(\sigma, \tau)[G] = \langle \sigma[G], \tau[G]\rangle$.

A few of the axioms of ZFC are easily verified for $M[G]$.

Lemma 10.14. *We have that $M[G]$ satisfies the axioms of extensionality, pairing, union, and foundation.*

Proof. Any transitive model satisfies extensionality so that is easily taken care of.

To check that $M[G]$ satisfies pairing, we must show that given $\sigma_1[G], \sigma_2[G]$ (where σ_1, σ_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 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.

As for foundation, an easy absoluteness argument shows every model satisfies foundation. ☐

One last thing to observe is the following.

Lemma 10.15. *The models M and $M[G]$ have the ordinals; that is $M \cap ON = M[G] \cap ON$.*

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

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 $\alpha = \mathrm{rk}(\alpha) = \mathrm{rk}(\tau[G]) \leq \mathrm{rk}(\tau)$. Since M is a model of ZFC, by absoluteness of the rank function, the $\mathrm{rk}(\tau) \in M$. Since M is transitive, $\mathrm{rk}(\tau[G])$ belongs to M as well. And this is just α. ☐

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 $\mathbb{1}$. We start truly taking advantage of G in the next section.

10.2. **The Forcing Relation.**

Definition 10.16. *The* forcing language *consists of the binary relation symbol* $\in$ *along with a constant symbol* τ *for each* $\tau \in M^{\mathbb{P}}$. *Let* $\phi(\tau_1, \dots \tau_n)$ *be formula of the forcing language, so that* $\tau_1, \dots \tau_n$ *all belong to* $M^{\mathbb{P}}$. *Let* $p \in \mathbb{P}$. *We say that* $p \Vdash \phi(\tau_1, \dots \tau_n)$ *if for* every $\mathbb{P}$*-generic filter* G *with* $p \in G$ *we have* $M[G] \models \phi(\tau_1[G], \dots \tau_n[G])$.

In order to make sense of this definition (and a few other things), let us take a breath and consider an example. Let us take $\mathbb{P}$ to be $\mathrm{Fn}(\omega, 2)$ the collection of finite functions whose domain is a subset of ω 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. Thus is 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

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

Then $\dot{g}[G] = \{\check{\langle m, n\rangle}[G] : p \in G, m \in \mathrm{dom}\, p, n \in \mathrm{ran}\, p\}$. Since $\check{\langle m, n\rangle}[G] = \langle m, n\rangle$ 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 clearly that $g(1) = 1$. In terms of forcing this is the same as saying

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

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

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

An important property of $\Vdash$ to observe is the following.

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

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

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

Theorem 10.18 (Forcing Theorem A). *If* $M[G] \models \phi(\tau_1[G], \dots \tau_n[G])$ *then there is a* $p \in \mathbb{P}$ *such that* $p \Vdash \phi(\tau_1, \dots, \tau_n)$.

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

The proofs of these theorems are quite long and technical, and when first learning the theory of forcing it may be worthwhile to take these theorems as black boxes. For now we will show how to use them to finish proving that $M[G]$ models of ZFC, and save the proofs for a later time.

Here's 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 a $\mathbb{P}$-generic filter 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 $q \Vdash \pi \in \sigma \wedge (\phi(\pi, \tau_1, \dots, \tau_n))$. $\square$

The forcing theorems are extremely important, and generally any forcing argument requires a heavy use of them. This we will see. They make it much easier to analyze the relation between M and $M[G]$. Indeed, our next use of them will be to finish proving Theorem 10.10.

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 π 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 $p \Vdash (\phi \in \sigma \wedge \phi(\pi, \sigma, \tau_1, \dots, \tau_n)$. So by definition of ρ, $\langle \pi, p\rangle \in \rho$. Since $p \in G$, $\pi[G] \in \rho[G]$. $\square$

Notice how in the above proof the ρ we constructed does not at all depend on what G actually is; this is one of the central tenets of forcing - that 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 $\phi(x, v, r, z_1, \dots z_n)$ 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 y in $M[G]$ so that $M[G] \vDash \phi(x, y, \sigma[G], \tau_1[G], \dots, \tau_n[G])$. We have to construct a name $\rho \in M^{\mathbb{P}}$ which witnesses replacement, ie so that

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

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

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

In fact, what we are applying here is a stronger looking version of replacement where we do not require the μ to be unique. In fact this version is implied by the regular version (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 ρ 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

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

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

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

□

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 $\mathcal{P}(\mathrm{dom}(\sigma) \times \mathbb{P})$, relatived to M. Let $\rho = S \times \{\mathbb{1}\}$.

Let us check that ρ 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 τ there is a $p \in G$ so that $p \Vdash \pi \in \sigma$ 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 $\pi[G] \in \tau[G]$. □

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 α 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. □

Finally notice that $M[G]$ satisfies the Axiom of Infinity since $\omega \in M[G]$. That gives us all of the axioms of ZFC, and so Theorem 10.10 is proved.

We give the proofs of the Forcing Theorems here.

The basic idea is to define another relation, $\Vdash^*$, which will not make any mention of generic sets and will clearly be absolute in M. We will show that Forcing Theorem A holds for this relation, and that it lines up with the original definition. The definition is by a somewhat complicated recursion. There are many ways to do this; we borrow the presentation from Kunen.

Definition 10.25. *Let $\mathbb{P}$ be a poset. Let $\phi(x_1, \dots, x_n)$ be a formula,$p \in \mathbb{P}$, and let $\tau_1, \dots \tau_n$ be $\mathbb{P}$-names. We define $p \Vdash^* \phi(\tau_1, \dots \tau_n)$ by recursion on the complexity of ϕ as follows.*

1. *$p \Vdash^* \tau_1 = \tau_2$ if and only if the following hold.*
    - A. *For all $\langle \pi_1, s_1 \rangle \in \tau_1$, the set*
    $$\{q : q \leq s_1 \to \exists \langle \pi_2, s_2 \rangle \in \tau_2 (q \leq s_2 \wedge q \Vdash^* \pi_1 = \pi_2)\}$$
    is dense below p.
    - B. *For all $\langle \pi_2, s_2 \rangle \in \tau_1$, the set*
    $$\{q : q \leq s_2 \to \exists \langle \pi_1, s_1 \rangle \in \tau_1 (q \leq s_1 \wedge q \Vdash^* \pi_1 = \pi_2)\}$$
    is dense below p.
2. *$p \Vdash^* \tau_1 \in \tau_2$ if and only if the set*
$$\{q : \exists \langle \pi, s \rangle \in \tau_2 (q \leq s \wedge q \Vdash^* \tau_1 = \pi)\}$$
is dense below p.
3. *$p \Vdash^* (\phi(\tau_1, \dots, \tau_n) \wedge \psi(\tau_1, \dots, \tau_n)$ if and only if*
$$p \Vdash^* \phi(\tau_1, \dots, \tau_n) \text{ and } p \Vdash^* \psi(\tau_1, \dots, \tau_n).$$
4. *$p \Vdash^* \neg\phi(\tau_1, \dots, \tau_n)$ if and only if there is no $q \leq p$ such that $q \Vdash^* \phi(\tau_1, \dots, \tau_n)$.*
5. *$p \Vdash^* \exists x \phi(x, \tau_1, \dots, \tau_n)$ if and only if the set*
$$\{r : \exists \sigma (r \Vdash^* \phi(\sigma, \tau_1, \dots, \tau_n))\}$$
is dense below p.

Lemma 10.26. *The following are equivalent:*

1. *$p \Vdash^* \phi(\tau_1, \dots, \tau_n)$.*
2. *$\forall r \leq p (r \Vdash^* \phi(\tau_1, \dots, \tau_n))$.*
3. *$\{r : \Vdash^* \phi(\tau_1, \dots, \tau_n)\}$ is dense below p.*

Proof. That (2) implies (1) is trivial, as is (2) implies (3).

Let us check the equivalence first in the case where ϕ is atomic (ie $\tau_1 \in \tau_2$ or $\tau_1 = \tau_2$). We don't even need to use the specifics of our definition, only that $p \Vdash^* \phi(\tau_1, \dots, \tau_n)$ is equivalent to a certain set being dense below p. For (1) implies (2), simply note that if a set is dense below p and $r \leq p$ then the set is also dense below r. For (3) implies (1) it is enough to show that if $D \subseteq \mathbb{P}$ and $\{r : D \text{ is dense below } r\}$ is dense below p then D is dense below p. And that is easy; let $q \leq p$. Let $r_1 \leq q$ with D dense below r_1. Then there is $r \leq r_1$ with $r \in D$.

Now the equivalence has been verified for atomic ϕ. The rest is by induction on the complexity of ϕ.

Say ϕ is $\psi \wedge \chi$. For (1) implies (2), we have $p \Vdash^* \phi$ if and only if $p \Vdash^* \psi$ and $p \Vdash^* \chi$. By induction (1) implies (2) for these formulas and so $\forall r \leq p (r \Vdash^* \psi)$

and $\forall r \leq p(r \Vdash^* \chi)$. So $\forall r \leq p(r \Vdash^* \psi$ and $r \Vdash^* \chi)$ which by definition gives $\forall r \leq p(r \Vdash^* \phi)$. Similarly for (3) implies (1), using the fact that the intersection of two dense open sets is dense.

Now say ϕ is $\neg\psi$. For (1) implies (2) we don't even need the induction; if there is no $q \leq p$ such that $q \Vdash^* \psi(\tau_1, \ldots, \tau_n)$ then for all $r \leq p$ there is no $q \leq r$ such that $q \Vdash^* \psi$. That is just the definition of $r \Vdash^* \phi$. For (3) implies (1), assume $D = \{r : r \Vdash^* \phi\}$ is dense below p but that $p \Vdash^* \phi$ fails. Since ϕ is $\neg\psi$, by definition we have some $q \leq p$ such that $q \Vdash^* \psi$. But then since (1) implies (2) for ψ we have that for all $r \leq q$, $r \Vdash^* \psi$. But this contradicts D being dense; if $r \leq q$ belongs to D then $r \nVdash^* \psi$.

Finally say that ϕ is $\exists x\psi$. We again don't even need the induction since as in the atomic case that definition only hinges on a certain set being dense. ☐

Theorem 10.27. *Let M be a countable transitive model of* ZFC, *$\mathbb{P} \in M$, and $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$. Let G be $\mathbb{P}$-generic over M. Then:*

1. *If $p \in G$ and $(p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M$, then $M[G] \vDash \phi(\tau_1[G], \ldots, \tau_n[G])$.*
2. *If $M[G] \vDash \phi(\tau_1[G], \ldots, \tau_n[G])$ then there is a $p \in G$ such that $(p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M$.*

Proof. This proof is going to be a long induction on the complexity of ϕ; and in fact each step of that induction will be another induction (using the definition of $\Vdash^*$) on the name rank of $\tau_1, \ldots, \tau_n$. So let us begin. For atomic formulas, $\Vdash^*$ is defined from absolute concepts using recursion on a well-founded relation ($\in$ restricted to $M^{\mathbb{P}}$) which is absolute for M. Thus $(\Vdash^*)^M$ is absolute for atomic formulas and we will omit the relativization to M in our notation in this part of the proof.

The first case is when ϕ is $\tau_1 = \tau_2$. We will check (1). Assume $p \in G$ and $p \Vdash^* \tau_1 = \tau_2$. We must show that $M[G] \vDash \tau_1[G] = \tau_2[G]$, which by absoluteness is the same as showing that $\tau_1[G] = \tau_2[G]$. We will show $\tau_1[G] \subseteq \tau_2[G]$; that $\tau_2[G] \subseteq \tau_1[G]$ is a symmetric argument. So let $x \in \tau_1[G]$. Then $x = \pi_1[G]$ where $\langle \pi_1, s_1 \rangle \in \tau_1$ and $s_1 \in G$. We must show that $\pi_1[G] \in \tau_2[G]$. As G is a filter, we can take $r \in G$ with $r \leq p$ and $r \leq s_1$. Since $r \leq p$ by the previous lemma we have $r \Vdash^* \tau_1 = \tau_2$. So by definition the set

$$D = \{q : q \leq s_1 \rightarrow \exists \langle \pi_2, s_2 \rangle \in \tau_2 (q \leq s_2 \wedge q \Vdash^* \pi_1 = \pi_2)\}$$

is dense below r. As G is a $\mathbb{P}$-generic filter, we get $q \leq r$ with $q \in D$. Since $q \leq s_1$, letting $\langle \pi_2, s_2 \rangle$ witness that $q \in D$ we have $q \leq s_2$ and hence $s_2 \in G$. We also have $q \Vdash^* \pi_1 = \pi_2$. Now $\pi_2[G] \in \tau_2[G]$ by definition, and $\pi_1[G] = \pi_2[G]$ by induction.

Now we check (2) for $\tau_1 = \tau_2$. Assume $M[G] \vDash \tau_1[G] = \tau_2[G]$ which is to say that $\tau_1[G] = \tau_2[G]$. To get the $p \in G$ we want we use a complicated dense set. Let D be the set of all the $r \in \mathbb{P}$ such that one of the following holds:

a. $r \Vdash^* \tau_1 = \tau_2$.
b. $\exists \langle \pi_1, s_1 \rangle \in \tau_1 (r \leq s_1 \wedge \forall \langle \pi_2, s_2 \rangle \in \tau_2 \forall q \in \mathbb{P}((q \leq s_2 \wedge q \Vdash^* \pi_1 = \pi_2) \rightarrow q \perp r))$.
c. $\exists \langle \pi_2, s_2 \rangle \in \tau_2 (r \leq s_2 \wedge \forall \langle \pi_1, s_1 \rangle \in \tau_1 \forall q \in \mathbb{P}((q \leq s_1 \wedge q \Vdash^* \pi_1 = \pi_2) \rightarrow q \perp r))$.

*C*laim 1. No $r \in G$ satisfies either (a) or (b). By symmetry we just check that (a) is not possible. Let $r \in G$ with $\langle \pi_1, s_1 \rangle$ as in (a). Then $s_1 \in G$ and hence $\pi_1[G] \in \tau_1[G]$; by hypothesis $\tau_1[G] = \tau_2[G]$ and so $\pi_1[G] \in \tau_2[G]$. By definition that means there is some $\langle \pi_2, s_2 \rangle \in \tau_2$ so that $s_2 \in G$ and $\pi_1[G] = \pi_2[G]$. Then, applying

(2) and the induction for $\pi_1 = \pi_2$ there is a $q_0 \in G$ with $q_0 \Vdash^* \pi_1 = \pi_2$. Now fix $q \in G$ with $q \leq q_0$ and $q \leq s_2$. Then by the previous lemma we have $q \Vdash^* \pi_1 = \pi_2$. So, by (a) we have $q \perp r$. But q and r both belong to G. Contradiction.

That claim given, we just have to show that D is dense, since $D \in M$ by absoluteness of the concepts used in its definition. For then an $r \in D \cap G$ would have to satisfy (a) and thus would be exactly what we are looking for to check (2). So fix $p \in \mathbb{P}$. We are looking for a member of D below it. If $p \Vdash^* \tau_1 = \tau_2$ holds already, then $p \in D$ and there is nothing further for us. So assume $p \Vdash^* \tau_1 = \tau_2$ fails. Then by definition of $\Vdash^*$ one of (A) or (B) fails. By symmetry say it is (A). Then the set given there is not dense below p, and so there is an $r \leq p$ and a $\langle \pi_1, s_1 \rangle \in \tau_1$ so that

$$(\dagger) \forall q \leq r (q \leq s_1 \wedge \forall \langle \pi_2, s_2 \rangle \in \tau_2 (\neg (q \Vdash^* \pi_1 = \pi_2 \wedge q \leq s_2))).$$

That means $r \leq s_1$. Let us see that r satisfies (b). Let $\langle \pi_2, s_2 \rangle \in \tau_2, q \leq s_2$ with $q \Vdash^* \pi_1 = \pi_2$. Then $q \Vdash r$ as any common extension of q and r would be contradictory to †. So $r \leq p$ and r satisfies (b) meaning that $r \in D$.

Whew. So we have established (1) and (2) for $\tau_1 = \tau_2$. Let us establish it for $\tau_1 \in \tau_2$. Suppose $p \in G$ with $p \Vdash^* \tau_1 \in \tau_2$. Then by definition the set

$$D = \{q : \exists \langle \pi, s \rangle \in \tau_2 (q \leq s \wedge q \Vdash^* \tau_1 = \pi)\}$$

is dense below p. By absoluteness $D \in M$ and so there is a $q \in G \cap D$. Fix $\langle \pi, s \rangle$ witnessing that $q \in D$, so that $q \leq s$ and $q \Vdash^* \pi = \tau_1$. Now $s \in G$ and $\langle \pi, s \rangle \in \tau_2$ and so by definition $\pi[G] \in \tau_2[G]$. Now $q \in G$ and $q \Vdash^* \pi = \tau_1$. Thus since we verified (1) and (2) for equality, we have $\pi[G] = \tau_1[G]$. Together this means that $\tau_1[G] \in \tau_2[G]$; so by absoluteness $M[G] \vDash \tau_1[G] \in \tau_2[G]$ which is just what we needed for (1).

Now to check (2) for $\tau_1[G] \in \tau_2[G]$. Assume $M[G] \vDash \tau_1[G] \in \tau_2[G]$, ie that $\tau_1[G] \in \tau_2[G]$. By definition, there is $\langle \pi, s \rangle \in \tau_2$ such that $s \in G$ and $\pi[G] = \tau_1[G]$. By (2) verified for equality there is an $r \in G$ such that $r \Vdash^* \pi = \tau_1$. Let $p \in G$ with $p \leq s$ and $p \leq r$. Then $\forall q \leq p (q \leq s \wedge q \Vdash^* \pi = \tau_1)$ which is more than we need to get $p \Vdash^* \tau_1 \in \tau_2$ from its definition.

So atomic formulas are handled. We can no longer ignore the relativization to M, since the $\exists$ stage of the recursive definition of $\Vdash^*$ is not absolute. We will drop mention of $\tau_1, \ldots, \tau_n$.

We check (1) for $\neg$. Inductively assume (1) and (2) hold for ϕ; we show we have it for $\neg\phi$. Assume $p \in G$ and that $(p \Vdash^* \neg\phi)^M$. We want to show that $M[G] \Vdash \neg\phi$. Assume otherwise; $M[G] \Vdash \phi$. Then by (2) for ϕ there is a $q \in G$ so that $(q \Vdash^* \phi)^M$. Let $r \leq p, q$. Then $r \Vdash^* \phi)^M$. But $r \leq p$ is absolute for M so by definition of $(p \Vdash^* \neg\phi)^M$ we should have $(r \Vdash^* \phi)^M$ failing; and we do not.

We check (2) for $\neg$. Suppose $M[G] \vDash \neg\phi$. Let

$$D = \{p : (p \Vdash^* \phi)^M \vee (p \Vdash^* \neg\phi)^M\}.$$

This set is the same as the set $\{p : p \Vdash^* \phi \vee p \Vdash^* \neg\phi\}$ definined inside of M; and one easily proves (in ZFC that this set is dense using the definition of $\Vdash^*$ for negations (for if p is given either there is some condition $q \leq p$ such that $q \Vdash^* \phi$ or every $q \leq p$ has $q \Vdash^* \phi$ failing, which means $p \Vdash^* \neg\phi$). So D is dense. Let $p \in D \cap G$. If $(p \Vdash^* \neg\phi)^M$ then we have what we needed for (2). Otherwise $(p \Vdash^* \phi)^M$ and so applying (1) for ϕ we would have $M[G] \Vdash \phi$, contradicting our assumption.

We check (1) for $\wedge$. Assume (1) and (2) hold for ϕ and ψ. We show we have (1) it for $\phi \wedge \psi$. Assume $p \in G$ and that $(p \Vdash^* \phi \wedge \psi)^M$. Then by definition of $\Vdash^*$ we have $p \Vdash^* \phi$ and $p \Vdash^* \psi$. By (1) $M[G] \vDash \phi$ and $M[G] \vDash \psi$. So $M[G] \vDash \phi \wedge \psi$. Exactly what we needed.

We check (2) for $\wedge$. Assume $M[G] \vDash \phi \wedge \psi$. Then $M[G] \vDash \phi$ and $M[G] \vDash \psi$. By (2) we have $p, q \in G$ such that $(p \Vdash^* \phi)^M$ and $(q \Vdash^* \psi)^M$. Since G is a filter there's $r \in G$ with $r \leq p, q$. By the previous lemma we have $(r \Vdash^* \phi)^M$ and $(r \Vdash^* \psi)^M$ and so by definition of $\Vdash^*$ we have $(r \Vdash^* \phi \wedge \psi)^M$.

We check (1) for $\exists$. Assume $p \in G$ and $(p \Vdash^* \exists x \phi(x))^M$. By relativizing the definition of $\Vdash^*$ at this step we have that the set

$$\{r : \exists \sigma \in M^{\mathbb{P}}(r \Vdash^* \phi(\sigma))^M\}$$

is desne below p. The set belongs to M so by genericity of G there is $r \in G$ and $\sigma \in M^{\mathbb{P}}$ with $(r \Vdash^* \phi(\sigma))^M$. By (1) for ϕ we have $M[G] \vDash (\phi(\sigma[G]))$ which of course gives $M[G] \vDash (\exists x \phi(x))$.

Finally we check (2) for $\exists$. Assume $M[G] \vDash (\exists x \phi(x))$. Then by definition of $M[G]$ (and $\vDash$) there is a $\sigma \in M^{\mathbb{P}}$. By (2) for ϕ there is $p \in G$ so that $(p \Vdash^* \phi(\sigma))^M$. So $\forall r \leq p((r \Vdash^* \phi(\sigma))^M)$ which is more than enough to conclude that $(p \Vdash^* \exists x \phi(x))^M$ by definition. $\square$

In particular, (1) of this theorem gives the following.

Corollary 10.28. *Let M be a countable transitive model of* ZFC *and $\mathbb{P} \in M$. Let $\tau_1, \ldots, \tau_n \in M^{\mathbb{P}}$. Then $p \Vdash \phi(\tau_1, \ldots, \tau_n)$ if and only if $(p \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M$.*

Proof. From right to left is just (1) of the previous theorem and the definition of $\Vdash$.

For left to right, say $p \Vdash \phi(\tau_1, \ldots, \tau_n)$. We want to show that $(p \Vdash^* \phi(\tau_1, \ldots, \tau_n)^M$. For this, by the lemma relativized to M we just have to show that the $\{r : (r \Vdash^* \phi(\tau_1, \ldots, \tau_n))^M\}$ is dense below p. If it weren't, there would be a $q \leq p$ so that for $\forall r \leq q(r \notin D)$. This gives $(q \Vdash^* \neg\phi(\tau_1, \ldots, \tau_n))^M$ by definition. Now applying the right to left of this corollary (already verified) we have $q \Vdash \neg\phi(\tau_1, \ldots, \tau_n)$. Let G be $\mathbb{P}$-generic over M with $q \in G$. So $M[G] \vDash \neg\phi(\tau_1[G], \ldots, \tau_n[G])$. But $q \leq p$ and so $p \in G$, so that by definition of $\Vdash$ (and the fact we assumed the left hand side) we have $M[G] \vDash \phi(\tau_1[G], \ldots, \tau_n[G])$. Contradiction. $\square$

That gives Forcing Theorem B immediately. And forcing Theorem A is immediate from Forcing Theorem B and (2) of the previous theorem. This ends the section of the notes lifted from Justin's version.

11. The independence of CH

As advertised we will prove the independence of CH from the axioms of ZFC. Recall that when we discussed formal proofs and model theory, that it is enough to construct two models of ZFC, one in which CH holds and the other in which it fails. To do this we will assume that ZFC is consistent in addition to assuming ZFC. However the assumption can be done away with by dealing with large enough fragments of ZFC, using the Reflection Theorem and using the fact that proofs are finite.

To show that ZFC $\nvdash A$ where A is a sentence in the language of set theory it is enough to show that for all finite fragments of T of ZFC, $T \nvdash A$. By the Reflection theorem and the Mostowski Collapse theorem, there is a countable transitive model M such that whenever $\mathbb{P} \in M$ and G is $\mathbb{P}$-generic over M, $M[G] \vDash T$. Essentially M needs to reflect enough set theory to prove the forcing theorems and construct the appropriate names for objects. So if by forcing we can make a model where A fails, then we know that $T + \neg A$ is consistent for every finite fragment T of ZFC. So in particular ZFC $\nvdash A$.

So in all of our forcing arguments we will just make the assumption that ZFC is consistent, 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 ω_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. So from the point of view of $M[G]$, the ω_1 in M is a countable ordinal. We will see two methods for arguing that ω_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.*

12. The consistency of CH

We wish to construct a model of ZFC + CH by forcing. Given a coutable 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, ie $p_1 \leq p_2$ if and only if $p_1 \supseteq p_2$.

To show that ω_1 is preserved we develop the notion of closure of a poset.

Definition 12.1. *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 12.2. *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 a $\mathbb{P}$-name $\dot{H}$ which is forced to be a function from ω onto ω_1^M. Let $n < \omega$, we claim that $D_n = \{p \in \mathbb{P} \mid p \Vdash \dot{f}(n) = \check{\alpha}$ 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 $i_G(\dot{f}(n)) = \alpha$. 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 be arbitrary. Given p_n let $p_{n+1} \in D_n$ with $p_{n+1} \leq p_n$ and record the value α_n witnessing $p_{n+1} \in D_n$. Let $p \leq p_n$ by the countable closure of $\mathbb{P}$. Let $\alpha = \sup \alpha_n$. Let G be $\mathbb{P}$-generic over M. Then in $M[G]$, $\operatorname{ran}(i_G(\dot{f}))$ is bounded by α, but this is a contradiction since it was supposed to be forced that f was onto. $\square$

A similar argument shows the following.

Lemma 12.3. *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. $\square$

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

Proof. Using the generic object G we define a list of $\omega_1^{M[G]} = \omega_1^M$-many subsets of ω. We then do a density argument to show that this list comprises all subsets of ω in $M[G]$. Work for the moment in $M[G]$. Let $g = \bigcup G$. Note that g is a function from ω_1 to 2. We define a collection of subsets of ω, $\{x_\alpha \mid \alpha < \omega_1\}$, 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 α such that $x = x_\alpha$. For this we will do a density argument. Work in M and let $x \subseteq \omega$. We claim $D_x = \{p \in \mathbb{P} \mid$ there is $\alpha < \omega_1$ such that for all n, $\chi_x(n) = p(\omega \cdot \alpha + n)$ for all $n < \omega\}$ is dense. (Here χ_x is the characteristic function of x.) Let $p \in \mathbb{P}$. Let $\operatorname{dom}(p) = \beta$. Let $\alpha > \beta$. It follows that for all $n < \omega$, $\omega \cdot \alpha + n \notin \operatorname{dom}(p)$. So we extend p to a condition p' in D_x where α is the witness.

It follows that in $M[G]$ the map $\alpha \mapsto x_\alpha$ is a surjection from ω_1 onto $\mathcal{P}(\omega)$. $\square$

So we have proved that CH is consistent.

13. 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 $\mathbb{P} = \{p \mid \text{there is } x \subseteq \omega_2 \text{ finite such that } p : x \to 2\}$ 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 13.1. *Suppose that $\mathbb{P}$ is a ccc poset. Whenever G is $\mathbb{P}$-generic over M and κ is an ordinal, $M \models$ "κ is a cardinal" if and only if $M[G] \models$ "κ is a cardinal".*

Proof. Let G be $\mathbb{P}$-generic over M and κ 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 $i_G(f)$ is a surjection from some $\alpha < \kappa$ onto κ. We fix a condition $p_0 \in G$ forcing this.

For every $\beta < \alpha$, the collection $D_\beta = \{p \in \mathbb{P} \mid p \text{ decides } \dot{f}(\beta)\}$ is dense below p_0, since $\dot{f}(\beta)$ is a name for an ordinal. So if we choose $A_\beta \subseteq D_\beta$ a maximal antichain, then there is a countable set of ordinals X_β such that whenever $p \in A_\beta$ there is an ordinal $\gamma \in A_\beta$ such that $p \Vdash \dot{f}(\beta) = \gamma$. It follows that $p_0 \Vdash \dot{f}(\beta) < \sup X_\beta$. But this means that $p_0 \Vdash \sup(\operatorname{ran}(\dot{f})) \leq \sup_{\beta<\alpha}(\sup X_\beta)$ and the right hand supremum is less than κ contradicting that p_0 forces that $\dot{f}$ is onto κ. $\square$

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

Let κ be a regular cardinal.

Definition 13.2. *A set $C \subseteq \kappa$ is* club *if it is unbounded in κ and for all $\alpha < \kappa$ if $C \cap \alpha$ is unbounded in α, then $\alpha \in C$.*

Lemma 13.3. *The collection of club subsets of κ form a κ-complete filter.*

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

Definition 13.4. *A set $S \subseteq \kappa$ is* stationary *if for every club C in κ, $S \cap C \neq \emptyset$.*

Lemma 13.5. *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 13.6. *If S is stationary in κ, then S is unbounded in κ.*

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 13.7 (The Δ-system lemma)**.** *Let X be a set of size ω_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 set $\{x_\alpha \mid \alpha \in I\}$ form a Δ-system with root r.

Proof. First note that tt is enough to show the lemma in the case $X = \omega_1$. Since for an arbitrary X of size ω_1 we can use a bijection with ω_1 to copy the problem. So let $\{x_\alpha \mid \alpha < \omega_1\}$ be a collection of finite subsets of ω_1.

We define a function $F : \mathrm{Lim}(\omega_1) \to \omega_1$ by $F(\alpha) = \max(x_\alpha \cap \alpha)$. Since each α is finite, we have $F(\alpha) < \alpha$ for all limit ordinals α. 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 δ, 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 γ_α of I for all $\alpha < \beta$. The set $\bigcup_{\alpha<\beta} x_{\gamma_\alpha}$ is countable and hence bounded in ω_1 by some ordinal $\eta < \omega_1$. Let γ_β be the least member of J greater than η.

Now we claim that $\{x_\alpha \mid \alpha \in I\}$ forms a Δ-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 γ_β, $x_{\gamma_\alpha} \subseteq \gamma_\beta$. 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. ☐

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 13.8. *$\mathbb{P}$ 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 Δ-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 Δ-system with root r.

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

We break the remaining proof into two pieces.

Lemma 13.9. *If G is $\mathbb{P}$-generic over M, then $M[G] \models 2^\omega \geq \omega_2$.*

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 $g(\omega \cdot \alpha + n) = 1$. We claim that for each $f \in (2^\omega)^M$ and each $\alpha < \omega_2$, the set $D_{f,\alpha} = \{p \mid \text{there is } n \text{ such that } f(n) \neq p(\omega \cdot \alpha + n)\}$ 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 so that it disagrees with $f(n)$. Since $G \cap D_{f,\alpha} \neq \emptyset$ for all $\alpha < \omega_2$ and $f \in (2^\omega)^M$, it follows that for all $\alpha < \omega_2$, $f_\alpha \notin M$. So $M[G] \models 2^\omega \geq \omega_2$. ☐

Lemma 13.10. *Let $\mathbb{P}$ be a poset and let $\dot{f}$ be a $\mathbb{P}$-name for a function from ω to 2. There is a sequence 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 $i_G(\dot{f}(n)) = h_n(p)$ 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. ☐

Lemma 13.11. *If $M \models 2^\omega \leq \omega_2$ and G is $\mathbb{P}$-generic, then $M[G] \models 2^\omega \leq \omega_2$.*

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 ω sequence of maximal antichains and an ω-sequence of elements of $(2^\omega)^M$. So we have atmost $(\omega_2^\omega)^\omega \cdot (2^\omega)^\omega \leq \omega_2$ objects. $\square$