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

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

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

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

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

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

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

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

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

Definition 8.2. A **signature** 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 $\mathbb{N} \setminus \{0\}$. The elements of $\mathcal{C} \cup \mathcal{F} \cup \mathcal{R}$ are the **non-logical symbols**.

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

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

Now we want to know when a structure has a given signature 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 $f^{\mathcal{M}} = f^{\mathcal{N}} \restriction M^k$ 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 in particular that in (1), $\mathcal{M}$ must be closed under the function $f^{\mathcal{N}}$ to be a substructure.

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

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

Definition 8.5. Let τ be a signature.

1. A **word** in FOL(τ) is a finite concatenation of logical symbols,
$$\neg \quad \wedge \quad \vee \quad \to \quad \forall \quad \exists \quad =$$
punctuation symbols,
$$, \quad (\quad)$$
and variables
$$v_0 \quad v_1 \quad v_2 \quad \ldots$$
as well as symbols coming from the signature τ: constant symbols $c \in \mathcal{C}$, function symbols $f \in \mathcal{F}$, and relation symbols $R \in \mathcal{R}$.
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 $t_1, \ldots t_n$ are terms, then $f(t_1, \ldots, t_n)$ 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.

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

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

If an occurrence of a variable is not bound then it is called **free**. When we write a formula 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 $\psi_1 \to \psi_2$, then $\mathcal{M} \models \phi(\vec{a})$ if and only if $\mathcal{M} \not\models \psi_1(\vec{a})$ or $\mathcal{M} \models \psi_2(\vec{a})$.
7. If ϕ is $\forall u\psi(\vec{v}, u)$, then $\mathcal{M} \models \psi(\vec{a})$ if and only if for all $b \in M$, $\mathcal{M} \models \psi(\vec{a}, b)$.
8. If ϕ is $\exists u\psi(\vec{v}, u)$, then $\mathcal{M} \models \psi(\vec{a})$ if and only if there exists $b \in M$ such that $\mathcal{M} \models \psi(\vec{a}, b)$.

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:

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

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 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 answer the question 'What is a proof?'. To do so we forget about structures altogether 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*:

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

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

DEFINITION 8.12. Let T be a theory and 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 $T \models \phi$ if every structure which satisfies T also satisfies ϕ.

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

DEFINITION 8.14. A theory T is **consistent** if there is no formula 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 $T \models \phi$, then $T \vdash \phi$.*

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

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

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

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

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

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

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

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

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

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

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

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 $\psi_0 \land \psi_1$, $\psi_0 \lor \psi_1$, $\psi_0 \to \psi_1$, $\neg\psi_0$ 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 $\vec{x} \in M^n$, $(M, \in) \vDash \phi$ if and only if ϕ holds.*

To save ourselves from writing $(M, \in) \vDash \phi$, 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. *$x = \varnothing$, x is a singleton, x is an ordered pair, $x = \{y, z\}$, $x = (y, z)$, $x \subseteq y$, x is transitive, x is an ordinal, x is a limit ordinal, x is a natural number, $x = \omega$.*
2. *$z = x \times y$, $z = x \setminus y$, $z = x \cap y$, $z = \bigcup x$, $z = \mathrm{ran}(x)$, $y = \mathrm{dom}(x)$.*
3. *R is a relation, f is a function, $y = f(x)$, $g = f \restriction x$.*

PROOF. Exercise. ⊣

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

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

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

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

THEOREM 9.4. *If κ is regular and uncountable, then H_κ is a transitive model of all of the axioms of* ZFC *except the power set axiom.*

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

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

1. *an M such that $M_0 \subseteq M$, $|M| \leq |M_0| \cdot \aleph_0$ and for all $\vec{a} \in M^n$, $\phi^M(\vec{a})$ if and only if $\phi(\vec{a})$ and*
2. *an ordinal 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 prove 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 for which $\exists x \phi_j(x, \vec{a})$ holds, all of the tuple's elements appear in some M_i and therefore there is a witness to $\exists x \phi_j(x, \vec{a})$ in M_{i+1}.

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

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

Definition 9.6. A model (P, E) is

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

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

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

Proof. Let (P, E) be a well-founded, extensional model. We define a map π on P by induction on E. Induction on E makes sense since E is well-founded. 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 $w\, E\, x$ and not $w\, E\, y$. Since $\pi(w) \in \pi(y)$, there is a $u\, E\, y$ such that $\pi(u) = \pi(w)$. This contradicts the minimality of the choice of z, since $\pi(u) = \pi(w) \in z$ and $u \neq w$.

To see that M and 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. $\dashv$

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

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

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

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

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

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

Since MA(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 $D_X = \{p \in \mathcal{P} : X \in \text{ran}(p)\}$ is dense, as is the set $E_\alpha = \{p \in \mathbb{P} : \alpha \in \text{dom}(p)\}$ for each $\alpha < \omega_1^M$. So a $\mathbb{P}$-generic filter G will intersect each of those sets, and will by the usual arguments yield a surjection $g : \omega_1^M \to \mathcal{P}(\omega)^M$. Thankfully, by the previous lemma, such a G exists. Unfortunately there is no reason to believe G is in M, and it is difficult to see how we would go about adding it. This is what we now learn: how to force a generic object which we can adjoin to M without doing too much damage to its universe.

Given any poset $\mathbb{P}$ in M, and a $\mathbb{P}$-generic filter G, the method of forcing will give us a way of creating a new countable transitive model $M[G]$ satisfying ZFC that extends M and contains G. Now just getting such a model is not enough. For in the example above the surjection $g : \omega_1 \to \mathcal{P}(\omega)$ defined from G was a mapping between the objects in M. But a priori it may well be that the model $M[G]$ has a different version of 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 $\{D \in M : D \text{ is dense}\}$ is exactly the same collection defined in M to be the collection of *all* dense subsets of $\mathbb{P}$. Unless

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

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

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

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

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

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

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

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 $x = \langle \sigma, p \rangle$ where σ 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 $\mathbb{P}$.

In analogy with the von Neumann hierarchy V_α, we may define

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

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

It is not hard to see that the notion of being a $\mathbb{P}$-name is absolute; that is, $M \vDash$ "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 any way of giving them a coherent meaning. But once we have a $\mathbb{P}$-generic filter G at hand, they can be given values.

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

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

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

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

One thing at least is not hard to see.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

To check that $M[G]$ satisfies pairing, we must show that given $\sigma_1[G], \sigma_2[G]$ (where 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 we need is precisely what $\mathrm{up}(\sigma_1, \sigma_2)$ provides.

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

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

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

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

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

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

10.2. The Forcing Relation.

DEFINITION 10.16. The **forcing language** consists of the symbols of first order logic, the binary relation symbol $\in$, and constant symbols 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 $p \Vdash \phi(\tau_1, \dots \tau_n)$ (read p **forces** φ) 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's take a breath and consider an example. Take $\mathbb{P}$ to be $\mathrm{Fn}(\omega, 2)$, the collection of finite functions whose domain is a subset of 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. Since G is $\mathbb{P}$-generic, we have for each $n \in \omega$ that $D_n \cap G$ is not empty. Thus as before we can define from G a function $g : \omega \to 2$ such that $g = \bigcup G$.

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

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

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

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

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

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

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

LEMMA 10.17. *If $p \Vdash \phi(\tau_1, \dots, \tau_n)$ 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 $M[G] \models \phi(\tau_1[G], \dots, \tau_n[G])$. $\dashv$

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

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

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

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

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

Proof. Let G be $\mathbb{P}$-generic with $p \in G$. Since $p \Vdash (\exists x \in \sigma)\phi(x, \tau_1, \ldots, \tau_n)$, by definition of $\Vdash$, $M[G] \vDash (\exists x \in \sigma[G])\phi(x, \tau_1[G], \ldots, \tau_n[G])$. So take $\pi[G] \in \sigma[G]$ such that we have $M[G] \vDash \phi(\pi[G], \tau_1[G], \ldots, \tau_n[G])$. 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 $r \Vdash \pi \in \sigma \wedge \phi(\pi, \tau_1, \ldots, \tau_n)$. 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 $q \Vdash \pi \in \sigma \wedge \phi(\pi, \tau_1, \ldots, \tau_n)$. ⊣

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

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

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

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

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

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 $p \Vdash \pi \in \sigma \wedge \phi(\pi, \sigma, \tau_1, \ldots, \tau_n)$. By definition of $\Vdash$ then we have that $\pi[G] \in \sigma[G]$, and $M[G] \vDash \phi(\pi[G], \sigma[G], \tau_1[G], \ldots \tau_n[G])$. 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 $M[G] \vDash \phi(\pi[G], \sigma[G], \tau_1[G], \ldots \tau_n[G])$. Applying Forcing Theorem A tells us that there is some $p \in G$ such that $p \Vdash \pi \in \sigma \wedge \phi(\pi, \sigma, \tau_1, \ldots, \tau_n)$. So by definition of ρ, $\langle \pi, p\rangle \in \rho$. Since $p \in G$, $\pi[G] \in \rho[G]$. ⊣

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: People living in M can reason out every aspect of $M[G]$ if they just imagined that some generic G existed.

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

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

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

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

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

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

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

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

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

⊣

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

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

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

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. ⊣

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

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

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

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

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

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

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

Clearly this relation on pairs is well-founded.

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

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

Note that in each case the name-rank of θ is less than one of σ, τ so the inductive definition makes sense.

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

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

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

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

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

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

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

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

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

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

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

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

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

if and only if

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

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

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

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

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

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

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

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

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

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

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

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

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

So

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

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

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

The third equivalence by inductive hypothesis.

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

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

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

Proof. Note once again that φ is a formula in the forcing* language.

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

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

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