Chapter 6. Mathematical logic: a super-rapid introduction and and some applications to set theory.*

6.1. First order (formal) languages. In this chapter we are going to take a brief glance at mathematical logic in order to explain what was meant at the end of §4.4. when we said that some things, for example (GCH), are not provable in the system ZF but neither are their contraries. This account of mathematical logic will help explain what it means to say that one can show that something is not demonstrable. The key to this is the relation between what is *provable* and what is *true*. This is neither a philosophical remark (nor a mystical one!), but rather something mathematical because by true we mean here *true in a model* of a collection of axioms. Put in another way what we are looking at here is the mathematical relationship between syntax and semantics. This subject is called is *model theory*. You can find fuller accounts of the material on model theory in this section in, for example, van Dalen's book, and I also recommend the books by Wilfrid Hodges, *A shorter model theory*, Cambridge University Press(?), 1996(?), and Bruno Poizat, *A course in model theory*, Springer, 2000 (or the original French version, Bruno Poizat, *Cours de théorie de modèles*, Nur al-Mantiq wal Ma'rifah, Lyons, 1985).

Definition. A language $\mathcal{L}$ for first order logic is composed of:
(i) a countably infinite set of variables v_0, v_1,... (also written x, y, z,... when we want);
(ii) a set of symbols for functions f, g,... of various "r-ities" – a symbol for a 0-ary function is a symbol for a constant c;
(iii) a set of symbols for relations R, S,... of various arities, and a special symbol $=$ (which is binary);
(iv) some logical connectives: &, or, $\longrightarrow$, $\neg$, $\perp$;
(v) quantifiers $\forall$, $\exists$;
(vi) punctuation symbols such as (,), ..

The *terms* of $\mathcal{L}$ are defined recursively
(i) Any variable v is a term;
(ii) if f is a symbol for an n-ary function and $t_0,\ldots, t_{n-1}$ are terms, then $f(t_0,\ldots,t_{n-1})$ is also a term.

We define the *free variables* of a term t, $\mathrm{FV}(t)$ by (i) $\mathrm{FV}(v) = \{v\}$ if v is a variable, and (ii) $\mathrm{FV}(f(t_0,\ldots,t_{n-1})) = \mathrm{FV}(t_0) \cup \ldots \mathrm{FV}(t_{n-1})$.

The *formulas* of $\mathcal{L}$ are defined recursively together with their free variables.
(i) If R is an n-ary relation symbol and $t_0,\ldots, t_{n-1}$ are terms, then $R(t_0,\ldots,t_{n-1})$ is a *formula*. $\mathrm{FV}(R(t_0,\ldots,t_{n-1})) = \mathrm{FV}(t_0) \cup \ldots \mathrm{FV}(t_{n-1})$. (These formulas are called *atomic formulas*.)
(ii) If ϕ and ψ are formulas, then $\phi \,\&\, \psi$, ϕ or ψ, $\phi \longrightarrow \psi$, $\neg\phi$, $\perp$ are also formulas. $\mathrm{FV}(\phi \,\&\, \psi) = \mathrm{FV}(\phi \text{ or } \psi) = \mathrm{FV}(\phi \longrightarrow \psi) = \mathrm{FV}(\phi) \cup \mathrm{FV}(\psi)$, $\mathrm{FV}(\neg\phi) = \mathrm{FV}(\phi)$ and $\mathrm{FV}(\perp) = \emptyset$.
(iii) If ϕ is a formula and $x \in \mathrm{FV}(\phi)$, then $\forall x\, \phi$ and $\exists x\, \phi$ is also a formula. $\mathrm{FV}(\forall x\, \phi) = \mathrm{FV}(\exists x\, \phi) = \mathrm{FV}(\phi) \setminus \{x\}$.

Thus the free variables of a formula ϕ are the variables "about which ϕ says something." A formula ϕ is a *sentence* if $\mathrm{FV}(\phi) = \emptyset$. Write $\phi(\bar{x})$ if $\bar{x}$ is the string $x_0,\ldots, x_{n-1}$ and ϕ is a formula such that $\mathrm{FV}(\phi) \subseteq \{x_0,\ldots,x_{n-1}\}$.

Note. We can define all of the connectives and quantifiers using solely or, $\neg$ and $\exists$ – and this will save us some time later on – if we define $\phi \,\&\, \psi = \neg(\neg\phi \text{ or } \neg\psi)$, $\phi \longrightarrow \psi = (\neg\phi \text{ or } \psi)$,

* version 1.0, 15/2/06

$\bot = \neg(\phi \text{ or } \neg\phi)$ and $\forall x\, \phi = \neg\exists x\, (\neg\phi)$.

We also need a definition of how to substitute a term for a variable in a term or a formula. We have to take a little care over this in order to preserve the sense of a formula, because if, for example, ϕ is $\exists y\, (y = x)$ and $t = f(y)$ we want that the result of the substitution of t for x is not $\exists y\, (y = f(y))$ but rather $\exists z\, (z = f(y))$.

So $s[t/x]$, $\phi[t/x]$, *the result of substituting t for x*, is defined by recursion:

$$
\begin{array}{ll}
v_i[t/x] & = t \text{ if } x = v_i, \text{ and} \\
 & = v_i \text{ otherwise;} \\
f(s_0, \ldots, s_{n-1})[t/x] & = f(s_0[t/x], \ldots, s_{n-1}[t/x]); \\
R(s_0, \ldots, s_{n-1})[t/x] & = R(s_0[t/x], \ldots, s_{n-1}[t/x]); \\
(\neg\phi)[t/x] & = \neg(\phi[t/x]); \\
\bot\, [t/x] & = \bot; \\
(\phi \,\&\, \psi)[t/x] & = \phi[t/x] \,\&\, \psi[t/x]; \\
(\phi \text{ or } \psi)[t/x] & = \phi[t/x] \text{ or } \psi[t/x]; \\
(\phi \longrightarrow \psi)[t/x] & = \phi[t/x] \longrightarrow \psi[t/x]; \\
(\forall y\, \phi)[t/x] & = \forall y \phi \text{ if } x = y, \\
 & = \forall y\, (\phi[t/x]) \text{ if } y \neq x \text{ and } y \notin \mathrm{FV}(s), \text{ and} \\
 & = \forall z\, ((\phi[z/y])[t/x]) \text{ if } y \neq x \text{ and } y \in \mathrm{FV}(s), \text{ where } z \\
 & \qquad \text{is a choice of something } z \notin \mathrm{FV}(s) \cup \mathrm{FV}(\phi); \\
(\exists y\, \phi)[t/x] & = \exists y \phi \text{ if } x = y, \\
 & = \exists y\, (\phi[t/x]) \text{ if } y \neq x \text{ and } y \notin \mathrm{FV}(s), \text{ and} \\
 & = \exists z\, ((\phi[z/y])[t/x]) \text{ if } y \neq x \text{ and } y \in \mathrm{FV}(s), \text{ where } z \\
 & \qquad \text{is a choice of something } z \notin \mathrm{FV}(s) \cup \mathrm{FV}(\phi).
\end{array}
$$

The relation $\vdash$ of syntactic derivation comes from thinking about trees of deductions with hypotheses as leaves and conclusions as the base, made from "primative" deductions.

Formally we define $\vdash$ by recursion. The definition is for formulas ϕ (a conclusion) by recursion on the structure of ϕ and for any set of formulas Γ (the hypotheses), at the same time.

To begin with we have that $\Gamma \vdash \phi$ if $\phi \in \Gamma$;

and: if $\Gamma \vdash \phi$ and $\Gamma \vdash \psi$,	then $\Gamma \vdash \phi \,\&\, \psi$;
if $\Gamma \vdash \phi \,\&\, \psi$,	then $\Gamma \vdash \phi$;
if $\Gamma \vdash \psi \,\&\, \psi$,	then $\Gamma \vdash \psi$;
or: if $\Gamma \vdash \phi$,	then $\Gamma \vdash \phi$ or ψ;
if $\Gamma \vdash \psi$,	then $\Gamma \vdash \phi$ or ψ;
if $\Gamma \vdash \psi$ or ψ, $\Gamma \cup \{\phi\} \vdash \chi$ and $\Gamma \cup \{\psi\} \vdash \chi$,	then $\Gamma \vdash \chi$;
implication: if $\Gamma \cup \{\phi\} \vdash \psi$,	then $\Gamma \vdash \phi \longrightarrow \psi$;
if $\Gamma \vdash \phi \longrightarrow \psi$ and $\Gamma \vdash \phi$,	then $\Gamma \vdash \psi$;
not/false: if $\Gamma \cup \{\phi\} \vdash \bot$,	then $\Gamma \vdash \neg\phi$;
if $\Gamma \vdash \bot$,	then $\Gamma \vdash \phi$ for any ϕ;
if $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$,	then $\Gamma \vdash \bot$;
if $\Gamma \cup \{\neg\phi\} \vdash \bot$,	then $\Gamma \vdash \phi$;
for all: if $\Gamma \vdash \phi(x)$ and $x \notin \bigcup\{\mathrm{FV}(\psi) \mid \psi \in \Gamma\}$,	then $\Gamma \vdash \forall x\, \phi(x)$;
if $\Gamma \vdash \forall x\, \phi(x)$ and t is a term,	then $\Gamma \vdash \phi[t/x]$;
there is: if $\Gamma \vdash \phi[t/x]$,	then $\Gamma \vdash \exists x\, \phi(x)$;
if $\Gamma \vdash \exists x\, \phi(x)$ and $\Gamma \cup \{\phi(y)\} \vdash \psi$ and $y \notin \mathrm{FV}(\psi)$,	then $\Gamma \vdash \psi$.
equality: if s is a term,	then $\Gamma \vdash s = s$;
if $\Gamma \vdash s = t$ and $\Gamma \vdash \phi[s/x]$,	then $\Gamma \vdash \phi[t/x]$.

and because in the definition of the introdution of *for all* we do not only say that "if $\Gamma \vdash \phi(x)$ and

$x \notin \bigcup\{\mathrm{FV}(\psi) \mid \psi$ is used in the proof of $\phi(x)$ of $\Gamma \,\&\, \psi \in \Gamma\}$, then $\Gamma \vdash \forall x\, \phi(x)$," we have to add that:

if $\Delta \subseteq \Gamma$ and $\Delta \vdash \phi$, then $\Gamma \vdash \phi$.

Definition. Γ is *consistent* if $\Gamma \nvdash \bot$. Equivalently Γ is *consistent* if there is no ϕ such that $\Gamma \vdash \phi \,\&\, \neg\phi$. We sometimes write $\mathrm{Con}(\Gamma)$ meaning Γ is consistent.

After all of these definitions I ought to give some examples – but that would take up time – so look for such in van Dalen.

Definition. Let $\mathcal{L}$ be a first order language, Γ a collection of sentences and ϕ a sentence in $\mathcal{L}$. d is a *derivation of* ϕ *from* Γ if d is a finite list (or *string*) of formulas, $\phi_0, \ldots, \phi_n$ for some $n < \omega$, with $\phi = \phi_n$ and such that for each $m \leq n$ we have either $\phi_m \in \Gamma$ or ϕ_m can be obtained by applying the basic clauses of the definition of $\vdash$ to $\{\phi_i \mid i < m\}$.

6.2. Models and the completeness and compactness theorems.

Let us write $\Gamma \Longrightarrow \phi$ for our informal mathematical notion that ϕ follows from Γ. (For example often in the course up to now we have shown that ZF $\Longrightarrow \phi$ for ϕ a sentence about sets.) It is clear from the definition of $\vdash$ that if $\Gamma \vdash \phi$, then $\Gamma \Longrightarrow \phi$. Now we shall introduce a notion of "follows semantically."

Definition. Let $\mathcal{L}$ be a first order language as in §6.1. $\mathcal{M}$ is an $\mathcal{L}$-structure if

$$\mathcal{M} = (M, f^{\mathcal{M}}, g^{\mathcal{M}}, \ldots, R^{\mathcal{M}}, S^{\mathcal{M}}, \ldots)$$

where M is a set, $f^{\mathcal{M}}$, $g^{\mathcal{M}}$,... are functions such that $f^{\mathcal{M}} : {}^nM \longrightarrow M$ if f is n-ary, and $R^{\mathcal{M}}$, $S^{\mathcal{M}}$,... are relations on M with $R^{\mathcal{M}} \subseteq {}^nM$ if R is n-ary.

Next we define when it is that a structure *thinks that a sentence is true*.

Definition. Let $\mathcal{L}$ be a first order language and $\mathcal{M}$ an $\mathcal{L}$-struture. Let ϕ be a sentence in $\mathcal{L}$. We define ϕ is true in $\mathcal{M}$, writing $\mathcal{M} \models \phi$, by recursion on the form of ϕ.

$$\begin{array}{ll}
\mathcal{M} \models R(a_0, \ldots, a_{n-1}) & \text{if and only if } (a_0, \ldots, a_{n-1}) \in R^{\mathcal{M}}; \\
\mathcal{M} \models \phi \,\&\, \psi & \text{if and only if } \mathcal{M} \models \phi \text{ and } \mathcal{M} \models \psi; \\
\mathcal{M} \models \phi \text{ or } \psi & \text{if and only if } \mathcal{M} \models \phi \text{ or } \mathcal{M} \models \psi; \\
\mathcal{M} \models \neg\phi & \text{if and only if } \mathcal{M} \not\models \phi; \\
\mathcal{M} \models \phi \longrightarrow \psi & \text{if and only if } \mathcal{M} \not\models \phi \text{ or } \mathcal{M} \models \psi; \\
\mathcal{M} \models \exists x\, \phi(x) & \text{if and only if there is some } a \in M \text{ and } \mathcal{M} \models \phi(a); \\
\mathcal{M} \models \forall x\, \phi(x) & \text{if and only if for all } a \in M \text{ we have } \mathcal{M} \models \phi(a).
\end{array}$$

Note. Let $\mathcal{L}$ be a first order language and $\mathcal{M}$ is an $\mathcal{L}$-structure. Then $\mathcal{M} \not\models \bot$ by the definition of $\models$.

For example, if $\mathcal{L} = \{e, ., {}^{-1}\}$, the *language of group theory*, and $\mathcal{M} \models (\forall x\, x.x^{-1} = e = x^{-1}.x) \,\&\, (\forall x\, x.e = x = e.x) \,\&\, (\forall x \forall y \forall z\, (x.y).z = x.(y.x))$ (so $\mathcal{M}$ is a group), then while it might or might not be that $\mathcal{M} \models \forall x \forall y\, x.y = y.x$, certainly $\mathcal{M} \not\models \exists x\, x.x^{-1} \neq e$.

Definition. $\Gamma \models \phi$ if for all structures $\mathcal{M}$ for the language $\mathcal{L}$ we have that if $\mathcal{M} \models \psi$ for each $\psi \in \Gamma$, then $\mathcal{M} \models \phi$.

Definition. $\mathcal{M}$ is a *model* of a set of sentences Γ for a language $\mathcal{L}$ if $\mathcal{M}$ is an $\mathcal{L}$-structure and $\mathcal{M} \models \phi$ for all $\phi \in \Gamma$. We write $\mathcal{M} \models \Gamma$.

It is clear by the definition of $\models$ that if $\Gamma \Longrightarrow \phi$ then $\Gamma \models \phi$. So we have the following.

Proposition. (*Soundness Theorem*) If $\Gamma \vdash \phi$, then $\Gamma \models \phi$.

Corollary. If Γ, a set of sentences, has a model, then Γ is consistent.

Proof. If Γ is not consistent, then $\Gamma \vdash \bot$, so $\Gamma \models \bot$. But there are no structures $\mathcal{M}$ such that $\mathcal{M} \models \bot$, by the definition of $\models$, so there are no structures $\mathcal{M}$ such that $\mathcal{M} \models \Gamma$.

But we also have the converse to the proposition.

Completeness Theorem. If $\Gamma \models \phi$, then $\Gamma \vdash \phi$.

This is a deeper fact, because it says there is a translation from an infinite procedure (confirming that each model of Γ, each of which could itself be infinite, is a model of ϕ as well) to a finite procedure (a derivation of ϕ from Γ). In order to prove the theorem it is convenient to introduce the following pair of definitions.

Definition. A collection of sentences Δ in a language $\mathcal{L}$ is a Henkin theory if whenever $\exists x\, \phi(x) \in \Delta$, there is a constant $c \in \mathcal{L}$ such that $\Delta \vdash \exists x\, \phi(x) \longrightarrow \phi(c)$. Thus in this situation c is a *witness* to the sentence $\exists x\, \phi(x)$.

Observation. If Δ, in a language $\mathcal{L}$, is a Henkin theory, and Δ' is a collection of sentences in the same language $\mathcal{L}$, $\Delta \subseteq \Delta'$, then Δ' is also a Henkin theory.

Definition. If Δ is a collection of sentences in a language $\mathcal{L}$ and $\Delta \subseteq \Delta'$, a collection of sentences in a language $\mathcal{L}'$ with $\mathcal{L} \subseteq \mathcal{L}'$, we say that Δ' is a conservative extension of Δ if $\Delta \vdash \phi$ if and only if $\Delta' \vdash \phi$ for each $\phi \in \mathcal{L}$. (And so Δ' is consistent if Δ is consistent, because $\bot \in \mathcal{L}$ and, hence, $\Delta' \vdash \bot$ if and only if $\Delta \vdash \bot$.)

Proposition. If Δ is a collection of sentences in a language $\mathcal{L}$, there is a collection of sentences Δ' in a language $\mathcal{L}'$ such that $\mathcal{L} \subseteq \mathcal{L}'$, $\overline{\overline{\mathcal{L}'}} = \overline{\overline{\mathcal{L}}} + \omega$, $\Delta \subseteq \Delta'$, Δ' is a Henkin theory and Δ' is a conservative extension of Δ.

Proof. Let Γ be any collection of sentences in a language $\mathcal{K}$. Add distinct constants c_ϕ for $\mathcal{K}$ for each $\phi(x)$ with free variable x in $\mathcal{K}$ to obtain $\mathcal{K}^* = \mathcal{K} \cup \{\, c_\phi \mid \phi(x) \in \mathcal{L} \,\}$. Let

$$\Gamma^* = \Gamma \cup \{\, \exists x\, \phi(x) \longrightarrow \phi(c_\phi) \mid \Gamma \vdash \exists x\, \phi(x) \,\}.$$

We shall prove that Γ^* is a conservative extension of Γ.

Suppose that $\Gamma^* \vdash \psi$ for some $\psi \in \mathcal{K}$. Then there is some $n < \omega$ and formulas ϕ_i for $i < n$ such that $\Gamma \cup \{\, \exists x\, \phi_i(x) \longrightarrow \phi_i(c_{\phi_i}) \mid i < n \,\} \vdash \psi$. Let $\Gamma^m = \Gamma \cup \{\, \exists x\, \phi_i(x) \longrightarrow \phi_i(c_{\phi_i}) \mid i < m \,\}$ for each $m \leq n$. We shall show that if $\Gamma^{m+1} \vdash \psi$ we have that $\Gamma^m \vdash \psi$, for any $m < n$, and, hence, by induction, that $\Gamma = \Gamma^0 \vdash \psi$. If $n = 0$ there is nothing to prove. So suppose that $n \neq 0$ and let $m < n$.

If $\Gamma^{m+1} = \Gamma^m \cup \{\, \exists x\, \phi_m(x) \longrightarrow \phi_m(c_{\phi_m}) \,\} \vdash \psi$, we have that $\Gamma^m \vdash (\exists x\, \phi_m(x) \longrightarrow \phi_m(c_{\phi_m})) \longrightarrow \psi$. So, we have that $\Gamma^m \vdash \exists y (\exists x\, \phi_m(x) \longrightarrow \phi_m(y)) \longrightarrow \psi$ by the definition of $\vdash$, specifically by the rule for the introdution existencial quantifier, because $c_{\phi_m} \notin \mathcal{K}$ and, hence, does not appear in the formulas ϕ and ψ. Thus $\Gamma^m \vdash (\exists x\, \phi_m(x) \longrightarrow \exists y\, \phi_m(y)) \longrightarrow \psi$, and so $\Gamma^m \vdash \psi$.

Now let $\Delta_0 = \Delta$ and $\mathcal{L}_0 = \mathcal{L}$, let $\Delta_{i+1} = \Delta_i^*$, $\mathcal{L}_{i+1} = \mathcal{L}_i^*$ for each $i < \omega$, and let $\Delta' = \bigcup \{\, \Delta_i \mid i < \omega \,\}$ and $\mathcal{L}' = \bigcup \{\, \mathcal{L}_i \mid i < \omega \,\}$. We have that Δ' is a conservative extension of Δ. Moreover Δ' is a Henkin theory because if $\exists x\, \phi(x) \in \Delta$ we have that $\exists x\, \phi(x) \in \Delta_i$ for some $i < \omega$ and hence $(\exists x\, \phi(x) \longrightarrow \phi(c_\phi)) \in \Delta_{i+1} \subseteq \Delta$.

Proof of the Completeness Theorem. Let Γ be a collection of sentences in a language $\mathcal{L}$, and

ϕ a formula in $\mathcal{L}$. In order to prove the theorem we show that if $\Gamma \not\vdash \phi$ there is a model of $\Gamma \cup \{\neg\phi\}$ (and so of $\Gamma \not\models \phi$).

By the proposition, let Γ' be a Henkin theory such that $\Gamma \cup \{\neg\phi\} \subseteq \Gamma'$ in a language $\mathcal{L}'$ with $\mathcal{L} \subseteq \mathcal{L}'$ and $\overline{\overline{\mathcal{L}'}} = \overline{\overline{\mathcal{L}}} + \omega$. Consider the set of collections of sentences Δ such that Δ is in $\mathcal{L}'$, Δ is consistent, and $\Gamma' \subseteq \Delta$, and order this set by $\Delta \leq \Delta^*$ if $\Delta \subseteq \Delta^*$. Each chain in the order has an upper bound, so we can use Zorn's Lemma to choose a theory $\Gamma_{\max}$ which is maximal such that $\Gamma_{\max}$ is in $\mathcal{L}'$, $\Gamma_{\max}$ is consistent, and $\Gamma' \subseteq \Gamma_{\max}$. Note that $\Gamma_{\max}$ is Henkin (by the observation). Also, if $\Gamma_{\max} \vdash \psi$, we have that $\psi \in \Gamma_{\max}$ by the maximality of $\Gamma_{\max}$; if $\exists x\, \psi(x) \in \Gamma_{\max}$, then we have a constant $c \in \mathcal{L}$ such that $\psi(c) \in \Gamma_{\max}$; if ψ or $\chi \in \Gamma_{\max}$ we have that $\psi \in \Gamma_{\max}$ or $\chi \in \Gamma_{\max}$; and $\neg\psi \in \Gamma_{\max}$ if and only if $\psi \notin \Gamma_{\max}$.

Now we use the terms themselves in the language $\mathcal{L}'$ as the elements of a model for $\Gamma_{\max}$. You should make certain here that you understand how we are using the terms, syntactic things, as objects with which to form a (semantic) structure.

Let $A = \{t \in \mathcal{L}' \mid t \text{ is a term in } \mathcal{L}'\}$. If f is a symbol for an n-ary function in $\mathcal{L}'$, let $f^{\mathcal{A}} : {}^{n}A \longrightarrow A$ be defined by $f^{\mathcal{A}}(t_0, \ldots, t_{n-1}) = f(t_0, \ldots, t_{n-1})$, a term as well. and if R is a symbol for an n-ary relation in $\mathcal{L}'$, define a subset $R^{\mathcal{A}}$ of ${}^{n}A$ by $(t_0, \ldots, t_{n-1}) \in R^{\mathcal{A}}$ if and only if $(t_0, \ldots, t_{n-1}) \in R$.

We are now close to the end of the proof, but it could be that there are terms s, t such that '$s = t$' $\in \Gamma_{\max}$, thus we have to take the quotient of $\mathcal{A}$ by this relation to get a structure for $\mathcal{L}'$. So let $s \sim t$ if '$s = t$' $\in \Gamma_{\max}$ for $s, t \in A$, and let $[s] = \{t \in A \mid t \sim s\}$ for $s \in A$. We form a structure $\mathcal{M}$ by taking $M = \{[s] \mid s \in A\}$, $f^{\mathcal{M}}([t_0], \ldots, [t_{n-1}]) = [f(t_0, \ldots, t_{n-1})]$ if f is an n-ary function symbol in $\mathcal{L}'$, and $([t_0], \ldots, [t_{n-1}]) \in R^{\mathcal{M}}$ if and only if $(t_0, \ldots, t_{n-1}) \in R^{\mathcal{A}}$ where R is an n-ary relation symbol $\mathcal{L}'$. We have that $\mathcal{M} = (M, f^{\mathcal{M}}, \ldots, R^{\mathcal{M}}, \ldots)$ is an L'-structure, and $\mathcal{M} \models \psi$ if and only if $\psi \in \Gamma_{\max}$ for $\psi \in \mathcal{L}'$ by the induction on the structure of sentences, using the observations that follow from the maximality of $\Gamma_{\max}$ in the last sentence of the first paragraph of this proof. Hence $\mathcal{N} = (M, f^{\mathcal{M}}, \ldots, R^{\mathcal{M}}, \ldots \mid f \in \mathcal{L})$ is an $\mathcal{L}$-structure and $\mathcal{N} \models \Gamma \cup \{\neg\phi\}$.

Corollary. (Existence of models). If Γ, a set of sentences, is consistent, then Γ has a model. Also, if Γ is in a language $\mathcal{L}$ and $\overline{\overline{\mathcal{L}}} = \kappa$, then Γ has a model $\mathcal{N}$ with $\overline{\overline{N}} = \omega + \kappa$.

Proof. If Γ does not have a model then we have that $\mathcal{M} \models \Gamma$ implies $\mathcal{M} \models \phi$ for every ϕ. Thus $\Gamma \models \phi$ for any ϕ, and so $\Gamma \models \perp$. Hence $\Gamma \vdash \perp$ by the theorem.

Moreover, by the proof of the theorem we get a model $\mathcal{N}$ of Γ whose elements are equivalence classes of terms in $\mathcal{L}'$. Thus $\overline{\overline{N}} \leq \overline{\overline{L}}'^{<\omega} = \overline{\overline{L}}' = \overline{\overline{L}} + \omega = \kappa + \omega$.

Corollary. (Compactness). If Γ is a set of sentences and every finite subset $\Delta \subseteq \Gamma$ has a model, then Γ also has a model.

Proof. If Γ does not have a model, then $\Gamma \vdash \perp$ by the "existence of models" corollary. But, for any ϕ, if $\Gamma \vdash \phi$ there is some finite $\Delta \subseteq \Gamma$ such that $\Delta \vdash \phi$, because proofs (syntactic deductions) are finitary. So there is some finite $\Delta \subseteq \Gamma$ such that $\Delta \vdash \perp$. Hence Δ has no models, a contradition to the hypothesis!

Example. Let $\mathcal{L} = \{0, +, ., s, <, exp, =\}$ be a first order language where *exp* is a binary realtion. Let $\mathbb{N} = (\omega, 0^{\mathbb{N}}, +^{\mathbb{N}}, .^{\mathbb{N}}, s^{\mathbb{N}}, <^{\mathbb{N}}, exp^{\mathbb{N}}, =^{\mathbb{N}})$ where $exp(x, y) = x^y$, and let Γ be the set of sentences that are true about arithmetic, *i.e.*, for $\mathbb{N}$ (or one could take Γ to be Peano's axioms for arithmetic). Add a new constant c to $\mathcal{L}$, obtaining $\mathcal{L}' = \mathcal{L} \cup \{c\}$, and consider

$$\Gamma' = \Gamma \cup \{s(s(\ldots(s(0)))) < c \mid n < \omega \ \&\ s(s(\ldots(s(0)))) \text{ is } s \text{ applied } n \text{ times to } 0\}.$$

The new sentences in $\Gamma' \setminus \Gamma$ say that $n < c$ for each $n < \omega$ because $n = 0 + 1 + 1 + \ldots + 1$ n times.

I claim that Γ' has a model. Proof: if $\Delta \subseteq \Gamma'$ is finite we have that $\Delta \setminus \Gamma$ is finite, so there is some $m_\Delta < \omega$ such that for each sentence of the form $n < c$ (*i.e.* of the form $s(s(\ldots(s(0)))) < c$ with n applications of s, formally) such that $n < c \in \Delta$ we have that $n \leq m_\Delta$. So Δ has a model, towit $\mathbb{N} = (\omega, 0^{\mathbb{N}}, +^{\mathbb{N}}, .^{\mathbb{N}}, s^{\mathbb{N}}, <^{\mathbb{N}}, exp^{\mathbb{N}}, =^{\mathbb{N}}, c_\Delta^{\mathbb{N}})$, where $c_\Delta^{\mathbb{N}}$ is any natural number greater than m_Δ. Thus by the Compactness Theorem Γ' has a model, and this model is a model for arithmetic which is non-standard because it has infinite natural numbers, and hence is not $\mathbb{N}$.

Example. ("Non-standard" analysis.) At the beginings of calculus, which it can reasonably be argued as the start of all modern mathematics, Newton and Liebniz and their followers talked and reasoned about infinitely small and infinitely large quantities in order to reach their results. In the 19^{th} century no one was able to supply a satisfactory formalization of these reasonings, and with time they fell out of favour, being replaced by Kronecker's ϵ-δ method which seemed "safer." Now, through the Compactness Theorem we can recover the possibility of using these original intuitions of Newton and Liebniz.

Let $\mathcal{L} = \{+, -, ., <, 0, 1\}$ and let T the set of true sentences about the reals with the expected structure, $(\mathbb{R}, +^{\mathbb{R}}, -^{\mathbb{R}}, .^{\mathbb{R}}, <^{\mathbb{R}}, 0^{\mathbb{R}}, 1^{\mathbb{R}})$, in this language. Let $\mathcal{L}' = \mathcal{L} \cup \{c\}$, where c is a new constant symbol, and let $\Gamma = T \cup \{0 < c < y \,\&\, y.(1{+}1\ldots{+}1) \mid 1{+}1{+}\ldots{+}1$ is the sum of 1 n times& $n < \omega\}$. As in the previous example Γ has a model, let us call it $\mathbb{R}^*$, in which there is some c such that $0 < c$ and $c < 1/n$ for each natural number n, that is such that c is infintessimal. Note that as $\forall x\, \exists y\, x.y = 1 \in T$ there are infinite elements in $\mathbb{R}^*$ as well.

Corollary to the Completeness Theorem**.** (*Extension by definitions.*) Suppose that $\mathcal{L}$ is a first order language, and Γ is a set of sentences in $\mathcal{L}$. (i) If $\phi(\bar{x})$ is a formula, let $\mathcal{L}' = \mathcal{L} \cup \{R_\phi\}$ and $\Gamma' = \Gamma \cup \{\forall \bar{x}\, (\phi(\bar{x}) \longleftrightarrow R_\phi(\bar{x}))\}$. Then for each $\psi \in \mathcal{L}$ we have $\Gamma' \models \psi$ if and only if $\Gamma \models \psi$. (ii) If $\phi(\bar{x}, y)$ is a formula and $\Gamma \vdash \forall \bar{x}\, \exists! y\, \phi(\bar{x}, y)$, let $\mathcal{L}' = \mathcal{L} \cup \{f_\phi\}$ and $\Gamma' = \Gamma \cup \{\forall \bar{x}\, \forall y\, (\phi(\bar{x}, y) \longleftrightarrow f_\phi(\bar{x}) = y)\}$. Then for each $\psi \in \mathcal{L}$ we have $\Gamma' \models \psi$ if and only if $\Gamma \models \psi$.

Proof. "$\Longleftarrow$" is trivial in each case. "$\Longrightarrow$": Let $\mathcal{M}$ be a model for Γ. It is trivial to expand $\mathcal{M}$ to a structure $\mathcal{M}'$ for $\mathcal{L}'$ which is a model, in case (i), for Γ' by interpreting R_ϕ to be the subset of $\mathcal{M}$ picked out by ϕ, *i.e.*, by taking $R_\phi^{\mathcal{M}'} = \{\bar{a} \in M \mid \mathcal{M} \models \phi(\bar{a})\}$, and, similarly, in case (ii) by taking $f_\phi^{\mathcal{M}'}(\bar{a}) =$ to be the unique b such that $\mathcal{M} \models \phi(\bar{a}, b)$. Now, if we have $\psi \in \mathcal{L}$ and $\Gamma' \models \psi$ then we have that $\mathcal{M}' \models \psi$ (by the Soundness Theorem), so $\mathcal{M} \models \psi$, and thus $\Gamma \models \psi$ (by the Completeness Theorem).

Note. By iterating we can make any finite sequence of extensions by definitions.

6.3. Models of fragments of set theory.

In this section we will look at models for various fragments of ZF. That is, we can show from the axioms of ZF that there are models for the fragments we consider. Actually each of the models we look at here are reasonably easy to define.

ZF$\setminus${Infinity} (= ZF-I).
$(V_\omega, \in) \models$ ZF - I. This is easy to see, checking one by one that the axioms of ZF apart from Infinity hold in $(V_\omega, \in)$. (It is also clear that $(V_\omega, \in) \not\models$ Infinity.) Hence ZF$\vdash$ Con(ZF-I).

ZF$\setminus${Replacement or Collection} (= Z).
Again we can easily check one by one that the axioms apart from replacement/collection hold in $(V_\lambda, \in)$ for any limit ordinal $\lambda > \omega$. So $(V_\lambda, \in) \models$ Z. On the other hand, as we saw in §1.7, $(V_{\omega+\omega}, \in) \not\models$ "Collection " because, for example, $\{\omega + n \mid n \in \omega\}$ is the image of a Function on the set ω, but is not a member of $V_{\omega+\omega}$. (Or because there are more than $\omega + \omega$-many non-order-isomorphic well-orders which are members of $V_{\omega+\omega}$, whence $V_{\omega+\omega} \not\models \Phi$ if Φ is the theorem which

says that every well-order is order-isomorphic to an ordinal.) (Summarizing: ZF$\vdash$ Con(Z).)

ZF$\setminus\{$Power ser$\}$ (= ZF-P).
If κ is a regular cardinal let $H_\kappa = \{x \mid \overline{\overline{\mathrm{TC}(x)}} < \kappa\}$. Again one can check that $H_\kappa \models$ ZF-P. Note that $H_{\omega_1} \not\models$ "power sets" because, for example, $V_{\omega+1} = \mathcal{P}(V_\omega) \notin H_{\omega_1}$ (equivalently, $V_{\omega+2} \not\subseteq H_{\omega_1}$). (Hence ZF$\vdash$ Con(ZF-P).)

It is reasonable to ask if, similarly, there are models other fragments of ZF. We have already seen in the first collection of exercises some of the dependencies one can prove between the axioms and this reduces somewhat the interesting questions. Nevertheless, we can still ask if there are models of ZF$\setminus\{$Foundation$\}$, and even if there are models of all of ZF. However in the next section we will see that Incompletenss Theorem shows that one cannot prove the existence of such models.

6.4. Incompleteness Theorem.

The *incompleteness theorem*, just as the completeness theorem, is a general theorem about collections of sentences in a first order language. It applies to a wide variety of collections of sentences (or *theories*) and doe not have anything specifically to do with set theory (apart from that ZF is one, amongst others, of the collections of sentences to which it applies).

Incompleteness Theorem. (Gödel, 1931) If T is a *reasonable* theory (in a language $\mathcal{L}$), then $T \nvdash \mathrm{Con}(T)$.

Linguistic note. The name *incompleteness* theorem comes from the fact that T is an *incomplete* theory, that is that it does not have the property that either it proves ϕ or it proves $\neg\phi$ for every sentence ϕ in $\mathcal{L}$. On the other hand, the *completeness* theorem gets its name because it shows that first order logical deduction is complete in the sense that if ϕ is true in each model of T, then there is a proof of ϕ from T in first order logic. Hence, although it sounds a little strange perhaps, there is no problem at all in saying for example that both the completeness theorem and the incompleteness theorem hold for such and such a theory T.

Reasonable here (in the statement of the incompleteness theorem): means (a) we can express Con(T) in $\mathcal{L}$, (b) T is consistent, (c) we can do enough combinatorics in T to code up logical manipulations within T and to decode them, and (d) T is recursive – where some explanation of what this last means needs to be given (informally: can be generated mechanically by an idealised computer).

We need (a) because obviously T can only prove sentences in $\mathcal{L}$, so without (a) the conclusion of the theorem would be incomprehensible. We also do need (b) in order to hope to be able to prove the theorem because if T is not consistent we have that $T\vdash \perp$, and so T can prove any sentence. Specifically, $T\vdash \mathrm{Con}(T)$. We explain (c) and (d) more completely below, but this vague formulation of (c) already gives a clue as to how one can hope to express Con(T) in $\mathcal{L}$.

Examples of 'reasonable' theories. If they are consistent, the collection of all true sentences about ($\mathbb{N}$, +, ., s, exp, 0, 1) and Peano arithmetic are reasonable. If they are consistent then ZF and ZF-I-P-Foundation are reasonable. Many theories of analysis in which one has the sine function, and so can define the natural numbers, are also reasonable if consistent.

Explication of (c) in the definition of reasonable. If T is reasonable we can prove from T that we can find representations of various syntactic things about $\mathcal{L}$ and T by terms in $\mathcal{L}$: the symbols of $\mathcal{L}$, the terms in $\mathcal{L}$, the formulas in $\mathcal{L}$ and the proofs from T in $\mathcal{L}$. If a is a piece of syntax we write $\#(a)$ for the term which represents a. At times we call these terms *codes*.

Also, we can find a formula $Pr(x, y)$ and a definable function $Subst(x, y)$ in $\mathcal{L}$ such that d is a Proof of ϕ of T if and only if $T\vdash Pr(\#(d), \#(\phi))$, and $\phi[a/x] = \psi$ if and only if $T\vdash Subst(\#(\phi), a) = \#(\psi)$.

First of all suppose that T is reasonable in this sense, and let us see how we can prove the theorem. Afterwards we will see how to do the coding for some specific theories.

Proof of the Incompleteness Theorem. Observe that if $\phi(x)$ is a formula we have

$$T\vdash \text{“}Subst(\#(\phi), \#(\phi)) = \#(\psi)\text{”}$$

if $\psi = \phi[\#(\phi)/x]$, that is, if ψ is the formula that we obtain if we substitute the term which represents ϕ itself for the free variable x of ϕ. Consider the formula $\neg\exists y\, Pr(y, Subst(z, z)), = \chi(z)$, let us say. Then, for any formula ϕ, we have that $T\vdash \chi(\#(\phi))$ if and only if there is no proof of $\phi(\#(\phi)$ from T. So, because χ itself is a formula, we have $T\vdash \chi(\#(\chi))$ if and only if there is no proof of $\chi(\#(\chi)$ from T.

But if $T\vdash \chi(\#(\chi))$ one has d proves $\chi(\#(\chi))$ for some proof d, and thus $T\vdash Pr(\#(d), \#(\chi(\#(\chi)))$. So $T\vdash \exists y Pr(y, \#(\chi(\#(\chi)))$. Hence $T\vdash \neg(\chi(\#(\chi)))$. Thus T is inconsistent.

And if $T\vdash \neg(\chi(\#(\chi)))$ we have that $T\vdash \exists y Pr(y, \#(\chi(\#(\chi)))$. Assuming that T is consistent there is a proof d such that $T\vdash Pr(\#(d), \#(\chi(\#(\chi)))$. So $T\vdash \chi(\#(\chi))$ and again we have that T is inconsistent.

If we write $\mathrm{Con}(T)$ for $\neg\exists y\, \mathrm{Pr}(\#(y), \#(z \neq z))$ and if we formalize the argumentation in T we have that $T\vdash \mathrm{Con}(T) \longrightarrow (\chi(\#(\chi))$. Hence $T\vdash \mathrm{Con}(T)$ implies $T\vdash (\chi(\#(\chi))$, and thus T is not consistent. (End of Proof of the Theorem.)

T must be recursive because we need to have that “d is a proof of ϕ from T” and “$\phi[a/x] = \psi$” are recursive, and so we have that there are formulas in $\mathcal{L}$ that define Pr and $Subst$, and because in order to do the formalization we also need to know that for assertions about proofs and substitutions we can use the formulas in $\mathcal{L}$ to write equivalent sentences in $\mathcal{L}$.

Codification.

Still to be done in this account of the incompleteness theorem:
- Say exactly what manipulations we need to be able to do in (c).
- Explain how to do the formalization of the argumentation in T (above I just say “one can”).
- Explain why I do this coding in a recursive way/explain more precisely why we need T to be recursive.
- Give references. (*E.g.*, Sam Buss's “Introduction to Proof Theory” in the “Handbook of Proof Theory” − also available from Buss's web-site at

`http://math.ucsd.edu/~sbuss/ResearchWeb/handbookII/index.html//`.)

Let $\mathcal{L} = \{ f^i \mid i < \omega \} \cup \{ R^i \mid i < \omega \} \cup \{=\}$. Because formulas are finite strings of symbols and proofs are finite trees of formulas we need that T gives us the ability to manipulate strings of codes for symbols and codes for strings of symbols. Thus one part of the definition of “reasonable” is that we can define functions f_n for each $n < \omega$ and π^n_m for each $m < n < \omega$ in $\mathcal{L}$ such that we can show from T that $\pi^n_m(f_n(x_0, \ldots, x_{n-1})) = x_m$. Thus f_n functions as if it allows us to make “n-tuples” and π^n_m functions as if it allows us to recover the m^{th} element of an “n-tuple” for each $n < \omega$ and $m < n$. In order to avoid confusion it is good to ask that $\mathrm{rge}(f_n) \cap \mathrm{rge}(f_m) = \emptyset$ if $m \neq n$. [*Do I* really *need this point below?*]

We define $\#(t)$, $\#(\phi)$ for terms t and formulas $\phi \in \mathcal{L}$ by induction on the structure of t, ϕ.

So let $\langle t_i \mid i < \omega \rangle$ be terms in $\mathcal{L}$. Let $\#(\&) = t_1$, $\#(\neg) = t_3$, $\#(\exists) = t_5$, $\#(() = t_7$, $\#()) = t_9$, $\#(,) = t_{11}$ $\#(=) = t_{13}$, $\#(v_i) = t_{2i}$, $\#(f^i) = t_{4i+15}$, and $\#(R^i) = t_{4i+17}$ for $i < \omega$.

We have already seen that $\#(v_i) = t_{2i}$. If $i < \omega$, f^i is a k-ary function symbol and $s_0, \ldots, s_{k-1}$ are terms in $\mathcal{L}$, we have $\#(f^i(v_{j_0}, \ldots, v_{j_{k-1}})) = f_{k+3}(t_{4i+15}, t_7, \#(s_{2j_0}), \ldots, \#(s_{2j_{k-1}}), t_9)$.

If $i < \omega$, R^i is a k-ary relation symbol and $s_0, \ldots, s_{k-1}$ are terms in $\mathcal{L}$ we have $\#(R^i(v_{j_0}, \ldots, v_{j_{k-1}}))$ $= f_{k+3}(t_{4i+17}, t_7, \#(s_{2j_0}), \ldots, \#(s_{2j_{k-1}}), t_9)$. If ϕ is a formula we have $\#(\neg\phi) = f_4(t_3, t_7, \#(\phi), t_9)$. If ϕ and ψ are formulas we have $\#(\phi \,\&\, \psi) = f_3(\#(\phi), t_1, \#(\psi))$. and if $\phi(x)$ is a formula free variable x we have $\#(\exists x \phi(x)) = f_2(t_5, \#(\phi(x)))$.

Thus for a formula $\phi \in \mathcal{L}$ we have $\#(\phi)$ is a term in $\mathcal{L}$. If d is a proof of ϕ from T, then d is a finite tree, hence a string of strings, and thus we can define $\#(d)$, a term in $\mathcal{L}$ as well.

Examples. In ZF(−I−P−Fnd) we can simply use $f_n(x_0, \ldots, x_{n-1}) = (x_0, \ldots, x_{n-1})$ and $t_i = i$ for $i < \omega$, because if $x_0, \ldots, x_{n-1} \in V_\omega$ we have that $(x_0, \ldots, x_{n-1}) \in V_\omega$.

In Peano arithmetic we need to do a little more work. $f_n(x_0, \ldots, x_{n-1}) = \ldots$.
Still to finish writing up this part of the proof.

6.5. Applications of the Incompleteness Theorem to ZF.

Theorem. If ZF is consistent we cannot prove, from ZF alone, that there is a model of ZF.

Proof. It is easy to see that ZF is "reasonable" because we can carry out manipulations of finite strings in ZF and the axioms are a recursive set. So the theorem is immediate from the Incompleteness Theorem. (*Hmmm. I should explain exactly how the axiomas are a set and, thus can be recursice or not.*)

Theorem. ZF$\vdash$ Con(ZF−{Foundation}) $\longrightarrow$ Con(ZF).

Proof. Let $\mathcal{U}$ be a model of ZF−{Foundation}. We can construct the hierarchy V_αs without using foundation – recall that we saw after the construction of the hierarchy of V_αs in §2.7 that the axiom of foundation is equivalent to saying that $\forall x\, \exists \alpha \in \mathrm{On}\ x \in V_\alpha$. So construct the hierarchy inside $\mathcal{U}$ and let $\mathcal{V}$ be the result. Using the fact, which can be shown simply by induction on the structure of formulas, that if $x \in \mathcal{V}$ and $\phi(y)$ is a formula, then $x \cap \{\, y \mid \phi(y)\,\} \cap \mathcal{V} = x \cap \{\, y \mid \phi^V(y)\,\} \cap \mathcal{U}$, where $\phi^V(y)$ is ϕ with all the quantifiers running over $\mathcal{V}$, we can check that $\mathcal{V} \models$ ZF. Thus Con(ZF−{Foundation}) implies Con(ZF). Hence we have shown that ZF$\vdash$ Con(ZF−{Foundation})∎ $\longrightarrow$ Con(ZF).

Corollary. We cannot show from ZF that there is a model of ZF−{Foundation}.

Proof. Immediate from the previous theorem and the incompleteness theorem.

6.6. Constructability and forcing. Now we are in a position to explain more exactly what we meant by the observations at the end of §4.3, for example, that we can neither prove AC from ZF nor that AC is false. We have:

Teorema. (Gödel, 1938) Con(ZF) $\Longrightarrow$ Con(ZF + AC + GCH).

Teorema. (Cohen, 1963) Con(ZF) $\Longrightarrow$ Con(ZF + ¬AC), and Con(ZF) $\Longrightarrow$ Con(ZF + AC + $2^{\aleph_0} = \aleph_{\alpha+1}$) for some ordinal α.

But expressing things in this way, as about consistency, perhaps gives a misleading impression about the subject matter of set theory in general and of these theorems in particular. It is better to understand these theorems as saying that if there is a model in which such and such properties are true there is another one in which so and so other properties are true. This is an example of a central theme in set theory, because the theory itself concerns, after the initial material of chapters 1-4, the relations betweeen models of the theory (ZF) and of things that hold (or occur) in specific collections of models of ZF. (Albeit, clearly, that set theory is interested in properties of specific models (collections of size 1) and in properties true in the collection of all models of ZF(C) - the syntactic consequences of the theory, by the completeness theorem.)

This should not, on reflection, come as a great surprise: for example, group theory is interested in the relations between groups as well as the properties of specific collections of groups, and the theory of topological spaces is interested in relations between space as well as the properties shared within specific collections of spaces.

Only one should remember that in these last examples the models of the theory are groups and spaces, respectively, while in the case of set theory the models are not called 'sets' but 'models of set theory' or, as we said in §1.4, 'universes' (models are universes because if $\mathcal{M} \models \text{ZF}$ we have $\mathcal{M} \models \forall x\, \exists \alpha \in \text{On}^{\mathcal{M}}\ x \in V_\alpha^{\mathcal{M}}$).

Constructability. The basic idea in the proof of Gödel's theorem is similar to the idea in the proof of the theorem at the end of §6.5 that if we have a model of ZF−Foundation we can obtain a model of ZF. This time we begin with an arbitrary model $\mathcal{M}$ of ZF and and define inside $\mathcal{M}$ a thinner model, L, with $L \subseteq \mathcal{M}$, $\text{On}^L = \text{On}^{\mathcal{M}}$ and, in this case, with ZF + AC + GCH. We do this by defining L as a hierarchy, like V, but at the successor levels $\alpha + 1$ adding only the subsets definable in the level α instead of all of the subsets of it.

Definition. By recursion on $\text{On}^{\mathcal{M}}$ define $L_0 = \emptyset$, $L_\lambda = \bigcup\{ L_\alpha \mid \alpha < \lambda\}$ for limit ordinals λ, and $L_{\alpha+1} = \{ x \subseteq L_\alpha \mid \exists \phi(y_0, \ldots, y_{n-1}, z) \in$ LST such that

$$x = \{ a \mid \exists b_0, \ldots, b_{n-1} \in L_\alpha (L_\alpha, \in \cap (L_\alpha \times L_\alpha)) \models \phi(b_0, \ldots, b_{n-1}, a) \} \}.$$

Let $L = \bigcup\{ L_\alpha \mid \alpha \in \text{On}^{\mathcal{M}} \}$.

Gödel's theorem is that for any $\mathcal{M}$ we have that $L^{\mathcal{M}} \models$ ZF + AC + GCH. And also that if $\mathcal{M}$, $\mathcal{N}$ are models of ZF with $\text{On}^{\mathcal{M}} \subseteq \text{On}^{\mathcal{N}}$, we have that $L^{\mathcal{N}}_{\text{On}^{\mathcal{M}}} = L^{\mathcal{M}}$, so, especially, $L \models V = L$. But the proof that the model L has these properties is considerably more complicated than the proof in §6.5. We will look at L and these proofs in the next chapter.

Forcing. In order to show that we can get from a model of ZF + AC + GCH to a model of ZF + AC in which, for example, CH fails we use a technique, forcing, in order to expand the model to get to, in this case, a model with more subsets of ω. The technique is very similar to making an algebraic extension to a field in algebra: we make a formal extension by an unknown and then take a quotient to obtain a field with the properties we want. (In fact, one can show that this is not only a good analogy, but also that there is a technical – category theory – sense in which forcing and taking algebraic extensions are instances of the same categorical construction.)

The final two chapters of Kunen's book[1] and Kanamori's book[2] are two good places for more information about forcing. (And they are good on 'next steps' in set theory more generally.) *I hope that we will get to something about forcing towards the end of the course.*

6.7. Models all of ZF.

[1] Ken Kunen, *Set theory: an introduction to independence proofs*, North-Holland, 1980.

[2] Aki Kanamori, *The Higher Infinite*, Springer, 1996.

Note that the consequence of the incompleteness theorem in §6.5 does not say that there are no models of ZF, only that one cannot show from the axioms of ZF alone that such models exist. We have already seen in §6.3 that $V_\lambda \models \mathrm{ZF} - \{\text{Collection}\}$ for any limit ordinal λ with $\omega < \lambda$ and that $V_\omega \models \mathrm{ZF} - \{\text{Infinity}\}$.

One can check by the same proofs the following.

Theorem. If κ is a cardinal which is strongly inaccessible, we have $V_\kappa \models \mathrm{ZF}$.

Observation. (Skolem's 'paradox'). If ZF is consistent then by the last corollary of §6.2 it has a model, let us say $\mathcal{M}$, which is *countable*. Thus $\mathcal{P}(\omega)^{\mathcal{M}}$, $\omega_1^{\mathcal{M}}$, ... are countable. This is not a real paradox because although we can enumerate, for example, $\mathcal{P}(\omega)^{\mathcal{M}}$ outside of the model $\mathcal{M}$ and see that this set, which is the set of all subsets of ω for "people" "living" inside $\mathcal{M}$, is countable, nevertheless inside $\mathcal{M}$ there is no bijection between ω and $\mathcal{P}(\omega)^{\mathcal{M}}$, and, so, inside $\mathcal{M}$ one cannot countably enumerate $\mathcal{P}(\omega)$.

Although, by the incompleteness theorem applied to ZF one cannot prove from ZF that there are any strongly inaccessible cardinals, the statement that such cardinals do exist is an interesting one and sometimes useful. Even though formally this (and similar suppositions, such as that there are weakly inaccessible cardinals) are strengthenings of ZF, in *practice* – the most important point – there seems no reason to suppose that they are much different from, for example, strengthening Peano arithmetic with the assumption that PA is not inconsistent. On the contrary, in fact, the exploration of suppositions of this type forms a large productive part of set theory and, so, a productive part of mathematics. (See, particularly, Kanamori's book.) We close this chapter with an example of this phenomenon concerning cardinal exponentiation.

Definition. For any $\delta < \omega_1$, the statement that a cardinal κ is $\kappa+\delta+1$-*strong* is a strengthening of saying that κ is weakly compact. (It means that there is an injection $i : V \longrightarrow \mathcal{M}$ such that $\mathcal{M} \subseteq V$, $\mathrm{On}^{\mathcal{M}} = \mathrm{On}^V$, if $\phi(x) \in LTC$ and $a \in V$ one has that $V \models \phi(a)$ if and only if $\mathcal{M} \models \phi(i(a))$, $i(\alpha) = \alpha$ for each $\alpha < \kappa$, $\kappa < i(\kappa)$ and $V_{\kappa+\delta+1} \subseteq \mathcal{M}$. Don't expect to understand now why this definition is more or less natural, nor even how to formulate it as a statement about sets rather than about classes as done here. See Kanamori's book for these details.)

The reason for introducing the definition is that we can use it to give the correct statement of the following theorem (stated at the end of §4.4) – of which the proof is a subtle argument using forcing.

Theorem. (Gitik-Magidor, 1989) Let $\omega \leq \delta < \omega_1$. Con(ZF + AC + $\exists\kappa$ s.t. κ is $\kappa+\delta+1$-strong) $\Longrightarrow$ Con(ZF + AC + $2^{\aleph_n} = \aleph_{n+1}$ for all $n < \omega$, and $2^{\aleph_\omega} = \aleph_{\omega+\delta+1}$.)