Logic

Alberto Santini

June 12, 2011

Contents

1 First order logic

1.1 First order languages

In first order logic we will use a particular language — that is, a set of symbols — and will usually deal with finite strings of those symbols that obey a precise syntax.

Definition 1.1. The symbols are usually divided in two main categories as follows:

Logical symbols Those are:

Parentheses (and)

Connectives $\neg$ and $\wedge$

Variables $v_1, v_2, \dots$

Equality symbol $=$

Parameter symbols Those are:

Existential quantifier $\exists$

Predicate symbols Symbols in the form $P_1^n, P_2^n, \dots$, for everery $n \in \mathbb{N}^+$; n is called the ariety of the symbol

Function symbols Symbols in the form $f_1^n, f_2^n, \dots$, for every $n \in \mathbb{N}$; again n is called the ariety of the symbol. The 0-ary functions are called **constants** and are denoted with the shorter notation $c_1, c_2, \dots$

If we call ar the function that assigns to each predicate and function its ariety, we usually call the triple

$$\sigma = (\{P_i^n\}, \{f_i^n\}, \text{ar}) \tag{1.1}$$

of all the predicate symbols, all the function symbols and the ariety function, the **signature** of the language.

Definition 1.2. We call a finite string of symbols an **expression**. An expression t is a:

First order term If it is either a variable v or in the form $f^n(v_1, \dots, v_n)$ where f^n is an n-ary function and $v_1, \dots, v_n$ are variables.

k-th order term If it is in the form $f^n(t_1, \dots, t_n)$ where f^n is an n-ary function and $t_1, \dots, t_n$ are terms of order $< k$.

In general we will say that t is a **term**, without regard for its order.

Definition 1.3. An **atomic formula** φ is an expression in the form $P^n(t_1, \dots t_n)$, where P^n is an n-ary predicate and $t_1, \dots, t_n$ are terms. We also include equality expressions $= (t_1, t_2)$ (or its common shorthand $t_1 = t_2$) among atomic formulas. An expression ϕ is a:

First order formula If it's in one of the following forms: $\neg\varphi_1$, $\varphi_1 \wedge \varphi_2$, $\exists v\ \varphi_1$, where φ_1, φ_2 are atomic formulas.

k-th order formula If it's in one of the following forms: $\neg\phi_1$, $\phi_1 \wedge \phi_2$, $\exists v\ \phi_1$, where ϕ_1, ϕ_2 are formulas of order $< k$.

In general we will say that ϕ is a **formula**, disregarding its order.

Definition 1.4. Given an atomic formula ϕ, we say that the variable v occurs **free** (o.f.) in ϕ iff it occurs in ϕ. Given a formula ψ in the form $\neg\phi$, we say that v o.f. in ψ iff it o.f. in ϕ. If ψ is in the form $\phi_1 \wedge \phi_2$, we say that v o.f. in ψ iff it occurs free in either ϕ_1 or ϕ_2. Finally, if ψ is in the form $\exists w\ \psi$, we say that v o.f. in ψ iff it o.f. in ϕ and it doesn't occur in the quantifier — e.g. it is a different symbol from w. We call a formula with no free variables a **sentence**.

Definition 1.5. We will use the following shorthands:

1. $v \neq w$ for $\neg(v = w)$
2. $\phi \vee \psi$ for $\neg(\neg\phi \ \wedge \ \neg\psi)$
3. $\phi \to \psi$ for $\neg\phi \ \vee \ \psi$
4. $\phi \leftrightarrow \psi$ for $\phi \to \psi \ \wedge \ \psi \to \phi$
5. $\forall v \ \phi$ for $\neg\exists v \ \neg\phi$

Maybe the most useful thing we can do with formulas is to give them a truth-value, since we're often interested in working with true statements. In order to do this, we will first need to give an interpretation of our language that can assign a meaning to the formulas.

Definition 1.6. A **structure** is a triple

$$\mathcal{S} = (S, \sigma, I) \tag{1.2}$$

Where S is a non-empty set called the **universe**, σ is the signature of some language and I is an **interpretation function** from the set of all predicative and functional symbols in σ to the power-set $\mathcal{P}(S)$. In particular, a predicative symbol P with $\text{ar}(P) = n$ is interpreted as a relation

$$I(P) \subseteq S^n \tag{1.3}$$

while a functional symbol f with $\text{ar}(f) = n$ is interpreted as a function

$$I(f) : S^n \to S \tag{1.4}$$

By convention, 0-ary functions — that is, constants — are simply assigned to particular elements of S. Also, to stress the fact that $\mathcal{S}$ depends on σ, we will call it a σ-structure.

Definition 1.7. We need a way to assign terms a semantic meaning. This is done replacing variables and their functions with objects of the universe; this process is called an **evaluation** of the term and uses an evaluation function μ that assigns elements of S to variables. Given a first order term t, if t is a variable v, then it's evaluation is just $\mu(v)$; if it is a function $f(v_1, \dots, v_n)$ then its evaluation is $(I(f) \circ \mu)\,(v_1, \dots, v_n)$. If t is a k-th order term $f(t_1, \dots, t_n)$ and called $s_1, \dots, s_n$ the evaluations of $t_1, \dots, t_n$, the evaluation of t is just $I(f)(s_1, \dots, s_n)$.

Definition 1.8. Fixed an evaluation μ, we are now ready to assign a **truth value** (that is: T for true, or F for false) to each formula.

Atomic formulas An atomic formula in the form $P(t_1, \dots, t_n)$ has truth value T iff the evaluations $(s_1, \dots, s_n) \in I(P) \subset S^n$. An atomic formula in the form $t_1 = t_2$ has truth value T iff their evaluations are equal in S: $s_1 = s_2$.

Formulas with connectives The truth value for formulas of the kind $\phi \wedge \psi$ or $\neg\phi$ is given by the truth tables:

ϕ	ψ	$\phi \wedge \psi$
T	T	T
T	F	F
F	T	F
F	F	F

ϕ	$\neg\phi$
T	F
F	T

Formulas with $\exists$ A formula of the kind $\exists v\ \phi$ is true iff there is an evaluation ν such that ϕ is true according to ν and ν and μ are the same for every variable but for v.

When we want to say that we are evaluating a formula ϕ with n free variables, by substituting its free variables with the elements $s_1 = \mu(v_1), \ldots, s_n = \mu(v_n)$ we simply write $\phi(s_1, \ldots, s_n)$.

Definition 1.9. If a sentence ϕ is true given the structure $\mathcal{S}$ (or, as it is usual to say, in the interpretation $\mathcal{S}$), no matter what the evaluation μ is, we say that $\mathcal{S}$ **satisfies** ϕ and we write $\mathcal{S} \models \phi$. If ϕ is true for every interpretation, then it's called a **tautology**.

1.2 Propositional logic

We will now give a way to infer a conclusion from some premises, in a purely syntactical way. In the following, the premise will be written above an horizontal straight line and the conclusion below it. To denote that ψ can be inferred by ϕ we will write $\phi \vdash \psi$.

Definition 1.10. Our main rule of inference will be the one of **modus ponens**:

$$\frac{\begin{array}{c}\phi \to \psi \\ \phi\end{array}}{\psi} \tag{1.5}$$

Definition 1.11. We fix a family Λ of formulas, called the **logical axioms**, and for other families of formulas Γ, we say that the **theorems** of Γ are the formulas that may be inferred by formulas in $\Lambda \cup \Gamma$ by using modus ponens a finite number of times. If ψ is a theorem of Γ, write $\Gamma \vdash \psi$. A **deduction** for ψ from Γ is a sequence $\langle \phi_1, \ldots, \phi_n \rangle$ of formulas such that

1. $\phi_n = \psi$
2. $\forall i \leq n$ either $\phi_i \in \Lambda \cup \Gamma$ or ϕ_i is a theorem of $\{\phi_1, \ldots, \phi_{i-1}\}$

When $\Gamma = \emptyset$ we simply write $\vdash \phi$.

Theorem 1.12. $\Gamma \vdash \psi$ *iff there is a deduction for* ψ *from* Γ.

We will now try to describe the family Λ of logical axioms that we'll deal with.

Definition 1.13. A formula ϕ is said to be a **generalization** of ψ if it is in the form

$$\forall v_1 \ \cdots \ \forall v_n \ \psi \tag{1.6}$$

for some $n \in \mathbb{N}$ and some variables $v_1, \ldots, v_n$.

Definition 1.14. Given a formula ϕ, a variable v and a term t, we say that t is **substitutable** for v in ϕ in any of the following cases:

1. ϕ is atomic
2. ϕ is in the form $\neg\psi$ and t is substitutable for v in ψ
3. ϕ is in the form $\psi_1 \wedge \psi_2$ and t is substitutable for v in both ψ_1 and ψ_2
4. ϕ is in the form $\exists w \ \psi$ and either
   (a) v doesn't occur free in ψ or
   (b) t is substitutable for v in ψ and w doesn't occur free in t (notice that it makes sense to say that a variable occurs free in a term since a term is, in particular, a formula)

If t is substitutable for v in ϕ, the formula obtained by actually substituting each occurrence of v with t is denoted by $\phi(t/x)$ and is called a **substitution instance** of ϕ.

Observation 1.15. Any substitution instance of a tautology is still a tautology.

Definition 1.16. The family Λ of logical axioms that we'll use consists of all the generalizations of the following kind of formulas:

1. Tautologies
2. Formulas of the form $\forall v \ \phi \to \phi(w/v)$ where w is substitutable for v in ϕ
3. Formulas of the form $(\forall v \ (\phi \to \psi)) \to ((\forall v \ \phi) \to (\forall v \ \psi))$ where v doesn't occur free in neither ϕ nor ψ
4. Formulas of the form $\phi \to (\forall v \ \phi)$ where v doesn't occur free in ϕ
5. Formulas of the form $v = v$
6. Formulas of the form $(v = w) \to (\phi \to \phi(w/v))$ where ϕ is atomic

Definition 1.17. The family Γ is said to be **consistent** if there is no formula ϕ such that both $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$. Otherwise we say that Γ is **inconsistent**.

Theorem 1.18. *Given two formulas ϕ, ψ such that $\Gamma \vdash \phi$ and $\Gamma \vdash \psi$, then $\Gamma \vdash (\phi \wedge \psi)$.*

Proof. Notice that $\phi \to (\psi \to (\phi \wedge \psi))$ is a tautology and so $\Gamma \vdash (\phi \to (\psi \to (\phi \wedge \psi)))$. Hence, by modus ponens we have that $\Gamma \vdash (\psi \to (\phi \wedge \psi))$ and so, using modus ponens again $\Gamma \vdash (\phi \wedge \psi)$. □

Theorem 1.19 (The explosion principle)**.** *Given an inconsistent family Γ, for every formula ψ it holds that $\Gamma \vdash \psi$.*

Proof. Let ϕ be the formula such that $\Gamma \vdash \phi$ and $\Gamma \vdash \neg\phi$. By theorem 1.18 it holds that $\Gamma \vdash (\phi \wedge \neg\phi)$. Now notice that $(\phi \wedge \neg\phi) \to \psi$ is a tautology for every formula ψ. So $\Gamma \vdash (\phi \wedge \neg\phi) \to \psi$, and hence by modus ponens $\Gamma \vdash \psi$. □

Theorem 1.20. *Given a formula ϕ such that $\Gamma \vdash \phi$ and a variable v that isn't free in any formula of Γ, then $\Gamma \vdash (\forall v\ \phi)$*

Proof. By definition of $\vdash$, to prove the theorem it is sufficient to prove that the family

$$\mathscr{F} = \{\phi \ :\ \Gamma \vdash (\forall v\ \phi)\} \tag{1.7}$$

contains $\Lambda \cup \Gamma$ and is closed by modus ponens, i.e. if ϕ and ψ are such that ψ and $\psi \to \phi$ are in $\mathscr{F}$ then also ϕ is in $\mathscr{F}$.

It contains the axioms: $\mathscr{F} \supseteq \Lambda$ Given a formula $\lambda \in \Lambda$, by construction of the axioms as generalizations, we have that $(\forall v\ \lambda) \in \Lambda$, so $\vdash (\forall v\ \lambda)$ and hence in particular $\Gamma \vdash (\forall v\ \lambda)$.

It contains the premises: $\mathscr{F} \supseteq \Gamma$ Given a formula $\gamma \in \Gamma$, by axiom rule 4 and since v doesn't occur free in γ by hypothesis, we have that $\gamma \to (\forall v\ \gamma)$ and so, by modus ponens $\Gamma \vdash (\forall v\ \gamma)$.

It is closed under modus ponens Given a formula ψ obtained by the modus ponens:

$$\begin{array}{c} \phi \to \psi \\ \phi \\ \hline \psi \end{array} \tag{1.8}$$

since $(\phi \to \psi)$ and ϕ are in $\Lambda \cup \Gamma$, by the previous steps we have that

$$\Gamma \vdash (\forall v\ (\phi \to \psi)) \tag{1.9}$$
$$\Gamma \vdash (\forall v\ \phi) \tag{1.10}$$

so by axiom rule 3 we get $(\forall v\ \phi) \to (\forall v\ \psi)$; using modus ponens with (1.9) we obtain

$$\Gamma \vdash ((\forall v\ \phi) \to (\forall v\ \psi)) \tag{1.11}$$

and using it again with (1.10) we finally get

$$\Gamma \vdash (\forall v\ \psi) \tag{1.12}$$

□

Observation 1.21. If $\Gamma \vdash (\gamma \to \phi)$ then $(\Gamma \cup \{\gamma\}) \vdash \phi$, just by definition of $\vdash$ and modus ponens.

Theorem 1.22 (Deduction theorem)**.** *The converse of the previous observation is true, that is: if* $(\Gamma \cup \{\gamma\}) \vdash \phi$ *then* $\Gamma \vdash (\gamma \to \phi)$*.*

Proof. We'll proceed step by step:

Step 1 The thesis holds if $\phi = \gamma$. In fact, trivially $\phi \to \phi$ is a tautology and so $\Gamma \vdash (\phi \to \phi)$.

Step 2 The thesis holds if $\phi \in \Lambda \cup \Gamma$. Just notice that $\phi \to (\gamma \to \phi)$ is a tautology and so by modus ponens we have that $\Gamma \vdash (\gamma \to \phi)$.

Step 3 The last way in which ϕ may be inferred by $\Gamma \cup \{\gamma\}$ is that it is given by modus ponens

$$\frac{\begin{array}{c}\psi \to \phi \\ \psi\end{array}}{\phi} \tag{1.13}$$

where $\Gamma \vdash \psi$. Now, by the previous steps we have that

$$\Gamma \vdash (\gamma \to \psi) \tag{1.14}$$

$$\Gamma \vdash (\gamma \to (\psi \to \phi)) \tag{1.15}$$

and so by theorem 1.18

$$\Gamma \vdash ((\gamma \to \psi) \wedge (\gamma \to (\psi \to \phi)) \tag{1.16}$$

Now notice that

$$((\gamma \to \psi) \wedge (\gamma \to (\psi \to \phi)) \to (\gamma \to \phi) \tag{1.17}$$

is a tautology and so, by modus ponens we have $\Gamma \vdash (\gamma \to \phi)$.

□

Theorem 1.23 (Reductio ad absurdum)**.** $\Gamma \cup \{\gamma\}$ *is inconsistent iff* $\Gamma \vdash \neg\gamma$*.*

Proof. If $\Gamma \vdash \neg\gamma$ then also $\Gamma \cup \{\gamma\} \vdash \neg\gamma$, but since trivially $\Gamma \cup \{\gamma\} \vdash \gamma$ we have realized the definition of inconsistency. On the other hand, if $\Gamma \cup \{\gamma\}$ is inconsistent and ϕ is a formula such that

$$\Gamma \cup \{\gamma\} \vdash \phi \text{ and } \Gamma \cup \{\gamma\} \vdash \neg\phi \tag{1.18}$$

then, by theorem 1.22 we have that

$$\Gamma \vdash (\gamma \to \phi) \text{ and } \Gamma \vdash (\gamma \to \neg\phi) \tag{1.19}$$

and so by theorem 1.18

$$\Gamma \vdash ((\gamma \to \phi) \wedge (\gamma \to \neg\phi)) \tag{1.20}$$

Now notice that
$$(\gamma \to \phi) \wedge (\gamma \to \neg\phi) \to \neg\gamma \tag{1.21}$$
is a tautology and so, applying modus ponens we get $\Gamma \vdash \neg\gamma$. □

Theorem 1.24 (Compactness theorem for propositional logic)**.** *If $\Gamma \vdash \phi$ then there is a finite subfamily $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \vdash \phi$.*

Proof. The thesis immediately follow by theorem 1.12 and by recalling that in the definition of deduction only finite many formulas may be involved. □

Corollary 1.25. *If Γ is inconsistent then there is a finite subfamily $\Gamma_0 \subseteq \Gamma$ that is inconsistent.*

Proof. If ϕ is the formula that realizes inconsistency, just apply the previous theorem to ϕ and $\neg\phi$ and take the union of the two resulting finite families. □

2 Basic model theory

To give a more meaningful version of the compactness theorem — and to explain why it is called so — we will now dig deeper into the semantics of first order languages.

2.1 Models, completeness and compactness theorems

We will start with a lot of positive results that are also fundamental in logic: the soundness, the completeness and finally the compactness theorem.

Definition 2.1. Given a language with signature σ and two σ-structures $\mathcal{S} = (S, \sigma, I)$ and $\mathcal{T} = (T, \sigma, J)$, a **$\sigma$-embedding** is an injective function $\eta : \mathcal{S} \to \mathcal{T}$ that preserves the interpretation of symbols, i.e.

1. $(s_1, \dots, s_n) \in I(P)$ iff $(\eta(s_1), \dots, \eta(s_n)) \in J(P)$ for all the predicate symbols P of ariety n and all $(s_1, \dots, s_n) \in S^n$
2. $\eta(I(f)(s_1, \dots, s_n)) = J(f)(\eta(s_1), \dots, \eta(s_n))$ for all the functional symbols f of ariety n and all $(s_1, \dots, s_n) \in S^n$; in particular, for constants we have that $\eta(I(c)) = J(c)$

A **σ-isomorphism** is a bijective σ-embedding. If $S \subseteq T$ and the inclusion is a σ-embedding, we say that $\mathcal{S}$ is a **substructure** of $\mathcal{T}$ (and that $\mathcal{T}$ is an **extension** of $\mathcal{S}$).

Definition 2.2. A σ-embedding $\eta : \mathcal{S} \to \mathcal{T}$ is called **elementary** if for every $n \geq 0$, for every n-ary formula ϕ and for every $(s_1, \ldots, s_n) \in S^n$, $\mathcal{S} \models \phi(s_1, \ldots, s_n)$ iff $\mathcal{T} \models \phi(\eta(s_1), \ldots, \eta(s_n))$. If $\mathcal{S}$ is a substructure of $\mathcal{T}$ and the inclusion is an elementary embedding, in particular we call $\mathcal{S}$ an **elementary substructure** of $\mathcal{T}$ and $\mathcal{T}$ an **elementary extension** of $\mathcal{S}$.

Definition 2.3. Given a σ-structure $\mathcal{S}$, we call its **full theory** $\mathrm{Th}(\mathcal{S})$ the collection of all formulas ϕ such that $\mathcal{S} \models \phi$. We say that two σ-structures $\mathcal{S}$ and $\mathcal{T}$ are equivalent if $\mathrm{Th}(\mathcal{S}) = \mathrm{Th}(\mathcal{T})$ and we write $\mathcal{S} \equiv \mathcal{T}$. This is the same to say that $\mathcal{S} \models \phi$ iff $\mathcal{T} \models \phi$.

Definition 2.4. Given a theory T, we call its **deductive closure** the collection of all formulas ϕ such that $T \vdash \phi$.

Theorem 2.5. *Two isomorphic σ-structures are equivalent.*

Definition 2.6. Given a language of signature σ we call a set T of formulas a **theory**. Given a σ-structure $\mathcal{M}$, we write $\mathcal{M} \models T$ if $\mathcal{M} \models \phi$ for all the formulas ϕ in T. In this case, we call $\mathcal{M}$ a **model** for T and we say that T is **satisfiable**. Also, given a formula ψ, we say that it is a **consequence** of T and we write $T \models \psi$ if $\mathcal{M} \models \psi$ for all the models $\mathcal{M}$ of T.

Now it's clear that what we called until now "axioms and premises" and denoted with $\Lambda \cup \Gamma$ are nothing but a theory (e.g. let's call it T). We have given some nice property to help us find out if $T \vdash \phi$ for some formula ϕ, but are now more interested in discovering if $T \models \phi$. So, let's state an important result that links those two notions — having in mind that in what follows we always consider theories T that contain at most countable formulas.

Theorem 2.7 (Soundness theorem)**.** *Given a theory T and a formula ϕ, if $T \vdash \phi$ then $T \models \phi$. That is to say, everything provable is true.*

Proof. By theorem 1.12 we have that there is a deduction $\langle \phi_1, \ldots, \phi_n \rangle$ for ϕ from T. We want to prove that whenever $T \models \phi, \ldots, T \models \phi_{n-1}$ then $T \models \phi_n$ — that is $T \models \phi$. But, by definition of deduction and since if $\phi \in T$ it's trivial that $T \models \phi$, without loss of generality we can just prove the case when ϕ is obtained by modus ponens from a formula in T:

$$\begin{array}{c} \psi \to \phi \\ \psi \\ \hline \phi \end{array} \tag{2.1}$$

where $T \models (\psi \to \phi)$ and $T \models \psi$. By definition we have that

$$\forall \mathcal{M} \text{ model of } T \quad \forall \mu \text{ evaluation} \quad \psi \to \phi \text{ is true} \tag{2.2}$$

$$\forall \mathcal{M} \text{ model of } T \quad \forall \mu \text{ evaluation} \quad \psi \text{ is true} \tag{2.3}$$

and so, by the truth table of $\to$

$$\forall \mathcal{M} \text{ model of } T \;\; \forall \mu \text{ evaluation} \quad \phi \text{ is true} \tag{2.4}$$

that means that $T \models \phi$. $\square$

Definition 2.8. A theory T is called **complete** if for every formula ϕ either $T \models \phi$ or $T \models \neg\phi$.

Theorem 2.9 (Lindenbaum's lemma)**.** *Given a consistent theory T it's possible to extend T to a theory $T^* = T \cup A$ that is consistent and complete.*

Proof. First of all notice that all the formulas deducible from T are at most countable, so it's possible to enumerate them — say as $\{\phi_1, \ldots, \phi_i, \ldots\}$. We define by induction the sets of formulas:

- $A_0 = \emptyset$;
- Given A_{i-1}, we consider the formula ϕ_i. If ϕ_i has free variables or if $T \cup A_{i-1} \vdash \phi_i$ or $T \cup A_{i-1} \vdash \neg\phi_i$ we define $A_i = A_{i-1}$, otherwise we define $A_i = A_i \cup \{\phi_i\}$.

In the end we let $A = \bigcup A_i$ and we'd like to prove that T^* is both consistent and complete.

Consistency Suppose T^* is inconsistent and let ϕ be a formula that realizes inconsistency, i.e.

$$T^* \vdash (\phi \;\wedge\; \neg\phi) \tag{2.5}$$

Of course it can't be the case that $T \vdash (\phi \;\wedge\; \neg\phi)$ since by hypothesis T is consistent, so some of the "new" formulas must be involved in the proof. Let's call them $\{\psi_{i_1}, \ldots, \psi_{i_n}\}$ in the same order as we added them. But then, applying reductio ad absurdum (theorem 1.23) we have that since

$$T \cup \{\psi_{i_1}, \ldots, \psi_{i_n}\} \tag{2.6}$$

is inconsistent, it must hold that

$$T \cup \{\psi_{i_1}, \ldots, \psi_{i_{n-1}}\} \vdash \neg\psi_{i_n} \tag{2.7}$$

which is impossible by construction of A.

Completeness Suppose T^* is incomplete and let ϕ_{i_h} be a formula that realizes incompleteness, i.e. it's not the case that $T^* \models \phi_{i_h}$ nor that $T^* \models \neg\phi_{i_h}$. By soundness theorem this means that neither $T^* \vdash \phi_{i_h}$ nor $T^* \vdash \neg\phi_{i_h}$ but this would mean that, in particular, it's not the case that $T \cup A_{i_{h-1}} \vdash \phi_{i_h}$ or that $T \cup A_{i_h - 1} \vdash \neg\phi_{i_h}$, which is impossible by construction.

$\square$

Theorem 2.10 (Henkin's theorem)**.** *If a theory T is consistent, then it is satisfiable by a model that is at most countable.*

Proof. We will divide this proof in various steps, since it is particularly long.

Step 1 First of all we add to the language of T a countable set of new constants $\{d_1, \ldots, d_i, \ldots\}$ and call T_0 the new theory that is obtained. Let's show the trivial fact that T_0 is also consistent. In fact, if it weren't and if ϕ were a formula that realizes inconsistency, simply notice that we don't lose generality in supposing that none of the newly added constants appear in the proof of $T_0 \vdash (\phi \wedge \neg\phi)$ and so T would be inconsistent itself. To see why we may assume that none of the new constants appear, just notice that at most a finite number of them, as well as a finite number of variables, can be in the proof and so it's always possible to pick some of the remaining variables and substitute them to the constants. By construction this preserves the validity of the formulas originally in T (including the logical axioms) and of the inference rules.

Step 2 Now that we have those new objects to build the model we're looking for, we start shaping it in order to have it have the properties we need. First of all we want the model to contain all the "objects" whose existence can be proved in T_0. So, for each of the at most countable formulas of T_0 with one free variable (say ϕ), we fix one of the new constants (say d) and we add to the theory the sentence

$$\exists x\ (\phi(x) \to \phi(d)) \tag{2.8}$$

We're simply attaching a name to the "objects" produced by existence proofs: if T_0 can prove that $\exists x\ (\phi(x))$, then we call d the object whose existence has been proved. We call this new extended theory T_1 and prove it is still consistent. First of all we enumerate the formulas of T_0 with one free variable as $\{\phi_1, \ldots, \phi_i, \ldots\}$. We give an inductive way of choosing the appropriate constant for each of them:

- For ϕ_1 we can chose any of the constants that does not appear in the axioms
- For a generic ϕ_i we add the further constraint that the constant must not appear in any of the formulas

$$\exists x\ (\phi_j(x) \to \phi(d_{i_j})) \text{ for all } j < i \tag{2.9}$$

After adding all this new sentences to the axioms we obtain T_1. If it were inconsistent and ϕ was a formula that realizes inconsistency, notice that only finitely many of those new axioms can appear in the proof of $T_1 \vdash (\phi \wedge \neg\phi)$. In particular, if none of them appear, it would mean that $T_0 \vdash (\phi \wedge \neg\phi)$, which is impossible since T_0 is consistent. Otherwise, let's call n the greatest index for which the axiom regarding ϕ_n appears in the proof. Similarly to what we did in step 1, we now replace the constant d_{i_n} with one of the infinite variables not occurring in the proof (say y). Notice that the validity of the replacement is ensured by the fact that the constant d_{i_n} does not appear in ϕ_n. Now, by theorem 1.23 we have that

$$T_0 \cup \{\exists x\ (\phi_j(x) \to \phi(d_{i_j}))\ :\ j < n\} \cup \{\exists x\ (\phi_n(x) \to \phi(y))\} \tag{2.10}$$

being inconsistent is equivalent to

$$T_0 \cup \{\exists x\ (\phi_j(x) \to \phi(d_{i_j}))\ :\ j < n\} \vdash \neg\exists x\ (\phi_n(x) \to \phi(y)) \tag{2.11}$$

and

$$\neg(\phi_n(x) \to \phi(y)) \leftrightarrow (\exists x\ \phi(x))\ \wedge\ \neg\phi(y) \tag{2.12}$$

So in particular we have that $\neg\phi(y)$ is provable; but this formula, by construction of the logical axioms, is equivalent to $\forall y\ \neg\phi(y)$ that is $\neg\exists y\ \phi(y)$ and, by substitution $\neg\exists x\ \phi(x)$. So we finally proved that

$$T_0 \cup \{\exists x\ (\phi_j(x) \to \phi(d_{i_j}))\ :\ j < n\} \tag{2.13}$$

is inconsistent since it proves both $\exists x\ \phi(x)$ and $\neg\exists x\ \phi(x)$. Now we can just repeat this process until we run out of "new" axioms in the proof and we're forced to admit that T_0 itself is inconsistent, so finding the absurd.

Step 3 Here we simply use theorem 2.9 to extend T_1 to a theory T_2 which is also complete.

Step 4 We're now ready to build a model $\mathcal{M}$ of T_0. We let the universe M be the countable set made up of all the constants of T_0 (note that this set is non-empty since it contains at least the constants we added in step 1) and the interpretation function I as follows:

- A predicative symbol P with $\mathrm{ar}(P) = n$ is interpreted as the relation $I(P) \subseteq M^n$ defined as

$$I(P)(t_1, \ldots, t_n)\ \leftrightarrow\ T_2 \vdash P(t_1, \ldots, t_n) \tag{2.14}$$

  Notice that this is well-defined since T_2 is consistent.
- A functional symbol f with $\mathrm{ar}(f) = n$ is interpreted simply as itself. That's to say that $I(f)(t_1, \ldots, t_n)$ is just $f(t_1, \ldots, t_n)$. There is a subtle point here, since we don't mean that it is the value of the function f evaluated in $(t_1, \ldots, t_n)$: it doesn't make sense to *evaluate* a functional symbol... it's just a symbol. We mean that the interpretation of $f(t_1, \ldots, t_n)$ is exactly the very same symbol — in other words, we could say that it is the string "$f(t_1, \ldots, t_n)$".
- In particular, a constant c is interpreted as the element $I(c) = c \in M$.

Step 5 We finally prove that $\mathcal{M}$ is a model for T_2. This concludes the proof since all the formulas of the deductive closure of T are in particular formulas of the deductive closure of T_2, so $\mathcal{M}$ will also satisfy T. Given a formula ϕ, we want to prove that $\mathcal{M} \models \phi \leftrightarrow T \vdash \phi$; by definition this means that, if $v_1, \dots, v_m$ are the free variables occurring in ϕ and $c_1, \dots, c_m$ are constants, then

$$\mathcal{M} \models \phi(t_1, \dots, t_m) \leftrightarrow T \vdash \phi(t_1, \dots, t_m) \tag{2.15}$$

We will prove this by induction on the complexity of ϕ.

Base case. If ϕ is atomic, then it is in the form $P(v_1, \dots, v_n)$ where P is a predicative symbol. When we operate the substitution we obtain $P(t_1, \dots, t_n)$. But then by definition of I we know that $I(P)$ is true (that is $\mathcal{M} \models P$) if and only if $T_2 \vdash P$.

Inductive step. Here we need to branch again our reasoning, since we have various kinds of non-atomic formulas:

- If ϕ is in the form $\neg\psi$, then by the truth table of $\neg$ we have that ϕ is true in $\mathcal{M}$ iff ψ is false in $\mathcal{M}$. By inductive assumption ψ is true in $\mathcal{M}$ iff $T_2 \vdash \psi$; but then by the completeness and consistency of T_2, ψ is false in $\mathcal{M}$ iff $T_2 \vdash \neg\psi$, that is iff $T_2 \vdash \phi$.
- In an analogous way (that would be very long to describe here and perhaps not very useful since it's just a technical proof that doesn't give many insights) it's possible to prove this result for ϕ in the forms $\psi_1 \wedge \psi_2$ and $\exists x\ \psi$.

□

Theorem 2.11 (Gödel's completeness theorem)**.** *Given a theory T and a formula ϕ, $T \vdash \phi$ iff $T \models \phi$. This adds to theorem 2.7 the converse part: everything true is provable.*

Proof. Given ϕ such that $T \models \phi$ we want to prove that $T \vdash \phi$. But suppose this is not the case and then consider the theory

$$T_0 = T \cup \{\neg\forall v_1 \ \cdots \ \forall v_n\ \phi\} \tag{2.16}$$

where $v_1, \dots, v_n$ are the free variables of ϕ. We want to show that T_0 is consistent. In fact, if it were inconsistent, by theorem 1.23 we would have that

$$T \vdash \neg\neg\forall v_1 \ \cdots \ \forall v_n\ \phi \tag{2.17}$$

that is to say

$$T \vdash \forall v_1 \ \cdots \ \forall v_n\ \phi \tag{2.18}$$

and so $T \vdash \phi$, which we ruled out by hypothesis. So T_0 is consistent and then, in virtue of theorem 2.10, it has a model $\mathcal{M}$. But now notice that on one hand ϕ is true in T, so is true in all models of T and in particular $\mathcal{M}$ is a model of T; hence $\mathcal{M} \models \phi$ and so $\mathcal{M} \models \forall v_1 \ \cdots \ \forall v_n\ \phi$. On the other hand also all formulas of T_0 are true in $\mathcal{M}$ and so in particular $\mathcal{M} \models \neg\forall v_1 \ \cdots \ \forall v_n\ \phi$. Here is the contradiction. □

Corollary 2.12. *T is consistent iff it is satisfiable.*

Proof. If T is not satisfiable, it doesn't have any model. So it's trivially true that every model of T is a model for $\phi \wedge \neg\phi$ and so by definition $T \models (\phi \wedge \neg\phi)$; by theorem 2.11 this means $T \vdash (\phi \wedge \neg\phi)$, so $T \vdash \phi$ and $T \vdash \neg\phi$, that is the definition of inconsistency. Analogously we prove the converse. □

Corollary 2.13 (Löwenheim-Skolem theorem)**.** *If a theory T has a model, then it has a model that is at most countable.*

Proof. If T is satisfiable, then by corollary 2.12 it is consistent and so, by theorem 2.10 it has a model that is at most countable. □

We are now ready to give a more interesting, semantic, version of the compactness theorem.

Theorem 2.14 (Compactness theorem for first order theories)**.** *T is satisfiable iff every finite subset T_0 of T is satisfiable.*

Proof. If T is satisfiable it's trivial that all its subsets are satisfiable. On the other hand, suppose that every finite subset of T is satisfiable but T is not. This means by corollary 2.12 that T is inconsistent. Let ϕ be a formula that realizes inconsistency and consider a proof of $\phi \wedge \neg\phi$. Since the proof is finite, there is a finite $T_0 \subseteq T$ that is inconsistent and hence not satisfiable, which is absurd. □

2.2 A topological version of the compactness theorem

We will now link the compactness theorem with the actual — topological — compactness of a particular space, called the Stone space.

Definition 2.15. Let $\mathcal{M} = (M, \sigma, I)$ be a σ-structure and consider a set $A \subseteq M$. Let σ_A be the signature obtained from σ, by adding a constant c_a for every element $a \in A$. If we extend I in a way such that $I(c_a) = a$ for all the $a \in A$, we can view $\mathcal{M}$ as a σ_A-structure. To distinguish the full theory $\text{Th}(\mathcal{M})$ for $\mathcal{M}$ as a σ-structure to the full theory for $\mathcal{M}$ as a σ_A-structure, we'll call the latter $\text{Th}_A(\mathcal{M})$. Now, let p be a set of formulas in the language of σ_A with n free variables. We say that p is an **n-type** if $p \cup \text{Th}_A(\mathcal{M})$ is satisfiable; moreover, p is a **complete n-type** if for every formula ϕ with n free variables either $\phi \in p$ or $\neg\phi \in p$, but not both. We will denote as $S_n^{\mathcal{M}}(A)$ the set of all complete n-types.

Observation 2.16. By the compactness theorem, we may safely replace the word satisfiable with "finitely satisfiable" in the previous definition.

Definition 2.17. In the notations of the previous definition, given an n-type p of σ_A, we say that $(m_1, \dots, m_n) \in M^n$ **realizes** p if for all $\phi \in p$ it holds that $\mathcal{M} \models \phi(m_1, \dots, m_n)$. If there is some n-uple that realizes p we say that p is **realized** in $\mathcal{M}$; otherwise we say that $\mathcal{M}$ **omits** p.

Definition 2.18. Given a σ-structure $\mathcal{M}$, $A \subseteq M$ and $m = (m_1, \dots, m_n) \in M^n$ we call $\mathrm{tp}(m; A)$ the complete n-type made up by all the formulas ϕ in the language of σ_A such that $\mathcal{M} \models \phi(m_1, \dots, m_n)$. If $A = \emptyset$ we just write $\mathrm{tp}(m)$.

Theorem 2.19. *Given a σ-structure $\mathcal{M}$, $A \subseteq M$ and an n-type p of σ_A, then it's always possible to find a model $\mathcal{M}'$ that realizes p such that $\mathcal{M}'$ is an elementary extension of $\mathcal{M}$.*

Definition 2.20. Given a σ-structure $\mathcal{M}$, $A \subseteq M$, for each formula ϕ consider

$$[\phi] = \{p \in S_n^{\mathcal{M}}(A) \ : \ \phi \in p\} \subseteq S_n^{\mathcal{M}}(A) \tag{2.19}$$

Notice that, given ϕ and ψ such that $\phi \vee \psi \in p$ one has either $\phi \in p$ or $\psi \in p$; hence

$$[\phi \vee \psi] = [\phi] \cup [\psi] \tag{2.20}$$

Analogously

$$[\phi \ \wedge \ \psi] = [\phi] \cap [\psi] \tag{2.21}$$

And finally, by definition of complete n-type, exactly one of ϕ and $\neg\phi$ is in p, so

$$[\phi] = S_n^{\mathcal{M}}(A) \setminus [\neg\phi] \tag{2.22}$$

Then, we call the **Stone topology**, the topology on $S_n^{\mathcal{M}}(A)$ generated by taking the sets $[\phi]$ as an open base. Notice that by (2.22) all the sets $[\phi]$ are both open and closed (usually said clopen), since they are the complement of the open set $[\neg\phi]$.

Theorem 2.21 (Why the Compactness theorem is called in this way)**.** *The Stone space $S_n^{\mathcal{M}}(A)$ is compact.*

Proof. We want to show that every cover of $S_n^{\mathcal{M}}(A)$ of basic open sets has a finite subcover. In fact, suppose there existed an infinite cover

$$C = \{[\phi_i] \ : \ i \in I\} \tag{2.23}$$

that doesn't admit any finite cover. Consider

$$\Gamma = \{\neg\phi_i \ : \ i \in I\} \tag{2.24}$$

We want to prove that $\Gamma \cup \mathrm{Th}_A(\mathcal{M})$ is satisfiable using the compactness theorem. Given a finite subset $I_0 \subset I$, since C doesn't have a finite subcover, we can always find an element

$$p \in S_n^{\mathcal{M}}(A) \setminus \bigcup_{i \in I_0} [\phi_i] \tag{2.25}$$

Now, by theorem 2.19, there is an elementary extension $\mathcal{M}'$ of $\mathcal{M}$ that realizes p and let $m' = (m'_1, \ldots, m'_n)$ be the realization. By definition $\mathcal{M} \models \mathrm{Th}_A(\mathcal{M})$ and so also $\mathcal{M}' \models \mathrm{Th}_A(\mathcal{M})$. Also, since $\mathcal{M}' \models p(m')$, by (2.25), (2.21), (2.22) and by recalling that the finite intersection of elements of a base of a topological space is still in the base, we have that

$$\mathcal{M}' \models \bigwedge_{i \in I_0} \neg\phi_i(m') \tag{2.26}$$

and so $\mathcal{M}' \models \Gamma$. We have shown that $\Gamma \cup \mathrm{Th}_A(\mathcal{M})$ is finitely satisfiable and so, by the compactness theorem it is satisfiable. Now, since $\Gamma \cup \mathrm{Th}_A(\mathcal{M})$ is satisfiable, we proved that Γ is an n-type and hence we can consider an elementary extension $\mathcal{M}''$ of $\mathcal{M}$ that realizes Γ via $m'' = (m''_1, \ldots, m''_n)$ — i.e. $\mathcal{M}'' \models \psi$ for all $\psi \in \Gamma$. Now, by definition of Γ this means that

$$\mathcal{M}'' \models \neg\phi_i \text{ for all } i \in I \tag{2.27}$$

Recalling the definition of tp, this means that

$$\neg\phi_i \in \mathrm{tp}(m''; A) \text{ for all } i \in I \tag{2.28}$$

and so

$$\mathrm{tp}(m''; A) \in [\neg\phi_i] \text{ for all } i \in I \tag{2.29}$$

Since $\mathrm{tp}(m''; A)$ is in every $[\neg\phi_i]$, it is in the intersection:

$$\mathrm{tp}(m''; A) \in \bigcap_{i \in I} [\neg\phi_i] \tag{2.30}$$

and by using (2.22) we get

$$\mathrm{tp}(m''; A) \in \bigcap_{i \in I} \left(S_n^{\mathcal{M}}(A) \setminus [\phi_i]\right) \tag{2.31}$$

so

$$\mathrm{tp}(m''; A) \in S_n^{\mathcal{M}}(A) \setminus \bigcup_{i \in I} [\phi_i] \tag{2.32}$$

which is absurd, since $C = \{[\phi_i] \ : \ i \in I\}$ is a cover of $S_n^{\mathcal{M}}(A)$. □

2.3 Incompleteness results

Our objective in this subsection is to prove the two famous incompleteness theorems by Gödel.

Definition 2.22. Given a formula ϕ we assign it its **Gödel number**, $\ulcorner \phi \urcorner$, according to the following rules: first of all we fix an integer to every symbol of the language

Symbol	Number
$\wedge$	1
$\neg$	3
$\exists$	5
(	7
)	9
$\in$	11
$=$	13
v_i	$2i$

To encode ϕ we regard it as a sequence of symbols $a_1 a_2 \cdots a_n$ with values $b_1, b_2, \ldots, b_n$ determined according to the preceding table; if we call $p_1, \ldots, p_n$ the first n prime numbers, then we define

$$\ulcorner \phi \urcorner = \prod_{k=1}^{n} p_k^{b_k} \tag{2.33}$$

In this way, because of the uniqueness of prime factorization, there is always an unique formula corresponding to a given Gödel number. Of course, it's also possible to encode sequences of formulas $\langle \phi_1, \ldots, \phi_n \rangle$ with finite sequences of integers $\langle \ulcorner \phi_1 \urcorner, \ldots, \ulcorner \phi_n \urcorner \rangle$.

Observation 2.23. With this formalism, modus ponens becomes a relation in $\mathbb{N}^3$:

$$\begin{aligned} Mp(\langle n, m \rangle, p) \ \leftrightarrow\ & n \text{ is the Gödel number of a formula in the form } \ulcorner \phi \to \psi \urcorner \text{ and} \\ & m = \ulcorner \phi \urcorner \text{ and} \\ & p = \ulcorner \psi \urcorner \end{aligned} \tag{2.34}$$

Observation 2.24. From previous observation it's easy to define a relation that express if a sequence of Gödel numbers $\underline{m} = \langle m_1, \ldots, m_k \rangle$ (we'll call, abusing the nomenclature, a Gödel number the whole sequence) encodes a proof for some other formula:

$$\begin{aligned} Proof(\underline{m}, p) \ \leftrightarrow\ & \underline{m} \text{ is the Gödel number of a proof of } \phi \text{ and} \\ & p = \ulcorner \phi \urcorner \end{aligned} \tag{2.35}$$

Theorem 2.25 (Gödel's first incompleteness theorem)**.** *If T is a consistent theory that contains arithmetics (this basically means that we're allowed to talk about Gödel numbers), then T is not complete.*

Proof. First of all notice that, by the completeness theorem we have to produce a formula ϕ such that $T \nvdash \phi$ and $T \nvdash \neg\phi$. So, we consider the formula $Q(x, y)$ defined as follows:

$$\begin{aligned} & y \text{ is the Gödel number of a formula } \phi \text{ and} \\ & \text{called } y' \text{ the Gödel number of } \phi(\ulcorner \phi \urcorner) \text{ then} \\ & \neg Proof(x, y') \end{aligned} \tag{2.36}$$

So, basically $Q(\underline{m}, \ulcorner\phi\urcorner)$ says that $\underline{m}$ is not the Gödel number of a proof of $\phi(\ulcorner\phi\urcorner)$. Now can quantify over a variable and define the formula $G(y)$ as

$$\forall x\ Q(x, y) \tag{2.37}$$

In this case $G(\ulcorner\phi\urcorner)$ simply says that there is no proof of $\phi(\ulcorner\phi\urcorner)$. Now we do trick, creating a self-referential statement and showing that it and its negation are both unprovable. The statement in question is

$$G(\ulcorner G\urcorner) \tag{2.38}$$

Of course this is well-defined, since G is a formula and so it makes sense to consider its Gödel number $\ulcorner G\urcorner$ and then to use it in a substitution of the free variable of G itself. Let's show that $G(\ulcorner G\urcorner)$ is not provable. In fact, if it were, let x be the Gödel number of a proof of $G(\ulcorner G\urcorner)$. Then we would have that $\neg Q(x, \ulcorner G\urcorner)$ is true and so

$$T \vdash \neg Q(x, \ulcorner G\urcorner) \tag{2.39}$$

On the other hand, saying that $T \vdash G(\ulcorner G\urcorner)$ simply means that

$$T \vdash \forall x\ Q(x, \ulcorner G\urcorner) \tag{2.40}$$

Hence the contradiction. We're left to show that $\neg G(\ulcorner G\urcorner)$ is not provable. In fact, first notice that for any x it can't be the case that $T \vdash \neg Q(x, \ulcorner G\urcorner)$: if this were the case, x it would be the Gödel number of a proof of $G(\ulcorner G\urcorner)$, but we just showed that $G(\ulcorner G\urcorner)$ is not provable. By consistency, if $\neg Q(x, \ulcorner G\urcorner)$ is not provable, its converse is (for any x) and so

$$T \vdash \forall x\ Q(x, \ulcorner G\urcorner) \tag{2.41}$$

Now, assume that $\neg G(\ulcorner G\urcorner)$ were provable. But $\neg G(\ulcorner G\urcorner)$ is the statement

$$\neg\forall x\ Q(x, \ulcorner G\urcorner) \tag{2.42}$$

and this is the same as

$$\exists x\ \neg Q(x, \ulcorner G\urcorner) \tag{2.43}$$

so if $\neg G(\ulcorner G\urcorner)$ is provable we have that

$$T \vdash \exists x\ \neg Q(x, \ulcorner G\urcorner) \tag{2.44}$$

that gives an absurd, in conjunction with 2.41. □

In what follows we will always assume that the theory T considered contains arithmetic, as in the hypothesis of the previous theorem. For example, ZF is such a theory.

Theorem 2.26 (Diagonalization lemma). *If $\phi(x)$ is a formula of the theory T in the only free variable x, then there is a sentence ψ such that*

$$T \vdash (\psi \ \leftrightarrow\ \phi(\ulcorner\psi\urcorner)) \tag{2.45}$$

Proof. Consider the formula $D(x)$ defined as

$$\begin{aligned}&\text{if } \chi \text{ is the formula such that } x = \ulcorner \chi \urcorner \text{ then}\\ &\phi(\chi(\ulcorner \chi \urcorner))\end{aligned} \tag{2.46}$$

[TODO] □

Theorem 2.27 (Tarski's undefinability theorem)**.** *There is no formula $true(x)$ of the theory T in the only free variable x that can tell wether the sentence with Gödel number x is a true sentence or not.*

Proof. In fact, by taking $\phi(x) = \neg true(x)$ and applying the diagonalization lemma we can find a sentence ψ such that

$$T \vdash (\psi \ \leftrightarrow \ \neg true(\ulcorner \psi \urcorner)) \tag{2.47}$$

So ψ is true exactly when $true(\ulcorner \psi \urcorner)$ is not. □

Theorem 2.28 (Gödel's second incompleteness theorem)**.** *[TODO]*

3 Models of ZF

3.1 Well-founded sets

We will now explore some interesting application of logic to set theory. For this section, it will be assumed that the reader is already familiar with basic set theory such as the ZF axioms, ordinal and cardinal numbers, cofinality, inaccessible cardinals, etc.

Definition 3.1. Let $\mathbb{ON}$ be the class of all ordinal numbers and define, by transfinite recursion $R(\alpha)$ for $\alpha \in \mathbb{ON}$:

1. $R(0) = 0$
2. $R(\alpha) = \mathcal{P}(R(\beta))$ if $\alpha = \beta + 1$
3. $R(\alpha) = \bigcup_{\beta<\alpha} R(\beta)$ if α is a limit ordinal

We call the class of **well-founded sets** as

$$\mathbb{WF} = \bigcup_{\alpha \in \mathbb{ON}} R(\alpha) \tag{3.1}$$

Theorem 3.2. *$R(\alpha)$ is transitive for all $\alpha \in \mathbb{ON}$; moreover, $\forall \beta < \alpha \quad R(\beta) \subset R(\alpha)$.*

Proof. We prove these results by transfinite induction:

Base case. If $\alpha = 0$, both results are trivial since $0 \in \mathbb{ON}$ and there is no $\beta < 0$.

α **successor.** If $\alpha = \beta + 1$, then $R(\alpha) = \mathcal{P}(R(\beta))$. This both shows that $R(\alpha)$ is transitive, since $R(\beta)$ is so by inductive hypothesis and that $R(\beta) \subset R(\alpha)$, which is enough (again by inductive hypothesis) to prove the second result.

α **limit.** In this case $R(\alpha)$ is transitive, since by inductive hypothesis is union of transitive sets. The second result is trivial by definition.

□

Definition 3.3. Given a set $x \in \mathbb{WF}$, notice that by definition the least $\alpha \in \mathbb{ON}$ such that $x \in R(\alpha)$ can't be a limit ordinal. So, we can define the **rank** of x as the least ordinal β such that $x \in R(\beta + 1)$. We'll denote it with $rk(x)$.

Theorem 3.4. *Given $\alpha \in \mathbb{ON}$ we have that*

$$R(\alpha) = \{x \in \mathbb{WF} \ : \ rk(x) < \alpha\} \tag{3.2}$$

Proof. By definition $rk(x) < \alpha$ is the same to say that $\exists \beta < \alpha$ such that $x \in R(\beta + 1)$. But since $R(\beta+1) = \mathcal{P}(R(\beta))$, by theorem 3.2 this is the same to say that $x \in R(\alpha)$. □

Theorem 3.5. *Given $y \in \mathbb{WF}$ and $x \in y$ then also $x \in \mathbb{WF}$ and $rk(x) < rk(y)$.*

Proof. Let $\alpha = rk(y)$, then $y \in R(\alpha + 1) = \mathcal{P}(R(\alpha))$ and so $x \in R(\alpha)$, that is $x \in \mathbb{WF}$ and moreover $rk(x) < \alpha = rk(y)$ by theorem 3.4. □

Theorem 3.6. *Given $y \in \mathbb{WF}$, $rk(y) = \sup\{rk(x) + 1 \mid x < y\}$.*

Proof. Let $\alpha = \sup\{rk(x)+1 \mid x < y\}$. Then, by theorem 3.5 $\alpha \leq rk(y)$. For the other side of the inequality, notice that since for all $x < y$ we have $rk(x) < \alpha$, then $y \subset R(\alpha)$, that is by definition $y \in R(\alpha + 1)$ and so $rk(y) \leq \alpha$. □

We proved those boring technical lemmata to get to an important result that gives a first justification of the importance of $\mathbb{WF}$: all the ordinal numbers are well-founded, as shown in the next theorem.

Theorem 3.7. *Given $\alpha \in \mathbb{ON}$, $\alpha \in \mathbb{WF}$ and moreover $rk(\alpha) = \alpha$.*

Proof. We prove this result by transfinite induction:

Base case. If $\alpha = 0$ by definition $\alpha \in \mathbb{WF}$ and its rank is 0.

α **successor or limit.** In this case, given $\beta \in \alpha$, since $\beta \in R(\beta + 1) \subset R(\alpha)$, we have that $\alpha \in R(\alpha+1)$ and so $\alpha \in \mathbb{WF}$. Moreover, applying theorem 3.6 we have that

$$\begin{aligned} rk(\alpha) &= \sup\{rk(\beta) + 1 \mid \beta < \alpha\} = && (3.3)\\ &= \sup\{\beta + 1 \mid \beta < \alpha\} = && (3.4)\\ &= \alpha && (3.5) \end{aligned}$$

which is true both if α is a successor or a limit ordinal.

□

Theorem 3.8. *If $x, y \in \mathbb{WF}$ then the following sets are also well-founded:*

$$\bigcup x, \quad \mathcal{P}(x), \quad \{x\}, \quad x \times y, \quad x \cup y, \quad x \cap y, \quad \{x, y\}, \quad \langle x, y \rangle, \quad x^y \tag{3.6}$$

Observation 3.9. The previous theorem just shows that all the most commonly used mathematical structures are well-funded. For example, $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ are, as are groups, rings, fields and topological spaces.

Until now, we've been able to see that most of the everyday mathematics takes place in $\mathbb{WF}$, so one can wonder if it's acceptable in ZF to just assume that $\mathbb{WF}$ is the whole Von Neumann universe $\mathbb{V} = \{x \mid x = x\}$ that is to say, that every set is well-founded. We will now show that the axioms of ZF imply $\mathbb{V} = \mathbb{WF}$ and that this statement can replace the axiom of foundation in ZF. It's also nice to notice that sets in $\mathbb{WF}$ constitute models for most of the mathematics we use; in particular:

- $R(\omega)$ is a model for $ZF - I$, that is Zermelo-Fraenknel set theory without the axiom of infinity;
- $R(\omega+\omega)$ is a model for the usual mathematical theory theory needed to build all the objects of observation 3.9, as follows by theorem 3.6 and the construction of those objects;
- $R(\kappa)$, where κ is an inaccessible cardinal, is a model for ZF, as we will prove soon.

Theorem 3.10. *If $\mathbb{V} = \mathbb{WF}$ then the axiom of foundation holds (in $ZF - F$, that is Zermelo-Fraenknel set theory without the axiom of foundation itself).*

Proof. Given a set $x \in \mathbb{WF}$ we'd like to prove the statement

$$\forall A\,((A \subset x \ \wedge \ \exists B\ (B \in A)) \ \rightarrow \ \exists B \ : \ (B \in A \ \wedge \ \neg\exists C \ : \ (C \in A \ \wedge \ C \in B))) \tag{3.7}$$

That is the same as claiming that every set admits an $\in$-minimal element. In fact, let $\alpha = \min\{rk(B) \mid B \in A\}$ and let B be the set that realizes $rk(B) = \alpha$. Then B is $\in$-minimal by theorem 3.5 and so the condition is satisfied. □

Theorem 3.11. $\mathbb{V} = \mathbb{WF}$ *holds in ZF.*

Proof. Suppose there existed a non-well-founded set $x \in \mathbb{V}$. This would mean that there is no $\alpha \in \mathbb{ON}$ such that $x \in R(\alpha)$. By the axiom of foundation we can suppose without loss of generality that x is $\in$-minimal. But then for all $y \in x$ by minimality we would have that $\exists \alpha \in \mathbb{ON}$ such that $y \in R(\alpha)$. Then it makes sense to consider

$$\beta = \sup\{rk(y) \mid y \in x\} \tag{3.8}$$

Now, $\forall y \in x$, $rk(y) \leq \beta$ and so $y \in R(\beta+1)$. But then $x \subset R(\beta+1)$ and so $x \in R(\beta+2)$, which is absurd. □

3.2 Building a model for ZF

Theorem 3.12. *Given $\alpha \in \mathbb{ON}$, if α is a limit ordinal strictly greater than ω, then $R(\alpha)$ is a model for $ZF - R$, that is the Zermelo-Fraenkel set theory without the axiom of replacement.*

Proof. We will show that the other axioms of ZF, different from the axiom of replacement, hold.

Set existence. Since $0 \in R(\alpha)$;

Extensionality. By transitivity of $R(\alpha)$ if two of its elements have the same elements, then they must be the same;

Foundation. Again by transitivity of $R(\alpha)$ is always possible to find an $\in$-minimal element;

Comprehension. Given $x \in R(\beta)$ and $y = \{z \in x \mid \phi\}$, where $\beta < \alpha$ and ϕ is a formula in which y doesn't occur free, then since $y \subseteq x \in R(\beta)$ we have that $y \subseteq R(\beta)$ by transitivity and so $y \in R(\beta + 1)$;

Pairing. Given $x \in R(\beta), y \in R(\gamma)$, where $\beta, \gamma < \alpha$, then $\{x, y\} \in R(\max\{\beta, \gamma\} + 1)$;

Union. Given $A \subseteq R(\beta)$, where $\beta < \alpha$, then $\bigcup A \in R(\beta + 1)$;

Infinity. Since $\omega \in R(\alpha)$;

Power set. Given $x \in R(\beta)$, where $\beta < \alpha$, then $\mathcal{P}(x) \in R(\beta + 2)$.

□

Corollary 3.13. *$ZF - R$ is satisfiable, since $R(\omega + \omega)$ is a model for it.*

Theorem 3.14. *Given an inaccessible cardinal $\kappa \neq \omega$ and an ordinal $\alpha < \kappa$*

$$|R(\alpha)| < \kappa \tag{3.9}$$

Proof. We proceed by transfinite induction:

Base case. If $\alpha = 0$, trivially $|R(\alpha)| = |0| = 0 < \kappa$.

α **successor.** If $\alpha = \beta + 1$ then

$$|R(\alpha)| = |\mathcal{P}(R(\beta)) = 2^{|R(\beta)|} < \kappa \tag{3.10}$$

by definition of inaccessible cardinal.

α **limit.** In this case, of course $|R(\alpha)| = \sup\{|R(\beta)| \mid \beta < \alpha\} \leq \kappa$. But, if we were in the case $|R(\alpha)| = \kappa$ then we would have a cofinal map from α into κ given by $\beta \mapsto |R(\beta)|$, that is absurd by inaccessibility.

□

Theorem 3.15. *Given an inaccessible cardinal κ, $R(\kappa)$ is a model for ZF.*

Proof. By theorem 3.12 we must just check that the axiom of replacement holds. So, given $A \in R(\kappa)$, we want to show that

$$B = \{y \mid \exists x \in A \ : \ \phi(x, y)\} \in R(\kappa) \tag{3.11}$$

where ϕ is a function-defining formula in which B doesn't occur free. First of all, notice that since κ is in particular a limit ordinal, then there is an $\alpha < \kappa$ such that $A \in R(\alpha)$ and so $A \subset R(\alpha)$. So by theorem 3.14 we have that $|A| \leq |R(\alpha)| < \kappa$. Now let

$$\beta = \sup\{rk(y) \mid y \in B\} \tag{3.12}$$

First of all notice that for all $y \in B$ we have $y \in R(\beta+1)$ and so, as usual, $B \in R(\beta+2)$. This means that we have to show that $\beta < \kappa$ to prove our thesis. So, call y_x the unique y such that $\phi(x, y)$ and assume — ad absurdum — that $\beta = \kappa$; in this case, $x \mapsto rk(y_x)$ would define a cofinal map between $|A|$ and κ, so $cf(\kappa) \leq |A|$. But by definition $cf(\kappa) = \kappa$ and we showed that $|A| < \kappa$, hence the absurd. □

Corollary 3.16. *If inaccessible cardinals exist then ZF is satisfiable. Moreover, since $ZF + \exists IC$ satisfies ZF, by Gödel second incompleteness theorem $\exists IC$ is not provable in ZF ($\exists IC$ is "an inaccessible cardinal exists").*