Cédric Milliet and Margaret Thomas

Lecture notes of

MODEL THEORY

Preliminary version

Master's course
Universität Konstanz, 2014

Contents

INTRODUCTION

CHAPTER 1

BASIC MODEL THEORY

1.1 Language, structures and morphisms

Definition 1.1 (language) A *language* is a set $\{(f, n_f), (r, n_r), c : f \in F, r \in R, c \in C\}$ consisting of three kinds of elements: function symbols f, relation symbols r and constant symbols c. Each function symbol f and relation symbol r come equipped with a natural number n_f and n_r respectively, called their *arity* (that will provide information on the size of their domains).

Examples 1.2

1. The *language of orderings*
$$L_{ord} = \{(\leqslant, 2)\}$$
consists of one binary relation symbol $\leqslant$.
2. The *language of semigroups*
$$L_{sgp} = \{(\times, 2)\}$$
consists of one binary function symbol $\times$.
3. The *language of monoids*
$$L_{mon} = \{(\times, 2), e\}$$
consists of one binary function symbol $\times$ and one constant symbol e.
4. The *language of groups*
$$L_{gp} = \{(\times, 2), ({}^{-1}, 1), e\}$$
consists of a binary function symbol $\times$, a unary function symbol ${}^{-1}$ and a constant symbol e.
5. The *language of ordered groups*
$$L_{ogp} = \{(\times, 2), ({}^{-1}, 1), (<, 2), e\}$$
consists of the language of groups together with one binary relation symbol $<$.
6. The *language of rings*
$$L_{ring} = \{(\times, 2), (+, 2), (-, 2), e_0, e_1\}$$
consists of three binary function symbols and two constant symbols.
7. The *language of fields*
$$L_{field} = \{(\times, 2), ({}^{-1}, 1), (+, 2), (-, 2), e_0, e_1\}$$
consists of the language of rings together with one unary function symbol.

Remark. One often omits to make these natural numbers precise when they are obvious from the context, and simply write a language $\{f, r, c : f \in F, r \in R, c \in C\}$.

Definition 1.3 (structure) Let $L = \{(f, n_f), (r, n_r), c : f \in F, j \in R, c \in C\}$ be a language. A *structure in the language* L (or L*-structure* for short) is given by
$$\left(M, f^M, r^M, c^M : f \in F, r \in R, c \in C\right)$$
where

- M is a **non-empty** set, called its *domain*,

- for every f in F, f^M is a function from M^{n_f} to M, called the *interpretation of f in M*,
- for every r in R, r^M is an n_r-ary relation on M (i.e. a subset of M^{n_r}), the *interpretation of r in M*,
- for every c in C, c^M is an element of M, called the *interpretation of c in M*.

$\left(f^M, r^M, c^M : f \in F, r \in R, c \in C\right)$ is called the *interpretation of the language L in M*, written L^M: an L-structure is written (M, L^M).

Examples 1.4 1. We shall write $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$ and $\mathbf{R}$ respectively for the set of natural numbers, integers, rationals and real numbers. $(\mathbf{N}, \leqslant^{\mathbf{N}})$, $(\mathbf{Z}, \leqslant^{\mathbf{Z}})$, $(\mathbf{Q}, \leqslant^{\mathbf{Q}})$ and $(\mathbf{R}, \leqslant^{\mathbf{R}})$ are structures in the language of orderings, where $\leqslant^{\mathbf{N}}$, $\leqslant^{\mathbf{Z}}$ and $\leqslant^{\mathbf{Q}}$ are the orderings induced by the natural ordering $\leqslant^{\mathbf{R}}$ on $\mathbf{R}$. Note that there are many other ways to interpret $\leqslant$: if $\triangleleft$ denotes the binary relation defined on $\mathbf{R}$ by putting $x \triangleleft y$ iff $xy \geqslant 2014$, then $(\mathbf{R}, \triangleleft)$ is also a structure in the language $\{\leqslant\}$, despite the fact that $\triangleleft$ is merely a binary relation and not even an order.

2. $(\mathbf{R}, +^{\mathbf{R}}, -^{\mathbf{R}}, \leqslant^{\mathbf{R}}, 0)$ is a structure in the language of ordered groups (where $+^{\mathbf{R}}$ denotes the usual addition, $-^{\mathbf{R}}$ the usual opposite function and $\leqslant^{\mathbf{R}}$ the canonical ordering on $\mathbf{R}$). Note that there are many other ways to endow the set $\mathbf{R}$ with a structure in the language of ordered groups.
3. A group G is naturally equipped with a structure in the language of groups $\{\times, ^{-1}, e\}$: $\times$ is interpreted by the group law, $^{-1}$ by the inverse operation and e by the neutral element. Similarly for monoids, semigroups, rings, fields in their respective languages.

Abusing notations, we sometimes identify a structure (M, L^M) with its domain M. Only when there is no ambiguity, we also simply write f instead of f^M, for the interpretation in M of a symbol f of the language. **In what follows, we assume that a language always contains the binary relation $=$ which is interpreted in every structure by the usual equality**.

Definition 1.5 (substructure, extension) Let L be a language, (M, L^M) and (N, L^N) two L-structures. (N, L^N) is an *L-substructure of* (M, L^M) (or a *substructure of* M for short when L is obvious from the context) if N is a subset of M and if L^N is 'the restriction of L^M to N', more precisely:

1. For every r in R, one has $r^N = r^M \cap N^{n_r}$.
2. For every f in F, one has $f^N = f^M\big|_{N^{n_f}}$.
3. For every c in C, one has $c^N = c^M$.

One also says that (M, L^M) is an *L-extension of* (N, L^N) (or simply an *extension* of M).

Remark. We write $(N, L^N) \subset (M, L^M)$ or $N \subset_L M$ when (N, L^N) is an L-substructure of (M, L^M). Note that this defines a reflexive, transitive, antisymmetric relation.

Examples 1.6 1. In the language of orderings $\{\leqslant\}$, the structures $(\mathbf{N}, \leqslant^{\mathbf{N}})$, $(\mathbf{Z}, \leqslant^{\mathbf{Z}})$ and $(\mathbf{Q}, \leqslant^{\mathbf{Q}})$ are substructures of $(\mathbf{R}, \leqslant^{\mathbf{R}})$. In fact, the substructures of $(\mathbf{R}, \leqslant^{\mathbf{R}})$ are precisely the subsets of $\mathbf{R}$ together with the induced ordering.

2. In the language of monoids $\{\times, e\}$, consider the structure $\mathbf{R}$ where $\times$ is interpreted by the addition and e by 0. Its substructures are precisely the subsets containing 0 that are closed under addition (i.e the submonoids). For instance $(\mathbf{N}, +, 0)$ is a substructure of $(\mathbf{R}, +, 0)$.
3. In the language of groups $\{\times, ^{-1}, e\}$, consider the additive structure of $\mathbf{R}$ (where $\times$ is interpreted by the addition, e by zero, and $^{-1}$ by the opposite function $-$). Its substructures are precisely the subsets containing zero that are closed under addition and opposite (i.e the subgroups). For instance whatever the L_{gp}-structure $L^{\mathbf{N}}$ on $\mathbf{N}$, $(\mathbf{N}, L^{\mathbf{N}})$ is not a substructure of $(\mathbf{R}, +, -, 0)$.
4. In the language of rings, the substructures of $(\mathbf{R}, +, \times, -, 0, 1)$ (with its natural ring structure) are precisely the subrings of $\mathbf{R}$ (exercise).
5. If (M, L^M) is an L-structure and $A \subset M$ a subset of M, then the *L-structure generated by A*, written $\langle A \rangle$, is by definition the L-substructure of M whose domain is the intersection of the domains of all the L-substructures that contain A (exercise: show that an intersection of substructures of M is again a substructure of M). The domain of $\langle A \rangle$ is the smallest subset of M containing A, the constants of L^M and closed under the functions of L^M (exercise).

Note that the notion of substructure depends on the language L.

Definition 1.7 (morphism, embedding, isomorphism) Let (M, L^M) and (N, L^N) be two L-structures.

1. A *morphism of L-structures* (or *L-morphism* for short, or even *morphism* when there is no ambiguity about the language) from M to N is a map $\sigma : M \longrightarrow N$ that preserves the language L, *i.e.* such that
   - for all constant symbols c, the equality $\sigma(c^M) = c^N$ holds,
   - for all relation symbols r and all a in M^{n_r}, then $a \in r^M$ implies $\sigma(a) \in r^N$,
   - for all function symbols f and a in M^{n_f}, then $\sigma(f^M(a)) = f^N(\sigma(a))$.
2. An *embedding of L-structures* (or *L-embedding* or simply *embedding*) from M to N is a morphism such that for all relation symbols r in R and for all a in M^{n_r}, $a \in r^M$ holds if and only if $\sigma(a) \in r^N$ holds. As the language contains equality, note that an embedding is always injective.
3. An *isomorphism of L-structures* (or *L-isomorphism*, or *isomorphism*) from M to N is a surjective embedding. An *L-automorphism* of M is an L-isomorphism from M to M.

Remark. If $\sigma : M \longrightarrow N$ is a morphism of L-structures, then $\sigma(M)$ is an L-substructure of N (and in particular an L-structure). If σ is in addition an L-embedding, then the map $\sigma : M \longrightarrow \sigma(M)$ is an isomorphism of L-structures (exercise).

Examples 1.8
1. In the language of orderings $\{\leqslant\}$, a morphism from $(\mathbf{R}, \leqslant^{\mathbf{R}})$ to $(\mathbf{R}, \leqslant^{\mathbf{R}})$ is an increasing map; an embedding from $\mathbf{R}$ to $\mathbf{R}$ is a strictly increasing map.
2. In the language of groups L_{gp}, a morphism from $(\mathbf{R}, +, -, 0)$ to $(\mathbf{R}, +, -, 0)$ is precisely a group morphism of the additive group of $\mathbf{R}$. An embedding is an injective group morphism, and an automorphism is a group automorphism. More generally, an L_{gp}-morphism between two groups G and H is precisely a group morphism from G to H.

1.2 Terms and formulas

We consider a fixed language $L = C \cup R \cup F$, and a fixed set V the elements of which are called *variables*.

Definition 1.9 (term) The set of *L-terms* is the smallest set containing the constant symbols, the variables, and such that if f is a function symbol of L and $t_1, \ldots, t_{n_f}$ are terms, then $ft_1 \cdots t_{n_f}$ is also a term.

Remarks.
1. An L-term is a finite word in the alphabet $C \cup V \cup F$.
2. Practically, an L-term is constructed inductively: one begins with variables and constant symbols and apply function symbols. The *complexity* $c(t)$ of term t is defined inductively as follows: variables and constant symbols have complexity 0, and a term $ft_1 \cdots t_{n_f}$ has complexity $1 + \max(c(t_1), \ldots, c(t_{n_f}))$.
3. A term is uniquely determined in the following sense: it is either a constant symbol, or a variable, or written in a unique way as $ft_1 \cdots t_{n_f}$ (and only one of these 3 possibilities holds).

Notations.
1. If t is a term and $x_1, \ldots, x_n$ are distinct variables, we write $t(x_1, \ldots, x_n)$ to indicate that **all** variables appearing in t (but possibly more) are among $x_1, \ldots, x_n$. Note that adding variables in between the brackets does not alter the term.
2. We may add parentheses to ease the reading and often write $f(t_1, \ldots, t_{n_f})$ for $ft_1 \cdots t_{n_f}$. If $n_f = 2$, we also write $(t_1 f t_2)$ (or even $t_1 t_2$ if there is no ambiguity on f) instead of $ft_1 t_2$.
3. (substitution in a term) If $t(x_1, \ldots, x_n)$ is a term and $t_1, \ldots, t_n$ are terms, one defines the term $t((t_1, \ldots, t_n))$ inductively on $c(t)$ by replacing in t every occurence of x_i by t_i.

Example 1.10 In the language of rings, $\times + xy - z1$ is a term (written also $\times(+(x, y), -(z, 1))$). For ease of reading, we prefer to write it $(x + y) \times (z - 1)$, which requires the use of parentheses. Note for instance that $(x + y) + z$ and $x + (y + z)$ are different terms (++xyz and +x+yz respectively).

Definition 1.11 (atomic formula, formula) An *atomic formula in the language L* is an expression of the form $r(t_1, \ldots, t_{n_r})$ where r is in R and $t_1, \ldots, t_{n_r}$ are L-terms. The set of *L-formulas* is the smallest set containing all atomic formulas and such that

1. if φ is a formula, then $\neg\varphi$ is also a formula;
2. if φ and ψ are formulas, then $\wedge\varphi\psi$ is a formula (written $(\varphi \wedge \psi)$);
3. if φ is a formula and x a variable, then $\exists x\varphi$ is a formula.

Remarks. 1. An L-formula is a finite word in the alphabet $L \cup V \cup \{\neg, \wedge, \exists\}$. $\neg$ is called the *negation* symbol, $\wedge$ the *conjunction* symbol and $\exists$ the *existential quantifier*.

2. As the language contains equality, for every term t_1 and t_2, the expression $t_1 = t_2$ is an atomic formula.
3. (uniqueness of reading) An L-formula is of one (and only one) of the following forms: a (unique) atomic formula $r(t_1, \ldots, t_{n_r})$, the negation $\neg\varphi$ of a (unique) formula φ, the conjunction $\wedge\varphi\psi$ of two unique formulas φ and ψ, or $\exists x\varphi$ for a unique variable x and formula φ.
4. We have used the notation $\wedge\varphi\psi$ instead of the usual $(\varphi \wedge \psi)$ to state the above uniqueness result without having to cope with parentheses. From now on, we shall use the usual notation with parentheses, and allow ourselves to add more parentheses around subformulas to ease the reading.
5. As with terms, formulas are constructed inductively, starting from atomic formulas and taking negations, conjunctions and existential quantifiers. The *complexity* $c(\varphi)$ of a formula φ is defined inductively: it equals 0 for atomic formulas, $1 + c(\varphi)$ for $(\neg\varphi)$ and $\exists x\varphi$, and $1 + \max(c(\varphi), c(\psi))$ for $\varphi \wedge \psi$.
6. $\exists x\varphi$ is a formula even if the variable x does not appear in φ.
7. As formulas are defined inductively, to show (or to define) that a given property P holds for every formula, one shows (or defines) that P holds for atomic formulas (which may require another induction on the complexity of terms); then one shows that if φ and ψ satisfy P, then so do $\varphi \wedge \psi$, $\neg\varphi$ and $\exists x\varphi$.

Notations. 1. If r is a binary relation symbol, we often write (xry) instead of $r(x, y)$.

2. If φ and ψ are formulas, we use the abbreviations

$$\varphi \vee \psi, \quad \varphi \to \psi, \quad \text{and} \quad \varphi \leftrightarrow \psi$$

respectively for $\neg(\neg\varphi \wedge \neg\psi)$, for $\psi \vee \neg\varphi$ and for $(\varphi \to \psi) \wedge (\psi \to \varphi)$. The symbol $\vee$ is called the *disjunction* symbol. $\neg$, $\wedge$ and $\vee$ are called *Boolean operations*.

3. We also write $\forall x\varphi$ for $\neg\exists x\neg\varphi$. The symbol $\forall$ is called the *universal quantifier*.

Remark. We could have introduced the symbols $\vee$ and $\forall$ earlier together with $\wedge$ and $\exists$ directly in the alphabet needed to build a formula. Our choice of the definition of a formula has the advantage of simplifying many definitions and proofs, but the drawback of breaking the symmetry between $\wedge$ and $\vee$ on one hand, and $\exists$ and $\forall$ on the other hand, for instance in the definition of the complexity of a formula.

Definition 1.12 (free variable, bounded variable) Let φ be a formula. The occurence of a variable in φ can be either *bounded* by a quantifier, or otherwise *free* (note that a variable can have both bounded and free occurences). The precise definition is by induction on the complexity $c(\varphi)$: if φ is atomic, every variable occurence is free. If $\varphi = \neg\psi$, the free occurences of a variable x in φ are the same as the free occurences of x in ψ. If $\varphi = \psi_1 \wedge \psi_2$, the free occurences of x in φ are the union of the free occurences of x in ψ_1 and ψ_2. If $\varphi = \exists x\psi$, then every occurence of x in φ is bounded, and the freeness of occurences of variables other than x in φ remains the same as in ψ.

Example 1.13 In the language $\{\leqslant\}$, consider the formula $\big(\exists x(x \leqslant y \wedge x \leqslant z)\big) \wedge x \leqslant y$: the variable x has three bounded occurences and a free one; all the occurences of y and z are free.

Definition 1.14 (sentence) A *sentence* is a formula in which every occurence of every variable is bounded.

Notation. If φ is a formula and $x_1, \ldots, x_n$ distinct variables, we shall write $\varphi(x_1, \ldots, x_n)$ to indicate that all variables having a free occurence in φ (but possibly more) are among $x_1, \ldots, x_n$. Note that adding additional variables in between the brackets does not alter the formula.

Definition 1.15 (substitution in a formula) Let $\varphi(x_1, \ldots, x_n)$ be a formula, and let $t_1, \ldots, t_n$ be terms. We define a formula $\varphi((t_1, \ldots, t_n))$ by replacing every **free** occurence of x_i by t_i.

Remark-Definition 1.16 (terms compatible with a formula) Note that if $\psi(y)$ is the formula $\exists x(x \neq y)$, then $\psi((x))$ is the formula $\exists x(x \neq x)$. In order to avoid unrequired interactions between bounded variables in $\varphi(x_1, \ldots, x_n)$ and variables in the terms $t_1, \ldots, t_n$ replacing $x_1, \ldots, x_n$, one should apply Definition 1.15 in the case when the variables occurences in $t_1, \ldots, t_n$ are free when being substituted in φ, that is when substitution does not change the number of bounded occurences of any variable. In that case, we say that the terms $t_1, \ldots, t_n$ are **compatible** with $\varphi(x_1, \ldots, x_n)$. For instance x is not compatible with $\exists x(x \neq y)$, but it is comptible with $\exists z(z \neq y)$. Note that $(x_1, \ldots, x_n)$ are always compatible with $\varphi(x_1, \ldots, x_n)$.

1.3 Interpretation of a term, satisfaction of a formula

Let L be a fixed language. Until now, an L-term t and an L-formula φ have been defined merely as strings of characters. We now define their meaning in a given L-structure (M, L^M).

Definition 1.17 (interpretation of a term at $\bar{a}$) Let $t(x_1, \ldots, x_n)$ be an L-term and $(a_1, \ldots, a_n)$ elements of M. The *interpretation* $t^M(a_1, \ldots, a_n)$ of the term t in M is an element of M defined inductively on the complexity of t: if t is a constant symbol c, then it is c^M. If t is a variable x_i, then it is a_i. If t is $ft_1 \ldots t_{n_f}$, then it is $f^M(t_1^M(a_1, \ldots, a_n), \ldots, t_{n_f}^M(a_1, \ldots, a_n))$.

Examples 1.18
1. In the language $\{+, e\}$, let $t(x, y, z)$ be the term $(x + y) + (z + e)$. Let a, b, c be three elements of the structure $(\mathbf{R}, +, 1)$. Then $x(a, b, c) = a$, $y(a, b, c) = b$ and $z(a, b, c) = c$ hence $(x + y)(a, b, c) = a + b$, so that $t(a, b, c)$ equals $a + b + c + 1$.
2. Note that the interpretation in M of a term $t(x_1, \ldots, x_n)$ defines a function t^M from M^n to M that maps $\bar{a}$ to $t^M(\bar{a})$. For instance, in the language of rings, for the natural ring structure on $\mathbf{R}$, these functions are precisely the polynomial functions having coefficients in $\mathbf{Z}$, the smallest L_{ring}-substructure of $\mathbf{R}$ (exercise).

Lemma 1.19 (Substitution Lemma for terms) *Let $\bar{x}$ be an n-tuple of variables, $t(\bar{x}), t_1(\bar{x}), \ldots, t_n(\bar{x})$ terms and $\bar{a}$ in M^n. Then one has*

$$t((t_1, \ldots, t_n))^M(\bar{a}) = t^M(t_1^M(\bar{a}), \ldots, t_n^M(\bar{a})).$$

Remark. The Substitution Lemma is a practical tool that provides a natural writing of the interpretation of a term of the form $t((t_1, \ldots, t_n))$ using the existing interpretation of the terms $t, t_1, \ldots, t_n$ instead of computing everything from the beginning. It is a kind of 'divide and conquer' algorithm. If t is the term $f(x_1, \ldots, x_n)$ for a function symbol f (hence has complexity 1), there is no point in invoking the Substitution Lemma to interpret $f((t_1, \ldots, t_n))$, which is simply $f(t_1, \ldots, t_n)$, whose interpretation is given by the inductive step of Definition 1.16; in that simple case, the Substitution Lemma and the Definition provide the same writing: $f^M(t_1^M(\bar{a}), \ldots, t_n^M(\bar{a}))$.

Proof. By induction on the complexity $c(t)$. If t is a variable x_i, then $t((t_1, \ldots, t_n))$ is t_i so one has

$$t((t_1, \ldots, t_n))^M(\bar{a}) = t_i(\bar{a}) = t^M(t_1^M(\bar{a}), \ldots, t_n^M(\bar{a})).$$

If t is a constant symbol c, then both sides equal c^M. If t is $fs_1(\bar{x}) \cdots s_m(\bar{x})$, then $t((t_1, \ldots, t_n))$ is the term $fs_1((t_1, \ldots, t_n)) \cdots s_m((t_1, \ldots, t_n))$, so one has

$$\begin{aligned} t((t_1, \ldots, t_n))^M(\bar{a}) &= f^M s_1^M((t_1, \ldots, t_n))(\bar{a}) \cdots s_m^M((t_1, \ldots, t_n))(\bar{a}) && \text{(def. of interpretation)} \\ &= f^M s_1^M(t_1^M(\bar{a}), \ldots, t_n^M(\bar{a})) \cdots s_m^M(t_1^M(\bar{a}), \ldots, t_n^M(\bar{a})) && \text{(by induction hyp.)} \\ &= t^M(t_1^M(\bar{a}), \ldots, t_n^M(\bar{a})). && \text{(def. of interpretation)} \end{aligned}$$

□

Definition 1.20 (satisfaction of a formula at $\bar{a}$) Let $\varphi(\bar{x})$ be a formula ($\bar{x}$ stands for $(x_1, \ldots, x_n)$) and $\bar{a} = (a_1, \ldots, a_n)$ in M^n. We define the fact that M *satisfies* $\varphi(\bar{a})$ inductively as follows:

1. If φ is the atomic formula $r(t_1(\bar{x}), \ldots, t_{n_r}(\bar{x}))$, then M satisfies $\varphi(\bar{a})$ if and only if $(t_1^M(\bar{a}), \ldots, t_{n_r}^M(\bar{a}))$ belongs to r^M.
2. If φ is the formula $\varphi_1 \wedge \varphi_2$, then M satisfies $\varphi(\bar{a})$ if and only if M satisfies both $\varphi_1(\bar{a})$ and $\varphi_2(\bar{a})$.
3. If φ is the formula $\neg\psi$, then M satisfies $\varphi(\bar{a})$ if and only if M does not satisfy $\psi(\bar{a})$.
4. If φ is the formula $\exists y \psi(\bar{x}, y)$, then M satisfies $\varphi(\bar{a})$ if and only if there exists b in M such that M satisfies $\psi(\bar{a}, b)$.

Notations. 1. We write $M \models \varphi(\bar{a})$ when M satisfies $\varphi(\bar{a})$. For a sentence σ, as there are no free variables involved, satisfaction does not depend on the tuple $\bar{a}$ so we write $M \models \sigma$. Note that M satisfies $\varphi(\bar{a})$ if and only if M satisfies the sentence $\varphi((\bar{a}))$ in the language L augmented with n new constant symbols interpreted as $a_1, \ldots, a_n$.

2. If $\varphi(x_1, \ldots, x_n)$ is a formula, we write $M \models \varphi$ as an abbreviation for $M \models \forall x_1 \cdots \forall x_n \varphi$ and we also say that M *satisfies the formula* φ.
3. If Λ is a set of formulas, we write $M \models \Lambda$ when M satisfies every formula in Λ, and we say that M *satisfies* Λ. We usually write Λ for a set of formulas, and Σ for a set of sentences.

Remarks. 1. One can check that M satisfies $(\varphi_1 \vee \varphi_2)(\bar{a})$ iff M satisfies $\varphi_1(\bar{a})$ or $\varphi_2(\bar{a})$, and that M satisfies $\forall y \psi(y, \bar{a})$ iff for every b in M, M satisfies $\psi(b, \bar{a})$.

2. By definition of $M \models \neg\varphi(\bar{a})$, one has either

$$M \models \varphi(\bar{a}) \quad \text{or otherwise} \quad M \models \neg\varphi(\bar{a}).$$

Examples 1.21 1. In the language of orderings, the sentence

$$\forall x \forall y \forall z\, [((x \leqslant y \wedge\ y \leqslant z) \to x \leqslant z) \wedge (x \leqslant x) \wedge ((x \leqslant y \wedge y \leqslant x) \to x = y) \wedge (x \neq y \to x \leqslant y \vee y \leqslant x)]$$

holds in an L_{ord}-structure $(M, \leqslant^M)$ if and only if $\leqslant^M$ is a linear ordering on M.

2. In the language of monoids, the sentence

$$\forall x \forall y \forall z\, [xy = yx \wedge (xy)z = x(yz) \wedge (\exists t(tx = e \wedge xt = e)) \wedge xe = x \wedge ex = x]$$

holds in an L_m-structure M if and only if M is an Abelian group (xy stands for $x \times y$).

3. Let $(M, \times, ^{-1}, e)$ be a group considererd as a a structure in the language of groups, and let us write $[x, y]$ for $x^{-1}y^{-1}xy$. Then the sentence

$$(\exists x \exists y [x, y] \neq e) \wedge (\forall x \forall y \forall z [[x, y], z] = e)$$

holds in M if and only if M is a nilpotent group of class 2.

4. In the language of rings, let $(M, +, \times, -, 0, 1)$ be a ring. Then the sentence

$$\forall x \forall y \forall z (z \neq 0 \to \exists t (x + yt + zt^2 = 0))$$

holds in M if every polynomial of degree 2 with coefficients in M has a root in M (here, t^2 stands for $t \times t$; as we work in a ring, $+$ is associative and we use the usual notations to simplify the writing of the formula without any ambiguity about its interpretation in M).

Lemma 1.22 (Substitution Lemma for formulas) *Let $\bar{x}$ be an n-tuple of variables, $\varphi(\bar{x})$ a formula, $t_1(\bar{x}), \ldots, t_n(\bar{x})$ terms compatible with φ (*i.e. *the occurences of $x_1, \ldots, x_n$ in every t_i are free when substituted in φ). Let $\bar{a}$ be an element of M^n. Then one has*

$$M \models \varphi((t_1, \ldots, t_n))(\bar{a}) \iff M \models \varphi(t_1^M(\bar{a}), \ldots, t_n^M(\bar{a})).$$

Proof. If t_i is x_i for every i, there is nothing to prove, so we assume that $x_1, \ldots, x_n$ have no bounded occurence in φ and proceed by induction on the complexity $c(\varphi)$. If φ is an atomic formula $r(s_1(\bar{x}), \ldots, s_m(\bar{x}))$, then $\varphi((t_1, \ldots, t_n))$ is the formula $r(s_1((t_1, \ldots, t_n)), \ldots, s_m((t_1, \ldots, t_n)))$

and we apply Lemma 1.19. If φ is the formula $\varphi_1(\bar{x}) \wedge \varphi_2(\bar{x})$, then $\varphi((t_1,\ldots,t_n))$ is $\varphi_1((t_1,\ldots,t_n)) \wedge \varphi_2((t_1,\ldots,t_n))$ and

$$\begin{aligned} M \models \varphi((t_1,\ldots,t_n))(\bar{a}) \quad &\iff M \models \varphi_1((t_1,\ldots,t_n))(\bar{a}) \text{ and } M \models \varphi_2((t_1,\ldots,t_n))(\bar{a}) \\ &\iff M \models \varphi_1(t_1^M(\bar{a}),\ldots,t_n^M(\bar{a})) \text{ and } M \models \varphi_2(t_1^M(\bar{a}),\ldots,t_n^M(\bar{a})) \\ &\iff M \models \varphi(t_1^M(\bar{a}),\ldots,t_n^M(\bar{a})) \end{aligned}$$

The argument is similar when φ is $\neg\psi$. If φ is the formula $\exists y\psi$ (where $\psi(y,\bar{x})$ and y does not appear among $x_1,\ldots,x_n$), then $\varphi((t_1,\ldots,t_n))$ is $\exists y(\psi((y,t_1,\ldots,t_n)))$ and the terms $(y,t_1,\ldots,t_n)$ are compatible with ψ, so we have

$$\begin{aligned} M \models \varphi((t_1,\ldots,t_n))(\bar{a}) \quad &\iff \text{there exists } b \in M \text{ s.t. } M \models \psi((y,t_1,\ldots,t_n))(b,\bar{a}) && \text{(by def. of } \models) \\ &\iff \text{there exists } b \in M \text{ s.t. } M \models \psi(b,t_1^M(\bar{a}),\ldots,t_n^M(\bar{a})) && \text{(by induction)} \\ &\iff M \models \varphi(t_1^M(\bar{a}),\ldots,t_n^M(\bar{a})). && \text{(by def. of } \models) \end{aligned}$$

□

Definition 1.23 (universal truth, logical equivalence)

1. A statement ψ is *universally true* if it is satisfied by every L-structure. An L-formula $\varphi(x_1,\ldots,x_n)$ is *universally true* if the statement $\forall x_1 \cdots \forall x_n \varphi$ is (note that this does not depend on the ordered list $\bar{x}$).
2. Two formulas φ and ψ are *logically equivalent* if the formula $\varphi \leftrightarrow \psi$ is universally true.

Notation. If φ is a universally true formula, we write $\models \varphi$ instead of $\{\} \models \varphi$.

We shall be interested in the complexity of formulas up to logical equivalence, focusing particularly on quantifiers. Here are the simplest forms.

Definition 1.24 (quantifier-free, existential, universal and prenex formulas)

1. A formula is *quantifier-free* if it is a Boolean combination of atomic formulas.
2. A formula is *existential* if it is of the form $\exists x_1 \cdots \exists x_n \varphi$ where φ is a quantifier free formula.
3. A formula is *universal* if it is of the form $\forall x_1 \cdots \forall x_n \varphi$ where φ is quantifier-free.
4. A formula is *prenex* if it is of the form $Q_1 x_1 \cdots Q_n x_n \varphi$ where $Q_1,\ldots,Q_n$ are quantifiers ($\exists$ or $\forall$) and φ is a quantifier-free formula.

Exercises 1.25

1. Every formula is logically equivalent to a prenex formula.
2. Every universal formula is logically equivalent to the negation of an existential formula.

1.4 Theories, models, semantic consequences and satisfiability

Let L be a fixed language and M an L-structure.

Definition 1.26 (theory) A *theory* (or *L-theory* if the language is not obvious from the context) is a set of sentences in the language L.

Definition 1.27 (theory of a structure) The *theory of M* is the set of all L-sentences satisfied by M.

Notation. We write $\mathrm{Th}(M)$ or $\Sigma(M)$ for the theory of M.

Definition 1.28 (model of a theory) Let Σ be a theory. We say that M is a *model of* Σ if M satisfies every sentence of Σ.

Definition 1.29 (semantic consequence) Let Λ be a set of formulas. If every structure that satisfies Λ also satisfies the formula φ, we say that φ is a *semantic consequence* of Λ and we write $\Lambda \models \varphi$.

Definition 1.30 (satisfiability) A set of L-formulas is *satisfiable* if it is satisfied by some L-structure.

Remark 1.31 (link between satisfiability and semantic consequence) For a sentence σ, one has $\Lambda \models \sigma$ if and only if $\Lambda \cup \{\neg\sigma\}$ is not satisfiable.

CHAPTER 2

SEMANTIC CONSEQUENCE, SYNTACTIC CONSEQUENCE AND THE COMPLETENESS THEOREM

The notion of satisfaction of a sentence σ in a structure M, using the interpretation of the language in M, provides a notion of *semantic* consequence (*i.e.* related to the *meaning* of sentences): we write $\Sigma \models \sigma$ if, given a set of sentences Σ, every structure that satisfies Σ also satisfies the sentence σ. In this Chapter, we define a notion of *syntactic* consequence (*i.e.* related to general *deduction rules* between sentences, regardless of their possible interpretations in a particular structure): we shall write $\Sigma \vdash \sigma$ if there is a *formal proof* of σ from the axiom system Σ. Gödel's Completeness Theorem (Theorem 2.24) asserts that these two notions of consequence actually coincide.

2.1 Logical axioms: tautologies, equality axioms and ∃ axioms.

We fix a given language L and introduce the logical axioms that will be used to define formal proofs. These axioms are of three kinds: tautologies, equality axioms and existential quantifier axioms.

Definition 2.1 (formula of sentential logic) Let S be a **countable** set whose elements are called *sentential variables* $a_1, a_2, \dots$. The set of *sentential formulas* is the smallest set containing the sentential variables and such that if B and C are sentential formulas, then $\wedge BC$ (written $B \wedge C$) and $\neg B$ are also sentential formulas.

A sentential formula is a finite word in the alphabet $S \cup \{\neg, \wedge\}$ (from which we define the symbols $\vee$, $\to$ and $\leftrightarrow$), constructed inductively: one begins with sentential formulas and apply $\wedge$ and $\neg$. The *complexity* $c(A)$ of a sentential formula is defined inductively: it is 0 for variables, $1 + c(A)$ for $\neg A$ and $1 + \max(c(A), c(B))$ for $A \wedge B$. We write $A(a_1, \dots, a_n)$ for a sentential formula where the sentential variables appearing in A are among $a_1, \dots, a_n$. Sentential variables are thought of as sentences having truth value either 0 or 1.

Definition 2.2 (truth function of a sentential formula) To every sentential formula A is associated a *truth function* f_A from $\{0,1\}^{\mathbf{N}}$ to $\{0,1\}$ that maps a choice $(t_1, t_2, \dots)$ for the truth values of all the sentential variables $(a_1, a_2, \dots)$ of S to a truth value $f_A(t_1, t_2, \dots)$ of A. The values of f_A are defined inductively on the complexity of A: if A is a sentential variable a_n, then $f_A(t_1, \dots, t_n, \dots) = 1$ if and only if $t_n = 1$. If A is $\neg B$, then $f_A = 1 - f_B$. If A is $B \wedge C$, then $f_A = f_B f_C$.

Definition 2.3 (sentential tautology) A sentential formula A is a *tautology* if $f_A = 1$.

Exercise 2.4 If A and B are sentential formulas, compute $f_{A\vee B}$, $f_{A\to B}$ and $f_{A\leftrightarrow B}$ and show that $A \vee \neg A$, $A \to (B \to A)$, $(A \wedge B) \to A$ and $(A \to B) \leftrightarrow (\neg B \to \neg A)$ are tautologies.

Given a sentential formula $A(a_1, \dots, a_n)$ and L-formulas $\varphi_1(\bar{x}), \dots, \varphi_n(\bar{x})$, we define $A(\varphi_1, \dots, \varphi_n)$ by replacing in A every occurence of a_i by $\varphi_i(\bar{x})$. This is an L-formula in free variables among $\bar{x}$, hence written $A(\varphi_1, \dots, \varphi_n)(\bar{x})$, and one can show by induction on the complexity of A that for every L-structure M and $\bar{a}$ in M,

$$(1) \qquad M \models A(\varphi_1, \dots, \varphi_n)(\bar{a}) \iff f_A(\varphi_1^M(\bar{a}), \dots, \varphi_n^M(\bar{a})) = 1,$$

where $\varphi^M(\bar{a})$ is the *truth value of* $\varphi(\bar{a})$ *in* M, defined by

$$\varphi^M(\bar{a}) = 1 \text{ if } M \models \varphi(\bar{a}), \quad \text{or} \quad \varphi^M(\bar{a}) = 0 \text{ if } M \not\models \varphi(\bar{a}).$$

Definition 2.5 (L-tautology) An *L-tautology* is an L-formula of the form $A(\varphi_1, \dots, \varphi_n)$ obtained from a sentential tautology $A(a_1, \dots, a_n)$. For instance, if φ and ψ are L-formulas, then $\varphi \vee \neg\varphi$, $\varphi \to (\psi \to \varphi)$ and $(\varphi \wedge \psi) \to \varphi$ are L-tautologies.

Lemma 2.6 *An L-tautology is universally true.*

Proof. Use (1). □

Definition 2.7 (equality axioms) The following sentences are called *equality axioms* in L.

1. $\forall x(x = x)$,
2. $\forall x \forall y(x = y \to y = x)$,
3. $\forall \bar{x} \forall \bar{y}(\bar{x} = \bar{y} \to f(\bar{x}) = f(\bar{y}))$ for every function f symbol in L.
4. $\forall \bar{x} \forall \bar{y}\big[(\bar{x} = \bar{y} \wedge r(\bar{x})) \to r(\bar{y})\big]$ for every relation symbol r in L.

Remark. Transitivity of equality follows from 4. applied to the relation symbol $=$.

Lemma 2.8 *Equality axioms are universally true.*

Proof. Immediate from the assumption that $=$ is always interpreted by usual equality. □

Definition 2.9 (existential quantifier axioms) The *existential quantifier axioms* are sentences of the form $\exists x\psi \leftrightarrow \exists x \neg\neg\psi$ where ψ is a formula, or of the form $\varphi((t, x_2, \dots, x_n)) \to \exists x_1 \varphi$ where $\varphi(x_1, \dots, x_n)$ is a formula and $t(x_1, \dots, x_n)$ a term such that the terms $t, x_2, \dots, x_n$ are **compatible** with φ.

Lemma 2.10 *Existential quantifier axioms are universally true.*

Proof. All the formulas involved are in free variables among $\bar{x}$. If $M \models \varphi((t, x_2, \dots, x_n))(\bar{a})$ holds for some L-structure M and $\bar{a} = (a_1, \dots, a_n)$ in M, then one has $M \models \varphi(t^M(\bar{a}), a_2, \dots, a_n)$ by the Substitution Lemma for formulas, so one has $M \models (\exists x_1 \varphi)(\bar{a})$. □

2.2 Deduction rules: modus ponens and generalisation rule

Definition 2.11 (deduction by modus ponens) Let φ_1, φ_2 and ψ be formulas. We say that ψ *is deduced by modus ponens from* φ_1 *and* φ_2 if φ_2 is the formula $\varphi_1 \to \psi$.

Definition 2.12 (deduction by generalisation) Let φ and ψ be formulas. We say that ψ *is deduced by generalisation from* φ if φ is of the form $\varphi_1 \to \varphi_2$ and ψ is the form $\varphi_1 \to \forall x \varphi_2$ for two formulas φ_1 and φ_2 and some variable x that has **no free occurence in** φ_1.

2.3 Formal proofs, syntactic consequences and coherence

Definition 2.13 (formal proof) Let Λ be a set of formulas, and φ a formula. A *formal proof of* φ *from* Λ is a finite sequence of formulas $\varphi_1, \dots, \varphi_n$ where φ_n is φ and such that for all $k \leqslant n$, either φ_k is in Λ, or φ_k is a logical axiom, or φ_k is deduced by modus ponens from two formulas φ_i and φ_j with $i < k$ and $j < k$, or φ_k is deduced by generalisation from φ_i with $i < k$. In that case, we say that Λ *proves* φ.

Remarks.
1. $\varphi(\bar{x})$ can have free variables $\bar{x}$. From the generalisation rule, it follows that if there is a formal proof of φ from Λ, then Λ actually proves $\forall \bar{x} \varphi$ (see exercise 2.15).
2. Reciprocally, if Λ proves a formula $\forall x \varphi$, then Λ also proves φ (see exercise 2.15).
3. As every proof involves finitely many formulas, if Λ proves φ, then there is a finite subset $\Lambda_0 \subset \Lambda$ such that Λ_0 proves φ.

Definition 2.14 (syntactic consequence) If there is a formal proof of φ from Λ, we say that φ is a *syntactic consequence* of Λ, and we write $\Lambda \vdash \varphi$.

Notations. We write $\vdash \varphi$ instead of $\{\} \vdash \varphi$, meaning that every set of formulas proves φ. If Λ_1 and Λ_2 are two sets of formulas, we write $\Lambda_1 \vdash \Lambda_2$ if $\Lambda_1 \vdash \varphi_2$ for every $\varphi_2 \in \Lambda_2$.

Exercise 2.15 Let $\varphi_1, \dots, \varphi_n, \varphi$ and ψ be formulas and Λ a set of formulas, possibly empty.

1. (conjunction) If $\Lambda \vdash \{\varphi_1, \dots, \varphi_n\}$, then $\Lambda \vdash \varphi_1 \wedge \cdots \wedge \varphi_n$.
2. (contrapositive) $\Lambda \vdash \varphi \to \psi$ if and only if $\Lambda \vdash \neg\psi \to \neg\varphi$.
3. (universal quantifier axiom) $\vdash \forall x_1 \varphi \to \varphi((t, x_2, \dots, x_n))$ where $\varphi(x_1, \dots, x_n)$, t is a term and $(t, x_2, \dots, x_n)$ are compatible with φ.
4. ($\forall$ rule) $\Lambda \vdash \varphi$ if and only if $\Lambda \vdash \forall x \varphi$.
5. (introduction of $\exists$) If x has no free occurence in ψ and $\Lambda \vdash \varphi \to \psi$, then $\Lambda \vdash \exists x \varphi \to \psi$.

Theorem 2.16 (a syntactic consequence is a semantic one) *If $\Lambda \vdash \varphi$, then $\Lambda \models \varphi$.*

Proof. Inductively on the length n of the proof. If $n = 1$, then φ is either in Λ, or a logical axiom, hence universally true by Lemmas 2.6, 2.8 and 2.10. In both cases, $\Lambda \models \varphi$. Assume that $\varphi_1, \dots, \varphi_{n-1}$ are semantic consequences of Λ. If φ_n is deduced by modus ponens from φ_i and $\varphi_j = \varphi_i \to \varphi_n$ and if M is a model of Λ, then $M \models \varphi_i$ and $M \models \varphi_j$, so $M \models \varphi_n$. If φ_n is deduced by the generalisation rule, it is of the form $\alpha \to \forall x \beta$ with $M \models \alpha \to \beta$, hence $M \models \forall x(\alpha \to \beta)$. Since x has no free occurence in α, one has $M \models \alpha \to \forall x \beta$. ☐

Definition 2.17 (coherence) A set of formulas Λ is *contradictory* if there is a formula φ such that $\Lambda \vdash \varphi$ and $\Lambda \vdash \neg\varphi$. Otherwise, it is *coherent*(1).

Remarks.
1. If Λ is contradictory, then Λ proves all formulas (use the tautology $A \to (\neg A \to B)$).
2. If Λ is contradictory, there is a finite subset $\Lambda_0 \subset \Lambda$ which is contradictory. In particular, Λ is coherent if and only if every finite subset of Λ is coherent.
3. The notion of coherence is syntactic. The corresponding semantic notion is satisfiability. A set Λ of formulas that is satisfiable is coherent, for otherwise it would prove $\exists x(x \neq x)$, so any model M of Λ would satisfy $\exists x(x \neq x)$ by Theorem 2.16.

Lemma 2.18 (Deduction Lemma) *If Λ is a set of formulas, φ a formula and σ a* ***sentence****, then*

$$\Lambda \cup \{\sigma\} \vdash \varphi \quad \textit{if and only if} \quad \Lambda \vdash \sigma \to \varphi.$$

Proof. If Λ proves $\sigma \to \varphi$, then $\Lambda \cup \{\sigma\}$ also does, so $\Lambda \cup \{\sigma\}$ proves φ by modus ponens. For the reverse implication, let $\varphi_1, \dots, \varphi_n$ be a proof of φ from $\Lambda \cup \{\sigma\}$. We show by induction on n that Λ proves $\sigma \to \varphi_i$ for every i. If $n = 1$, then φ_1 is either σ or in Λ or a logical axiom. In the first case, the result follows from the tautology $\sigma \to \sigma$. In the two last cases, one has $\Lambda \vdash \varphi_1$, as well as the tautology $\varphi_1 \to (\varphi \to \varphi_1)$, from which $\Lambda \vdash \varphi \to \varphi_1$. Induction step: if φ_n is a logical axiom or in $\Lambda \cup \{\sigma\}$, we conclude as before. If φ_n is deduced by modus ponens from φ_i and $\varphi_j = \varphi_i \to \varphi_n$, then $\Lambda \vdash \sigma \to \varphi_i$ and $\Lambda \vdash \sigma \to (\varphi_i \to \varphi_n)$. Using the tautology $((A \to B) \wedge (A \to (B \to C))) \to (A \to C)$, it follows that $\Lambda \vdash \sigma \to \varphi_n$. If φ_n is deduced by generalisation, it is of the form $\alpha \to \forall x \beta$, and we have $\Lambda \vdash \sigma \to (\alpha \to \beta)$ by the induction hypothesis, that is $\Lambda \vdash \sigma \wedge \alpha \to \beta$, hence $\Lambda \vdash \sigma \wedge \alpha \to \forall x \beta$ by generalisation (since x has no free occurence in either σ or α), that is $\Lambda \vdash \sigma \to (\alpha \to \forall x \beta)$. ☐

Here is the syntactic analogue to Remark 1.31.

Corollary 2.19 (link between coherence and syntactic consequence) *Let Λ be a set of formulas, and σ a sentence. Λ proves σ if and only if $\Lambda \cup \{\neg\sigma\}$ is contradictory.*

Proof. If Λ proves σ, then so does $\Lambda \cup \{\neg\sigma\}$, so $\Lambda \cup \{\neg\sigma\}$ is contradictory. Reciprocally, if $\Lambda \cup \{\neg\sigma\}$ is contradictory, then $\Lambda \cup \{\neg\sigma\}$ proves every formula and in particular σ, so Λ proves $\neg\sigma \to \sigma$ by the Deduction Lemma. From the tautology $(\neg A \to A) \to A$, one deduces that Λ proves σ. ☐

(1) Margaret Thomas made me realise that the wording *cohérent/contradictoire* is a French convention, whereas the English one is *consistent/inconsistent*. Since I unfortunately began to use coherent/contradictory, I shall keep this wording for sake of consistency.

2.4 A coherent theory has a model

Let L be a fixed language and Σ an L-theory. If Σ has a model M, then Σ is coherent. The purpose of this section is to show the converse: a coherent theory has a model.

Definition 2.20 (Henkin witnesses of a theory) If C is a set of constant symbols of L, we say that C is a set of *Henkin witnesses* for Σ if, for all formulas $\varphi(x)$ (with at most one free variable x), there is some c in C such that Σ proves $\exists x\varphi \to \varphi((c))$.

We shall show

1. that adding constant symbols in the language L does not affect the coherence of Σ,
2. if Σ is coherent, how to build a coherent $L \cup C$-theory Σ_C containing Σ such that C is a set of Henkin witnesses for Σ_C,
3. how to build a model of Σ using the constants in C.

We will restrict to the case where L is a **countable** language and the set of variables V is countable.

Lemma 2.21 (adding one constant symbol does not affect coherence) *Let $\varphi(x)$ be an L-formula and c a constant symbol that is not in L. If $\Sigma \vdash \varphi((c))$ in $L \cup \{c\}$, then $\Sigma \vdash \varphi$ in L.*

Proof. Let $\varphi_1, \dots, \varphi_n$ be an $L \cup \{c\}$-proof of $\varphi((c))$ from Σ. Let y be a variable that does not appear in any φ_i, and let ψ_i be obtained from φ_i by replacing any occurence of c by y. For each $k \leqslant n$, if φ_k is a logical axiom in $L \cup \{c\}$ (tautology, equality axiom, existential quantifier axiom), then ψ_k is a logical axiom in L of the same kind. If φ_k is deduced by modus ponens from φ_i and φ_j, then ψ_k is deduced by modus ponens from ψ_i and ψ_j. If φ_k is $\alpha \to \forall z\beta$ with φ_i equal to the formula $\alpha \to \beta$, then ψ_k is $\gamma \to \forall z\delta$ with ψ_i equal to the formula $\gamma \to \delta$. It follows that $\psi_1, \dots, \psi_n$ is an L-proof of ψ_n from Σ. But ψ_n is $\varphi((y))$, so $\Sigma \vdash \forall y\varphi((y))$ by Exercise 2.15.4. As x is compatible with $\varphi((y))(y)$ and as $\varphi((y))((x))$ is precisely φ, one has $\Sigma \vdash \varphi$ by Exercise 2.15.3. □

Corollary 2.22 *If Σ is a coherent L-theory, it is a coherent $L \cup C$-theory for any set C of constants.*

Proof. If Σ is contradictory as an $L \cup C$-theory, then there are finitely many constants $c_1, \dots, c_n$ in C such that Σ is contradictory as an $L \cup \{c_1, \dots, c_n\}$-theory (and we may choose n minimal as Σ is a coherent L-theory). It follows that Σ proves any $L \cup \{c_1, \dots, c_n\}$-formula, and in particular $\varphi((c_n))$ where $\varphi(x)$ is $x \neq x$. By the previous Lemma, Σ proves $x \neq x$ in the language $L \cup \{c_1, \dots, c_{n-1}\}$, hence is contradictory, a contradiction to n being minimal. □

Lemma 2.23 (Henkin's completion) *Let Σ be a coherent L-theory and C a countable set of constant symbols. There exists an $L \cup C$-theory Σ_C that is coherent and contains Σ, such that C is a set of Henkin witnesses for Σ_C.*

Proof. Let $(c_n)_{n \geqslant 1}$ be an enumeration of C. As $L \cup C$ and V are countable, the set of $L \cup C$-formulas is also countable, so let $(\varphi_n)_{n \geqslant 1}$ be an enumeration of those $L \cup C$-formulas having at most one free variable. Relabelling the variables, we may write x_n for the free variable of φ_n if it exists, or pick any variable in V that we write x_n otherwise, so that one can write $\varphi_n(x_n)$. We build a theory Σ_n inductively starting with $\Sigma_0 = \Sigma$ and setting

$$\Sigma_{n+1} = \Sigma_n \cup \{\exists x_n \varphi_n \to \varphi_n((c_{f(n)}))\}$$

where $f(n)$ is the smallest natural number such that $c_{f(n)}$ appears in none of the finitely many formulas of $\Sigma_n \setminus \Sigma$ that use constant symbols of C. We then define Σ_C to be $\bigcup_{n \geqslant 1} \Sigma_n$. The theory Σ_C contains Σ and has C as a set of Henkin witnesses by construction. We claim that Σ_C is coherent, that is, that Σ_n is a coherent $L \cup C$-theory for every n by induction on n. It is true for $n = 0$ by Corollary 2.22. If Σ_{n+1} is contradictory, then one has

$$\Sigma_n \vdash \neg(\exists x_n \varphi_n \to \varphi_n((c_{f(n)}))),$$

hence

$$\Sigma_n \vdash \exists x_n \varphi_n \wedge \neg\varphi_n((c_{f(n)})),$$

and, by Lemma 2.21

$$\Sigma_n \vdash \exists x_n \varphi_n \wedge \forall x_n \neg\varphi_n(x_n),$$

that is

$$\Sigma_n \vdash \exists x_n \varphi_n \wedge \neg\exists x_n \neg\neg\varphi_n(x_n),$$

so, using the existential quantifier axiom, Σ_n proves $\exists x_n \varphi_n \wedge \neg\exists x_n \varphi_n(x_n)$ and is contradictory. □

Completeness theorem 2.24 (Gödel, 1930) *Let Σ be a theory in a* ***countable*** *language L (using a countable set of variables). The theory Σ is coherent if and only if it has a model.*

Proof. By Lemma 2.23, there is a countable language $L_H \supset L$ and a coherent L_H-theory $\Sigma_C \supset \Sigma$ such that the set of **all** constant symbols C of L_H is a set of Henkin witnesses for Σ_C. We first build a maximal theory with these properties. Let $(\sigma_n)_{n\geqslant 1}$ be an enumeration of all the L_H-sentences. We define Σ_n by induction on n by putting $\Sigma_0 = \Sigma_C$ and

$$\Sigma_{n+1} = \Sigma_n \cup \{\sigma_n\} \text{ if } \Sigma_n \cup \{\sigma_n\} \text{ is coherent, or } \Sigma_{n+1} = \Sigma_n \cup \{\neg\sigma_n\} \text{ otherwise.}$$

Note that if Σ_n is coherent and $\Sigma_n \cup \{\sigma_n\}$ is contradictory, then $\Sigma_n \vdash \neg\sigma_n$ so $\Sigma_n \cup \{\neg\sigma_n\}$ is coherent. It follows that Σ_{n+1} is coherent. Putting $\Sigma_H = \bigcup_{n\geqslant 0} \Sigma_n$, one has

(1) Σ_H is a coherent L_H-theory,
(2) Σ_H contains Σ_C, hence has C as a set of Henkin witnesses,
(3) Σ_H is *complete*, *i.e.* for all L_H-sentence σ, either σ or $\neg\sigma$ is in Σ_H (so $\sigma \in \Sigma_H$ iff $\Sigma_H \vdash \sigma$).

Note that if c is a constant symbol, then Σ_H proves $\exists x(x = c)$, so there must exist a constant symbol d different from c by construction (see Lemma 2.23) such that Σ_H proves $d = c$. We define the relation $\sim$ on C by

$$c \sim d \iff c = d \in \Sigma_H.$$

Claim 1 $\sim$ *is an equivalence relation on the set C of constant symbols.*

Proof of Claim 1. (reflexivity) If $c = c$ is not in Σ_H, then $c \neq c$ is, by (2.4). But then Σ_H proves $\exists x(x \neq x)$ by the existential quantifier axiom, a contradiction with the first equality axiom.
(symmetry) If $c = d$ is in Σ_H but not $d = c$, then $d \neq c$ hence $(c = d) \wedge (d \neq c)$ are in Σ_H, a contradiction with the second equality axiom.
(transitiviy) Similary using the fourth equality axiom and the Remark after Definition 2.7. □

Claim 2 $C/\sim$ *is an L_H-structure.*

Proof of Claim 2. We write M for $C/\sim$ and for every c in C, we write $\tilde{c}$ for the class of c modulo $\sim$. We define

$$c^M = \tilde{c}.$$

For every n-ary relation symbol r, we define,

$$(\tilde{c}_1, \ldots, \tilde{c}_n) \in r^M \iff r(c_1, \ldots, c_n) \in \Sigma_H.$$

This is well-defined since if $c_1 \sim d_1, \ldots, c_n \sim d_n$ and $r(c_1, \ldots, c_n) \in \Sigma_H$ hold, then the fourth equality axiom implies $r(d_1, \ldots, d_n) \in \Sigma_H$, so that the definition of r^M does not depend on the choice of representatives $c_1, \ldots, c_n$ for the classes $\tilde{c}_1, \ldots, \tilde{c}_n$. Note that for the relation symbol $=$, the interpretation $=^M$ coincide with equality on M. If f is an m-ary function symbol, then for all constant symbols $c_1, \ldots, c_m$, the sentence $\exists x f(c_1, \ldots, c_m) = x$ is in Σ_H (by the first equality axiom and the existential quantifier axiom), so there is a constant symbol c such that the sentence $f(c_1, \ldots, c_m) = c$ is in Σ_H (and the sentence $f(d_1, \ldots, d_m) = d$ is in Σ_H for all $d_i \sim c_i$ and $d \sim c$ by the third equality axiom). We thus define

$$f^M(\tilde{c}_1, \ldots, \tilde{c}_m) = \tilde{d} \iff f(c_1, \ldots, c_m) = d \in \Sigma_H.$$

□

Claim 3 (interpretation of a term) *Let t be a term and c a constant symbol. Then*

$$M \models t = c \iff t = c \in \Sigma_H.$$

Proof of Claim 3. By induction on the complexity $c(t)$. It is true for constants by definition of $\sim$, and if t is the term $ft_1 \cdots t_n$, by (2) there exist constant symbols $c_1, \ldots, c_n$ such that $t_i = c_i \in \Sigma_H$ for all i, so that we have $t_i^M = c_i^M$ by the induction hypothesis. It follows that

$$M \models t = c \iff f^M t_1^M \cdots t_p^M = c^M \iff f^M c_1^M \cdots c_p^M = c^M \iff f(c_1, \ldots, c_p) = c \in \Sigma_H.$$

By the third equality axiom, one has $f(c_1, \ldots, c_p) = c \in \Sigma_H$ if and only if $f(t_1, \ldots, t_p) = c \in \Sigma_H$. □

Claim 4 (satisfaction of a sentence) *For a sentence σ, one has*

$$M \models \sigma \iff \sigma \in \Sigma_H.$$

Proof of Claim 4. By induction on the complexity of σ. If $r(t_1, \ldots, t_n)$ is an atomic sentence and $c_1, \ldots, c_n$ are constant symbols such that $t_i = c_i \in \Sigma_H$ for every i, then $t_i^M = c_i^M$ by Claim 3, so

$$M \models r(t_1, \ldots, t_n) \iff (t_1^M, \ldots, t_n^M) \in r^M \iff (c_1^M, \ldots, c_n^M) \in r^M \iff r(c_1, \ldots, c_n) \in \Sigma_H.$$

By the fourth equality axiom, $r(c_1, \ldots, c_n) \in \Sigma_H$ if and only if $r(t_1, \ldots, t_n) \in \Sigma_H$.
If σ is the sentence $\alpha \wedge \beta$, then

$$M \models \alpha \wedge \beta \iff M \models \alpha \text{ and } M \models \beta \iff \alpha \in \Sigma_H \text{ and } \beta \in \Sigma_H \implies \alpha \wedge \beta \in \Sigma_H.$$

Since Σ_H is complete, the converse of the last implication also holds.
If σ is $\neg\alpha$, the result holds again by completeness of Σ_H.
If σ is $\exists x \alpha$ for a formula $\alpha(x)$ of lower complexity, then

$$M \models \sigma \iff \text{there exists } c \in M \text{ such that } \alpha((c)) \in \Sigma_H.$$

By the existential quantifier axiom and modus ponens, this implies $\exists x \alpha \in \Sigma_H$. Conversely, if $\exists x \alpha \in \Sigma_H$, then there is a constant symbol c such that $\alpha((c)) \in \Sigma_H$ by (2) and modus ponens. □

Claim 5 *The L-theory Σ has a model.*

Proof of Claim 5. By the previous Claim, M is an L_H-structure that is a model of Σ_H. As $L \subset L_H$, the restriction L^M of L_H^M to the language L provides a natural interpretation for L in M. As $\Sigma \subset \Sigma_H$, the structure (M, L^M) is a model of Σ. □ □

Remarks.

1. This shows in particular that $\Sigma \models \sigma$ if and only iff $\Sigma \vdash \sigma$.
2. We have shown that Σ has a model M that embeds into $\mathbf{N}$, *i.e.* that is countable. In particular, the L_{field}-theory of the real numbers $\mathbf{R}$ with its natural structure has a countable model $\mathbf{F}$ that satisfies all the L_{field}-sentences satisfied by $\mathbf{R}$.
3. (The hypotheses 'L and V are countable' can be removed assuming the *Axiom of Choice*, one of whose equivalent formulations asserts that any set can be *well ordered*: under this assumption, instead of enumerating formulas $(\varphi_n)_{n \geqslant 1}$ and building the theories Σ_n inductively on the natural number n as is done in Lemma 2.23 and Theorem 2.24, one chooses a *well-ordering* $(\varphi_\alpha)_{\alpha \geqslant 1}$ on the set of formulas and builds Σ_α by *transfinite induction* (see Chapter 4).)

Corollary 2.25 (Compactness Theorem) *Let L be a countable language, and Σ a theory using countably many variables. Then Σ has a model if and only if every finite subset $\Sigma_0 \subset \Sigma$ has a model.*

Proof. If M is a model of Σ, then M is also a model of every (finite) subset of Σ. The converse is the important direction. One has

$$\begin{aligned} \Sigma \text{ has a model} &\iff \Sigma \text{ is coherent} \\ &\iff \text{every finite subset of } \Sigma \text{ is coherent} \\ &\iff \text{every finite subset of } \Sigma \text{ has a model.} \end{aligned}$$ □

Remark. The Compactness Theorem is a semantic corollary of the Completeness Theorem.

Corollary 2.26 *Let L be a countable language and Σ a theory using countably many variables. If Σ has finite models of arbitrary large cardinalities, then Σ has an infinite model.*

Proof. For every natural number $n \geqslant 1$, let M_n be a model of Σ having size at least n. Let $C = \{c_n : n \geqslant 1\}$ be a set of new constant symbols and let $\overline{L} = L \cup C$. Let

$$\overline{\Sigma} = \Sigma \cup \{c_n \neq c_m : n \neq m\}.$$

If $\overline{\Sigma}_0 \subset \overline{\Sigma}$ is a finite subset, then there is a natural number k and a finite $\Sigma_0 \subset \Sigma$ of size at most k such that $\overline{\Sigma}_0 \subset \Sigma_0 \cup \{c_n \neq c_m : n \neq m,\ n, m \leqslant k\}$. If follows that M_k is a model of $\overline{\Sigma}_0$ in the language $\overline{L}$ (one interprets the constants $\{c_n : n \leqslant k\}$ by pairwise distinct elements of M_k and the constants $\{c_n : n > k\}$ by any element for $n > k$). As this holds for any finite subset $\overline{\Sigma}_0 \subset \overline{\Sigma}$, the theory $\overline{\Sigma}$ has a model $(M, \bar{L}^M)$ by Corollary 2.25. M must be infinite since the interpretations of the constants of C are pairwise disjoint, and M is also a model of Σ. □

We finish by giving an explanation for the name of the Compactness Theorem 2.25. Let L be a language and $\mathcal{S}_L$ (or simply $\mathcal{S}$) the space of complete satisfiable L-theories using variables in V (recall that Σ is *complete* if for all L-sentence σ, either σ or $\neg\sigma$ is in Σ). We provide $\mathcal{S}$ with a topology by defining a basis of open sets. For any L-sentence σ, we define the *basic open set* $[\sigma]$ by

$$[\sigma] = \Big\{\Sigma \in \mathcal{S} : \sigma \in \Sigma\Big\}.$$

For any two L-sentences σ and τ, using completeness and satisfiability of any $\Sigma \in \mathcal{S}$, one has

$$[\sigma] \cap [\tau] = \Big\{\Sigma \in \mathcal{S} : \sigma \in \Sigma \text{ and } \tau \in \Sigma\Big\} = \Big\{\Sigma \in \mathcal{S} : \sigma \wedge \tau \in \Sigma\Big\} = [\sigma \wedge \tau].$$

It follows that the set of basic open sets is closed under finite intersections and does form a basis of open sets. By definition, an open subset of $\mathcal{S}$ is of the form $\bigcup_{\sigma\in\Sigma}[\sigma]$ for any set Σ of L-sentences. Also, using the completeness of any $\Sigma \in \mathcal{S}$ again, one has

$$\mathcal{S} \setminus [\sigma] = \Big\{\Sigma \in \mathcal{S} : \sigma \notin \Sigma\Big\} = \Big\{\Sigma \in \mathcal{S} : \neg\sigma \in \Sigma\Big\} = [\neg\sigma].$$

It follows that any basic open set $[\sigma]$ is *clopen* (*i.e.* both closed and open) and that a closed subset of $\mathcal{S}$ is of the form $\bigcap_{\sigma\in\Sigma}[\sigma]$ for any set Σ of L-sentences. We write $\mathcal{S}(\Sigma)$ this closed subset, which is the subset of $\mathcal{S}$ whose elements contain Σ.

Remark. With these notations, an L-theory Σ is satisfiable if and only if $\mathcal{S}(\Sigma)$ is not empty.

Proof. If $\mathcal{S}(\Sigma)$ is not empty, there is a satisfiable theory that contains Σ, so Σ is satisfiable. If Σ has a model M, then Th(M) is a complete satisfiable theory that contains Σ hence belongs to $\mathcal{S}(\Sigma)$. □

Corollary 2.27 *If L and V are countable, $\mathcal{S}$ is a compact Hausdorff topological space.*

Proof. If Σ_1 and Σ_2 are two distinct elements of $\mathcal{S}$, there must be a sentence σ in $\Sigma_1 \setminus \Sigma_2$. As Σ_2 is complete, one has $\neg\sigma \in \Sigma_2$, so that $[\sigma]$ and $[\neg\sigma]$ are disjoint neighbourhoods of Σ_1 and Σ_2 respectively. This shows that $\mathcal{S}$ is Hausdorff. To show that $\mathcal{S}$ is compact, let $\bigcap_{\sigma\in\Sigma}[\sigma] = \mathcal{S}(\Sigma)$ be an empty intersection of closed sets. By the above remark, Σ is not satisfiable. By the Compactness Theorem, there is a finite subset $\Sigma_0 \subset \Sigma$ that is not satisfiable, so that $\bigcap_{\sigma\in\Sigma_0}[\sigma]$ is empty. □

Remarks.
1. An example. In the language of fields L_{field}, the space $S_{L_{field}}$ is a set, some elements of which are: the theory of the field $\mathbf{Q}$ of rationals, the theory of $\mathbf{R}$, the theory of the field $\mathbf{C}$ of complex numbers, the theory of the finite field $\mathbf{F}_{p^n}$ for every n etc.
2. $\mathcal{S}(\Sigma)$ is a closed subset of $\mathcal{S}$ hence compact for the induced topology.
3. One could have similarly defined a topology on the space of **coherent** complete L-theories and showed this space to be compact Hausdorff using the simple fact that a contradictory theory has a contradictory finite subset (without invoking Gödel's Completeness Theorem). Gödel's Completeness Theorem asserts that this latter topological space coincides with the former space $\mathcal{S}$ of **satisfiable** complete theories.

CHAPTER 3

THE COMPACTNESS THEOREM

The Compactness Theorem states that a countable theory Σ has a model provided that every finite subset $\Sigma_0 \subset \Sigma$ has a model M_{Σ_0}. It is a semantic theorem that we derived from the Completeness Theorem using the fact that a formal proof involves only finitely many formulas. We shall construct a model of Σ built by an *ultraproduct* of the models M_{Σ_0}, the ultraproduct construction being an important tool to build a structure M out of a family of structures M_i by controlling the theory of M in terms of the theories of M_i. This will provide another proof of the Compactness Theorem that does not rely on the syntactic notions defined in Chapter 2.

3.1 Filter, ultrafilter

Definition 3.1 (filter) Let I be a set. A *filter* on I is a **non-empty** set $\mathcal{F}$ consisting of subsets of I such that

1. the empty set is not an element of $\mathcal{F}$,
2. (finite intersection property) if J and K are in $\mathcal{F}$, then so is $J \cap K$,
3. (extension property) if J is in $\mathcal{F}$, then so is any bigger $K \supset J$.

Remark. These are properties that one would expect from very large subsets of I.

Definition 3.2 (generated filter) Let I be a set and $\mathcal{A}$ a set of subsets of I. If, for all finitely many $A_1, \ldots, A_n$ in $\mathcal{A}$, the intersection $A_1 \cap \cdots \cap A_n$ is non-empty, then the set of all $B \subset I$ such that there exists n and $A_1, \ldots, A_n$ in $\mathcal{A}$ with $A_1 \cap \cdots \cap A_n \subset B$ is a filter on I called the *filter generated by* $\mathcal{A}$.

Examples 3.3

1. (trivial filters) For any non-empty subset $J \subset I$, the set singleton $\{J\}$ generates a filter. The filters of this kind are called the *principal filters* on I, or the trivial ones.
2. (Fréchet filter) For any **infinite** set I, the set of *cofinite subsets* of I (those whose complement in I is finite) is called the *Fréchet filter* on I.
3. (filter of neighbourhoods) Let X be a topological space and x a point of X. A subset V of X is called a *neighbourhood of* x if there exists an open set O that contains x such that $O \subset V$. The set $\mathcal{V}(x)$ of neighbourhoods of x is a filter on X. If the topology on X is generated by a basis of open sets $\mathcal{B}$, then $\mathcal{V}(x)$ is generated by those elements of $\mathcal{B}$ that contain x.
4. (filter of conegligible sets) Let X be a set, $\mathcal{A}$ an *algebra of subsets of* X (*i.e.* closed under finite intersections and taking complements) and μ a non-zero *finitely additive measure* on $(X, \mathcal{A})$. That is, μ is a map from $\mathcal{A}$ to $[0, +\infty]$ such that $\mu(\emptyset)$ is zero and

$$\mu(A \cup B) = \mu(A) + \mu(B) \text{ for all pairwise disjoint } A \text{ and } B \text{ in } \mathcal{A}.$$

An element A of $\mathcal{A}$ *has comeasure zero* if $\mu(X \setminus A)$ is zero. A subset $B \subset X$ is *conegligible* if there exists an $A \in \mathcal{A}$ having comeasure zero such that $A \subset B$. The set of elements of $\mathcal{A}$ having comeasure zero is closed under finite intersections and does not contain the empty set as μ is non zero, hence generates a filter on X that corresponds to all conegligible subsets of X. In particular, if $(X, \mathcal{B}, \mu)$ is a complete measure space, the conegligible elements of $\mathcal{B}$ form a filter on X.

Definition 3.4 (ultrafilter) An *ultrafilter* $\mathcal{U}$ *on* I is a filter that is maximal for inclusion.

Lemma 3.5 (characterisation of ultrafilters) *Let $\mathcal{F}$ be a filter on I. $\mathcal{F}$ is an ultrafilter if and only if for every subset $J \subset I$, either J or its complement $I \setminus J$ belongs to $\mathcal{F}$*

Proof. Assume that for every $J \subset I$ either J or $I \setminus J$ belongs to $\mathcal{F}$. Let $\mathcal{F}' \supset \mathcal{F}$ be a filter on I and let $J \in \mathcal{F}'$. If $I \setminus J$ is in $\mathcal{F}$, then $J \cap (I \setminus J)$ is in $\mathcal{F}'$, a contradiction. So J is in $\mathcal{F}$ and $\mathcal{F}$ is maximal. Conversely, if $\mathcal{F}$ is an ultrafilter, let $J \subset I$. If $I \setminus J$ is not in $\mathcal{F}$ (in particular J is not empty), then we claim that $\mathcal{A} = \mathcal{F} \cup \{J\}$ generates a filter: as $\mathcal{F}$ is closed under finite intersection, it is enough to show that $F \cap J$ is non-empty for any F in $\mathcal{F}$. But if $F \cap J = \emptyset$, then $F \subset I \setminus J$, so $I \setminus J$ is in $\mathcal{F}$, a contradiction. We have shown that $\mathcal{F} \cup \{J\}$ generates a filter, so J is in $\mathcal{F}$ by maximality of $\mathcal{F}$. □

Example 3.6 The filter generated by a singleton $\{x\}$ of I is a principal ultrafilter. Reciprocally, every principal ultrafilter is generated by a singleton.

Remark. An ultrafilter $\mathcal{U}$ induces a measure μ on the (σ-)algebra of all subsets of X defined for every $Y \subset X$ by putting

$$\mu(Y) = 1 \text{ if } Y \in \mathcal{U} \text{ or } \mu(Y) = 0 \text{ if } Y \notin \mathcal{U}.$$

Lemma 3.7 (characterisation of non-trivial ultrafilters) *Let I be an infinite set and $\mathcal{U}$ an ultrafilter on I. $\mathcal{U}$ is non-principal if and only if it contains all cofinite subsets of I.*

Proof. If $\mathcal{U}$ contains all cofinite subsets of I, it cannot be principal for otherwise it would contain a singleton $\{x\}$ (the generating set), but also its cofinite complement $X \setminus \{x\}$, hence the empty set, a contradiction. If $\mathcal{U}$ is non-principal, $\mathcal{U}$ does not contain any singleton so $\mathcal{U}$ contains the complement of any singleton by Lemma 3.5, hence any finite intersection of such sets, that is, any cofinite set. □

Lemma 3.8 (obtainment of ultrafilters) *Every filter $\mathcal{F}$ on I can be extended to an ultrafilter on I.*

Proof. Let $\mathcal{C}$ be the set of all filters on I extending $\mathcal{F}$. Together with inclusion, $\mathcal{C}$ is a partially ordered set. We shall use Zorn's Lemma, one of the equivalent formulations of the Axiom of Choice, to show that $\mathcal{C}$ has a maximal element.

> **Zorn's Lemma 3.9** *Any non-empty partially ordered set $\mathcal{C}$ that is* inductive *(that is, each of whose totally ordered subset has an upper bound in $\mathcal{C}$) has a maximal element.*

Let $\mathfrak{F} = \{\mathcal{F}_j : j \in J\}$ be a totally ordered subset of $\mathcal{C}$. We write $\mathcal{F}_J$ for $\bigcup_{j\in J} \mathcal{F}_j$ and claim that $\mathcal{F}_J$ is an upper bound of $\mathfrak{F}$ in $\mathcal{C}$. The set $\mathcal{F}_J$ is a set of subsets of I, contains all the elements of $\mathcal{F}$, and of $\mathcal{F}_j$ for every $j \in J$, does not contain the empty set (for otherwise one $\mathcal{F}_j$ would contain it) and satisfies the extension property (for any element of $\mathcal{F}_J$ belongs to a given F_j that satisfies the extension property). If A and B are elements of $\mathcal{F}_J$, say $A \in \mathcal{F}_j$ and $B \in \mathcal{F}_k$, as $\mathfrak{F}$ is totally ordered, one has for example $\mathcal{F}_j \subset \mathcal{F}_k$, so that $A \cap B \in \mathcal{F}_k$ hence $A \cap B \in \mathcal{F}_J$. This shows that $\mathcal{F}_J$ is a filter on J extending $\mathcal{F}$ and all elements of $\mathfrak{F}$. By Zorn's Lemma, $\mathcal{C}$ has a maximal element, which is an ultrafilter extending $\mathcal{F}$. □

Remark. This extension is hardly ever unique. Choosing one is choosing a precise notion of a 'very large' subset of I.

3.2 Cartesian product, reduced product, ultraproduct

Definition 3.10 (Cartesian product of structures) Let $(M_i)_{i\in I}$ be a family of L-structures. The *product of* $(M_i)_{i\in I}$ is the L-structure (M, L^M) such that

1. $M = \prod_i M_i$,
2. $c^M = (c^{M_i})_{i\in I}$ for every constant symbol c,
3. $f^M\left(a^1, \ldots, a^n\right) = (f^{M_i}(a_i^1, \ldots, a_i^n))_{i\in I}$ for every n-ary function symbol f and $a^1, \ldots, a^n$ in $\prod_i M_i$,

4. $(a^1,\dots,a^n)\in r^M \iff \Big((a^1_i,\dots,a^n_i)\in r^{M_i}$ for all i in $I\Big)$, for all n-ary relation symbol r in R and for all $a^1,\dots,a^n$ in $\prod_i M_i$.

Remarks. 1. If there exists a constant symbol c in L, then $\prod_i M_i$ is non-empty: being given the family $(M_i)_{i\in I}$, we are also given the family $(c^{M_i})_{i\in I}$. In general though, and in the case where the index set I is infinite, one may need the Axiom of Choice to ensure that $\prod_i M_i$ is non-empty.

2. For the equality symbol, one has $(a_i)_{i\in I} =^M (b_i)_{i\in I}$ if and only if $a_i =^{M_i} b_i$ for all $i\in I$, so $=^M$ is the usual equality in $\prod_i M_i$.
3. For every coordinate i_0, the i_0th projection $\prod_i M_i \longrightarrow M_{i_0}$ is a morphism.

Example 3.11 Let us consider $\mathbf{R}$ as an L_{ogp}-structure. The L_{ogp}-structure on $\mathbf{R}^n$ is obtained by interpreting e_0 as $(0,\dots,0)$, $+$ as coordinatewise addition, $-$ as coordinatewise inverse, and $\leqslant$ as $(a_1,\dots,a_n)\leqslant^{\mathbf{R}^n}(b_1,\dots,b_n)$ if and only if $a_i\leqslant^{\mathbf{R}} b_i$ for all i. Similarly for the infinite product $\mathbf{R}^{\mathbf{N}}$.

Lemma 3.12 (equivalence relation induced by a filter) *Let $(M_i)_{i\in I}$ be a family of L-structures, $\mathcal{F}$ a filter on I and let $\sim_{\mathcal{F}}$ be the relation on $\prod_i M_i$ defined by*

$$(a_i)_{i\in I}\sim_{\mathcal{F}}(b_i)_{i\in I}\iff \Big\{i\in I : a_i=b_i\Big\}\in\mathcal{F}.$$

Then $\sim_{\mathcal{F}}$ is an equivalence relation that is compatible with any n-ary function f^M and n-ary relation r^M, that is,

$$\text{the equivalences}\quad a^1\sim_{\mathcal{F}} b^1,\dots,a^n\sim_{\mathcal{F}} b^n \quad\text{imply}\quad f^M(a^1,\dots,a^n)\sim_{\mathcal{F}} f^M(b^1,\dots,b^n),\ \text{and}$$

$$a^1\sim_{\mathcal{F}} b^1,\dots,a^n\sim_{\mathcal{F}} b^n\quad\text{imply}\quad \{i\in I:(a^1_i,\dots a^n_i)\in r^{M_i}\}\in\mathcal{F}\iff\{i\in I:(b^1_i,\dots b^n_i)\in r^{M_i}\}\in\mathcal{F}.$$

We write $\prod_{\mathcal{F}} M_i$ for the quotient space modulo $\sim_{\mathcal{F}}$, and $a_{\mathcal{F}}$ for the equivalence class of any $a\in\prod_i M_i$.

Proof. As I belongs to $\mathcal{F}$, the relation $\sim_{\mathcal{F}}$ is reflexive. Symmetry follows from symmetry of equality. If $a\sim_{\mathcal{F}} b$ and $b\sim_{\mathcal{F}} c$, one has

$$\Big\{i\in I:a_i=b_i\Big\}\cap\Big\{i\in I:b_i=c_i\Big\}\subset\Big\{i\in I:a_i=c_i\Big\},\ \text{hence}\ \Big\{i\in I:a_i=c_i\Big\}\in\mathcal{F},$$

so $a\sim_{\mathcal{F}} c$, and $\sim_{\mathcal{F}}$ is transitive.
If $a^1\sim_{\mathcal{F}} b^1,\dots,a^n\sim_{\mathcal{F}} b^n$, one has

$$\Big\{i\in I:a^1_i=b^1_i\Big\}\cap\dots\cap\Big\{i\in I:a^n_i=b^n_i\Big\}\subset\Big\{i\in I:f^{M_i}(a^1_i,\dots,a^n_i)=f^{M_i}(b^1_i\ \dots,b^n_i)\Big\}\in\mathcal{F},$$

so $f^M(a^1,\dots,a^n)\sim_{\mathcal{F}} f^M(b^1,\dots,b^n)$. Similarly, one has

$$\Big\{i\in I:a^1_i=b^1_i\Big\}\cap\dots\cap\Big\{i\in I:a^n_i=b^n_i\Big\}\cap\Big\{i\in I:(a^1_i,\dots,a^n_i)\in r^{M_i}\Big\}\subset\Big\{i\in I:(b^1_i,\dots,b^n_i)\in r^{M_i}\Big\},$$

which proves the last statement. □

Remark. If $\mathcal{F}$ is the Fréchet filter on an infinite set I, the relation $a\sim_{\mathcal{F}} b$ holds if and only if $a_i=b_i$ for all but finitely many i in I. In the case where $\mathcal{U}$ is an ultrafilter on I, the relation $a\sim_{\mathcal{U}} b$ holds if and only if $a_i=b_i$ holds for almost every i in I (relatively to the measure $\mu_{\mathcal{U}}$).

Definition 3.13 (reduced product of structures) Let $(M_i)_{i\in I}$ be a family of L-structures, M their product, and $\mathcal{F}$ a filter on I. The *reduced product of* $(M_i)_{i\in I}$ is the L-structure $(M_{\mathcal{F}}, L^{M_{\mathcal{F}}})$ such that

1. $M_{\mathcal{F}}=\prod_{\mathcal{F}} M_i$,
2. $c^{M_{\mathcal{F}}}=\Big(c^M\Big)_{\mathcal{F}}=((c^{M_i})_{i\in I})_{\mathcal{F}}$ for every constant symbol c,
3. $f^{M_{\mathcal{F}}}\Big(a^1_{\mathcal{F}},\dots,a^n_{\mathcal{F}}\Big)=\Big(f^M(a^1,\dots,a^n)\Big)_{\mathcal{F}}$ for every n-ary function symbol and $a^1,\dots,a^n$ in $\prod_i M_i$,
4. $(a^1_{\mathcal{F}},\dots,a^n_{\mathcal{F}})\in r^{M_{\mathcal{F}}}\iff\Big\{i\in I:(a^1_i,\dots,a^n_i)\in r^{M_i}\Big\}\in\mathcal{F}$ for every n-ary relation symbol r.

In the particular case where all the structures M_i are equal to N, one writes $N^{\mathcal{F}}$ instead of $\prod_{\mathcal{F}} N$, and call $N^{\mathcal{F}}$ a *reduced power of* N.

Remarks. 1. By Lemma 3.12, the definitions of $f^{M_\mathcal{F}}(a^1_\mathcal{F},\ldots,a^n_\mathcal{F})$ and $r^{M_\mathcal{F}}(a^1_\mathcal{F},\ldots,a^n_\mathcal{F})$ do not depend on the choice of a representative for every $a^j_\mathcal{F}$, so $M_\mathcal{F}$ is well-defined.

2. The projection $\prod_i M_i \longrightarrow \prod_\mathcal{F} M_i$ is an L-morphism.
3. If $\mathcal{F}$ is the trivial filter $\{I\}$, then $\sim_\mathcal{F}$ is equality on $\prod_i M_i$ so $\prod_\mathcal{F} M_i$ equals $\prod_i M_i$.
4. If $\mathcal{F}$ is a principal filter generated by $\{J\}$ for some $J \subset I$, then $\prod_\mathcal{F} M_i$ is isomorphic to $\prod_{j\in J} M_j$. In particular, for the filter $\mathcal{F}_{i_0}$ generated by the point $\{i_0\}$, $\prod_{\mathcal{F}_{i_0}} M_i$ is isomorphic to M_{i_0} (exercise).

Exercise 3.14 (An example: reduced product of rings) Consider $\mathbf{R}$ with its L_{ring}-structure again. The L_{ring}-structure $\mathbf{R}^\mathbf{N}$ is the natural ring structure on the Cartesian power of $\mathbf{R}$. Let $\mathcal{F}$ be a filter on $\mathbf{N}$. Show that there is an ideal I of $\mathbf{R}^\mathbf{N}$ such that the reduced power $\mathbf{R}^\mathcal{F}$ is precisely the quotient ring $\mathbf{R}^\mathbf{N}/I$. Show that if $\mathcal{F}$ is an ultrafilter, then the ideal I is maximal. Conversely, show that for every ideal I of $\mathbf{R}^\mathbf{N}$, there is a filter $\mathcal{F}$ on $\mathbf{N}$ such that $\mathbf{R}^\mathbf{N}/I$ equals $\mathbf{R}^\mathcal{F}$.

Definition 3.15 (ultraproduct of structures) Let $(M_i)_{i\in I}$ be a family of L-structures, M their product, and $\mathcal{U}$ an ultrafilter on I. The structure $\prod_\mathcal{U} M_i$ is called the *ultraproduct of* $(M_i)_{i\in I}$. In the particular case where every M_i equals N, the structure $N^\mathcal{U}$ is called an *ultrapower of* N.

3.3 Satisfaction in an ultraproduct

Łos' Theorem 3.16 (satifaction in an ultraproduct) *Let $(M_i)_{i\in I}$ be a family of L-structures, M their product, $\mathcal{U}$ an ultrafilter on I and $M_\mathcal{U}$ the ultraproduct $\prod_\mathcal{U} M_i$.*

1. *Let t be an L-term with no variable occurences. Then*
$$t^{M_\mathcal{U}} = \left(t^M\right)_\mathcal{U} = \left((t^{M_i})_{i\in I}\right)_\mathcal{U}.$$
2. *Let σ be a sentence. Then*
$$\prod\nolimits_\mathcal{U} M_i \models \sigma \textit{ if and only if } \left\{i \in I : M_i \models \sigma\right\} \in \mathcal{U}.$$

Remarks. 1. Let $t(x_1,\ldots,x_n)$ be a term, $\varphi(x_1,\ldots,x_n)$ a formula and $a^1,\ldots,a^n$ elements of $\prod_i M_i$. Adding to the language n new constant symbols $c_1,\ldots,c_n$ that we interpret as $a^1_i,\ldots,a^n_i$ in M_i and hence as $a^1_\mathcal{U},\ldots,a^n_\mathcal{U}$ in $\prod_\mathcal{U} M_i$, one has
$$t^{M_\mathcal{U}}(a^1_\mathcal{U},\ldots,a^n_\mathcal{U}) = \left(t^{M_i}(a^1_i,\ldots,a^n_i)\right)_\mathcal{U}, \text{ and}$$
$$\prod\nolimits_\mathcal{U} M_i \models \varphi(a^1_\mathcal{U},\ldots,a^n_\mathcal{U}) \text{ if and only if } \left\{i \in I : M_i \models \varphi(a^1_i,\ldots,a^n_i)\right\} \in \mathcal{U}.$$

2. It follows that $\prod_\mathcal{U} M_i$ satisfies $\varphi(a^1_\mathcal{U},\ldots,a^n_\mathcal{U})$ if and only if M_i satisfies $\varphi(a^1_i,\ldots,a^n_i)$ for almost every $i \in I$ (with respect to the measure $\mu_\mathcal{U}$).

Proof. We show that $t^{M_\mathcal{U}}$ equals $(t^M)_\mathcal{U}$ by induction on the complexity of t. If t is a constant symbol, this follows from the definition of $c^{N_\mathcal{U}}$. If t is the term $ft_1\cdots t_n$, where the terms $t_1,\ldots,t_n$ have lower complexity, then
$$t^{M_\mathcal{U}} = f^{M_\mathcal{U}}(t_1^{M_\mathcal{U}},\ldots,t_n^{M_\mathcal{U}}) = f^{M_\mathcal{U}}((t_1^M)_\mathcal{U},\ldots,(t_n^M)_\mathcal{U}) = (f^M(t_1^M,\ldots,t_n^M))_\mathcal{U} = (t^M)_\mathcal{U}.$$
Let us show the second point of Łos' Theorem by induction on the complexity of σ. If σ is the atomic sentence $r(t_1,\ldots,t_n)$, then 2. follows from the definition of $r^{M_\mathcal{U}}$. If σ is the sentence $\sigma_1 \wedge \sigma_2$, then
$$\begin{aligned}\prod\nolimits_\mathcal{U} M_i \models \sigma &\iff \left\{i \in I : M_i \models \sigma_1\right\} \in \mathcal{U} \text{ and } \left\{i \in I : M_i \models \sigma_1\right\} \in \mathcal{U}\\ &\iff \left\{i \in I : M_i \models \sigma_1\right\} \cap \left\{i \in I : M_i \models \sigma_2\right\} \in \mathcal{U}\\ &\iff \left\{i \in I : M_i \models \sigma_1 \text{ and } M_i \models \sigma_2\right\} \in \mathcal{U}.\\ &\iff \left\{i \in I : M_i \models \sigma\right\} \in \mathcal{U}.\end{aligned}$$

If σ is the sentence $\exists x\varphi$ for some formula $\varphi(x)$, then $\prod_{\mathcal{U}} M_i$ satisfies σ if and only if there exists some $a_{\mathcal{U}}$ in $\prod_{\mathcal{U}} M_i$ such that $\prod_{\mathcal{U}} M_i$ satisfies $\varphi(a)$. Let c be a new constant symbol, and define its interpretation in M_i to be a_i. It follows that $\prod_{\mathcal{U}} M_i$ satisfies $\varphi(a)$ in L if and only if it satisfies the sentence $\varphi((c))$ in $L \cup \{c\}$ (note that $\varphi((c))$ and $\varphi(x)$ have the same complexity). By the induction hypothesis, one has

$$\prod_{\mathcal{U}} M_i \models \varphi((c)) \iff \{i \in I : M_i \models \varphi((c))\} \in \mathcal{U} \iff \{i \in I : M_i \models \varphi(a_i)\} \in \mathcal{U}.$$

As $\{i \in I : M_i \models \varphi(a_i)\} \subset \{i \in I : M_i \models \exists x\varphi\}$, the latter set belongs to $\mathcal{U}$. To show the reverse implication, let $J = \{i \in I : M_i \models \exists x\varphi\}$ be in $\mathcal{U}$. Using the Axiom of Choice, one may choose an element $(a_i)_{i\in I}$ in $\prod_i M_i$ such that for every i in J, one has $M_i \models \varphi(a_i)$, and a_i is arbitrary in M_i for $i \in I \setminus J$. It follows that the set $\{i \in I : M_i \models \varphi(a_i)\}$ contains J and hence is in $\mathcal{U}$, and we finish as previously, adding one new constant symbol c to the language and interpreting c^{M_i} by a_i, hence c^{M_U} by $(a_i)_{\mathcal{U}}$.

Note that all of the above holds if $\mathcal{U}$ is merely a filter on I. If σ is the sentence $\neg\tau$, then

$$\prod_{\mathcal{U}} M_i \models \sigma \iff \prod_{\mathcal{U}} M_i \not\models \tau \iff \Big\{i \in I : M_i \models \tau\Big\} \notin \mathcal{U}.$$

Since $\mathcal{U}$ is an ultrafilter, one has

$$\Big\{i \in I : M_i \models \tau\Big\} \notin \mathcal{U} \iff \Big\{i \in I : M_i \not\models \tau\Big\} \in \mathcal{U},$$

so that one has $\prod_{\mathcal{U}} M_i \models \sigma$ if and only if $\Big\{i \in I : M_i \models \neg\tau\Big\} \in \mathcal{U}$. □

Example 3.17 (model of non standard analysis) Let $\mathbf{R}$ be considered as an L_{ring}-structure, and let $\mathcal{F}$ be a filter on $\mathbf{N}$ and $\mathcal{U}$ an ultrafilter on $\mathbf{N}$. By Exercise 3.14, the reduced power $\mathbf{R}^{\mathcal{F}}$ is a ring and the ultrapower $\mathbf{R}^{\mathcal{U}}$ is a field. The latter statement can be deduced again by Łos' Theorem: $\mathbf{R}^{\mathcal{U}}$ has the same L_{ring}-theory as $\mathbf{R}$: it is a field of characteristic 0, every polynomial with coefficient in $\mathbf{R}^{\mathcal{U}}$ and odd degree has a root in $\mathbf{R}^{\mathcal{U}}$. The map $i : \mathbf{R} \longrightarrow \mathbf{R}^{\mathcal{U}}$ that maps a real number x to the element $(x, x, x, \dots)_{\mathcal{U}}$ is an L_{ring}-embedding, so that $\mathbf{R}$ can be seen as a subfield of $\mathbf{R}^{\mathcal{U}}$. One can define a ordering $\leqslant$ on $\mathbf{R}^{\mathcal{U}}$ by setting

$$a \leqslant b \iff b - a \text{ is a square} \iff \mathbf{R}^{\mathcal{U}} \models \varphi(a, b),$$

where φ is the formula $\exists z(y - x = z^2)$. As φ defines a dense linear ordering on $\mathbf{R}$ that is compatible with the field structure of $\mathbf{R}$, properties which are expressible by a L_{ring}-sentence, it follows from by Łos' Theorem that $\leqslant$ also defines a dense linear ordering on $\mathbf{R}^{\mathcal{U}}$ that is compatible with the field structure on $\mathbf{R}^{\mathcal{U}}$ and extends the natural ordering on $\mathbf{R}$ (i.e. such that $x \leqslant y$ implies $i(x) \leqslant i(y)$ for all real numbers x and y). If the ultrafilter $\mathcal{U}$ is principal, then $\mathbf{R}^{\mathcal{U}}$ is isomorphic to $\mathbf{R}$. If $\mathcal{U}$ is non-principal (*i.e.* contains every cofinite subset of $\mathbf{N}$), then $\mathbf{R}^{\mathcal{U}}$ has *infinitesimal numbers i.e.* elements ε that satisfy $0 < \varepsilon < x$ for every real number $x > 0$, for instance

$$\varepsilon = \left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\right)_{\mathcal{U}},$$

as $0 < \dfrac{1}{n} < x$ holds for cofinitely many n in $\mathbf{N}$. It also has *infinite numbers i.e.* elements ω satisfying $\omega > x$ for every real number x, for instance

$$\omega = \frac{1}{\varepsilon} = (1, 2, 3, 4, \dots)_{\mathcal{U}}.$$

Note that as $\mathbf{R}^{\mathcal{U}}$ is a field, every non-zero element has a unique multiplicative inverse. As $\varepsilon \cdot \omega = 1$, we may write $\omega = \varepsilon^{-1}$ without any ambiguity.

Corollary 3.18 (Compactness Theorem) *Let Σ be a theory each of whose finite subset $\Sigma_0 \subset \Sigma$ has a model. Then Σ has a model.*

Proof. In the particular case where the language is countable, let $(\sigma_n)_{n\geqslant 1}$ be an enumeration of Σ and let M_n be a model of $\{\varphi_1, \ldots, \varphi_n\}$ for every natural number n, so that if σ is in Σ, one has $M_n \models \sigma$ for all but finitely many $n \in \mathbf{N}$. Then, for any ultrafilter $\mathcal{U}$ on $\mathbf{N}$ extending the Fréchet filter, the ultraproduct $\prod_{\mathcal{U}} M_n$ is a model of Σ by Łos' Theorem.
General case. Let I be the set of all finite subsets of Σ, and, for every i in I, let M_i be a model of i. For every sentence σ in Σ, let $J(\sigma)$ be the subset of I each of whose elements contain σ, so that $M_i \models \sigma$ as soon as $i \in J(\sigma)$. For any ultrafilter $\mathcal{U}$ extending $\mathcal{F} = \{J(\sigma) : \sigma \in \Sigma\}$ (note that $J(\sigma_1) \cap \cdots \cap J(\sigma_n)$ contains $\{\sigma_1, \ldots, \sigma_n\}$ hence is never empty, so $\mathcal{F}$ generates a filter that can be extended to an ultrafilter), the ultraproduct $\prod_{\mathcal{U}} M_i$ is a model of Σ by Łos' Theorem. $\square$

CHAPTER 4

ENUMERATION AND SIZE OF INFINITE SETS

4.1 Ordinal numbers

Let X be a set. A binary relation $\leqslant$ on X is called an *ordering*, or *partial ordering*, if it is reflexive antisymmetric and transitive. A binary relation $<$ on X is called a *strict ordering* if it is antireflexive (*i.e.* if $x \not< x$ for any x) and transitive. Every ordering $\leqslant$ on X induces a natural strict ordering on X, written $<$, defined by $x < y$ if and only if $x \leqslant y$ and $x \neq y$. Conversely, every strict ordering $<$ on X induces an ordering, written $\leqslant$, defined by $x \leqslant y$ if and only if $x < y$ or $x = y$. An ordering $\leqslant$ on X is *linear* if for every x and y in X, either $x \leqslant y$ or $y \leqslant x$ hold.

Definition 4.1 (well-ordering) A *well-ordering* on X is an ordering such that every non-empty subset $Y \subset X$ has a *least element* (*i.e.* an element $y \in Y$ such that $y \leqslant z$ for all $z \in Y$).

Remark. A well-ordering on X is a linear ordering on X as every pair $\{x, y\}$ has a least element. Conversely, a linear ordering of X is a well-ordering of X if and only if there is no infinite strictly decreasing chain of elements of X.

Definition 4.2 (transitive set) The set X is *transitive* if every element of X is a subset of X.

Remark. Transitivity of X ensures that if $y \in X$ and $z \in y$, then $z \in X$. In particular, if $\in$ defines a (strict) ordering on X, then X is transitive.

Definition 4.3 (ordinal number) The set X is an *ordinal number* if it is transitive and $\in$ is a strict ordering on X that induces a well-ordering on X.

We shall denote ordinal numbers by Greek letters $\alpha, \beta, \gamma, \ldots$.

Lemma 4.4 1. *$\emptyset$ is an ordinal, written* 0.
2. *If β is an ordinal and $\alpha \in \beta$, then α is an ordinal.*
3. *If β is an ordinal, β equals the set $[0, \beta)$ of ordinals α such that $\alpha \in \beta$.*
4. *If α and β are distinct ordinals, then $\alpha \subset \beta$ iff $\alpha \in \beta$.*

Proof. 1. An ordinal number has been defined by universal properties, hence hold of the empty set.
2. If β is an ordinal, $\alpha \in \beta$ and $\gamma \in \alpha$, for any $\delta \in \gamma$, the elements γ and δ belong to β by transitivity of the set β. It follows that $\delta \in \alpha$ by transitivity of $\in$, so the set α is transitive. If a, b and c are elements of α, then a, b, c are elements of β by transitivity of β; if $a \in b$ and $b \in c$, then $a \in c$ by transitivity of $\in$, so $\in$ is transitive on α. As the relation $\in$ on α is the restriction of the relation $\in$ on β, $\in$ is a well-ordering on α.
3. Follows from 2.
4. One direction follows from the transitivity of β. Conversely, if $\alpha \subset \beta$ and $\alpha \neq \beta$, then $\beta \setminus \alpha$ is non-empty hence has a least element γ. We claim that $\alpha = \{x \in \beta : x \in \gamma\}$: the inclusion $\{x \in \beta : x \in \gamma\} \subset \alpha$ follows from the minimality of γ and conversely, if $x \in \alpha$, one has either $\gamma \in x$ (hence $\gamma \in \alpha$, a contradiction) or $\gamma = x$ (same contradiction again) or $x \in \gamma$. On the other hand, one also has $\gamma = \{x \in \beta : x \in \gamma\}$, and it follows that $\alpha = \gamma \in \beta$ (by definition of γ). □

We define a strict ordering $<$ on the class of all ordinals by setting for any two ordinals α and β,

$$\alpha < \beta \text{ if and only } \alpha \in \beta.$$

It follows from Lemma 4.4 that the corresponding partial ordering is

$$\alpha \leqslant \beta \text{ if and only if } \alpha \subset \beta.$$

Theorem 4.5 ($\leqslant$ is a linear ordering on the class of ordinals) *For any two ordinals α and β, one has*

$$\textit{either } \alpha \leqslant \beta, \ \textit{ or } \beta \leqslant \alpha.$$

Proof. $\gamma = \alpha \cap \beta$ is transitive (every element of $\alpha \cap \beta$ is a subset of $\alpha \cap \beta$), and well-ordered by $\in$: it is an ordinal satisfying $\gamma \subset \alpha$ and $\gamma \subset \beta$. If these inclusions are both strict, then $\gamma \in \alpha \cap \beta$ by Lemma 4.4, so $\gamma \in \gamma$, a contradiction. One thus has either $\alpha \cap \beta = \alpha$ or $\alpha \cap \beta = \beta$. □

Remarks.

1. (the class of ordinals is well-ordered) For every non-empty set A of ordinals, let $\bigcap A$ be the intersection of all $\alpha \in A$. Being an intersection of ordinals, $\bigcap A$ is an ordinal, and a lower bound of A. If $\bigcap A$ were not an element of A, one would have $\bigcap A < a$ by Theorem 4.5 (hence $\bigcap A \in a$) for every a in A, hence $\bigcap A \in \bigcap A$, a contradiction: $\bigcap A$ is the least element of A, and
$$\inf A = \min A = \bigcap A.$$
2. (every non-empty set of ordinals has an upper bound) For every set A of ordinals, $\bigcup A$ is an ordinal. If $x \in \bigcup A$, then $x \in \alpha$ for some $\alpha \in A$. If $y \in x$, then $y \in \alpha$ since α is an ordinal, so $\bigcup A$ is transitive. Any three elements of $\bigcup A$ must belong to one $\alpha \in A$ (since every three ordinals are linearly ordered by Theorem 4.5), so $\in$ is a transitive relation in $\bigcup A$. If B is a non-empty set of elements of $\bigcup A$, then $\bigcap B$ is the least element of B by the above remark. So $\bigcup A$ is an upper bound of A (in the class of ordinals), and the least such, as $\beta < \bigcup A$ implies $\beta \in \alpha$ for some $\alpha \in A$. This shows
$$\sup A = \bigcup A.$$

Theorem 4.6 *Every well-ordered set is isomorphic (as an ordered set) to a unique ordinal number.*

Definition 4.7 (successor ordinal, limit ordinal) For every ordinal α, the set $\alpha \cup \{\alpha\}$ is also an ordinal, written $\alpha + 1$, called the *successor ordinal of* α. If λ is not the successor of any ordinal (and not 0), λ is called a *limit ordinal* and λ is the set $\bigcup_{\alpha<\lambda} \alpha$.

Note that if $\beta = \{\alpha : \alpha < \beta\}$ holds for every ordinal, but $\lambda = \bigcup\{\alpha : \alpha < \lambda\}$ holds if and only if λ is a limit ordinal.

Theorem 4.8 (transfinite induction) *Assume that $\mathcal{C}$ is the class of ordinals such that*

1. 0 *is an element of* $\mathcal{C}$*.*
2. *if α is an element of $\mathcal{C}$ then $\alpha + 1$ also.*
3. *if λ is a limit ordinal such that α is an element of $\mathcal{C}$ for every $\alpha < \lambda$, then λ is an element of $\mathcal{C}$.*

Then $\mathcal{C}$ is the class of all ordinals.

Proof. If there is an ordinal γ that is not an element of $\mathcal{C}$, as γ is well-ordered, there is a least $\beta \leqslant \gamma$ that is not an element of $\mathcal{C}$. Either γ is a successor ordinal $\gamma = \alpha + 1$, but then α hence $\alpha + 1$ belong to $\mathcal{C}$ by 2, a contradiction; of γ is a limit ordinal. As $\gamma = [0, \lambda)$, one must have $\gamma \in \mathcal{C}$ by 3, a contradiction; or γ is 0, contradicting 1. □

Hence, practically, the class of ordinals is constructed by transfinite induction, starting from $0 = \emptyset$, and applying successor steps

$$1 = 0 + 1, \quad 2 = 1 + 1, \quad 3 = 2 + 1 \quad \textit{etc.}$$

then a limit step

$$\omega = \{1, 2, 3, \dots\},$$

then successor steps again

$$\omega+1, \quad \omega+2=(\omega+1)+1, \quad \omega+3=(\omega+2)+1 \quad etc.$$

then a limit step

$$\omega+\omega=\omega\cdot 2=\{\omega+1,\omega+2,\omega+3,\dots\}.$$

The first ordinal numbers are $0,1,2,\dots,n,\dots,\omega,\omega+1,\omega+2,\dots,\omega+n,\dots,\omega 2,\omega 2+1,\dots,\omega 3,\dots,$ $\omega n,\dots,\omega^2,\dots,\omega^n,\dots,\omega^\omega,\omega^\omega+1,\dots,\omega^\omega+\omega,\dots,\omega^\omega+\omega^n,\dots,\omega^\omega 2,\dots,\omega^\omega n,\dots,\omega^{\omega+1},\dots,$ where the operations $\alpha+\beta$, $\alpha\beta$ and α^β are defined for all α by transfinite induction on β:

$$\begin{aligned}
&\alpha+0=0, && \alpha\cdot 0=0, && \alpha^0=1,\\
&\alpha+(\beta+1)=(\alpha+\beta)+1, && \alpha(\beta+1)=\alpha\beta+\alpha, && \alpha^{\beta+1}=\alpha^\beta\cdot\alpha,\\
&\alpha+\lambda=\sup\{\alpha+\beta:\beta<\lambda\}, && \alpha\lambda=\sup\{\alpha\beta:\beta<\lambda\}, && \alpha^\lambda=\sup\{\alpha^\beta:\beta<\lambda\}.
\end{aligned}$$

Theorem 4.9 (Zermelo) *Every set X can be well-ordered: there exists an ordinal β such that X is the set $\{x_\alpha:\alpha<\beta\}$.*

4.2 Cardinal numbers

Definition 4.10 (having the same cardinal) Two sets X and Y have the *same cardinal*, which we write $|X|=|Y|$, if there exists a one-to-one map from X onto Y.

This defines an equivalence relation on the class of sets.

Definition 4.11 (having smaller cardinal) The set X has *smaller cardinal* than Y, which we write $|X|\leqslant|Y|$ if there exists an injective map from X to Y.

This defines a reflexive transitive relation on the class of sets.

Theorem 4.12 (the relation $\leqslant$ is a partial order modulo the cardinal relation, Cantor-Bernstein) *If $|X|\leqslant|Y|$ and $|Y|\leqslant|X|$, then $|X|=|Y|$.*

Proof. We begin with the particular case when $Y\subset X$. Let $f:X\to Y$ an injective map. We define inductively $X_0=X$, $X_{n+1}=f(X_n)$, and $Y_0=Y$, $Y_{n+1}=f(Y_n)$. Note that $Y_0\subset X_0$ hence $Y_n\subset X_n$ for all n. Note also $f(X_n\setminus Y_n)=X_{n+1}\setminus Y_{n+1}$. Let $g:X\to Y$ be the function defined by

$$g(x)=f(x)\text{ if }x\in\bigcup_{n\geqslant 0}X_n\setminus Y_n,\ \text{ or }g(x)=x\text{ otherwise.}$$

If $g(x)=g(y)$, then either only one of x and y (say x) belong to some $X_n\setminus Y_n$, hence $f(x)=y$ (then $f(x)\in X_{n+1}\setminus Y_{n+1}$, a contradiction), or both x and y belong to $\bigcup X_n\setminus Y_n$, hence $f(x)=f(y)$, or none of them belong to $\bigcup X_n\setminus Y_n$. In every case, $x=y$, so g is injective. On the other hand, if y is in Y. Either $y\in f(\bigcup X_n\setminus Y_n)$, so $y=f(x)$ for some x in $\bigcup X_n\setminus Y_n$, hence $y=g(x)$. Or $y\notin f(\bigcup X_n\setminus Y_n)$; but $y\in Y$, hence does not belong to $X\setminus Y$, and does not belong to any $X_n\setminus Y_n$. It follows that $y=g(y)$, so g is surjective.
General case. If $f:X\to Y$ and $g:Y\to X$ are injective, then $g(Y)\subset X$ and $g\circ f:X\to g(Y)$ is an injective map, so we may apply the first case to find a bijection between $g(Y)$ and X. As X and $g(Y)$ are in bijection via g, this concludes the proof. □

We write $|X|<|Y|$ if and only if $|X|\leqslant|Y|$ and $|Y|\neq|X|$.

Theorem 4.13 (the relation $\leqslant$ is non-trivial, Cantor) *For any set X, one has $|X|<|\mathcal{P}(X)|$.*

Proof. Let f be any function from X to $\mathcal{P}(X)$. Let us show that f is not onto. Let Y be the subset $Y=\{x\in X:x\notin f(x)\}$ of X. If Y is in the range of f, there is an element $a\in X$ with $Y=f(a)$. If $a\in Y$, then $a\notin f(a)$, a contradiction. If $a\notin Y$, then $a\in f(a)$, a contradiction also. It follows that f is not onto, so that $|X|\neq|\mathcal{P}(X)|$. On the other hand, the map $x\mapsto\{x\}$ is an injective map from X to $|\mathcal{P}(X)|$ so $|X|\leqslant|\mathcal{P}(X)|$. □

Definition 4.14 (cardinal number) An ordinal number α is called a *cardinal* if it is the least ordinal β such that $|\alpha| = |\beta|$, *i.e.* if α satisfies $|\beta| \neq |\alpha|$ for all $\beta \in \alpha$.

We shall use $\kappa, \lambda, \mu, \dots$ to denote cardinal numbers. Finite ordinals are cardinals, and ω is the least infinite cardinal.

Definition 4.15 (cardinal of a set) Every set X is in bijection with a unique cardinal number called its *cardinal*, written $|X|$.

Proof. By Zermelo's Theorem, there is an ordinal α that is in bijection with X. Among the ordinals $\beta \in \alpha + 1$ that satisfy this property, there is a least one for $\in$, say λ. We claim that λ is an ordinal number, for any $\beta \in \lambda$ in bijection with λ would be in bijection with X, contradicting the minimality of λ. If λ and μ are two cardinal numbers in bijection with X, they are in bijection, so $\lambda \notin \mu$ and $\mu \notin \lambda$ by definition of a cardinal, hence $\mu = \lambda$. □

Remarks. 1. The notation can $|X| = |Y|$ means either that X and Y are in bijection, or that X and Y have the same cardinal number. These statement are equivalent.4.10.
2. If λ is a cardinal, then $|\lambda| = \lambda$. If λ and μ are cardinal numbers, the notation $|\lambda| < |\mu|$ means either that $\lambda \in \mu$, or that there is an injection from λ to μ, but no bijection. These statement are equivalent.

Definition 4.16 (operation on cardinals) The arithmetic operations on cardinals are defined by:

$$\kappa + \lambda = |A \cup B| \quad \text{where } |A| = \kappa,\ |B| = \lambda,\ A, B \text{ disjoint},$$
$$\kappa \cdot \lambda = |A \times B| \quad \text{where } |A| = \kappa,\ |B| = \lambda,$$
$$\kappa^{\lambda} = |A^B| \quad \text{where } |A| = \kappa,\ |B| = \lambda,$$

and are independent of the choice of the sets A and B.

Exercises 4.17 1. If $|X| = \kappa$, then $|\mathcal{P}(X)| = 2^{\kappa}$.
2. $+$ and $\cdot$ are associative, commutative and $\cdot$ is distributive over $+$.
3. $(\kappa\lambda)^{\mu} = \kappa^{\mu}\lambda)^{\mu}$.
4. $\kappa^{\lambda+\mu} = \kappa^{\lambda}\kappa^{\mu}$.
5. $(\kappa^{\lambda})^{\mu} = \kappa^{\lambda\mu}$.
6. If $\kappa \leqslant \lambda$, then $\kappa^{\mu} \leqslant \lambda^{\mu}$.
7. If $0 < \lambda \leqslant \mu$, then $\kappa^{\mu} \leqslant \kappa^{\mu}$.
8. $\kappa^0 = 1$; $1^{\kappa} = 1$; $0^{\kappa} = 0$ if $\kappa > 0$.

Lemma 4.18 *If A is a set of cardinals, then* $\sup A$ *is also a cardinal.*

Proof. $\sup A$ is an ordinal according to the previous section. If $\beta < \sup A$ is an ordinal, there exists a cardinal $\lambda \in A$ such that $\beta \in \lambda$, so $|\beta| < |\lambda|$ by definition of a cardinal. It follows that $|\beta| < \sup A$. □

Definition 4.19 (aleph numbers) By Cantor's Theorem, for any cardinal λ, there exists a cardinal $\kappa > \lambda$. The set of cardinal numbers μ such that $\lambda < \mu \leqslant \kappa$ is thus non-empty, and has a least element (that does not depend on λ) that we write λ^+ and call the *successor cardinal* of λ. Using Lemma 4.18, we define an increasing enumeration $\aleph$ of cardinals by putting

$$\aleph_0 = \omega, \quad \aleph_{\alpha+1} = \aleph_{\alpha}^{+}, \quad \text{and} \quad \aleph_{\lambda} = \sup\left\{\aleph_{\beta} : \beta < \lambda\right\} \text{ for a limit ordinal } \lambda.$$

$\aleph_{\lambda}$ is called a *limit cardinal.*

Lemma 4.20 *Every infinite cardinal number is of the form $\aleph_{\alpha}$ for some unique ordinal α.*

Proof. Let λ be an infinite cardinal and consider the map $i : \lambda \longrightarrow \aleph_{\lambda}$ that maps α to $\aleph_{\alpha}$. Note that $\alpha < \beta$ implies $\aleph_{\alpha} < \aleph_{\beta}$ for any ordinals α and β (by transfinite induction on β). In particular, the map i is well-defined, and injective, so that $\aleph_{\lambda+1} > \lambda$. It follows that the class $\{\alpha \text{ ordinal} : \lambda < \aleph_{\alpha}\}$ is non-empty and has a least element β that satisfies $\aleph_{\beta} > \lambda$. As λ is infinite, β cannot be 0. Nor can β be a limit ordinal, one has $\beta = \alpha + 1$ so that $\aleph_{\alpha} \leqslant \lambda < \aleph_{\alpha+1}$, hence $\lambda = \aleph_{\alpha}$. □

CHAPTER 5

MORE MODEL THEORY

5.1 Elementary substructures, elementary extensions

Two L-structures N and M are called *elementarily equivalent*, which we write $N \equiv M$, if they have the same L-theory. This defines an equivalence relation on the class of all L-structures. Recall that N is an L-*substructure* of M, written $N \subset_L M$, if N is a subset of M containing all the interpretations of constants and closed under the interpretations of functions, and L^N is the restriction of L^M to N. We saw in the exercise sheets that if $\varphi(\bar{x})$ is a quantifier-free L-formula, $N \subset_L M$ are two L-structures, and $\bar{a}$ is a tuple in N, then one has

$$N \models \varphi(\bar{a}) \iff M \models \varphi(\bar{a}).$$

Definition 5.1 (elementary substructure, elementary extension) Let N and M be L-structures. N is an *elementary substructure* of M, written $N \prec M$, if N is a substructure of M and for every L-formula $\varphi(\bar{x})$ and tuple $\bar{a}$ in N, one has

$$N \models \varphi(\bar{a}) \iff M \models \varphi(\bar{a}).$$

One also says that M is an *elementary extension* of N.

Remarks.
1. This defines a reflexive, transitive, antisymmetric relation on the class of L-structures.
2. If M is an L-structure and A a subset of M, we write $L \cup A$ for the language obtained by adding a constant symbol for every element of A and define the $L \cup A$-structure M_A to be $(M, L^M \cup A)$, obtained from M by interpreting any m in A by m. If N is an L-substructure of M, then N is an elementary substructure of M iff the $L \cup N$-structures N_N and M_N are elementarily equivalent.
3. If $N \prec K$, $M \prec K$ and $N \subset_L M$, then $N \prec M$.

Recall that an L-embedding $\sigma : N \to M$ between two L-structures M and N is a map that preserves the language L.

Lemma 5.2 (characterisation of embeddings) *Let M, N be two structures and $\sigma : N \to M$ a map.*

1. *σ is an embedding if and only if, for every quantifier-free formula $\varphi(\bar{x})$ and tuple $\bar{a}$ in N,*

$$N \models \varphi(\bar{a}) \iff M \models \varphi(\sigma(\bar{a})).$$

2. *If σ is an isomorphism, then, for **every** formula $\varphi(\bar{x})$ and every tuple $\bar{a}$ in N, one has*

$$N \models \varphi(\bar{a}) \iff M \models \varphi(\sigma(\bar{a})).$$

Proof. 1. Assume that the equivalence holds and let c be a constant symbol, f an n-ary function symbol and r an n-ary relation symbol. Taking the atomic formula $x = c$ and $a = c^N$, one has $\sigma(c^N) = c^M$. Taking the atomic formula $f(x_1, \ldots, x_n) = x_{n+1}$ and $\bar{b} = (b_1, \ldots, b_n)$ and $\bar{a} = (\bar{b}, f^N(\bar{b}))$ in N, one has $f^M(\sigma(\bar{b})) = \sigma(f^N(\bar{b}))$. Taking the atomic formula $r(x_1, \ldots, x_n)$ and any $\bar{a} = (a_1, \ldots, a_n)$ in N, one has $\bar{a} \in r^N$ if and only if $\sigma(\bar{a}) \in r^M$. It follows that σ is an embedding. Conversely, if σ is an embedding, then the equivalence holds for atomic formulas: this can be shown first for a formulas of the form $x = t(\bar{y})$ inductively on the complexity of the term t, and then for a quantifier-free formula φ by induction on $c(\varphi)$.

2. If σ is an isomorphism, we show the equivalence by induction on the complexity of formulas. It holds for atomic formulas by 1. It holds for a formula $\neg\psi$ or $\varphi \wedge \psi$ by the induction hypothesis. If $\varphi(\bar{x})$ is the formula $\exists y \psi(y, \bar{x})$ (we assume without loss of generality that y does not occur in $\bar{x}$), then one has

$$N \models \varphi(\bar{a}) \iff \text{there exists } b \in N \text{ with } N \models \psi(b, \bar{a}) \iff \text{there exists } b \in N \text{ with } M \models \psi(\sigma(b), \sigma(\bar{a})),$$

which is equivalent to $M \models \varphi(\sigma(\bar{a}))$. □

Definition 5.3 (elementary embedding) Let M, N be two structures and $\sigma : N \to M$ a map. σ is an *elementary embedding* if for every formula $\varphi(\bar{x})$ and every tuple $\bar{a}$ in N, one has

$$N \models \varphi(\bar{a}) \iff M \models \varphi(\sigma(\bar{a})). \quad (2)$$

Remark. An elementary embedding is an embedding.

Examples 5.4 1. An ismorphism is an elementary embedding.

2. If M is an L-structure and $M^{\mathcal{U}}$ an ultrapower of M, the map $M \to M^{\mathcal{U}}$ that sends an element x to the class $(x, \dots, x, \dots)_{\mathcal{U}}$ is an elementary embedding by Łos' Theorem.

Lemma 5.5 *Let M, N be two structures and $\sigma : N \to M$ an embedding. σ is elementary if and only if $\sigma(N)$ is an elementary substructure of M.*

Proof. As σ is an embedding, then N and $\sigma(N)$ are isomorphic. By Lemma 5.2, for every formula $\varphi(\bar{x})$ and tuple $\bar{a}$ in N, one has $N \models \varphi(\bar{a}) \iff \sigma(N) \models \varphi(\sigma(\bar{a}))$.
On the other hand, $\sigma(N) \prec M$ is equivalent to $\sigma(N) \models \varphi(\sigma(\bar{a})) \iff M \models \varphi(\sigma(\bar{a}))$. □

Lemma 5.6 (Tarski-Vaught test) *Let M be a structure and $N \subset_L M$ a substructure. If, for every L-formula $\varphi(x, \bar{y})$ and tuple $\bar{a}$ in N, whenever one has $M \models \exists x \varphi(x, \bar{a})$, there exists b in N such that $M \models \varphi(b, \bar{a})$, then N is an elementary substructure of M.*

Proof. We show that the equivalence $N \models \varphi(\bar{a}) \iff M \models \varphi(\bar{a})$ holds for every tuple $\bar{a}$ in N by induction on the complexity of formulas. As N is a substructure of M, the equivalence holds for atomic formulas. If it holds for φ and ψ, it also holds for $\neg\varphi$ and $\varphi \wedge \psi$. If φ is of the form $\exists x \psi(x, \bar{y})$, then

$$\begin{aligned} N \models \varphi(\bar{a}) &\iff \text{there exists } b \text{ in } N \text{ with } N \models \psi(b, \bar{a}) \\ &\iff \text{there exists } b \text{ in } N \text{ with } M \models \psi(b, \bar{a}) \\ &\implies \text{there exists } b \text{ in } M \text{ with } M \models \psi(b, \bar{a}) \\ &\iff M \models \varphi(\bar{a}), \end{aligned}$$

and the missing implication is precisely the hypothesis. □

Theorem 5.7 (upward Löwenheim-Skolem's Theorem) *Let M be an infinite L-structure and κ a cardinal number. There is an elementary extension K of M such that $|K| \geqslant \kappa$.*

Proof. Let $\Sigma(M)$ be the $L \cup M$-theory of M (sometimes called the elementary diagram of M) and D a set of new constant symbols of cardinality κ. Consider the $L \cup M \cup D$-theory

$$\Sigma = \Sigma(M) \cup \Gamma \quad \text{where} \quad \Gamma = \{c \neq d : c, d \text{ distinct elements of } D\}.$$

Any finite subset $\Sigma_0 \subset \Sigma$ is the union of a finite subset of $\Sigma(M)$ and a finite subset of Γ involving constants symbols belonging to a finite set $D_0 \subset D$. Choosing finitely many distinct elements $(m_k)_{k \in D_0}$ of M, one defines an $L \cup M \cup D$-structure on M by considering $(M, L^M \cup M \cup (m_k)_{k \in D})$ where m_k is arbitrary chosen in M for $k \in D \setminus D_0$. It follows that $(M, L^M \cup M \cup (m_k)_{k \in D})$ is a model of Σ_0. By the Compactness Theorem, Σ has a model $(K, L^K \cup M^K \cup D^K)$. As K is given with interpretations of the constant symbols, there is a map $i : D \longrightarrow D^K \subset K$ sending c to c^K. As K satisfies Γ, the map i is injective, so $|K| \geqslant \kappa$. As K satisfies $\Sigma(M)$, one has, for every formula $\varphi(\bar{x})$ and every tuple $\bar{a} \in M$,

$$M \models \varphi(\bar{a}) \text{ (in } L) \iff M \models \varphi((\bar{a})) \text{ (in } L \cup M) \iff K \models \varphi((\bar{a})) \text{ (in } L \cup M) \iff K \models \varphi(\bar{a}) \text{ (in } L),$$

so the L-structure (K, L^K) is an elementary extension of M. □