A Crèche Course in Model Theory

Domenico Zambella
Università di Torino
ORCID: 0000-0003-1141-2898

`Branch: nuovo 2024-07-02 14:42:06+02:00`

Chapter 1

Preliminaries and notation

This chapter introduces the syntax and semantic of first order logic. We assume that the reader has at least some familiarity with first order logic.

The definitions of terms and formulas we give in Section 1.3 and 1.5 are more formal than is required in subsequent chapters. Our main objective is to convince the reader that a rigorous definition of language and truth is possible. However, the actual details of such a definition are not relevant for our purposes.

1.1 Tuples

A sequence is a function $a : I \to A$ whose domain is a linear order $I, <_I$. We may use the notation $a = \langle a_i : i \in I \rangle$ for sequences. A tuple is a sequence whose domain is an ordinal, say α, then we write $a = \langle a_i : i < \alpha \rangle$. When α is finite, we may also write $a = a_0, \ldots, a_{\alpha-1}$ The domain of the tuple a, the ordinal α, is denoted by $|a|$ and is called the length of a. If a is surjective, it is said to be an enumeration of A.

If $J \subseteq I$ is a subset of the domain of the sequence $a = \langle a_i : i \in I \rangle$, we write $a_{\restriction J}$ for the restriction of a to J. When J is well ordered by $<_I$, e.g. when a is a tuple or when J is finite, we identify $a_{\restriction J}$ with a tuple. This is the tuple $\langle a_{j_k} : k < \beta \rangle$ where $\langle j_k : k < \beta \rangle$ is the unique increasing enumeration on J.

⚠ Sometimes (i.e. not always) we may overline tuples or sequences as mnemonic. When a tuple $\bar{c}$ is introduced, we write c_i for the i-th element of $\bar{c}$. and $c_{\restriction J}$ for the restriction of $\bar{c}$ to $J \subseteq |\bar{c}|$. Note that the bar is dropped for ease of notation.

The set of tuples of elements of length α is denoted by A^α. The set of tuples of length $< \alpha$ is denoted by $A^{<\alpha}$. For instance, $A^{<\omega}$ is the set of all finite tuples of elements of A. When α is finite we do not distinguish between A^α and the α-th Cartesian power of A. In particular, we do not distinguish between A^1 and A.

If $a, b \in A^\alpha$ and h is a function defined on A, we write $h(a) = b$ for $h(a_i) = b_i$. We often do not distinguish between the pair $\langle a, b \rangle$ and the tuple of pairs $\langle a_i, b_i \rangle$. The context will resolve the ambiguity.

Note that there is a unique tuple of length 0, the empty set $\varnothing$, which in this context is called empty tuple. Recall that by definition $A^0 = \{\varnothing\}$ for every set A. Therefore, even when A is empty, A^0 contains the empty string.

We often concatenate tuples. If a and b are tuples, we write $a\,b$ for the concatenation or, when emphasis is required, $a^\frown b$.

1.2 Structures

A (first order) language L (also called signature) is a triple that consists of

1. a set L_{fun} whose elements are called function symbols;

2. a set L_{rel} whose elements are called relation symbols;
3. a function that assigns to every $f \in L_{\text{fun}}$, respectively $r \in L_{\text{rel}}$, non-negative integers n_f and n_r that we call arity of the function, respectively relation, symbol. We say that f is an n_f-ary function symbol, and similarly for r. A 0-ary function symbol is also called a constant.

Warning: it is customary to use the symbol L to denote both the language and the set of formulas (to be defined below) associated to it. We denote by $|L|$ the cardinality of $L_{\text{fun}} \cup L_{\text{rel}} \cup \omega$. Note that, by definition, $|L|$ is always infinite.

A (first order) structure M of signature L (for short L-structure) consists of

1. a set that we call the domain or support and denote by the same symbol M used for the structure as a whole;
2. a function that assigns to every $f \in L_{\text{fun}}$ a total map $f^M : M^{n_f} \to M$;
3. a function that assigns to every $r \in L_{\text{rel}}$ a relation $r^M \subseteq M^{n_r}$.

We call f^M, respectively r^M, the interpretation of f, respectively r, in M.

Recall that, by definition, $M^0 = \{\varnothing\}$. Therefore the interpretation of a constant c is a function that maps the unique element of M^0 to an element of M. We identify c^M with $c^M(\varnothing)$.

We may use the word model as a synonym for structure. But beware that, in some contexts, the word is used to denote a particular kind of structure.

If M is an L-structure and $A \subseteq M$ is any subset, we write $L(A)$ for the language obtained by adding to L_{fun} the elements of A as constants. In this context, the elements of A are called parameters. There is a canonical expansion of M to an $L(A)$-structure that is obtained by setting $a^M = a$ for every $a \in A$.

1 Example The language of additive groups consists of the following function symbols:

1. a constant (that is, a function symbol of arity 0) $\quad 0$
2. a unary function symbol (that is, of arity 1) $\quad -$
3. a binary function symbol (that is, of arity 2) $\quad +$.

In the language of multiplicative groups the three symbols above are replaced by 1, ${}^{-1}$, and $\cdot$ respectively. Any group is a structure in either of these two signatures with the obvious interpretation. Needless to say, not all structures with these signatures are groups.

The language of (unitary) rings contains all the symbols above except ${}^{-1}$. The language of ordered rings also contains the binary relation symbol $<$.

The following example is less straightforward. The reason for the choice of the language of vector spaces will become clear in Example 9 below.

2 Example Let F be a field. The language of vector spaces over F, which we denote by L_F, extends that of additive groups by a unary function symbol k for every $k \in F$.

Recall that a vector space over F is an abelian group M together with a function $\mu : F \times M \to M$ satisfying some properties (that we assume to be well-known, see Example 6). To view a vector space over F as an L_F-structure, we interpret the group symbols in the obvious way and each $k \in F$ as the function $\mu(k, \text{-})$.

The languages in Examples 1 and 2, with the exception of that of ordered rings, are functional languages, that is, $L_{\mathrm{rel}} = \varnothing$. In what follows, we consider two important examples of relational languages, that is, languages where $L_{\mathrm{fun}} = \varnothing$.

3 Example The language of strict orders only contains a binary relation symbol, usually denoted By $<$. The language of graphs, too, only contains a binary relation symbol (for which there is no standard notation).

1.3 Terms

Let V be an infinite set whose elements we call variables. We use the letters x, y, z, etc. to denote variables or tuples of variables. We rarely refer to V explicitly, and we always assume that V is large enough for our needs.

We fix a signature L for the whole section.

4 Definition A term is a finite sequence of elements of $L_{\mathrm{fun}} \cup V$ that are obtained inductively as follows

0. every variable, intended as a tuple of length 1, is a term;

i. if $f \in L_{\mathrm{fun}}$ and $t_1, \ldots, t_{n_f}$ are terms, then $f\, t_1, \ldots, t_{n_f}$ is a term.

We say L-term when we need to specify the language L.

Note that any constant f, intended as a tuple of length 1, is a term (by i, the term f is obtained concatenating $n_f = 0$ terms and prefixing by f). Terms that do not contain variables are called closed terms.

The intended meaning of, for instance, the term $+ + x\, y\, z$ is $(x + y) + z$. The first expression uses prefix notation; the second uses infix notation. When convenient, we informally use infix notation and add parentheses to improve legibility and avoid ambiguity.

The following lemma shows that prefix notation allows to write terms unambiguously without using parentheses.

5 Lemma (Unique legibility of terms) Let a be a sequence of terms. Suppose a can be obtained both by concatenating the terms $t_1, \ldots, t_n$ and by concatenating the terms $s_1, \ldots, s_m$. Then $n = m$ and $s_i = t_i$.

Proof. By induction on $|a|$. If $|a| = 0$ than $n = m = 0$ and there is nothing to prove. Suppose the claim holds for tuples of length k and let $a = a_1, \ldots, a_{k+1}$. Then a_1 is the first element of both t_1 and s_1. If a_1 is a variable, say x, then t_1 and s_1 are the term x and $n = m = 1$. Otherwise a_1 is a function symbol, say f. Then $t_1 = f\,\bar{t}$ and $s_1 = f\,\bar{s}$, where $\bar{t}$ and $\bar{s}$ are obtained by concatenating the terms $t'_1, \ldots, t'_p$ and $s'_1, \ldots, s'_p$. Now apply the induction hypothesis to $a_2, \ldots, a_{k+1}$ and to the terms $\bar{t}\,t_2 \ldots t_n$ and $\bar{s}\,s_2 \ldots s_m$. □

If $x = x_1, \ldots, x_n$ is a tuple of distinct variables and $s = s_1, \ldots, s_n$ is a tuple of terms, we write $t[x/s]$ for the sequence obtained by replacing x by s coordinatewise. Proving that $t[x/s]$ is indeed a term is a tedious task that can be safely skipped.

If t is a term and $x_1, \ldots, x_n$ are (tuples of) variables, we write $t(x_1, \ldots, x_n)$ to declare that the variables occurring in t are among those that occur in $x_1, \ldots, x_n$. When a term has been presented as $t(x, y)$, we write $t(s, y)$ for $t[x/s]$.

Finally, we define the interpretation of a term in a structure M. We begin with closed terms. These are interpreted as 0-ary functions, i.e. as elements of the structure.

6 Definition Let t be a closed $L(M)$ term. The interpretation of t, denoted by t^M, is defined by induction on the syntax of t as follows

i. if $t = f\,t_1 \ldots t_{n_f}$, where $f \in L_{\text{fun}}$, then $t^M = f^M(t_1^M, \ldots, t_{n_f}^M)$.

Note that in i we have used Lemma 5 in an essential way. In fact this ensures that the sequence $t_1, \ldots, t_{n_f}$ uniquely determines the terms $t_1^M, \ldots, t_{n_f}^M$.

The inductive definition above is based on the case $n_f = 0$, that is, the case where f a constant, or a parameter. When $t = c$, a constant, $t_1 \ldots t_{n_f}$ is the empty tuple, and so $t^M = c^M(\varnothing)$, which we abbreviate as c^M. In particular, if $t = a$, a parameter, then $t^M = a^M = a$.

Now we generalize the interpretation to all (not necessarily closed) terms. If $t(x)$ is a term, we define $t^M(x) : M^{|x|} \to M$ to be the function that maps a to $t(a)^M$.

1.4 Substructures

In the working practice, a *substructure* is a subset of a structure that is closed under the interpretation of the functions in the language. But there are a few cases when we need the following formal definition.

7 Definition Fix a signature L and let M and N be two L-structures. We say that M is a substructure of N, and write $M \subseteq N$, if

1. the domain of M is a subset of the domain of N
2. $f^M = f^N \restriction M^{n_f}$ for every $f \in L_{\text{fun}}$
3. $r^M = r^N \cap M^{n_r}$ for every $f \in L_{\text{rel}}$.

Note that when f is a constant 2 becomes $f^M = f^N$, in particular the substructures of N contains at least all the constants of N.

If a set $A \subseteq N$ is such that

1. $f^N[A^{n_f}] \subseteq A$ for every $f \in L_{\text{fun}}$

then there is a unique substructure $M \subseteq N$ with domain A, namely, the structure with the following interpretation

2. $f^M = f^N \restriction A^{n_f}$ (which is a good definition by the assumption on A);
3. $r^M = r^N \cap A^{n_r}$.

It is usual to confuse subsets of N that satisfy 1 with the unuque substructure they support.

It is immediate to verify that the intersection of an arbitrary family of substructures of N is a substructure of N. Therefore, for any given $A \subseteq N$ we may define the substructure of N generated A as the intersection of all substructures of N that contain A. We write $\langle A \rangle_N$. The following easy proposition gives more concrete representation of $\langle A \rangle_N$

8 Lemma The following hold for every $A \subseteq N$

1. $\langle A \rangle_N = \left\{ t^N \ : \ t \text{ a closed } L(A)\text{-term} \right\}$
2. $\langle A \rangle_N = \left\{ t^N(a) \ : \ t(x) \text{ an } L\text{-term and } a \in A^{|x|} \right\}$
3. $\langle A \rangle_N = \bigcup_{n \in \omega} A_n$, where $A_0 = A$
$$A_{n+1} = A_n \cup \left\{ f^N(a) \ : \ f \in L_{\text{fun}},\ a \in A_n^{n_f} \right\}.$$

9 Example Let L be the language of groups. Let N be a group, which we consider as an L-structure in the natural way. Then the substructures of N are exactly the subgroups of N and $\langle A \rangle_N$ is the group generated by $A \subseteq N$. A similar claim is true when L_F is the signature of vector spaces over some fixed field F. The choice of the language is more or less fixed if we want that the algebraic and the model theoretic notion of substructure coincide.

1.5 Formulas

Fix a language L and a set of variables V as in Section 1.3. A formula is a finite sequence of symbols in $L_{\text{fun}} \cup L_{\text{rel}} \cup V \cup \{\doteq, \perp, \neg, \vee, \exists\}$. The last set contains the logical symbols that are called respectively

$\doteq$	equality	$\perp$	contradiction	$\neg$	negation
$\vee$	disjunction	$\exists$	existential quantifier.		

Syntactically, $\doteq$ behaves like a binary relation symbol. So, for convenience set $n_{\doteq} = 2$. However $\doteq$ is considered as a logic symbol because its semantic is fixed (it is always interpreted in the diagonal).

The definition below uses the prefix notation which simplifies the proof of the unique legibility lemma. However, in practice we always we use the infix notation: $t \doteq s$, $\varphi \vee \psi$, etc.

10 Definition A formula is any finite sequence is obtained with the following inductive procedure

0. if $r \in L_{\text{rel}} \cup \{\doteq\}$ and t is a tuple obtained concatenating n_r terms then $r\,t$ is a formula. Formulas of this form are called atomic;

i. if φ e ψ are formulas then the following are formulas: $\bot$, $\neg\,\varphi$, $\vee\,\varphi\,\psi$, and $\exists x\,\varphi$, for any $x \in V$.

We use L to denote both the language and the set of formulas. We write L_{at} for the set of atomic formulas and L_{qf} for the set of quantifier-free formulas i.e. formulas where $\exists$ does not occur.

The proof of the following is similar to the analogous lemma for terms.

11 Lemma (Unique legibility of formulas) Let a be a sequence of formulas. Suppose a can be obtained both by the concatenation of the formulas $\varphi_1, \dots, \varphi_n$ or by the concatenation of the formulas $\psi_1, \dots, \psi_m$. Then $n = m$ and $\varphi_i = \psi_i$.

A formula is closed if all its variables occur under the scope of a quantifier. Closed formulas are also called sentences. We will do without a formal definition of *occurs under the scope of a quantifier* which is too lengthy. An example suffices: all occurrences of x are under the scope a quantifiers in the formula $\exists x\,\varphi$. These occurrences are called bounded. The formula $x \doteq y \wedge \exists x\,\varphi$ has free (i.e., not bound) occurrences of x and y.

Let x is a tuple of variables and t is a tuple of terms such that $|x| = |t|$. We write $\varphi[x/t]$ for the formula obtained substituting t for all free occurrences of x, coordinatewise.

We write $\varphi(x)$ to declare that the free variables in the formula φ are all among those of the tuple x. In this case we write $\varphi(t)$ for $\varphi[x/t]$.

We will often use without explicit mention the following useful syntactic decomposition of formulas with parameters which we state without proof.

12 Lemma For every formula $\varphi(x) \in L(A)$ there is a formula $\psi(x\,;z) \in L$ and a tuple of parameters $a \in A^{|z|}$ such that $\varphi(x) = \psi(x\,;a)$.

Just as a term $t(x)$ is a name for a function $t(x)^M : M^{|x|} \to M$, a formula $\varphi(x)$ is a name for a subset $\varphi(x)^M \subseteq M^{|x|}$ which we call the subset of M defined by $\varphi(x)$. It is also very common to write $\varphi(M)$ for the set defined by $\varphi(x)$. In general sets of the form $\varphi(M)$ for some $\varphi(x) \in M$ are called definable.

13 Definitionof truth For every formula φ with variables among those of the tuple x we define $\varphi(x)^M$ by induction as follows

o1. $(\doteq t\, s)(x)^M = \Big\{a \in M^{|x|} \ : \ t^M(a) = s^M(a)\Big\}$

o2. $(r\, t_1 \ldots t_n)(x)^M = \Big\{a \in M^{|x|} \ : \ \langle t_1^M(a), \ldots, t_n^M(a)\rangle \in r^M\Big\}$

i0. $\bot(x)^M = \varnothing$

i1. $(\neg\xi)(x)^M = M^{|x|} \smallsetminus \xi(x)^M$

i2. $(\vee\ \xi\, \psi)(x)^M = \xi(x)^M \cup\ \psi(x)^M$

i3. $(\exists y\, \varphi)(x)^M = \bigcup\limits_{a\in M} (\varphi[y/a])(x)^M$

Condition i2 assumes that ξ and ψ are uniquely determined by $\vee\,\xi\,\psi$. This is a guaranteed by the unique legibility o formulas, Lemma 11. Analogously, o1 e o2 assume Lemma 5.

The case when x is the empty tuple is far from trivial. Note that $\varphi(\varnothing)^M$ is a subset of $M^0 = \{\varnothing\}$. Then there are two possibilities either $\{\varnothing\}$ or $\varnothing$. We will read them as two truth values: True and False, respectively. If $\varphi^M = \{\varnothing\}$ we say that φ is true in M, if $\varphi^M = \varnothing$, we say that φ is false M and we write $M \vDash \varphi$, respectively $M \nvDash \varphi$. In words we may say that M models φ, respectively M does not model φ. It is immediate to verify that

$$\varphi(M) = \big\{a \in M^{|x|} : M \vDash \varphi(a)\big\}.$$

Note that usually, we say *formula* when, strictly speaking, we mean *pair* that consists of a formula and a tuple of variables. Such pairs are interpreted in definable sets (cfr. Definition 13). In fact, if the tuple of variables were not given, the arity of the corresponding set is not determined.

In some contexts we also want to distinguish between two sorts of variables that play different roles. Some are placeholder for parameters, some are used to define a set. In the the first chapters this distinction is only a clue for the reader, in the last chapters it is an essential part of the definitions.

14 Definition A partitioned formula is a triple $\varphi(x\,;z)$ consisting of a formula and two tuples of variables such that the variables occurring in φ are all among x, z.

We use a semicolon to separate the two tuples of variables. Typically, z is the placeholder for parameters and x runs over the model.

1.6 Yet more notation

Now we abandon the prefix notation in favor of the infix notation. We also use the following logical connectives as abbreviations

$\top$	stands for	$\neg\bot$	tautology
$\varphi \wedge \psi$	stands for	$\neg[\neg\varphi \vee \neg\psi]$	conjunction
$\varphi \rightarrow \psi$	stands for	$\neg\varphi \vee \psi$	implication
$\varphi \leftrightarrow \psi$	stands for	$[\varphi \rightarrow \psi] \wedge [\psi \rightarrow \varphi]$	bi-implication
$\varphi \not\leftrightarrow \psi$	stands for	$\neg[\varphi \leftrightarrow \psi]$	exclusive disjunction
$\forall x\, \varphi$	stands for	$\neg\exists x \neg\, \varphi$	universal quantifier

We agree that $\rightarrow$ e $\leftrightarrow$ bind less than $\wedge$ e $\vee$. Unary connectives (quantifiers and negation) bind stronger then binarary connectives. For example

$$\exists x\, \varphi \wedge \psi \ \rightarrow\ \neg\xi \vee \vartheta \quad \text{reads as} \quad \Big[[\exists x\, \varphi] \wedge \psi\Big] \ \rightarrow\ \Big[[\neg\xi] \vee \vartheta\Big]$$

We say that $\forall x\, \varphi(x)$ and $\exists x\, \varphi(x)$ are the universal, respectively, existential closure of $\varphi(x)$. We say that $\varphi(x)$ holds in M when its universal closure is true in M. We say that $\varphi(x)$ is consistent in M when its existential closure is true in M.

The semantic of conjunction and disjunction is associative. Then for any finite set of formulas $\{\varphi_i : i \in I\}$ we can write without ambiguities

$$\bigwedge_{i\in I} \varphi_i \qquad\qquad \bigvee_{i\in I} \varphi_i$$

When $x = x_1, \ldots, x_n$ is a tuple of variables we write $\exists x\, \varphi$ or $\exists x_1, \ldots, x_n\, \varphi$ for $\exists x_1 \ldots \exists x_n\, \varphi$. Thesre is a first order sentences that say that $\varphi(M)$ has at least n elements (also, no more than, or exactly n). It is convenient to use the following abbreviations.

$\exists^{\geq n} x\, \varphi(x)$	stands for	$\exists x_1, \ldots, x_n \Big[\bigwedge\limits_{1\leq i\leq n} \varphi(x_i) \ \wedge \bigwedge\limits_{1\leq i<j\leq n} x_i \dot{\neq} x_j \Big].$
$\exists^{\leq n} x\, \varphi(x)$	stands for	$\neg\exists^{\geq n+1} x\, \varphi(x)$
$\exists^{= n} x\, \varphi(x)$	stands for	$\exists^{\geq n} x\, \varphi(x) \ \wedge\ \exists^{\leq n} x\, \varphi(x)$

15 Exercise Let M be an L-structure and let $\psi(y), \varphi(x, y) \in L$. For each of the following conditions, write a sentence true in M exactly when

a. $\psi(M) \ \in\ \{\varphi(a, M) : a \in M\}$;

b. $\{\varphi(a, M) : a \in M\}$ contains at least two sets;

c. $\{\varphi(a, M) : a \in M\}$ contains only sets that are pairwise disjoint.

16 Exercise Let M be a structure in a signature that contains a symbol r for a binary relation. Write a sentence φ such that

a. $M \vDash \varphi$ if and only if there is $A \subseteq M$ such that $r^M \ \subseteq\ A \times \neg A$.

Remark: φ assert an asymmetric version of the property below

b. $M \vDash \psi$ if and only if there is $A \subseteq M$ such that $r^M \ \subseteq\ (A \times \neg A) \cup (\neg A \times A)$.

Assume M is a graph, what required in b is equivalent to saying that M is a *bipartite graph*, or equivalently that it has *chromatic number* 2 i.e., we can color the vertices with 2 colors so that no two adjacent vertices share the same color.

Chapter 2

Theories and elementarity

2.1 Logical consequences

A theory is a set $T \subseteq L$ of sentences. We write $M \vDash T$ when $M \vDash \varphi$ for every $\varphi \in T$. If $\varphi \in L$ is a sentence we write $T \vdash \varphi$ when

$$M \vDash T \quad \Rightarrow \quad M \vDash \varphi \qquad \text{for every } M.$$

In words, we say that φ is a logical consequence of T or that φ follows from T. If S is a theory $T \vdash S$ has a similar meaning. If $T \vdash S$ and $S \vdash T$ we say that T and S are logically equivalent. We may say that T axiomatizes S (or vice versa).

We say that a theory is consistent if it has a model. With the notation above, T is consistent if and only if $T \nvdash \perp$.

The closure of T under logical consequence is the set $\mathrm{ccl}(T)$ which is defined as follows:

$$\mathrm{ccl}(T) \quad = \quad \Big\{\varphi \in L \ : \ \text{sentence such that } T \vdash \varphi\Big\}$$

If T is a finite set, say $T = \{\varphi_1, \ldots, \varphi_n\}$ we write $\mathrm{ccl}(\varphi_1, \ldots, \varphi_n)$ for $\mathrm{ccl}(T)$. If $T = \mathrm{ccl}(T)$ we say that T is closed under logical consequences.

The theory of M is the set of sentences that hold in M and is denoted by $\mathrm{Th}(M)$. More generally, if $\mathcal{K}$ is a class of structures, $\mathrm{Th}(\mathcal{K})$ is the set of sentences that hold in every model in $\mathcal{K}$. That is

$$\mathrm{Th}(\mathcal{K}) \quad = \quad \bigcap_{M \in \mathcal{K}} \mathrm{Th}(M)$$

The class of all models of T is denoted by $\mathrm{Mod}(T)$. We say that $\mathcal{K}$ is axiomatizable if $\mathrm{Mod}(T) = \mathcal{K}$ for some theory T. If T is finite we say that $\mathcal{K}$ is finitely axiomatizable. To sum up

$$\begin{aligned}
\mathrm{Th}(M) \quad &= \quad \Big\{\varphi \ : \ M \vDash \varphi\Big\} \\
\mathrm{Th}(\mathcal{K}) \quad &= \quad \Big\{\varphi \ : \ M \vDash \varphi \text{ for all } M \in \mathcal{K}\Big\} \\
\mathrm{Mod}(T) \quad &= \quad \Big\{M \ : \ M \vDash T\Big\}
\end{aligned}$$

1 Example Let L be the language of multiplicative groups. Let T_{g} be the set containing the universal closure of following three formulas

1. $(x \cdot y) \cdot z \ = \ x \cdot (y \cdot z);$
2. $x \cdot x^{-1} \ = \ x^{-1} \cdot x \ = \ 1;$
3. $x \cdot 1 \ = \ 1 \cdot x \ = x.$

Then T_{g} axiomatizes the theory of groups, i.e. $\mathrm{Th}(\mathcal{K})$ for $\mathcal{K}$ the class of all groups. Let φ be the universal closure of the following formula

$$z \cdot x \ = \ z \cdot y \ \to \ x \ = \ y.$$

As φ formalizes the cancellation property then $T_{\mathrm{g}} \vdash \varphi$, that is, φ is a logical consequence of T_{g}. Now consider the sentence ψ which is the universal closure of

4. $x \cdot y \ = \ y \cdot x.$

So, commutative groups model ψ and non commutative groups model $\neg\psi$. Hence neither $T_{\mathrm{g}} \vdash \psi$ nor $T_{\mathrm{g}} \vdash \neg\psi$. We say that T_{g} does not decide ψ.

Note that even when T is a very concrete set, $\mathrm{ccl}(T)$ may be more difficult to grasp. In the example above T_{g} contains three sentences but $\mathrm{ccl}(T_{\mathrm{g}})$ is an infinite set containing sentences that code theorems of group theory yet to be proved.

2 Remark The following properties say that ccl is a finitary closure operator.

1. $T \subseteq \mathrm{ccl}(T)$ (extensive)
2. $\mathrm{ccl}(T) \ = \ \mathrm{ccl}\big(\mathrm{ccl}(T)\big)$ (idempotent)
3. $T \subseteq S \ \Rightarrow \ \mathrm{ccl}(T) \subseteq \mathrm{ccl}(S)$ (increasing)
4. $\mathrm{ccl}(T) \ = \ \bigcup\big\{\mathrm{ccl}(S) \ : \ S \text{ finite subset of } T\big\}.$ (finitary)

Properties 1-3 are easy to verify while 4 requires the compactness theorem.

In the next example we list a few algebraic theories with straightforward axiomatization.

3 Example We write T_{ag} for the theory of abelian groups which contains the universal closure of following

a1. $(x+y)+z \ = \ y+(x+z);$

a2. $x+(-x) \ = \ 0;$

a3. $x+0 \ = \ x;$

a4. $x+y \ = \ y+x.$

4 Example The theory T_r of (unitary) rings extends T_{ag} with

a5. $(x \cdot y) \cdot z \;=\; x \cdot (y \cdot z)$;

a6. $1 \cdot x \;=\; x \cdot 1 \;=\; x$;

a7. $(x + y) \cdot z \;=\; x \cdot z + y \cdot z$;

a8. $z \cdot (x + y) \;=\; z \cdot x + z \cdot y$.

The theory of commutative rings T_{cg} contains also com of examples 1.

5 Example The theory of ordered rings T_{or} extends T_{cr} with

o1. $x < z \;\rightarrow\; x + y < z + y$;

o2. $0 < x \wedge 0 < z \;\rightarrow\; 0 < x \cdot z$.

6 Example The axiomatization of the theory of vector spaces is less straightforward. Fix a field F. The language L_F extends the language of additive groups with a unary function for every element of F. The theory of vector fields over F extends T_{ag} with the following axioms (for all $h, k, l \in F$)

m1. $h\,(x + y) \;=\; h\,x + h\,y$

m2. $l\,x \;=\; h\,x + k\,x$, where $l \;=\; h +_F k$

m3. $l\,x \;=\; h\,(k\,x)$, where $l \;=\; h \cdot_F k$

m4. $0_F\,x \;=\; 0$

m5. $1_F\,x \;=\; x$

The symbols 0_F and 1_F denote the zero and the unit of F. The symbols $+_F$ and $\cdot_F$ denote the sum and the product in F. These are not part of L_F, they are symbols we use in the metalanguage.

7 Example Recall from Example 3 that we represent a graph with a symmetric irreflexive relation. Therefore theory of graphs contains the following two axioms

1. $\neg r(x, x)$;
2. $r(x, y) \rightarrow r(y, x)$.

Our last example is a trivial one.

8 Example Let L be the empty language The theory of infinite sets is axiomatized by the sentences $\exists^{\geq n} x\ (x = x)$ for all positive integer n.

9 Exercise Prove that $\operatorname{ccl}(\varphi \vee \psi) = \operatorname{ccl}(\varphi) \cap \operatorname{ccl}(\psi)$.

10 Exercise Prove that $T \cup \{\varphi\} \vdash \psi$ then $T \vdash \varphi \rightarrow \psi$.

11 Exercise Prove that $\operatorname{Th}(\operatorname{Mod}(T)) = \operatorname{ccl}(T)$.

2.2 Elementary equivalence

The following is a fundamental notion in model theory.

12 Definition We say that M and N are elementarily equivalent if

ee. $N \vDash \varphi \quad \Leftrightarrow \quad M \vDash \varphi,$ for every sentence $\varphi \in L$.

In this case we write $M \equiv N$. More generally, we write $M \equiv_A N$ and say that M and N are elementarily equivalent over A if the following hold

a. $A \subseteq M \cap N$

ee'. equivalence ee above holds for every sentence $\varphi \in L(A)$.

The case when A is the whole domain of M is particularly important.

13 Definition When $M \equiv_M N$ we say that M is an elementary substructure of N and write $M \preceq N$.

The following lemma shows that the use of the term *substructure* in the definition above is appropriate.

14 Lemma If M and N are such that $M \equiv_A N$ and A is the domain of a substructure of M then A is also the domain a substructure of N and the two substructures coincide.

Proof. Let f be a function symbol and ler r be a relation symbol. It suffices to prove that $f^M(a) = f^N(a)$ for every $a \in A^{n_f}$ and that $r^M \cap A^{n_r} = r^N \cap A^{n_r}$.

If $b \in A$ is such that $b = f^M a$ then $M \vDash fa = b$. So, from $M \equiv_A N$, we obtain $N \vDash fa = b$, hence $f^N a = b$. This proves $f^M(a) = f^N(a)$.

Now let $a \in A^{n_r}$ and suppose $a \in r^M$. Then $M \vDash ra$ and, by elementarity, $N \vDash ra$, hence $a \in r^N$. By symmetry $r^M \cap A^{n_r} = r^N \cap A^{n_r}$ follows. □

It is not easy to prove that two structures are elementary equivalent. A direct verification is unfeasible even for the most simple structures. It will take a few chapters before we are able to discuss concrete examples.

We generalize the definition of $\mathrm{Th}(M)$ to include parameters

$$\mathrm{Th}(M/A) \quad = \quad \Big\{\varphi \ : \ \text{sentence in } L(A) \text{ such that } \ M \vDash \varphi\Big\}.$$

The following proposition is immediate

15 Proposition For every pair of structures M and N and every $A \subseteq M \cap N$ the following are equivalent

a. $M \equiv_A N$;

b. $\mathrm{Th}(M/A) = \mathrm{Th}(N/A)$;

c. $M \vDash \varphi(a) \quad \Leftrightarrow \quad N \vDash \varphi(a)$ for every $\varphi(x) \in L$ and every $a \in A^{|x|}$.

d. $\varphi(M) \cap A^{|x|} \ = \ \varphi(N) \cap A^{|x|}$ for every $\varphi(x) \in L$.

If we restate a and c of the proposition above when $A = M$ we obtain that the following are equivalent

a'. $M \preceq N$;

d'. $\varphi(M) \quad = \quad \varphi(N) \cap M^{|x|} \qquad$ for every $\varphi(x) \in L$.

Note that c' extends to all definable sets what Definition 7 requires for a few basic definable sets.

16 Example Let G be a group which we consider as a structure in the multiplicative language of groups. We show that if G is simple and $H \preceq G$ then also H is simple. Recall that G is simple if all its normal subgroups are trivial, equivalently, if for every $a \in G \smallsetminus \{1\}$ the set $\{gag^{-1} : g \in G\}$ generates the whole group G.

Assume H is not simple. Then there are $a, b \in H$ such that b is not the product of elements of $\{hah^{-1} : h \in H\}$. Then for every n

$$H \models \neg\exists x_1, \ldots, x_n \ \left(b \ = \ x_1 a x_1^{-1} \cdots \ x_n a x_n^{-1}\right)$$

By elementarity the same hold in G. Hence G is not simple.

17 Exercise Let $A \subseteq M \cap N$. Prove that $M \equiv_A N$ if and only if $M \equiv_B N$ for every finite $B \subseteq A$.

18 Exercise Let $M \preceq N$ and let $\varphi(x) \in L(M)$. Prove that $\varphi(M)$ is finite if and only if $\varphi(N)$ is finite and in this case $\varphi(N) = \varphi(M)$.

19 Exercise Let $M \preceq N$ and let $\varphi(x, z) \in L$. Suppose there are finitely many sets of the form $\varphi(a, N)$ for some $a \in N^{|x|}$. Prove that all these sets are definable over M.

20 Exercise Consider $\mathbb{Z}^n$ as a structure in the additive language of groups with the natural interpretation. Prove that $\mathbb{Z}^n \not\equiv \mathbb{Z}^m$ for every positive integers $n \neq m$. Hint: in $\mathbb{Z}^n$ there are at most 2^n elements that are not congruent modulo 2.

2.3 A nonstandard example

This section concerns an example that is useful to look at in some detail to get some familiarity with the notion of elementary substructure. The example is culturally interesting, because it formalises rigorously the notions of infinity and infinitesimal, which were used in Newton and Leibniz's time to develop real analysis. These notions were not well defined - in fact, they were inconsistent.

It was only in the mid 19th century that mathematicians of Weiertsrass' generation developed the notion of limit, thus providing rigorous grounds for the development of analysis.

Nonstandard analysis was developed by Abraham Robinson in the 1950s. Robinson found a way to formalise the ideas of infinity and infinitesimals through the concept of elementary extension.

The notation that follows will be used throughout this section.

The language we use contains

1. X, a relation symbol of arity n, for every $n \in \omega$ and every $X \subseteq \mathbb{R}^n$;

2. f, a function symbol of arity n, for every $n \in \omega$ and every $f : \mathbb{R}^n \to \mathbb{R}$.

The standard model of real analysis is $\mathbb{R}$ with the natural interpretation of the symbols in our language, so we use the same symbol for an element of the language and its interpretation in $\mathbb{R}$. This is not an abuse of notation: in this case, elements and interpretations coincide.

Suppose $\mathbb{R}$ has a proper elementary extension ${}^*\mathbb{R}$. The existence of such an extension will be proved later. The interpretations of the symbols f and X in ${}^*\mathbb{R}$ will be denoted by ${}^*f$ and ${}^*X$, respectively. The elements of ${}^*\mathbb{R}$ are called hyperreals, the elements of $\mathbb{R}$ are called standard (hyper)reals, those in ${}^*\mathbb{R} \setminus \mathbb{R}$ are called nonstandard (hyper)reals.

It is easy to verify that ${}^*\mathbb{R}$ is an ordered field. This is because the operations sum and product are in the language, and so is the order relation. The property of being an ordered field can be expressed via a set of sentences that are true in $\mathbb{R}$, and hence also in ${}^*\mathbb{R}$.

A hyperreal c is said to be infinitesimal if $|c| < \varepsilon$ for every positive standard ε. A hyperreal c is infinite if $k < |c|$ for every standard k; the hyperreal c if finite otherwise. Hence if c is infinite, then c^{-1} is infinitesimal. Clearly, all standard reals are finite, and 0 is the only infinitesimal standard real.

The proof of the following lemma is left to the reader.

21 Fact Infinitesimals are closed under sum, product and multiplication by standard reals.

We begin with an existence proof.

22 Lemma There are infinite nonstandard hyperreals and nonzero infinitesimals. Moreover, for every finite hyperreal c there is a unique standard b such that $b - c$ is infinitesimal.

Proof. If $c \in {}^*\mathbb{R}$ in infinite then c^{-1} is infinitesimal and if c is a nonzero infinitesimal, then c^{-1} is infinite. Then the first claim follows from the second one. Let c be finite. The set $\{a \in \mathbb{R} : c < a\}$ is a nonempty set of (standard) reals that is bounded from below. Let $b \in \mathbb{R}$ the least upper bound. We show that $b - c$ is a infinitesimal. Assume for a contradiction that $\varepsilon < |b - c|$ for some standard positive ε. Then $c < b - \varepsilon$, or $b + \varepsilon < c$, depending on whether $c < b$ or $b < c$. Both cases contradict that b is an infimum. This proves the existence of b.

To prove uniquenes, suppose that $b' \in \mathbb{R}$ and $b' - c$ is also infinitesimal. Then $b - b'$ is also infinitesimal. As 0 is the only standard infinitesimal $b = b'$. □

The existence of infinite non standard hyperreals shows that ${}^*\mathbb{R}$ is not archimedean: standard integers are not cofinal in ${}^*\mathbb{R}$. However, by elementarity, the nonstandard integers ${}^*\mathbb{N}$ are cofinal in ${}^*\mathbb{R}$: from the perspective of an inhabitant of ${}^*\mathbb{R}$, the latter is a normal archimedean field.

Dedekind completeness is another fundamental property of $\mathbb{R}$ that does not hold in ${}^*\mathbb{R}$. An ordered set is Dedekind complete if every subset that is bounded above has a least upper bound. Then ${}^*\mathbb{R}$ is not Dedekind complete: the set of infinitesimals is bounded (both above and below), but, by Lemma 21, it does not have a least upper bound. It follows that Dedekind completeness is *not* a first-order property.

However, it sometimes happens that properties that hold of all sets in the standard model hold in ${}^*\mathbb{R}$ for definable sets only.

We define the following equivalence relation on ${}^*\mathbb{R}$: we write $a \approx b$ if $|a-b|$ is infinitesimal. The fact that this is an equivalence relation follows from Fact 21. The equivalence class of c is called monad. By Lemma **??** if c is a finite hyperreal, there is a unique real in the monad of c.

Hyperreals that are not finite are said to be infinite. If c is finite, the unique standard real in the monad of c is called standard part of c and it is denoted by $\mathrm{st}(c)$.

In the following lemma, the expressions on the left can be formalised as first-order sentences, and therefore they hold in $\mathbb{R}$ if and only if they hold in ${}^*\mathbb{R}$.

23 Proposition For every $f : \mathbb{R} \to \mathbb{R}$, for every $a, l \in \mathbb{R}$ the following equivalences hold.

a. $\lim\limits_{x\to+\infty} fx = +\infty \quad \Leftrightarrow \quad {}^*f(c)$ is positive and infinite for every infinite $c > 0$

b. $\lim\limits_{x\to+\infty} fx = l \quad \Leftrightarrow \quad {}^*f(c) \approx l$ for every infinite $c > 0$

c. $\lim\limits_{x\to a} fx = +\infty \quad \Leftrightarrow \quad {}^*f(c)$ is positive and infinite for every $c \approx a \neq c$

d. $\lim\limits_{x\to a} fx = l \quad \Leftrightarrow \quad {}^*f(c) \approx l$ for every for every $c \approx a \neq c$.

Proof. We prove part d and we leave the other parts as an exercise. For $\Rightarrow$, assume that the left-hand side of the equivalence d holds and we write it as a first-order sentence:

1. $$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \Big[0 < |x-a| < \delta \;\;\to\;\; |fx - l| < \varepsilon\Big].$$

By assumption, the formula 1 holds in $\mathbb{R}$, or, equivalently, in ${}^*\mathbb{R}$. In what follows, we use certain abbreviations which we assume the reader can translate into first-order formulas. For consistency with standard notation in analysis, we use the Greek letters ε and δ as variables. The symbols $\dot\varepsilon$ e $\dot\delta$ denote parameters.

We now check that ${}^*f(c) \approx l$ holds for every $c \approx a \neq c$, that is, $|{}^*fc - l| < \dot\varepsilon$ for every positive standard $\dot\varepsilon$. Fix a standard positive $\dot\varepsilon$ and let $\dot\delta \in \mathbb{R}$ be a standard real obtained from 1. By elementarity, we have

$$ {}^*\mathbb{R} \vDash \forall x \Big[0 < |x-a| < \dot\delta \;\;\to\;\; |fx - l| < \dot\varepsilon\Big],$$

where $\dot\varepsilon$ and $\dot\delta$ are now parameters. If $a \approx c \neq a$, then $0 < |c-a| < \dot\delta$ certainly holds (because $\dot\delta$ is standard). Hence 1 gives $|{}^*fc - l| < \dot\varepsilon$.

For $\Leftarrow$, assume that 1 is false, that is, assume that, in $\mathbb{R}$,

2. $$\exists \varepsilon > 0\, \forall \delta > 0\, \exists x \Big[0 < |x-a| < \delta \;\;\wedge\;\; \varepsilon \leq |fx - l|\Big],$$

We want to show that ${}^*f(c) \not\approx l$ for some $c \neq a$ infinitely close to a. Fix a witness $\dot\varepsilon$ of this formula in $\mathbb{R}$ — so $\dot\varepsilon$ is a standard real. Now, by elementarity we have

$$ {}^*\mathbb{R} \vDash \forall \delta > 0\, \exists x \Big[0 < |x-a| < \delta \;\;\wedge\;\; \dot\varepsilon \leq |fx - l|\Big].$$

Let $\dot\delta$ be an arbitrary infinitesimal. Then

$$^*\mathbb{R} \vDash \exists x \left[0 < |x - a| < \dot\delta \ \wedge \ \dot\varepsilon \leq |fx - l|\right].$$

Any c that witnesses the truth of this formula in $^*\mathbb{R}$ is such that $c \approx a \neq c$ and, simultaneously, $\varepsilon \leq |^*fc - l|$. But $\dot\varepsilon$ was chosen to be standard, so $^*fc \not\approx l$. $\square$

The following corollary is immediate.

24 Corollary For every $f : \mathbb{R} \to \mathbb{R}$ the following are equivalent

a. f is continuous;

b. $^*f(a) \approx {}^*f(c)$ for every pair of finite hyperreals such that $c \approx a$.

In Corollary 24, it is important to restrict c to *finite* hyperreals, otherwise we get a stronger property.

25 Proposition For every $f : \mathbb{R} \to \mathbb{R}$, the following are equivalent

a. f is uniformly continuous;

b. $^*f(a) \approx {}^*f(b)$ for every pair of hyperreals such that $a \approx b$.

Proof. We prove a⇒b. Recall that f is uniformly continuous if

1. $$\mathbb{R} \vDash \forall \varepsilon > 0\, \exists \delta > 0\, \forall x, y \left[|x - y| < \delta \ \rightarrow \ |fx - fy| < \varepsilon\right].$$

Assume a and let $a \approx b$. We wnat to show that $|^*f(a) - {}^*f(b)| < \dot\varepsilon$ for every positive standard $\dot\varepsilon$. Given a standard positive $\dot\varepsilon$, let $\dot\delta$ be a standard real obtained from the validity of 1 in $\mathbb{R}$. Now elementarity gives

2. $$^*\mathbb{R} \vDash \forall x, y \left[|x - y| < \dot\delta \ \rightarrow \ |fx - fy| < \dot\varepsilon\right].$$

In particular,

$$^*\mathbb{R} \vDash \ |a - b| < \dot\delta \ \rightarrow \ |fa - fb| < \dot\varepsilon.$$

Since $a \approx b$, we have $|a - b| < \dot\delta$ for any standard $\dot\delta$. Therefore $|^*f(a) - {}^*f(b)| < \dot\varepsilon$.

To prove b ⇒ a we negate a

3. $$\mathbb{R} \vDash \exists \varepsilon > 0\, \forall \delta > 0\, \exists x, y \left[|x - y| < \delta \ \wedge \ \varepsilon \leq |fx - fy|\right].$$

We want $a \approx b$ such that $\dot\varepsilon \leq |^*f(a) - {}^*f(b)|$ for some positive standard $\dot\varepsilon$. Let $\dot\varepsilon$ be a standard real that witnesses the truth of 3 in $\mathbb{R}$. Elementarity gives

$$^*\mathbb{R} \vDash \forall \delta > 0\, \exists x, y \left[|x - y| < \delta \ \wedge \ \dot\varepsilon \leq |fx - fy|\right].$$

So we can fix an arbitrary infinitesimal $\dot\delta > 0$ and get $a, b \in {}^*\mathbb{R}$ such that

$$^*\mathbb{R} \vDash |a - b| < \dot\delta \ \wedge \ \dot\varepsilon \leq |fa - fb|.$$

Since $\dot\delta$ is infinitesimal, we have $a \approx b$, as required. $\square$

The following proposition is an immediate consequence of Proposition 23.

26 **Proposition** For every unary function f and every standard a, the following are equivalent.

a. f is differentiable in a;

b. $\operatorname{st}\left(\dfrac{f(a)-f(a+h)}{h}\right)$ exists and is constant for all infinitesimal $h \neq 0$.

27 **Exercise** Every definable (possibly with parameters) subset of ${}^*\mathbb{R}$ that is bounded from above has a least upper bound.

28 **Exercise** Prove that if the function $f : \mathbb{R} \to \mathbb{R}$ is injective, then ${}^*fa$ is nonstantard whenever a is nonstandard.

29 **Exercise** Prove that the following are equivalent for every subset $X \subseteq \mathbb{R}$

1. X is a finite set;
2. ${}^*X = X$.

30 **Exercise** Prove that the following are equivalent for every subset $X \subseteq \mathbb{R}$

1. X is an open set in the usual topology on $\mathbb{R}$;
2. $b \approx a \in X \;\;\Rightarrow\;\; b \in {}^*X$ for every $b \in {}^*\mathbb{R}$.

31 **Exercise** Prove that the following are equivalent for every subset $X \subseteq \mathbb{R}$

1. X is closed in the usual topology on $\mathbb{R}$;
2. $a \in {}^*X \;\;\Rightarrow\;\; \operatorname{st} a \in X$ for every finite $a \in {}^*\mathbb{R}$.

32 **Exercise** Prove that the following are equivalent for every subset $X \subseteq \mathbb{R}$

1. X is bounded and closed in the usual topology on $\mathbb{R}$;
2. for every $b \in {}^*X$ there is an $a \in X$ such that $a \approx b$.

33 **Exercise** For what sets $X \subseteq \mathbb{R}$ does the following hold for every $a, b \in {}^*\mathbb{R}$?

$$b \approx a \in {}^*X \;\Rightarrow\; b \in {}^*X.$$

34 **Exercise** Prove that $|\mathbb{R}| \;\leq\; |{}^*\mathbb{Q}|$, that is, that the cardinality of ${}^*\mathbb{Q}$ is at least that of the continuum. (Hint: define an injective function $f : \mathbb{R} \to {}^*\mathbb{Q}$ by choosing a nonstandard rational in the monad of every standard real.)

35 **Exercise** Prove that every proper elementary extension of $\mathbb{Q}$ in the language of rings plus a relation symbol for every subset of $\mathbb{Q}$ has infinities and infinitesimals.

2.4 Embeddings and isomorphisms

Here we prove that isomorphic structures are elementarily equivalent and a few related results.

36 Definition An embedding of M into N is an injective total map $h : M \hookrightarrow N$ such that

1. $\quad a \in r^M \quad \Leftrightarrow \quad ha \in r^N \qquad$ for every $r \in L_{\text{rel}}$ and $a \in M^{n_r}$;
2. $\quad h\, f^M(a) \quad = \quad f^N(h\, a) \qquad$ for every $f \in L_{\text{fun}}$ and $a \in M^{n_f}$.

Note that when $c \in L_{\text{fun}}$ is a constant 2 reads $h\, c^M = c^N$. Therefore $M \subseteq N$ if and only if $\text{id}_M : M \to N$ is an embedding.

An surjective embedding is an isomorphism or, when domain and codomain coincide, an automorphism.

Condition 1 above and the assumption that h is injective can be summarized in the following

1'. $\quad M \vDash r(a) \quad \Leftrightarrow \quad N \vDash r(ha) \qquad$ for every $r \in L_{\text{rel}} \cup \{\doteq\}$ and every $a \in M^{n_r}$.

Note also that, by straightforward induction on syntax, from 2 we obtain

2' $\quad h\, t^M(a) \quad = \quad t^N(h\, a) \qquad$ for every term $t(x)$ and every $a \in M^{|x|}$.

Combining these two properties and a straightforward induction on the syntax give

3. $\quad M \vDash \varphi(a) \quad \Leftrightarrow \quad N \vDash \varphi(ha) \qquad$ for every $\varphi(x) \in L_{\text{qf}}$ and every $a \in M^{|x|}$.

Recall that we write L_{qf} for the set of quantifier-free formulas. It is worth noting that when $M \subseteq N$ and $h = \text{id}_M$ then 3 becomes

3' $\quad M \vDash \varphi(a) \quad \Leftrightarrow \quad N \vDash \varphi(a) \qquad$ for every $\varphi(x) \in L_{\text{qf}}$ and for every $a \in M^{|x|}$.

In words this is summarized by saying that the truth of quantifier-free formulas is preserved under sub- and superstructure.

Finally, we prove that first order truth is preserved under isomorphism. We say that a map $h : M \to N$ fixes $A \subseteq M$ (pointwise) if $\text{id}_A \subseteq h$. An isomorphism that fixes A is also called an A-isomorphism.

37 Theorem If $h : M \to N$ is an isomorphism then for every $\varphi(x) \in L$

$$M \vDash \varphi(a) \quad \Leftrightarrow \quad N \vDash \varphi(ha) \text{ for every } a \in M^{|x|}$$

In particular, if h is an A-isomorphism then $M \equiv_A N$.

Proof. We proceed by induction on the syntax of $\varphi(x)$. When $\varphi(x)$ is atomic # holds by 3 above. Induction for the Boolean connectives is straightforward so we only need to consider the existential quantifier. Assume as induction hypothesis that

$$M \vDash \varphi(a, b) \quad \Leftrightarrow \quad N \vDash \varphi(ha, hb) \text{ for every } a \in M^{|x|} \text{ and } b \in M.$$

We prove that the equivalence in the theorem holds for the formula $\exists y\, \varphi(x, y)$.

$$\begin{aligned} M \vDash \exists y\, \varphi(a, y) \quad &\Leftrightarrow \quad M \vDash \varphi(a, b) \text{ for some } b \in M \\ &\Leftrightarrow \quad N \vDash \varphi(ha, hb) \text{ for some } b \in M \qquad \text{(by induction hypothesis)} \\ &\Leftrightarrow \quad N \vDash \varphi(ha, c) \text{for some } c \in N \qquad (\Leftarrow \text{ by surjectivity}) \\ &\Leftrightarrow \quad N \vDash \exists y\, \varphi(ha, y). \end{aligned}$$

□

38 Corollary If $h : M \to N$ is an isomorphism then for every $\varphi(x) \in L$.

$$h[\varphi(M)] \quad = \quad \varphi(N)$$

We can now give a few very simple examples of elementarily equivalent structures.

39 Example Let L be the language of strict orders. Consider intervals of $\mathbb{R}$ (or in $\mathbb{Q}$) as structures in the natural way. The intervals $[0,1]$ and $[0,2]$ are isomorphic, hence $[0,1] \equiv [0,2]$ follows from Theorem 37. Clearly, $[0,1]$ is a substructure of $[0,2]$. However $[0,1] \not\preceq [0,2]$, in fact the formula $\forall x\,(x{\leq}1)$ holds in $[0,1]$ but is false in $[0,2]$. This shows that $M \subseteq N$ and $M \equiv N$ does not imply $M \preceq M$.

Now we prove that $(0,1) \preceq (0,2)$. By Exercise 17 above, it suffices to verify that $(0,1) \equiv_B (0,2)$ for every finite $B \subseteq (0,1)$. This follows again by Theorem 37 as $(0,1)$ and $(0,2)$ are B-isomorphic for every finite $B \subseteq (0,1)$.

For the sake of completeness we also give the definition of homomorphism.

40 Definition A homomorphism is a total map $h : M \to N$ such that

1. $a \in r^M \quad \Rightarrow \quad ha \in r^N$ for every $r \in L_{\text{rel}}$ and $a \in M^{n_r}$;
2. $h\, f^M(a) \quad = \quad f^N(h\, a)$ for every $f \in L_{\text{fun}}$ and $a \in M^{n_f}$.

Note that only one implication is required in 1.

41 Exercise Prove that if $h : N \to N$ is an automorphism and $M \preceq N$ then $h[M] \preceq N$.

42 Exercise Let L be the empty language. Let $A, D \subseteq M$. Prove that the following are equivalent

1. D is definable over A;
2. either D is finite and $D \subseteq A$, or $\neg D$ is finite and $\neg D \subseteq A$.

Hint: as structures are plain sets, every bijection $f : M \to M$ is an automorphism.

43 Exercise Prove that if $\varphi(x)$ is an existential formula and $h : M \hookrightarrow N$ is an embedding then

$$M \vDash \varphi(a) \quad \Rightarrow \quad N \vDash \varphi(ha) \qquad\qquad \text{for every } a \in M^{|x|}.$$

Recall that existential formulas as those of the form $\exists y\, \psi(x,y)$ for $\psi(x,y) \in L_{\text{qf}}$. Note that Theorem **??** proves that the property above characterizes existential formulas.

44 Exercise Let M be the model with domain $\mathbb{Z}$ in the language of additive groups which is interpreted in the natural way. Prove that there is no existential formula $\varphi(x)$ such that $\varphi(M)$ is the set of odd integers. Hint: use Exercise 43.

45 Exercise Let $\mathbb{Q}^+$ be the multiplicative group of positive rationals. Let M be the subgroup of the numbers of the form n/m for some odd integers m and n. Prove that $M \preceq \mathbb{Q}^+$. Hint: use the fundamental theorem of arithmetic and reason as in Example 39.

46 **Exercise** Prove that, in the language of strict orders, $\mathbb{R} \setminus \{0\} \preceq \mathbb{R}$ and $\mathbb{R} \setminus \{0\} \not\preceq \mathbb{R}$.

2.5 Quotient structure

The content of this section is mainly technical and only required later in the course. Its reading may be postponed.

If E is an equivalence relation on N we write $[c]_E$ for the equivalence class of $c \in N$. We use the same symbol for the equivalence relation on N^n defined as follow: if $a = a_1, \ldots, a_n$ and $b = b_1, \ldots, b_n$ are n-tuples of elements of N then $a \mathrel{E} b$ means that $a_i \mathrel{E} b_i$ holds for all i. It is easy to see that $b_1, \ldots, b_n \in [a, ..., a_n]_E$ if and only if $b_i \in [a_i]_E$ for all i. Therefore we use the notation $[a]_E$ for both the equivalence class of $a \in N^n$ and the tuple of equivalence classes $[a_1]_E, \ldots, [a_n]_E$.

47 **Definition** We say that the equivalence relation E on a structure N is a congruence if for every $f \in L_{\text{fun}}$

c1. $$a \mathrel{E} b \quad \Rightarrow \quad f^N a \mathrel{E} f^N b;$$

When E is a congruence on N we write N/E for the a structure that has as domain the set of E-equivalence classes in N and the following interpretation of $f \in L_{\text{fun}}$ and $r \in L_{\text{rel}}$:

c2. $$f^{N/E}[a]_E \quad = \quad \left[f^N a\right]_E;$$

c3. $$[a]_E \ \in \ r^{N/E} \quad \Leftrightarrow \quad [a]_E \cap r^N \neq \varnothing.$$

We call N/E the quotient structure.

By c1 the quotient structure is well defined. The reader will recognize it as a familiar notion by the following proposition (which is not required in the following and requires the notion of homomorphism, see Definition 40. Recall that the kernel of a total map $h : N \to M$ is the equivalence relation E such that

$$a \mathrel{E} b \quad \Leftrightarrow \quad ha = hb$$

for every $a, b \in N$.

48 **Proposition** Let $h : N \to M$ be a surjective homomorphism and let E be the kernel of h. Then there is an isomorphism k that makes the following diagram commute

$$\begin{tikzcd}
N \arrow[r, "h"] \arrow[d, "\pi"'] & M \\
N/E \arrow[ur, "k"] &
\end{tikzcd}$$

where $\pi : a \mapsto [a]_E$ is the projection map.

Quotients clutter the notation with brackets. To avoid the mess, we prefer to reason in N and tweak the satisfaction relation. Warning: this is not standard (though it is what everyone informally does).

Recall that in model theory, equality is not treated as a all other predicates. In fact, the interpretation of equality is fixed to always be the identity relation. In a few contexts is convinient to allow any congruence to interprete equality. This allows to work in N while thinking of N/E.

We define $N/E \overset{*}{\vDash}$ to be $N \vDash$ but with equality interpreted with E. The proposition below shows that this is the same thing as the regular truth in the quotient structure, $N/E \vDash$.

49 Definition For t_2, t_2 closed terms of $L(N)$ define

1* $\quad N/E \overset{*}{\vDash} t_1 = t_2 \quad \Leftrightarrow \quad t_1^N \; E \; t_2^N$

For t a tuple of closed terms of $L(N)$ and $r \in L_{\mathrm{rel}}$ a relation symbol

2* $\quad N/E \overset{*}{\vDash} r\,t \quad \Leftrightarrow \quad t^N \; E \; a \quad$ for some $a \in r^N$

Finally the definition is extended to all sentences $\varphi \in L(N)$ by induction in the usual way

3* $\quad N/E \overset{*}{\vDash} \neg\varphi \quad \Leftrightarrow \quad$ not $N/E \overset{*}{\vDash} \varphi$

4* $\quad N/E \overset{*}{\vDash} \varphi \wedge \psi \quad \Leftrightarrow \quad N/E \overset{*}{\vDash} \varphi$ and $N/E \overset{*}{\vDash} \psi$

5* $\quad N/E \overset{*}{\vDash} \exists x\, \varphi(x) \quad \Leftrightarrow \quad N/E \overset{*}{\vDash} \varphi(a) \quad$ for some $a \in N$.

Now, by induction on the syntax of formulas one can prove $\overset{*}{\vDash}$ does what required. In particular, $N/E \overset{*}{\vDash} \varphi(a) \leftrightarrow \varphi(b)$ for every $a \; E \; b$.

50 Proposition Let E be a congruence relation of N. Then the following are equivalent for every $\varphi(x) \in L$

1. $\quad N/E \quad \overset{*}{\vDash} \quad \varphi(a)$;
2. $\quad N/E \quad \vDash \quad \varphi([a]_E)$.

2.6 Completeness

A theory T is maximally consistent if it is consistent and there is no consistent theory S such that $T \subset S$. Equivalently, T contains every sentence φ consistent with T, that is, such that $T \cup \{\varphi\}$ is consistent. Clearly a maximally consistent theory is closed under logical consequences.

A theory T is complete if $\mathrm{ccl}T$ is maximally consistent. Concrete examples will be given in the next chapters as it is not easy to prove that a theory is complete.

51 Proposition The following are equivalent

a. T is maximally consistent;

b. $T = \mathrm{Th}(M)$ for some structure M;

c. T is consistent and $\varphi \in T$ or $\neg\varphi \in T$ for every sentence φ.

Proof. To prove a⇒b, assume that T is consistent. Then there is $M \vDash T$. Therefore $T \subseteq \mathrm{Th}(M)$. As T is maximally consistent $T = \mathrm{Th}(M)$. Implication b⇒c is immediate. As for c⇒a note that if $T \cup \{\varphi\}$ is consistent then $\neg\varphi \notin T$ therefore $\varphi \in T$ follows from c. □

The proof of the proposition below is is left as an exercise for the reader.

52 Proposition The following are equivalent

a. T is complete;

b. there is a unique maximally consistent theory S such that $T \subseteq S$;

c. T is consistent and $T \vdash \text{Th}(M)$ for every $M \vDash T$;

d. T is consistent and either $T \vdash \varphi$ or $T \vdash \neg\varphi$ for every sentence φ;

e. T is consistent and $M \equiv N$ for every pair of models of T.

53 Exercise Prove that the following are equivalent

a. T is complete;

b. for every sentence φ, o $T \vdash \varphi$ o $T \vdash \neg\varphi$ but not both.

By contrast prove that the following are *not* equivalent

a. T is maximally consistent;

b. for every sentence φ, o $\varphi \in T$ o $\neg\varphi \in T$ but not both.

Hint: consider the theory containing all sentences where the symbol $\neg$ occurs an even number of times. This theory is not consistent as it contains $\perp$.

54 Exercise Prove that if T has exactly 2 maximally consistent extension T_1 and T_2 then there is a sentence φ such that $T, \varphi \vdash T_1$ and $T, \neg\varphi \vdash T_2$. State and prove the generalization to finitely many maximally consistent extensions.

2.7 The Tarski-Vaught test

There is no natural notion of *smallest* elementary substructure containing a set of parameters A. The downward Löwenheim-Skolem, which we prove in the next section, is the best result that holds in full generality. Given an arbitrary $A \subseteq N$ we shall construct a model $M \preceq N$ containing A that is small in the sense of cardinality. The construction selects one by one the elements of M that are required to realise the condition $M \preceq N$. Unfortunately, Definition 13 supposes full knowledge of the truth in M and it may not be applied during the construction. The following lemma comes to our rescue with a property equivalent to $M \preceq N$ that only mention the truth in N.

55 Lemma (Tarski-Vaught test) For every $A \subseteq N$ the following are equivalent

1. A is the domain of a structure $M \preceq N$;
2. for every formula $\varphi(x) \in L(A)$, with $|x| = 1$,
   $N \vDash \exists x\, \varphi(x) \quad \Rightarrow \quad N \vDash \varphi(b)$ for some $b \in A$.

Proof. 1⇒2

$$\begin{aligned} N \vDash \exists x\, \varphi(x) \quad &\Rightarrow \quad M \vDash \exists x\, \varphi(x) \\ &\Rightarrow \quad M \vDash \varphi(b) \qquad \text{for some } b \in M \\ &\Rightarrow \quad N \vDash \varphi(b) \qquad \text{for some } b \in A. \end{aligned}$$

2⇒1 Firstly, note that A is the domain of a substructure of N, that is, $f^N a \in A$ for every $f \in L_{\text{fun}}$ and every $a \in A^{n_f}$. In fact, this follows from 2 with $fa = x$ for $\varphi(x)$.

Write M for the substructure of N with domain A. By induction on the syntax we prove that for every $\xi(x) \in L$

$$M \vDash \xi(a) \;\;\Leftrightarrow\;\; N \vDash \xi(a) \qquad \text{for every } a \in M^{|x|}.$$

If $\xi(x)$ is atomic the claim follows from $M \subseteq N$ and the remarks underneath Definition 36. The case of Boolean connectives is straightforward, so only the existential quantifier requires a proof. So, let $\xi(x)$ be the formula $\exists y\, \psi(x, y)$ and assume the induction hypothesis holds for $\psi(x, y)$

$$\begin{aligned} M \vDash \exists y\, \psi(a, y) \;\;&\Leftrightarrow\;\; M \vDash \psi(a, b) && \text{for some } b \in M \\ &\Leftrightarrow\;\; N \vDash \psi(a, b) && \text{for some } b \in M \\ &\Leftrightarrow\;\; N \vDash \exists y\, \psi(a, y). \end{aligned}$$

The second equivalence holds by induction hypothesis, in the last equivalence we use 2 for the implication $\Leftarrow$. □

2.8 Downward Löwenheim-Skolem

The main theorem of this section was proved by Löwenheim at the beginning of the last century. Skolem gave a simpler proof immediately afterwards. At the time, the result was perceived as paradoxical.

A few years earlier, Zermelo and Fraenkel provided a formalization of set theory in a first order language. The downward Löwenheim-Skolem theorem implies the existence of an infinite countable model M of set theory: this is the so-called Skolem paradox. The existence of M seems paradoxical because, in particular, a sentence that formalises the axiom of power set holds in M. Therefore M contains an element b which, in M, is the set of subsets of the natural numbers. But the set of elements of b is a subset of M, and therefore it is countable.

In fact, this is not a contradiction, because the expression *all subsets of the natural numbers* does not have the same meaning in M as it has in the real world. The notion of cardinality, too, acquires a different meaning. In the language of set theory, there is a first order sentence that formalises the fact that b is uncountable: the sentence says that there is no bijection between b and the natural numbers. Therefore the bijection between the elements of b and the natural numbers (which exists in the real world) does not belong to M. The notion of equinumerosity has a different meaning in M and in the real world, but those who live in M cannot realise this.

56 Theorem (Downward Löwenheim-Skolem) Let N be an infinite structure and fix some set $A \subseteq N$. Then there is a structure M of cardinality $\leq |L(A)|$ such that $A \subseteq M \preceq N$.

Proof. Set $\lambda = |L(A)|$. Below we construct a chain $\langle A_i : i < \omega \rangle$ of subsets of N. The chain begins at $A_0 = A$. Finally we set $M = \bigcup_{i<\omega} A_i$. All A_i will have cardinality $\leq \lambda$ so $|M| \leq \lambda$ follows.

Now we construct A_{i+1} given A_i. Assume as induction hypothesis that $|A_i| \leq \lambda$. Then $|L(A_i)| \leq \lambda$. For some fixed variable x let $\langle \varphi_k(x) \;:\; k < \lambda \rangle$ be an enumeration of the formulas in $L(A_i)$ that are consistent in N. For every k pick $a_k \in N$ such that $N \vDash \varphi_k(a_k)$. Define $A_{i+1} = A_i \cup \{a_k : k < \lambda\}$. Then $|A_{i+1}| \leq \lambda$ is clear.

We use the Tarski-Vaught test to prove $M \preceq N$. Suppose $\varphi(x) \in L(M)$ is consistent in N. As finitely many parameters occur in formulas, $\varphi(x) \in L(A_i)$ for some i. Then $\varphi(x)$ is among the formulas we enumerated at stage i and $A_{i+1} \subseteq M$ contains a solution of $\varphi(x)$. □

We will need to adapt the construction above to meet more requirements on the model M. To better control the elements that end up in M it is convenient to add one element at the time (above we add λ elements at each stage). We need to enumerate formulas with care if we want to complete the construction by stage λ.

Second proof of the downward Löwenheim-Skolem Theorem. From set theory we know that there is a bijection $\pi : \lambda^2 \to \lambda$ such that $j, k \leq \pi(j,k)$ for all $j, k < \lambda$. Suppose we have defined the sets A_j for every $j \leq i$ and let $\langle \varphi_k^j(x) : k < \lambda \rangle$ be an enumeration of the consistent formulas of $L(A_j)$. Let $j, k \leq i$ be such that $\pi(j,k) = i$. Let b be a solution of the formula $\varphi_k^j(x)$ and define $A_{i+1} = A_i \cup \{b\}$.

We use Tarski-Vaught test to prove $M \preceq N$. Let $\varphi(x) \in L(M)$ be consistent in N. Then $\varphi(x) \in L(A_j)$ for some j. Then $\varphi(x) = \varphi_k^j$ for some k. Hence a witness of $\varphi(x)$ is enumerated in M at stage $\pi(j,k) + 1$. □

57 Exercise Assume L is countable and let $M \preceq N$ have arbitrary (large) cardinality. Let $A \subseteq N$ be countable. Prove there is a countable model K such that $A \subseteq K \preceq N$ and $K \cap M \preceq N$ (in particular, $K \cap M$ is a model). Hint: adapt the construction used to prove the downward Löwenheim-Skolem Theorem.

58 Exercise The language contains only $<$, the relation of strict order. Prove that there is a countable ordinal α that is an elemenrary substructure of ω_1, the first uncountable ordinal.

2.9 Elementary chains

An elementary chain is a chain $\langle M_i : i < \lambda \rangle$ of structures such that $M_i \preceq M_j$ for every $i < j < \lambda$. The union (or limit) of the chain is the structure with as domain the set $\bigcup_{i<\lambda} M_i$ and as relations and functions the union of the relations and functions of M_i. It is plain that all structures in the chain are substructures of the limit.

59 Lemma Let $\langle M_i : i \in \lambda \rangle$ be an elementary chain of structures. Let N be the union of the chain. Then $M_i \preceq N$ for every i.

Proof. By induction on the syntax of $\varphi(x) \in L$ we prove

$$M_i \vDash \varphi(a) \;\Leftrightarrow\; N \vDash \varphi(a) \qquad \text{for every } i < \lambda \text{ and every } a \in M_i^{|x|}$$

As remarked in 3' of Section 2.4, the claim holds for quantifier-free formulas. Induction for Boolean connectives is straightforward so we only need to consider the existential quantifier

$$\begin{aligned} M_i \vDash \exists y\, \varphi(a,y) \;&\Rightarrow\; M_i \vDash \varphi(a,b) \qquad &&\text{for some } b \in M_i. \\ &\Rightarrow\; N \vDash \varphi(a,b) \qquad &&\text{for some } b \in M_i \subseteq N \end{aligned}$$

where the second implication follows from the induction hypothesis. Vice versa

$$N \vDash \exists y\, \varphi(a,y) \ \Rightarrow\ N \vDash \varphi(a,b) \qquad \text{for some } b \in N$$

Without loss of generality we can assume that $b \in M_j$ for some $j \geq i$ and obtain

$$\Rightarrow\ M_j \vDash \varphi(a,b) \qquad \text{for some } b \in M_j$$

Now apply the induction hypothesis to $\varphi(x,y)$ and M_j

$$\Rightarrow\ M_j \vDash \exists y\, \varphi(a,y)$$

$$\Rightarrow\ M_i \vDash \exists y\, \varphi(a,y)$$

where the last implication holds because $M_i \preceq M_j$. □

60 Exercise Let $\langle M_i : i \in \lambda \rangle$ be an chain of elementary substructures of N. Let M be the union of the chain. Prove that $M \preceq N$ and note that Lemma 59 is not required.

61 Exercise Give an alternative proof of Exercise 57 using the downward Löwenheim-Skolem Theorem (instead of its proof). Hint: construct two countable chains of countable models such that $K_i \cap M \subseteq M_i \preceq N$ and $A \cup M_i \subseteq K_{i+1} \preceq N$. The required model is $K = \bigcup_{i \in \omega} K_i$. In fact it is easy to check that $K \cap M = \bigcup_{i \in \omega} M_i$.

Chapter 3

Ultraproducts

In these notes we only use ultraproducts to prove the compactness theorem. Since a syntactic proof of the compactness theorem is also given, this chapter is, strictly speaking, not required. However, the importance of ultraproducts transcends its application to model theory.

3.1 Filters and ultrafilters

The material in this section will appear again in a more general setting in Chapter **??**.

Let I be any set. A filter on I (or a filter of $\mathcal{P}(I)$), is a non empty set $F \subseteq \mathcal{P}(I)$ such that for every $a, b \in \mathcal{P}(I)$

f1 $\qquad a \in F$ and $a \subseteq b \quad\Rightarrow\quad b \in F$

f2 $\qquad a \in F$ and $b \in F \quad\Rightarrow\quad a \cap b \in F$

A filter F is proper if $F \neq \mathcal{P}(I)$, equivalently if $\varnothing \notin F$. Otherwise it is improper. By f1 above, F is proper if and only if $\varnothing \notin F$. A filter F is principal if $F = \{a \subseteq I \ : \ b \subseteq a\}$ for some set $b \subseteq I$. When this happens, we say that $\{b\}$ generates F. When I is finite every filter F is principal and generated by $\{\cap F\}$. Non-principal filters on I exist as soon as I is infinite. In fact, if I is infinite it is easy to check that the following is a filter

$$F \quad = \quad \Big\{ a \subseteq I \ : \ a \ \text{ is cofinite in } \ I \Big\}.$$

where cofinite (in I) means that $I \smallsetminus a$ is finite. This is called the Fréchet filter on I. The Fréchet filter is the minimal non-principal filter.

A proper filter F is maximal if there is no proper filter H such that $F \subset H$. A proper filter F is an ultrafilter if for every $a \subseteq I$

$$a \notin F \quad\Rightarrow\quad \neg a \in F \qquad\qquad\qquad \text{where } \neg a \ = \ I \smallsetminus a\,.$$

Below we prove that the ultrafilters are exactly the maximal filters.

Let $B \subseteq \mathcal{P}(I)$. The filter generated by B is the intersection of the filters that contain B. It is easy to check that the intersection of a family of filters is a filter, so the notion is well-defined. The following easy proposition gives a workable characterization of the filter generated by a set.

1 Proposition The filter generated by B is $\Big\{ a \subseteq I \ : \ \bigcap C \subseteq a \ \text{ for some finite } \ C \subseteq B \Big\}$.

We say that B has the finite intersection property if $\bigcap C \neq \varnothing$ for every finite $C \subseteq B$. The following proposition is immediate.

2 Proposition The following are equivalent:

1. the filter generated by B is non principal;
2. B has the finite intersection property.

A proper filter F is prime if for every $a, b \subseteq I$.

$$a \cup b \in F \quad \Rightarrow \quad a \in F \quad \text{or} \quad b \in F$$

Prime filters coincide with maximal filters. However, in Chapter **??**, we introduce a more general context where primality is distinct from maximality.

3 Proposition For every filter F on I, the following are equivalent:

1. F is a maximal filter;
2. F is a prime filter;
3. F is an ultrafilter.

Proof. 1⇒2. Suppose $a, b \notin F$, where F is maximal. We claim that $a \cup b \notin F$. By maximality, there is a $c \in F$ such that $a \cap c = b \cap c = \varnothing$. Therefore $(a \cup b) \cap c = \varnothing$. Hence $a \cup b \notin F$.

2⇒3. It suffices to note that $I = a \cup \neg a \in F$.

3⇒1. If $a \notin F$ then $\neg a \in F$, hence no proper filter contains $F \cup \{a\}$. □

4 Proposition Let $B \subseteq \mathcal{P}(I)$ have the finite intersection property. Then B is contained in a maximal filter.

Proof. First we prove that the union of a chain of subsets of $\mathcal{P}(I)$ with the finite intersection property has the finite intersection property. Let $\mathcal{B}$ be such a chain and suppose for a contradiction that $\bigcup \mathcal{B}$ does not have the finite intersection property. Fix a finite $C \subseteq \bigcap \mathcal{B}$ such that $\bigcap C = \varnothing$. As C is finite, we have $C \subseteq B$ for some $B \in \mathcal{B}$. Hence B does not have the finite intersection property, which is a contradiction.

Now apply Zorn's lemma to obtain a $B \subseteq \mathcal{P}(I)$ which is maximal among the sets with the finite intersection property. It is immediate that B is a filter. □

5 Exercise Let I be infinite. Prove that every non-principal ultrafilter on I contains Fréchet's filter. Show that this does not hold for plain filters.

6 Exercise Prove that all principal ultrafilters are generated by a singleton.

3.2 Direct products

In this and in the next section $\langle M_i : i \in I \rangle$ is a sequence of L-structures. (We are abusing of the word sequence, since I is only a set.) The direct product of this sequence is a structure denoted by

$$N \quad = \quad \prod_{i \in I} M_i$$

and defined by conditions 1-3 below. If $M_i = M$ for all $i \in I$, we say that N is a direct power of M and denote it by M^I.

The domain of N is the set containing all functions

1. $$\hat{a} \ : \ I \ \to \ \bigcup_{i\in I} M_i$$
$$\hat{a} \ : \ i \ \mapsto \ \hat{a}i \ \in \ M_i$$

We do not distinguish between tuples of elements of N and tuple-valued functions. For instance, the tuple $\hat{a} = \langle \hat{a}_1 \dots \hat{a}_n \rangle$ is identified with the function $\hat{a} : i \mapsto \hat{a}i = \langle \hat{a}_1 i, \dots, \hat{a}_n i \rangle$. On a first reading of what follows, it may help to pretend that all functions and relations are unary.

The interpretation of $f \in L_{\text{fun}}$ is defined as follows:

2. $$(f^N \hat{a})i \ = \ f^{M_i}(\hat{a}i) \qquad \text{for all } i \in I.$$

The interpretation of $r \in L_{\text{rel}}$ is the product of the relations r^{M_i}, that is, we define

3. $$\hat{a} \in r^N \ \Leftrightarrow \ \hat{a}i \in r^{M_i} \qquad \text{for all } i \in I.$$

The following proposition is immediate.

7 Proposition If $=, \wedge, \forall, \exists$ are the only logical symbol that occur in $\varphi(x) \in L$, then for every $\hat{a} \in N^{|x|}$

$\sharp$ $$N \vDash \varphi(\hat{a}) \ \Leftrightarrow \ M_i \vDash \varphi(\hat{a}i) \qquad \text{for all } i \in I.$$

Proof. By induction on syntax. First note that we can extend 2 to all terms $t(x)$ as follows:

2′. $$(t^N \hat{a})i \ = \ t^{M_i}(\hat{a}i) \qquad \text{for all } i \in I.$$

Combining 3 and 2′ gives that for every $r \in L_{\text{rel}} \cup \{=\}$ and every L-term $t(x)$

$$N \vDash rt\hat{a} \ \Leftrightarrow \ M_i \vDash rt\hat{a}i \qquad \text{for all } i \in I.$$

This shows that $\sharp$ holds for $\varphi(x)$ atomic. Induction for the connectives $\wedge$, $\forall$ ed $\exists$ is immediate. □

A consequence of Proposition 7 is that a direct product of groups, rings or vector spaces is a structure of the same sort. However, a product of fields is not a field.

3.3 Łoś's Theorem

Let $\langle M_i : i \in I \rangle$ be a sequence of structures. We (temporary) denote by N' the following structure. The domain and the interpretation of the function symbols is as in the structure N defined in 1, respectively 2, of the previous section. As interpretation of $r \in L_{\text{rel}}$ we take

3′. $$\hat{a} \in r^{N'} \ \Leftrightarrow \ \big\{ i \in I \ : \ \hat{a}i \in r^{M_i} \big\} \in F.$$

Let F be a filter on I. We define the following congruence on N' (see Definition 47)

$$\hat{a} \sim_F \hat{c} \ \Leftrightarrow \ \big\{ i \in I \ : \ \hat{a}i = \hat{c}i \big\} \in F.$$

To check that $\sim_F$ is indeed a congruence, we first need to check that it is an equivalence relation. Reflexivity and symmetry are immediate, and transitivity follows from f2 in Section 3.1. Then we check that $\sim_F$ is compatible with the functions of L, that is, that c1 of Definition 47 is satisfied. This follows from f1 in Section 3.1.

For brevity, we write N'/F for $N'/{\sim_F}$ and $[\hat{a}]_F$ for $[\hat{a}]_{\sim_F}$. We write $N'/F \overset{*}{\vDash} \varphi(\hat{a})$ for $N'/F \vDash \varphi([\hat{a}]_F)$. A more formal interpretation of $\overset{*}{\vDash}$ is given in Definition 49.

The structure N'/F is called the reduced product of the structures $\langle M_i : i \in I \rangle$ or, when $M_i = M$ for all $i \in I$, the reduced power of M. When F is an ultrafilter we say ultraproduct, respectively ultrapower.

The difference between N and N' is highlighted in Exercise 12. For neater notation, below we write N for N'.

The following proposition is almost tautological. Note that it is a special case of Łoš's Theorem below (but it is does not require F to be an *ultra*filter).

8 Proposition Let $r \in L_{\text{rel}}$. Let $t(x)$ be tuple of terms of length n_r, the arity of r. Then for every $\hat{a} \in N^{|x|}$ and every filter F the following are equivalent

1. $N/F \overset{*}{\vDash} r\, t(\hat{a})$;
2. $\{i \ :\ M_i \vDash r\, t(\hat{a}i)\} \ \in\ F$.

Proof. 1⇒2 Assume 1. Then, by 2* of Definition 49, there is a $\hat{b} \sim t^N(\hat{a})$ such that $N \vDash r\, \hat{b}$. Therefore $\{i : M_i \vDash r\, \hat{b}i\}$ is in F. But $\hat{b} \sim t(\hat{a})$ means that $\{i \ :\ \hat{b}i = t(\hat{a}i)\}$ is in F, hence 2 follows.

2⇒1 Is clear. □

9 Łoś's Theorem Let $\varphi(x) \in L$ and let F be an ultrafilter on I. Then for every $\hat{a} \in N^{|x|}$ the following are equivalent:

1. $N/F \ \overset{*}{\vDash}\ \varphi(\hat{a})$ (see Definition 49);
2. $\{i \ :\ M_i \vDash \varphi(\hat{a}i)\} \ \in\ F$.

Proof. We proceed by induction on the syntax of $\varphi(x)$. If $\varphi(x)$ is equality, then equivalence holds by definition of $\sim$. If $\varphi(x)$ is of the form $r\, t(x)$ for some tuple of terms $t(x)$ and $r \in L_{\text{rel}}$ then 1⇔2 is Proposition 8.

We prove the inductive step for the connectives $\neg$, $\wedge$, and the quantifier $\exists$. We begin with $\neg$. This is the only place in the proof where the assumption that F is an *ultra*filter is required. By the inductive hypothesis,

$$N/F \ \overset{*}{\vDash}\ \neg\varphi(\hat{a}) \quad\Leftrightarrow\quad \{i \ :\ M_i \ \vDash\ \varphi(\hat{a}i)\} \ \notin\ F$$

So, as F is an *ultra*filter

$$\Leftrightarrow\quad \{i \ :\ M_i \ \vDash\ \neg\varphi(\hat{a}i)\} \ \in\ F\,.$$

Now consider $\wedge$. Assume inductively that the equivalence 1⇔2 holds for $\varphi(x)$ and $\psi(x)$. Then

$$N/F \ \overset{*}{\vDash}\ \varphi(\hat{a}) \wedge \psi(\hat{a}) \quad\Leftrightarrow\quad \{i \ :\ M_i \vDash \varphi(\hat{a}i)\} \ \in\ F \ \text{ and } \ \{i \ :\ M_i \vDash \psi(\hat{a}i)\} \ \in\ F\,.$$

As filters are closed under intersection, we obtain

$$\Leftrightarrow\quad \{i \ :\ M_i \ \vDash\ \varphi(\hat{a}i) \wedge \psi(\hat{a}i)\} \ \in\ F\,.$$

Finally, consider $\exists y$. Assume inductively that the equivalence 1⇔2 holds for $\varphi(x, y)$. Then

$$N/F \ \overset{*}{\vDash}\ \exists y\, \varphi(\hat{a}, y) \quad\Leftrightarrow\quad N/F \ \overset{*}{\vDash}\ \varphi(\hat{a}, \hat{b}) \qquad\qquad \text{for some } \hat{b} \in N$$

$$\Leftrightarrow \quad \{\, i \;:\; M_i \vDash \varphi(\hat{a}i, \hat{b}i) \,\} \;\in\; F \qquad \text{for some } \hat{b} \in N.$$

We claim this is equivalent to

$$\Leftrightarrow \quad \{\, i \;:\; M_i \vDash \exists y\, \varphi(\hat{a}i, y) \,\} \;\in\; F.$$

The $\Rightarrow$ direction is trivial. For $\Leftarrow$, we choose as $\hat{b}$ a sequence that picks a witness of $M_i \vDash \exists y\, \varphi(\hat{a}i, y)$ if it exists, and some arbitrary element of M_i otherwise. □

Let a^I denote the element of M^I that has constant value a. The following is an immediate consequence of Łoś's theorem.

10 Corollary Let $N = M^I$. For every $a \in M$

$$N/F \overset{*}{\vDash} \varphi(a^I) \quad \Leftrightarrow \quad M \vDash \varphi(a)\,.$$

We often identify M with its image under the embedding $h : a \mapsto [a^I]_F$, and say that M^I/F is an elementary extension of M.

The following corollary is an immediate consequence of the compactness theorem that we prove in the next chapter but here we give a direct proof.

11 Corollary Every infinite structure has a proper elementary extension.

Proof. Let M be an infinite structure and let F be a *non-principal* ultrafilter on ω. It suffices to show that $h[M]$, the image of the embedding defined above, is a *proper* substructure of M^ω/F. As M is infinite, there is an injective function $\hat{d} \in M^\omega$. Then for every $a \in M$ the set $\{i : \hat{d}\, i = a\}$ is either empty or a singleton and, as F is non principal, it does not belong to F. So, by Łoš Theorem, we have $M^\omega/F \overset{*}{\vDash} \hat{d} \neq a^I$ for every $a \in M$, that is, $[\hat{d}\,]_F \notin h[M]$. □

12 Exercise Let N be the structure defined in the previous section. Prove that if $M_i = M$ for all $i \in I$ then $N'/F = N/F$. More generally, prove that if for every $r \in L_{\text{rel}}$, either $r^{M_i} \neq \varnothing$ for all i or $r^{M_i} = \varnothing$ for all i then $N'/F = N/F$.

13 Exercise Consider $\mathbb{N}$ as a structure in the language of strict orders. Let F be a non-principal ultrafilter on ω. Prove that in $\mathbb{N}^\omega$ there is a sequence $\langle \hat{a}_i : i \in \omega \rangle$ such that $\mathbb{N}^\omega/F \overset{*}{\vDash} \hat{a}_{i+1} < \hat{a}_i$.

14 Exercise Let I be the set of integers $i > 1$. For $i \in I$, let $\mathbb{Z}_i$ denote the additive group of integers modulo i, and let N denote the product $\prod_{i \in I} \mathbb{Z}_i$. Prove that, if F is a non-principal ultrafilter on I,

1. for some F, N/F does not contain any element of finite order;
2. for some F, N/F has some elements of order 2;
3. for all F, N/F contains an element of infinite order;
4. for some F, $N/F \vDash \forall x\, \exists y\; my = x$ for every integer $m > 0$;
5. for some F, N/F contains an element $\hat{a}$ such that $N/F \vDash \forall x\; mx \neq \hat{a}$ for every positive integer m.

Programma d'esame

Capitolo 1

Esercizio 15

Capitolo 2

Esercizi 17, 18, 19, 20

Lemma 22, Proposition 23, Proposition 25

Esercizi 27, 30, 31, 32, 33, 34, 35

Esecizio 54

Esercizi 44, 45

Lemma 55

Theorem 56

Esercizi 57, 58