Chapter III

The compactness theorem

We will introduce in this and the next chapter two crucial concepts: the "compactness theorem" and the notion of "quantifier elimination", which are among the most important tools and methods of model theory.

1 - The compactness theorem

Let $\mathcal{L}$ be a language, and let T be an $\mathcal{L}$-theory. A basic question is whether T has a model or not (the theory $\exists x \exists y\ (x=y) \wedge \neg(x=y)$ has no model for instance). Although this is a hard question in general, at least it is only a question of finite theories: this is a striking feature of the use of first order logic.

Theorem ("Compactness theorem"). $\mathcal{L}$ language, T an $\mathcal{L}$-theory. There exist a model of T if and only if, for any finite subset $T_0 \subset T$, there exists a model of T_0. (One says that T is finitely-satisfiable)

Remark. Why "compactness"? It is clearly reminiscent of the criterion

" if $(C_i)_{i\in I}$ is a family of closed sets in X s.t. $\forall J \subset I$, J finite, $\bigcap_{j\in J} C_j \neq \emptyset$, then $\bigcap_{i\in I} C_i \neq \emptyset$"

for compactness of a topological space X.

In fact, there is a proof where the compactness theorem is literally saying that a certain space ("the space of types") is compact.

Example. (1) Let $\mathcal{L}_{gp} = (\cdot, e)$ be the language of groups and T_{gp} the $\mathcal{L}_{gp}$-theory of groups.

Claim. There is no $T \supset T_{gp}$ such that a group G is a model of T if and only if

G is a torsion group.

Proof- Assume such a T exists. Let $\mathcal{L}^\circ = \mathcal{L}_{gp} \cup \{\underline{c}\}$ where $\underline{c}$ is a new constant symbol. Any $\mathcal{L}_{gp}$-formula is an $\mathcal{L}^\circ$-formula. For $n \geq 1$, let ϕ_n be the $\mathcal{L}^\circ$-sentence "$\neg c^n = e$", and let T° be the $\mathcal{L}^\circ$-theory $T \cup \{\phi_n \mid n \geq 1\}$.

Then T° is finitely satisfiable: if $S \subset T^\circ$ is finite, then $S \cap \{\phi_n \mid n \geq 1\} = \{\phi_i \mid i \in I\}$ for some finite set I of integers. Let

$$G = \mathbb{Z}/(\max(I)+1)\mathbb{Z}.$$

Then $G \models T$ (it is a torsion group) and if we view G as an $\mathcal{L}^\circ$-structure by interpreting $\underline{c}$ as $1 \in G$, then $G \models \phi_n$ for $n \leq \max(I)$.

By the Compactness Theorem, we conclude that T° has a model G. But then $G \models T$ so G is torsion (by assumption), whereas the element c_G

of G satisfies $c_G^n \neq e$ for all n, which is a contradiction.

(2) Let $\mathcal{L}_{peano} = (\underset{2}{+}, \underset{2}{\cdot}, \underset{1}{s}, \underline{0})$ be the language of <u>Peano arithmetic</u>. The "standard" structure is $\mathbb{N}$ with s interpreted as the function $n \longmapsto n+1$ and $+, \cdot, 0$ as expected.

The sentences $\phi_1, \ldots, \phi_6, \{\mathrm{Ind}(\psi)\}$ below define the theory T_{peano} of Peano arithmetic:

$$
\begin{aligned}
\phi_1 &: \quad \forall x,\ s(x) \neq 0 \\
\phi_2 &: \quad \forall x,\ (x \neq 0 \longrightarrow \exists y,\ s(y) = x) \\
\phi_3 &: \quad \forall x,\ x + 0 = x \\
\phi_4 &: \quad \forall x, \forall y,\ x + s(y) = s(x+y) \\
\phi_5 &: \quad \forall x,\ x \cdot 0 = 0 \\
\phi_6 &: \quad \forall x, \forall y,\ x \cdot s(y) = x \cdot y + x
\end{aligned}
$$

$$
\mathrm{Ind}(\phi): \quad \forall w_1 \ldots \forall w_k, \Big[\phi(0,(w_i)) \wedge \big(\forall v,\ \phi(v,(w_i)) \longrightarrow \phi(s(v),(w_i))\big) \longrightarrow \forall x, \phi(x,(w_i))\Big]
$$

for all formulas $\phi(v_1, w_1, \ldots, w_k)$, $k \geq 1$.

Note that $\mathbb{N} \models T_{peano}$. However, there are many more models!

In particular:

<u>Fact:</u> (1) one can define an order in a model of T_{peano}: let $\phi(x,y): \exists z,\ x+z=y$.
Then for any model M of T_{peano}, the set

$$\phi(M) \subset M^2$$

"is" an order relation, and $\phi(\mathbb{N})$ is the usual order.

(2) there is a model M of T_{peano} and an $a \in M$ such that for every <u>integer</u> $n \in \mathbb{N}$

$$M \models \big(\underbrace{s(s(s(\cdots s(0))\cdots))}_{n \text{ times } s} \leq a \big)$$

Here, (1) is elementary, and (2) follows from the Compactness Theorem: consider the new language $\mathcal{L}^\circ = \mathcal{L}_{peano} \cup \{c\}$, c being a new constant, and the $\mathcal{L}^\circ$-theory

$$T^\circ = T_{peano} \cup \{ \underbrace{s(s(\cdots(s(0)\cdots)) \leq c}_{n \text{ times, and really to be replaced by } \text{"}\exists z,\ x+z=c\text{"}} \}$$

Then as one can guess, we check easily that T° is finitely-satisfiable, hence satisfiable, and we get what we want from a model M of T° and the element $a = \underline{c}_M \in M$. (One says that such an M is a "nonstandard" model of arithmetic.)

2. Proof of the Compactness Theorem, I

There are quite a few proofs of the theorem. I will sketch one which is logically interesting, and give the full proof using ultrafilters, which is maybe the most useful and enlightening for non-logicians.

(Note: both proofs involve some form of the Axiom of Choice, though not the strongest.)

Sketch of proof 1. This hinges on a key result of first order logic:

Theorem [Gödel: the Completeness Theorem]

$\mathcal{L}$ language, T an $\mathcal{L}$-theory, ϕ an $\mathcal{L}$-sentence. Then $T \models \phi$ (i.e, ϕ holds in all models of T; one says also that ϕ is a logical consequence of T) if and only if ϕ can be proved from T (i.e. there is a finite sequence of formulas $\psi_1, \ldots, \psi_m$, ending at $\psi_m = \phi$, where ψ_{i+1} is either taken from T or is a "deduction" from the previous ψ_j's using rules like "from ψ_j, ψ_k one can deduce $\psi_j \wedge \psi_k$"...).

Whatever rigorous definition is used, it is not hard to show that a proof of ϕ from T implies that $T \models \phi$, and the point (which uses some version of choice) is the converse.

Now, assuming the Completeness Theorem, we can prove Compactness as follows: suppose that

T is not satisfiable. Then it follows that

$$T \models (\phi \wedge \neg \phi)$$

for ϕ any $\mathcal{L}$-sentence. By Completeness, there is a proof of $\phi \wedge \neg \phi$ from T; now by definition, the proof is finite, so it uses only finitely many sentences from T, say those from $T_0 \subset T$ finite. But then

$$T_0 \models (\phi \wedge \neg \phi)$$

so T_0 is not satisfiable.

3 - Proof of Compactness, II

In the second proof, we will actually construct a model of a finitely satisfiable theory. We do this using ultraproducts.

Definition - X any set

(1) [Filter] A filter $\mathcal{F}$ on X is a set of subsets of X with the following conditions:

(F1) $\emptyset \notin \mathcal{F}$, $\mathcal{F} \neq \emptyset$

(F2) If A, B are in $\mathcal{F}$, so is $A \cap B$

(F3) If $A \in \mathcal{F}$, $B \supset A$, then $B \in \mathcal{F}$

[Ex. (1) X topological space, $x \in X$, $\mathcal{F}$

the set of neigorhoods of x in X

(2) $X = \mathbb{N}$, $\mathcal{F} = \{$ complements of $I \subset \mathbb{N}$ finite $\}$

(2) [Ultrafilter] A filter $\mathcal{F}$ on X is an ultra filter if for all $A \subset X$, either A or its complement is in $\mathcal{F}$ (not both because of (F1)).

If there exists $x \in X$ such that $\mathcal{F} = \{A \mid x \in A\}$ then $\mathcal{F}$ is a principal ultrafilter.

Remark. (1) A filter $\mathcal{F}$ is a principal ultrafilter if and only if it contains a set $\{x\}$ with a single element.

(2) If $\mathcal{F}_0$ is a set of subsets of X satisfying

just (F1), (F2) then

$$\mathcal{F} = \{ A \subset X \mid A \supset B_1 \cap \dots \cap B_m \text{ for some } m \geq 1, \ B_i \in \mathcal{F}_0 \}$$

is a filter, called the filter <u>generated</u> by $\mathcal{F}_0$.

(3) Originally, filters were invented by H. Cartan to define the most general form of limit in topo-logy: given a filter $\mathcal{F}$ on a topological space X, an element $x \in X$ is a limit of $\mathcal{F}$ if and only if $\mathcal{F} \supset \{\text{neighborhoods of } x\}$.

<u>Example</u>. $X = \mathbb{R}$, $(x_n)_{n \in \mathbb{N}}$ sequence in $\mathbb{R}$, $\mathcal{F} = \{ A \subset \mathbb{R} \mid A \text{ contains all } x_n \text{ except finitely many} \}$

Then is a filter on $\mathbb{R}$ and $\mathcal{F}$ has limit $x \in \mathbb{R}$ if and only if $\lim\limits_{n \to \infty} x_n = x$. (Indeed, for all $\varepsilon > 0$, we must have $]x - \varepsilon, x + \varepsilon[\ \in \mathcal{F}$, etc...)

(4) In particular, one can prove: X is compact $\iff$ every ultrafilter on X converges (and X is

Hausdorff).

Of special importance is the following result:

Proposition – Let $\mathcal{F}$ be a filter on X. There exists an ultrafilter $\hat{\mathcal{F}} \supset \mathcal{F}$; in fact,

$$\mathcal{F} = \bigcap_{\substack{\hat{\mathcal{F}} \supset \mathcal{F} \\ \text{ultrafilter}}} \hat{\mathcal{F}}.$$

Proof – Apply Zorn's Lemma to the set $\mathcal{O}$ of filters $\mathcal{G} \supset \mathcal{F}$, ordered by inclusion (of sets of subsets of X); note that $\mathcal{F} \in \mathcal{O}$ and $\bigcup_{\mathcal{G} \in \mathcal{Y}} \mathcal{G} \in \mathcal{O}$ if $\mathcal{Y} \subset \mathcal{O}$ is linearly-ordered, so there is $\hat{\mathcal{F}} \supset \mathcal{F}$ in $\mathcal{O}$ maximal. If $A \subset X$ but $A \notin \hat{\mathcal{F}}$, then

$$\mathcal{G} = \{ B \subset X \mid A \cup B \in \hat{\mathcal{F}} \}.$$

Then $\mathcal{G}$ is a filter on X (exercise!) and
$\phi \notin \mathcal{G}$ because $A \notin \hat{\mathcal{F}}$

$\hat{\mathcal{F}} \subset \mathcal{G}$ (by (F3): if $B \in \hat{\mathcal{F}}$ then $A \cup B \supset B$ so $A \cup B \in \hat{\mathcal{F}}$); so $\hat{\mathcal{F}} = \mathcal{G}$ by maximality, so $X - A \in \hat{\mathcal{F}}$ (since $(X-A) \cup A = X \in \hat{\mathcal{F}}$, so

$X - A \in \mathcal{G}$).

$\square$

Definition - [Ultraproduct]

$\mathcal{L}$ language, I set, $(M_i)_{i\in I}$ family of $\mathcal{L}$-structures with $M_i \neq \emptyset$.

Let $\mathcal{F}$ be a ultrafilter on I. Then the ultraproduct $\hat{\prod}_{\mathcal{F}} M_i$ is the $\mathcal{L}$-structure defined as follows:

(1) $M = \hat{\prod}_{\mathcal{F}} M_i = \prod_{i\in I} M_i \Big/ \sim$, where

$(m_i) \sim (n_i) \iff \{ i \in I \mid m_i = n_i \} \in \mathcal{F}$

(2) If $\underline{c}$ is a constant symbol then

$\underline{c}_M =$ class of $(\underline{c}_{M_i})_{i\in I}$ modulo $\sim$

(3) If $(\underline{r}, n)$ (resp. $(\underline{f}, n)$) is a relation (resp. function)

then $\underline{r}_M = \{$ classes of $(x_1, \dots, x_n) \in (\prod M_i)^n$

such that $\{ i \in I \mid (x_{1,i}, \dots, x_{n,i}) \in \underline{r}_{M_i} \} \in \mathcal{F} \}$.

(resp. $\underline{f}_M$ (class of $x_1, \dots, x_n$) = class of $(f(x_i))_{i\in I}$)

(Intuitively: $\mathcal{F}$ corresponds to sets containing "almost all" $i \in I$; so a binary relation $\leq$ will hold between $x = (x_i)$ and $y = (y_i)$ if " $x_i \leq y_i$ for almost all i ".)

The key property is the following:

Theorem. (Łoś) Let $\mathcal{L}$, I, $(M_i)_{i \in I}$ and $\mathcal{F}$ be as above. For any $\mathcal{L}$-formula $\phi(x_1, \dots, x_n)$ and $a \in \left(\widehat{\prod_{\mathcal{F}}} M_i\right)^n$, we have $a \in \phi\left(\widehat{\prod_{\mathcal{F}}} M_i\right)$ if and only if $\{ i \in I \mid M_i \models \phi \} \in \mathcal{F}$.

In particular, if T is an $\mathcal{L}$-theory and $M_i \models T$ for all i, then $\widehat{\prod_{\mathcal{F}}} M_i \models \phi$.

Sketch of proof. We leave some fussy details as exercises... Clearly we must proceed by induction on the construction of sentences. Let $M = \widehat{\prod_{\mathcal{F}}} M_i$.

(1) A term $t(v_1, \dots, v_m)$ is interpreted in M as the function

$$t_M : (x_1, \dots, x_m) \longmapsto \text{class of } \left(t_{M_i}(x) \right)_{i \in I}$$

(2) Consider a term "$t_1(v) = t_2(v)$" and $a \in M$.

Write a as the class of $(a_i)_{i \in I}$. Then $t_j(v)$ is (by (1)) the class of $\left(t_{j, M_i}(a_i) \right)$ so

$$M \models \left(t_1(a) = t_2(a) \right)$$

if and only if $\{ i \in I \mid t_{1,M_i}(a_i) = t_{2,M_i}(a_i) \}$ is in $\mathcal{F}$. Similarly for general terms.

(3) For formulas $\underline{R}(t_1, \dots, t_m)$: exercice

(4) If the property holds for φ, ψ then it does for $\varphi \wedge \psi$: for a = class of (a_i)

$$\begin{aligned} M \models (\varphi \wedge \psi)(a) &\overset{\text{definition of } \models}{\iff} M \models \varphi(a) \text{ and } M \models \psi(a) \\ &\overset{\text{"induction"}}{\iff} \{ i \in I \mid M_i \models \varphi(a_i) \} \in \mathcal{F} \text{ and} \\ &\qquad \{ i \in I \mid M_i \models \psi(a_i) \in \mathcal{F} \} \\ &\iff \{ i \in I \mid M_i \models \varphi(a_i) \text{ and } M_i \models \psi(a_i) \} \in \mathcal{F} \end{aligned}$$

(for $\Leftarrow$ observe that if $A \cap B \in \mathcal{F}$ then $A \in \mathcal{F}$ and $B \in \mathcal{F}$)

(5) If the property holds for ϕ, it does for $\neg\phi$: for a = class of (a_i)

$$M \models \neg\phi(a) \iff \{i \in I \mid M_i \models \phi(a_i)\} \notin \mathcal{F}$$

def. of $\models$ + "induction"

$$\iff \{i \in I \mid M_i \models \neg\phi(a_i)\} \in \mathcal{F}$$

$\mathcal{F}$ ultrafiltre

(6) For $\exists x, \phi(x)$:

$$M \models \exists x, \phi(x) \iff \exists\, a = \text{class of } (a_i),\ M \models \phi(a)$$

"induction"

$$\iff \exists (a_i),\ \{i \in I \mid M_i \models \phi(a_i)\} \in \mathcal{F}$$

$$\iff \{i \in I \mid M_i \models \exists x, \phi(x)\} \in \mathcal{F}$$

($\Rightarrow$: by definition; $\Leftarrow$: for i s.t. $M_i \models \exists x, \phi(x)$, pick $a_i \in \phi(M_i)$; for the others, pick $a_i \in M_i$ arbitrarily, using the assumption $M_i \neq \emptyset$; then $(a_i) \in \prod M_i$ has the desired property.)

$\square$

Remark. Suppose $\mathcal{F}$ is principal: $\mathcal{F} = \{ I \mid i_0 \in I\}$ for some $i_0 \in I$. Then the map $(x_i) \xmapsto{\pi} x_{i_0}$ gives $\widehat{\prod_{\mathcal{F}}} M_i \xrightarrow{\pi} M_{i_0}$, and this is a <u>bijection</u> (with inverse $x \longmapsto (x_i)$ where $x_{i_0} = x$ and x_j = anything in M_j for $j \neq i_0$), so the ultraproduct is just M_{i_0}.

However, when $\mathcal{F}$ is <u>not</u> principal, then $\widehat{\prod_{\mathcal{F}}} M_i$ can be very interesting.

<u>Proof of the Compactness Theorem</u>. We assume that T is finitely-satisfiable, and infinite (otherwise there is nothing to do).

Let $$I = \{ T_0 \subset T \mid T_0 \text{ finite} \} ;$$
for each $T_0 \in I$, there exists by assumption a model M_{T_0} of T_0. Now for each $T_0 \in I$, let
$$A_{T_0} = \{ T_1 \in I \mid T_1 \supset T_0 \} ,$$
and $\mathcal{F}_0 = \{ A_{T_0} \mid T_0 \in I \}$. Then $\mathcal{F}_0$ satisfies (F1), (F2), e.g. $A_{T_0} \cap A_{T_0'} = A_{T_0 \cup T_0'}$.
($\emptyset \notin \mathcal{F}_0$) ($A, B \in \mathcal{F}_0 \Rightarrow A \cap B \in \mathcal{F}_0$)

So $$\mathcal{F} = \{ A \subset I \mid \exists\, m, T_1, \ldots, T_m,\ A \supset A_{T_1} \cap \ldots \cap A_{T_m} \}$$
is a filter on I containing $\mathcal{F}_0$. Let $\hat{\mathcal{F}}$ be

a ultrafilter containing $\mathcal{F}$.

<u>Claim:</u> the ultraproduct $\hat{\prod}_{\mathcal{F}} M_i$ is a model of T.

Indeed, let $\phi \in T$. Since $\{\phi\} \in I$, we have

$$M_{\{\phi\}} \models \phi \qquad \text{(assumption)}$$

so $\{T_0 \in I \mid M_{T_0} \models \phi\} \supset \{T_0 \in I \mid \phi \in T_0\}$

$$= A_{\{\phi\}}.$$

But then this set $\{T_0 \in I \mid M_{T_0} \models \phi\}$ <u>must</u> be in $\hat{\mathcal{F}}$ (since otherwise the complement is in $\hat{\mathcal{F}}$, and then $A_{\{\phi\}} \cap \{T_0 \in I \mid M_{T_0} \not\models \phi\}$ would be in $\hat{\mathcal{F}}$, but it is the empty set), so by Łoś's Theorem $\hat{\prod}_{\mathcal{F}} M_i \models \phi$.

$\square$

<u>Remark.</u> Ultraproducts can be used to construct models in many other circumstances. Those are usually only interesting in the case of a non-principal

ultrafilter. Note then for instance the following: if X is an infinite set, then

$$\mathcal{F}_0 = \{\text{complements of finite sets in } X\}$$

is a filter, and any ultrafilter $\mathcal{F}$ containing $\mathcal{F}_0$ is non-principal (otherwise, $A = \{x_0\} \in \mathcal{F}$ as well as $B = \{x \neq x_0\}$, so $A \cap B = \emptyset$ would be in $\mathcal{F}$).

More generally a filter is an ultrafilter if and only if it contains no finite set.

Example. One also often uses ultraproducts as pure "set-theoretic" constructions without invoking model theory explicitly.

For instance, let $\mathcal{F}$ be a non-principal ultrafilter on $\mathbb{N}$; consider the "ultrapower"

$$R = \widetilde{\prod_{\mathcal{F}}} \mathbb{R} \subset \mathbb{R}^{\mathbb{N}}.$$

Then one can check that R is a field (because $(x_n) \sim (y_n) \iff (x_n - y_n) \sim (0)$ and

$I = \{ (x_n) \sim (0) \} \subset \mathbb{R}^{\mathbb{N}}$ is an ideal of the "usual" product $\mathbb{R}^{\mathbb{N}}$, so R is a ring; and if $(x_n) \notin I$ then let

$$\begin{cases} y_n = \dfrac{1}{x_n}, & x_n \neq 0 \\ y_n = 1, & x_n = 0. \end{cases}$$

Then $x_n y_n = 1$ for $x_n \neq 0$, so

$$(x_n)(y_n) = 1 \mod I$$

so R is a field), that it totally ordered by

$$(x_n) \leq (y_n) \iff \{ n \mid x_n \leq y_n \} \in \mathcal{F}.$$

Furthermore the diagonal map

$$j \begin{cases} \mathbb{R} \longrightarrow R \\ t \longmapsto \text{class of } (t, t, \dots) \end{cases}$$

is <u>injective</u>, so R is an extension of $\mathbb{R}$.

As such it has many unusual elements:

(i) the class of $(n)_{n \geq 0}$ is "infinitely large"

(ii) the class of $\left(1/_{n+1} \right)_{n \geq 0}$ is "infinitesimal".

etc...

Nevertheless, R shares many properties of $\mathbb{R}$, (which is due to the implicit model-theoretic perspective).

We finish this section with an important formalization of this last point:

<u>Proposition</u>. Let $\mathcal{L}$ be a language, M an $\mathcal{L}$-structure, I a set and $\mathcal{F}$ an ultrafilter on I. Let $\hat{M} = \hat{\prod_{\mathcal{F}}} M$ be the ultrapower of M.

(1) The map $j : M \longrightarrow \hat{M}$ given by $x \longmapsto$ class of $(x, x, \dots)$ is <u>injective</u>.

(2) In fact, j is an <u>elementary embedding</u>:

$$M \models \phi(a_1, \dots, a_k) \iff \hat{M} \models \phi(j(a_1), \dots, j(a_k))$$

for any $\mathcal{L}$-formula $\phi(v_1, \dots, v_k)$ and $(a_i) \in M^k$.

(<u>Intuitively</u>: anything that can be "proved" in $\hat{M}$, using first order logic and the language $\mathcal{L}$, about elements of $j(M)$, can be proved already in M).

<u>Proof</u>:- If $j(x) = j(y)$ then

$$\{ i \in I \mid x = y \} \in \mathcal{F},$$

so $\{ i \in I \mid x = y \} \neq \emptyset$, hence $x = y$.

The second statement is just Łoś's Theorem.

$\square$