Applications of the Compactness Theorem

Annalisa Conversano

Massey University Auckland

New Zealand Mathematics Research Institute

January 8-13, 2022

The plan

(1) Understanding the statement and how to use it

(2) Henkin's proof and amazing applications

(3) Applications to Number Theory

(X, τ) is a **topological space** when X is a set and $\tau \subset \mathcal{P}(X)$ s.t.

- $\emptyset, X \in \tau$
- $A_i \in \tau \Rightarrow \bigcup_i A_i \in \tau$
- $A, B \in \tau \Rightarrow A \cap B \in \tau$

τ is a **topology** on X & its elements are called **open** sets.

Closed sets are the complements of open sets.

A subset of X is **compact** when every open covering of it has a finite subcovering. Equivalently, any collection of closed subsets with the finite intersection property has nonempty intersection.

Example. The Euclidean topology on $\mathbb{R}^n$ is generated by the open balls $B_r(p) = \{x \in \mathbb{R}^n : d(p, x) < r\}$.

– $\mathbb{R}^n$ is not compact: $\{B_n(0) : n \in \mathbb{N}\}$ is an open covering

– $\{\mathbb{R}^n \setminus B_n(0) : n \in \mathbb{N}\}$ has the finite intersection property

$X \subset \mathbb{R}^n$ is compact $\Leftrightarrow$ X is closed and bounded.

The following are equivalent (in ZF):

(1) The compactness theorem for first-order logic:
If every finite subset of a first-order theory T has a model, then T has a model [finitely satisfiable (f.s.) $\Rightarrow$ satisfiable]

(2) **Gödel completeness theorem:** $T \models \varphi \iff T \vdash \varphi$
(φ is true wherever T is $\iff$ there is a proof of φ from T)

(3) The space of L-structures (L-theories) is compact.

(4) Any product of compact Hausdorff spaces is compact.

(5) The Stone-Čech compactification theorem.

(6) Ideals in a Boolean algebra can be extended to prime ideals.

(7) (Banach-Alaoglu) The closed unit ball in the continuous dual space of any normed space is weak-* compact.

A **language** (or signature, or vocabulary) is a set L of symbols:

- **relation** symbols R with associated arity $a(R) \in \mathbb{N}^{>0}$
- **function** symbols F with associated arity $a(F) \in \mathbb{N}$

  function symbols of arity 0 are called **constant** symbols.

Given a language $L = \{R, \ldots, F, \ldots\}$, $\mathcal{A} = (A, R^{\mathcal{A}}, \ldots, F^{\mathcal{A}}, \ldots)$ is an L-**structure** when:

- A is a nonempty set, the **underlying set** (or universe) of $\mathcal{A}$;
- $R^{\mathcal{A}} \subseteq A^m$, the **interpretation** of R in $\mathcal{A}$, $m = a(R)$;
- $F^{\mathcal{A}} \colon A^n \to A$ the **interpretation** of F in $\mathcal{A}$, $n = a(R)$;

  the interpretation of a constant $c \in L$ is some $c^{\mathcal{A}} \in A$.

Example. Let $L = \{<, +, \times, 0, 1\}$, where $<$ is a binary relation symbol, $+$ and $\times$ are binary function symbols, 0 and 1 are constant symbols. An example of L-structure is

$\mathcal{A} = (\mathbb{Z}, <^{\mathcal{A}}, +^{\mathcal{A}}, \times^{\mathcal{A}}, 0^{\mathcal{A}}, 1^{\mathcal{A}})$, with standard interpretations.

With an abuse of notation, we will write $\mathcal{A} = (\mathbb{Z}, <, +, \times, 0, 1)$

Let $L = \{\times, i, 1\}$, $\times$ binary function symbol, i unary function symbol, 1 constant symbol. Some L-structures:

$\mathcal{A} = (\mathrm{GL}_5(\mathbb{Z}), \times^{\mathcal{A}}, i^{\mathcal{A}}, 1^{\mathcal{A}})$, $\times^{\mathcal{A}}(X, Y) = XY$, $i^{\mathcal{A}}(X) = X^{-1}$,$1^{\mathcal{A}} = I_5$

$\mathcal{B} = (\mathbb{Z}, \times^{\mathcal{B}}, i^{\mathcal{B}}, 1^{\mathcal{B}})$, $\times^{\mathcal{B}}(x, y) = x + y$, $i^{\mathcal{B}}(x) = -x$, $1^{\mathcal{B}} = 0$

$\mathcal{C} = (\mathbb{Z}, \times^{\mathcal{C}}, i^{\mathcal{C}}, 1^{\mathcal{C}})$, $\times^{\mathcal{C}}(x, y) = \lfloor ye^x \rfloor$, $i^{\mathcal{C}}(x) = x + 7$, $1^{\mathcal{C}} = -23$

Let $L = \{E\}$, E binary relation symb. A graph is an L-structure.

Let $L = \{<, +, \times, 0, 1\}$ as before. Some other L-structures:

- $\mathcal{A} = (\mathbb{Z}, <, +, \times, 0, 1)$
- $\mathcal{B} = (\mathbb{Q}, <, +, \times, 0, 1)$
- $\mathcal{C} = (\mathbb{R}, <, +, \times, 0, 1)$
- $\mathcal{D} = (\mathbb{N}, <, +, \times, 0, 1)$

Fix $\mathrm{Var} = \{x_i : i \in I\} \supset \{x, y, z\}$ a set of symbols, the **variables**.

L-**terms** are defined inductively as follows:

(i) each **variable** is an L-term;

(ii) if $F \in L$ is a n-ary function symbol & $t_1, \ldots, t_n$ are L-terms, then $\boldsymbol{F(t_1, \ldots, t_n)}$ is an L-term.

Examples. $L = \{+, \times, 0, 1\}$. We will write $xy + 3$ for the term $+(\times(x, y), +(+(1, 1), 1))$ and $2x + y$ for $+(+(x, x), y)$.

Given an L-structure $\mathcal{A}$ and an L-term $t = t(\vec{x})$, $\vec{x} = (x_1, \ldots, x_m)$, set $\boxed{t^{\mathcal{A}} \colon A^m \to A}$ as follows:

(i) $t = x_i \;\Rightarrow\; t^{\mathcal{A}}(a) = a_i$ for $a = (a_1, \ldots, a_m) \in A^m$;

(ii) $t = F(t_1, \ldots, t_n) \;\Rightarrow\; t^{\mathcal{A}}(a) = F^{\mathcal{A}}(t_1^{\mathcal{A}}(a), \ldots, t_n^{\mathcal{A}}(a))$, $a \in A^m$.

Example. If $\mathcal{A}$ is the ring of the integers, and $t = 2x + y$, then $t^{\mathcal{A}} \colon \mathbb{Z}^2 \to \mathbb{Z}$ is given by $t^{\mathcal{A}}(a, b) = 2a + b$

The **atomic** L-formulas are:

(i) $\boldsymbol{t_1 = t_2}$, where t_1 and t_2 are L-terms;

(ii) $\boldsymbol{R(t_1, \ldots, t_m)}$, R is a m-ary rel sym & $t_1, \ldots, t_m$ are L-terms.

L-**formulas** are defined inductively as follows:

- each **atomic** L-formula is an L-formula;
- if φ, ψ are L-formulas, then so are $\boldsymbol{\neg\varphi, \varphi \wedge \psi, \varphi \vee \psi}$;
- if φ is an L-form & $x \in \mathrm{Var}$, then $\boldsymbol{\exists x \varphi, \forall x \varphi}$ are L-formulas.

Notation. We write $\boldsymbol{\varphi \to \psi}$ for $\neg\varphi \vee \psi$

Example. $L = \{<, +, \times, 0, 1\}$

$$\forall xy\, \exists z\, ((x < y) \to (x < z < y))$$

$$(\exists x(x \leqslant 0)) \wedge (x^2 \neq 5)$$

A L-**sentence** is a L-formula in which all occurrences of variables are bound (no free variables). A L-**theory** is a set of L-sentences.

Given $\mathcal{A}$ an L-structure, set $L_A = L \cup \{\underline{c} : c \in A\}$ and $\underline{c}^{\mathcal{A}} = c$.

For any L_A sentence σ, define $\boxed{\mathcal{A} \models \sigma}$ (σ **is true** in $\mathcal{A}$, or σ holds in $\mathcal{A}$, or $\mathcal{A}$ satisfies σ) as follows:

(i) $\boxed{\mathcal{A} \models t_1 = t_2}$ iff $\boxed{t_1^{\mathcal{A}} = t_2^{\mathcal{A}}}$ (t_1, t_2 variable-free L_A-terms)

(ii) $\boxed{\mathcal{A} \models R(t_1, \ldots, t_m)}$ iff $\boxed{(t_1^{\mathcal{A}}, \ldots, t_m^{\mathcal{A}}) \in R^{\mathcal{A}}}$ (R is a m-ary relation symbol & $t_1, \ldots, t_m$ variable-free L_A-terms)

(iii) $\boxed{\mathcal{A} \models \neg\varphi}$ iff $\boxed{\mathcal{A} \not\models \varphi}$

(iii) $\boxed{\mathcal{A} \models \varphi \wedge \psi}$ iff $\boxed{\mathcal{A} \models \varphi \quad \& \quad \mathcal{A} \models \psi}$

(iv) $\boxed{\mathcal{A} \models \varphi \vee \psi}$ iff $\boxed{\mathcal{A} \models \varphi \quad \text{or} \quad \mathcal{A} \models \psi}$

(v) $\boxed{\mathcal{A} \models \exists x \varphi(x)}$ iff $\boxed{\mathcal{A} \models \varphi(\underline{a}) \quad \text{for some } a \in A}$

(vi) $\boxed{\mathcal{A} \models \forall x \varphi(x)}$ iff $\boxed{\mathcal{A} \models \varphi(\underline{a}) \quad \text{for all } a \in A}$

Given a set of L-sentences T (that is, an L-**theory**), we say that the L-structure $\mathcal{A}$ is a **model** of T whenever $\mathcal{A} \models \sigma$ for all $\sigma \in T$.

Notation: $\boxed{\mathcal{A} \models T}$

Examples

(1) $L = \{<\}$, $<$ binary relation symbol.

$$\sigma_1 := \forall x\, \neg(x < x) \quad \wedge \quad \forall xyz\, ((x < y \wedge y < z) \rightarrow x < z)$$

$$\sigma_2 := \forall xy\, (x < y \vee x = y \vee y < x)$$

$$\sigma_3 := \forall xy\, ((x < y) \rightarrow \exists z\, (x < z < y))$$

$$\sigma_4 := \forall x \exists yz\, (y < x < z)$$

$$\mathcal{A} \models \sigma_1 \quad \Longleftrightarrow \quad \mathcal{A} \text{ is a } \textbf{partial order}.$$

$$\mathcal{A} \models \{\sigma_1, \sigma_2\} \quad \Longleftrightarrow \quad \mathcal{A} \text{ is a } \textbf{linear order}.$$

$$\mathcal{A} \models DLO = \{\sigma_i\}_{1 \leqslant i \leqslant 4} \Longleftrightarrow \mathcal{A} \text{ is a } \textbf{dense l.o. w/o endpoints}.$$

Ex: Write a sentence for "every element has a unique successor"

(2) $L = \{<, +, 0, -\}$, $<$ binary relation symbol,
$+$ binary, 0 const, $-$ unary function symbols.

$\varphi_1 := \forall xyz\ ((x+y)+z) = x+(y+z))$

$\varphi_2 := \forall x\ ((x+0 = x = 0+x)\ \wedge\ (x+(-x) = (-x)+x = 0))$

$\varphi_3 := \forall xy\ (x+y = y+x)$

$\varphi_4 := \forall xyz\ (x < y \to ((x+z) < (y+z)))$

$\psi_n := \forall x\ (nx = 0 \to x = 0)$

$$\mathcal{A} \models \{\varphi_1, \varphi_2\} \iff \mathcal{A} \text{ is a } \textbf{group}.$$

$$\mathcal{A} \models \{\varphi_1, \varphi_2, \varphi_3\} \iff \mathcal{A} \text{ is an } \textbf{abelian group}.$$

$$\mathcal{A} \models \{\varphi_i, \sigma_1, \sigma_2\}_{1\leqslant i\leqslant 4} \iff \mathcal{A} \text{ is an } \textbf{ordered abelian group}.$$

$$\mathcal{A} \models \{\varphi_1, \varphi_2, \varphi_3, \psi_n\}_{n\in\mathbb{N}^{>0}} \iff \mathcal{A} \text{ is a } \textbf{torsion-free abelian group}.$$

ordered $\Rightarrow$ torsion-free: $\{\varphi_i, \sigma_1, \sigma_2\}_{1\leqslant i\leqslant 4} \models \{\psi_n\}_{n\in\mathbb{N}}$

$T_1 \models T_2$ means “Every model of T_1 is a model of T_2"

We say that T_2 **follows from** T_1 or T_2 **is a consequence** of T_1.

An L-theory T is **maximal** when for each L-sentence σ either $\sigma \in T$ or $\neg\sigma \in T$.

Given $\mathcal{A}$ an L-structure, key max theories (**over *L* and *L*$_A$**) are:

- ***Th*($\mathcal{A}$)** $= \{\sigma : \sigma$ is an L-sentence & $\mathcal{A} \models \sigma\}$
- ***ED*($\mathcal{A}$)** $= \{\sigma : \sigma$ is an L_A-sentence & $\mathcal{A} \models \sigma\}$

Example. $L = \{+, \times, 0, 1\}$, $\mathcal{A} = (\mathbb{R}, +, \times, 0, 1)$

$$\exists x\ (x^2 = 2) \in Th(\mathcal{A})$$

$$\exists x\ (x^2 = \pi) \in ED(\mathcal{A}) \setminus Th(\mathcal{A})$$

Clearly, $\mathcal{A} \models Th(\mathcal{A})$ and $\mathcal{A} \models ED(\mathcal{A})$.

If $\mathcal{B} \models ED(\mathcal{A})$, then $\mathcal{A}$ embeds elementarily into $\mathcal{B}$ through the map $\quad \underline{c}^{\mathcal{A}} \mapsto \underline{c}^{\mathcal{B}} \quad$ for all $c \in A$

T f.s. L-theory $\Rightarrow$ there is a maximal f.s. L-theory $T' \supseteq T$.

Proof.

Order $I = \{T' : T' \text{ f.s. } L\text{-theory }, T' \supseteq T\}$ by inclusion.

If $C \subseteq I$ is a chain, then $T_C := \bigcup\{\Sigma : \Sigma \in C\}$ is an upper bound.

By Zorn's lemma, there is $T' \in I$ maximal w.r.t. the partial order.

T f.s. L-theory, φ L-sentence $\Rightarrow$ $T \cup \{\varphi\}$ or $T \cup \{\neg\varphi\}$ f.s.

Proof.

If $T \cup \{\varphi\}$ is not f.s. there is a finite $\Delta \subseteq T$, $\Delta \models \neg\varphi$.

If $\Sigma \subseteq T$ is finite, then $\Sigma \cup \Delta$ is finite and has a model $\mathcal{M}$,

$\mathcal{M} \models \Delta \cup \Sigma \cup \{\neg\varphi\}$. So $T \cup \{\neg\varphi\}$ is f.s. □

Therefore T' is a maximal L-theory. □

Using the compactness theorem

Assume we know that

> If every finite subset of T has a model, then T has a model.

Goal:

Getting a mathematical object $\mathcal{A}$

Strategy:

Find a language L and a L-theory T such that

(1) $\mathcal{A}$ is a model of T (or $\mathcal{A}$ embeds in any model of T) &

(2) you can find a model for any finite subset of T

A partial order $(P, <)$ can be extended to a linear one

Let $L = \{\prec\} \cup \{\underline{a} : a \in P\}$, $\prec$ binary relation symbol.

Let T be the L-theory given by:

(1) $\prec$ is a linear ordering;

(2) $\underline{a} \prec \underline{b}$ whenever $a < b$ [that is, add $ED(P, <)$];

Given a finite $T_0 \subset T$, let $P_0 \subset P$ be the finite subset of P corresponding to the constants in T_0 from (2).

$<$ can be extended to a linear $<'$ on P_0 by induction on $|P_0|$.

$(P_0, <')$ is a model of T_0. By compactness, T has a model $\mathcal{A}$.

Define $<'$ on P as:

$$a <' b \iff \mathcal{A} \models \underline{a} \prec \underline{b}$$

$<'$ is a linear ordering on P extending $<$

If every finite subset of a graph Γ is k-colorable, then Γ is k-colorable

Let $L = \{E, u_1, \ldots, u_k\} \cup \{\underline{v} : v \in \Gamma\}$, E binary relation symbol, u_i unary relation symbols.

Let T be the L-theory given by:

(1) $\forall x \forall y\, (\neg E(x,x) \wedge (E(x,y) \to E(y,x)))$ [optional]

(2) $E(\underline{v}, \underline{w})$ whenever there is an edge between v & w in Γ

(3) $\forall x\, (u_1(x) \vee \cdots \vee u_k(x))$ [each vertex is colored]

(4) $$\forall x \forall y \left(\bigwedge_{i=1}^{k} (u_i(x) \wedge u_i(y) \to \neg E(x,y) \right)$$ [k-coloring]

Given a finite $T_0 \subset T$, let $\Delta = \{v_1, \ldots, v_n\}$ where $\underline{v}_1, \ldots \underline{v}_n$ are the constants in T_0 from (2). Then Δ is the universe of a model of T_0 because it is k-colorable by assumption.

By compactness, T has a model $\mathcal{A}$. A k-coloring on Γ is:

$$v \text{ has color } i \;\Leftrightarrow\; \underline{v}^{\mathcal{A}} \in u_i^{\mathcal{A}}$$

Every torsion-free abelian group G can be ordered

Let $L = \{<, +\} \cup \{\underline{a} : a \in G\}$, $<$ binary relation symbol, $+$ binary function symbol.

Let T be the L-theory given by:

(1) the axioms of ordered abelian groups;

(2) $ED(G, +)$: axioms of the group operation on the elem of G;

Given a finite $T_0 \subset T$, let $H = \langle a_1, \ldots, a_n \rangle$ where $\underline{a}_1, \ldots \underline{a}_n$ are the constant symbols in T_0 from (2). H is a finitely generated subgroup of G (torsion-free), so $H \cong \mathbb{Z}^k$, where $k \leqslant n$.
So $(H, <_{lex}) \models T_0$.

By compactness, T has a model $\mathcal{A}$. Set $G' = \{\underline{a}^{\mathcal{A}} : a \in G\}$.

$(G', <^{\mathcal{A}}, +^{\mathcal{A}})$ is an ordered abelian group, group-isomorphic to G.

Exercises 1.

(1) Fix a language L and find an L-theory T whose models are:

(i) the divisible abelian groups,

(ii) the fields of characteristic 0,

(iii) the algebraically closed fields,

(iv) the bounded metric spaces with diameter D,

(v) the vector spaces over the field $\mathbb{K}$.

(2) Prove (i)–(iv) are not finitely axiomatizable. How about (v)?

(3) Let $L = \{+, \times, 0, 1\}$. Fix $d \in \mathbb{N}^{>0}$. Write an L-sentence Φ_d such that for any field $\mathcal{K}$, $\mathcal{K} \models \Phi_d$ if and only if every injective polynomial map $f\colon K \to K$ with degree at most d is surjective (generalize it to $\Phi_{n,d}$ and $f\colon K^n \to K^n$ where each coordinate function has degree at most d)

The compactness theorem

If every finite subset of a first-order L-theory T has a model, then T has a model.

Main idea of Henkin's proof: Building a model of T by adding enough constants to the language so that every element of the model will be named by a constant symbol.

Summary

A **language** is a set L of symbols:

- **relation** symbols R with associated arity $a(R) \in \mathbb{N}^{>0}$
- **function** symbols F with associated arity $a(F) \in \mathbb{N}$

Given a language $L = \{R, \ldots, F, \ldots\}$, $\mathcal{A} = (A, R^{\mathcal{A}}, \ldots, F^{\mathcal{A}}, \ldots)$ is an L-**structure** when:

- A is a nonempty set, the **underlying set** (or universe) of $\mathcal{A}$;
- $R^{\mathcal{A}} \subseteq A^m$, the **interpretation** of R in $\mathcal{A}$, $m = a(R)$;
- $F^{\mathcal{A}} \colon A^n \to A$ the **interpretation** of F in $\mathcal{A}$, $n = a(R)$;

L-**terms** are defined inductively as follows:

(i) each **variable** is an L-term;

(ii) if $F \in L$ is a n-ary function symbol & $t_1, \ldots, t_n$ are L-terms, then $\boldsymbol{F(t_1, \ldots, t_n)}$ is an L-term.

The **atomic** L-formulas are:

(i) $\boldsymbol{t_1 = t_2}$, where t_1 and t_2 are L-terms;

(ii) $\boldsymbol{R(t_1, \ldots, t_m)}$, R is a m-ary rel sym & $t_1, \ldots, t_m$ are L-terms.

A L-**sentence** is a L-formula in which all occurrences of variables are bound (no free variables). A L-**theory** is a set of L-sentences.

A **model** of an L-theory T is an L-structure where all sentences of T are true. Notation: $\mathcal{A} \models T$.

Main steps of Henkin's proof

> If every finite subset of a first-order L-theory T has a model, then T has a model.

(1) Add constants to L and sentences to T to get a f.s. L^*-theory T^* with the **witness property**:
for any L^*-formula $\varphi(x)$ there is $c \in L^*$ s.t.

$$T^* \models (\exists x\ \varphi(x)) \rightarrow \varphi(c)$$

(2) Extend T^* to a **maximal** f.s. L^*-theory T' (with the w.p.):

for any L^*-sentence σ either $\sigma \in T'$ or $\neg\sigma \in T'$

(3) Define a model of T' on a quotient of the constants in L^*.

T finitely satisfiable L-theory $\Rightarrow$ T is satisfiable.

(1) For any L-formula $\varphi(x)$, let c_φ be a new constant symbol and

$$\Theta_\varphi : \ (\exists x\, \varphi(x)) \ \rightarrow \ \varphi(c_\varphi)$$

Ex. $L = \{+, \times, 0, 1\}$ $\qquad \varphi(x): \ x^2 = 5$

$\Theta_\varphi : \ \exists x (x^2 = 5) \rightarrow (c_\varphi^2 = 5)$ $\qquad \mathcal{M} \overset{?}{\models} \Theta_\varphi$

$\mathcal{A} = (\mathbb{R}, +, \times, 0, 1, c_\varphi^{\mathcal{A}} = \sqrt{5}) \models \Theta_\varphi$

$\mathcal{B} = (\mathbb{R}, +, \times, 0, 1, c_\varphi^{\mathcal{B}} = -\sqrt{5}) \models \Theta_\varphi$

$\mathcal{C} = (\mathbb{Q}, +, \times, 0, 1, c_\varphi^{\mathcal{C}} = \frac{1}{2}) \models \Theta_\varphi$

T finitely satisfiable L-theory $\Rightarrow$ T is satisfiable.

(1) For any L-formula $\varphi(x)$, let c_φ be a new constant symbol and

$$\Theta_\varphi : \ (\exists x\, \varphi(x)) \ \rightarrow \ \varphi(c_\varphi)$$

Set $L_1 := L \cup \{c_\varphi : \varphi(x) \text{ is an } L\text{-formula}\}$ and

$T_1 := T \cup \{\Theta_\varphi : \varphi(x) \text{ is an } L\text{-formula}\}$. T_1 is finitely satisfiable:

If $\Delta \subseteq T_1$ is finite, then $\Delta = \Delta_0 \cup \{\Theta_{\varphi_1}, \ldots, \Theta_{\varphi_n}\}$, $\Delta_0 \subseteq T$.

Take $\mathcal{A} \models \Delta_0$. If $\mathcal{A} \models \exists x\, \varphi_i(x)$, take $a_i \in A$ s.t. $\mathcal{A} \models \varphi_i(a_i)$ and

set $c_{\varphi_i}^{\mathcal{A}'} = a_i$. Otherwise, let $c_{\varphi_i}^{\mathcal{A}'}$ be any $a \in A$. Clearly $\mathcal{A}' \models \Delta$.

Iteretating: $L \subseteq L_1 \subseteq L_2 \subseteq \ldots$, $T \subseteq T_1 \subseteq T_2 \subseteq \ldots$ Set

$$L^* := \bigcup L_i \qquad T^* := \bigcup T_i$$

T^* is f.s. & for any L^*-formula $\varphi(x)$ there is $c \in L^*$ s.t.

$$T^* \models (\exists x\, \varphi(x)) \ \rightarrow \ \varphi(c)$$

T f.s. L-theory $\Rightarrow$ there is a maximal f.s. L-theory $T' \supseteq T$.

Proof.

Order $I = \{T' : T' \text{ f.s. } L\text{-theory }, T' \supseteq T\}$ by inclusion.

If $C \subseteq I$ is a chain, then $T_C := \bigcup\{\Sigma : \Sigma \in C\}$ is an upper bound.

By Zorn's lemma, there is $T' \in I$ maximal w.r.t. the partial order.

T f.s. L-theory, φ L-sentence $\Rightarrow$ $T \cup \{\varphi\}$ or $T \cup \{\neg\varphi\}$ f.s.

Proof.

If $T \cup \{\varphi\}$ is not f.s. there is a finite $\Delta \subseteq T$, $\Delta \models \neg\varphi$.

If $\Sigma \subseteq T$ is finite, then $\Sigma \cup \Delta$ is finite and has a model $\mathcal{M}$,

$\mathcal{M} \models \Delta \cup \Sigma \cup \{\neg\varphi\}$. So $T \cup \{\neg\varphi\}$ is f.s. $\square$

Therefore T' is a maximal L-theory. $\square$

(3) Let T' be a f.s. maximal L^*-theory extending T^*.

We can show that T' has a model $\mathcal{A}$:

Let $C \subseteq L^*$ be the set of constant symbols of L^*.

For $c, d \in C$ define $c \sim d \iff T' \models c = d$

$\sim$ is an equivalence relation because T' is f.s. & maximal.

Set $\boxed{A = C/\sim}$ and for any $c \in C$, set $c^* \in A$ and $\boxed{c^{\mathcal{A}} = c^*}$

T maximal f.s. L-theory, $\Delta \models \sigma$, $\Delta \subseteq T$ finite $\Rightarrow \sigma \in T$

Proof.

$\sigma \notin T \Rightarrow \neg\sigma \in T \Rightarrow \Delta \cup \{\neg\sigma\} \subseteq T$ finite & unsatisfiable.

Contradiction. □

For any $R \in L^*$ n-ary relation symbol, set

$$R^{\mathcal{A}} = \{(c_1^*, \ldots, c_n^*) \in A^n : R(c_1, \ldots, c_n) \in T'\}$$

$R^{\mathcal{A}}$ is well-defined: $\quad c_i \sim d_i \;\Rightarrow\; c_i = d_i \in T'$

So if $\vec{c} \sim \vec{d}$ then $\quad R(\vec{c}) \in T' \;\Leftrightarrow\; R(\vec{d}) \in T'$

For any $F \in L^*$ n-ary function symbol, and any $c_1, \ldots, c_{n+1} \in C$:

$$F^{\mathcal{A}}(c_1^*, \ldots, c_n^*) = c_{n+1}^* \iff F(c_1, \ldots, c_n) = c_{n+1} \in T'$$

$F^{\mathcal{A}}$ is well-defined: $\quad c_i \sim d_i \Rightarrow c_i = d_i \in T'$

As $\emptyset \models \exists x\, (F(c_1, \ldots, c_n) = x)$ & T' has the witness property,

there is $c_{n+1} \in C$ such that $F(c_1, \ldots, c_n) = c_{n+1} \in T'$.

Similarly, $F(d_1, \ldots, d_n) = d_{n+1} \in T'$ & $c_{n+1} \sim d_{n+1}$.

So $\mathcal{A} = (A, c^*, \ldots, R^{\mathcal{A}}, \ldots, F^{\mathcal{A}}, \ldots)$ is an L^*-structure.

$\mathcal{A} \models T'$:

> For any L^*-formula $\varphi(x_1, \ldots, x_n)$ and $c_1, \ldots, c_n \in C$
>
> $$\mathcal{A} \models \varphi(\vec{c}) \quad \Leftrightarrow \quad \varphi(\vec{c}) \in T'$$

Proof.

By induction on the complexity of φ (using that T' has the witness property and is maximal finitely satisfiable). $\square$

T finitely satisfiable L-theory $\Rightarrow$ T is satisfiable.

Corollary. Let T be an L-theory & σ an L-sentence.

$T \models \sigma \implies$ there is a finite $T_0 \subseteq T$ such that $T_0 \models \sigma$

Proof.

If not, for each finite $T_0 \subseteq T$, $T_0 \cup \{\neg\sigma\}$ has a model.

Therefore, $T \cup \{\neg\sigma\}$ is finitely satisfiable.

By compactness, $T \cup \{\neg\sigma\}$ is satisfiable. Contradiction. □

Example: Algebraically Closed Fields.

Algebraically closed fields

Let $L = \{+, \times, 0, 1\}$ and T be the L-theory of fields.

(That is, $+$ is an abelian group operation with identity 0, $0 \neq 1$, $\times$ is an abelian group operation on the non-zero elements with identity 1, left and right distribution laws of $\times$ with respect to $+$).

$\varphi_n := \forall u_1 \ldots u_n \, \exists x \, (x^n + u_1 x^{n-1} + \cdots + u_n = 0)$

$ACF := T \cup \{\varphi_n : n \geqslant 1\}$

$ACF_0 := ACF \cup \{n1 \neq 0 : n \geqslant 1\}$, $ACF_p := ACF \cup \{p1 = 0\}$

Let $\mathcal{F}$ be an L-structure. Then

$\mathcal{F} \models ACF \quad \Leftrightarrow \quad \mathcal{F}$ is an **algebraically closed field**.

$\mathcal{F} \models ACF_0 \Leftrightarrow \mathcal{F}$ acf of **characteristic 0**

$\mathcal{F} \models ACF_p \Leftrightarrow \mathcal{F}$ acf of **characteristic p**

Algebraically closed fields

Let $L = \{+, \times, 0, 1\}$ and T be the L-theory of fields.

(That is, $+$ is an abelian group operation with identity 0, $0 \neq 1$, $\times$ is an abelian group operation on the non-zero elements with identity 1, left and right distribution laws of $\times$ with respect to $+$).

$\varphi_n := \forall u_1 \dots u_n \, \exists x \, (x^n + u_1 x^{n-1} + \cdots + u_n = 0)$

$ACF := T \cup \{\varphi_n : n \geqslant 1\}$

$ACF_0 := ACF \cup \{n1 \neq 0 : n \geqslant 1\}$, $ACF_p := ACF \cup \{p1 = 0\}$

Let $\mathcal{F}$ be an L-structure. Then

$\mathcal{F} \models ACF \quad \Leftrightarrow \quad \mathcal{F}$ is an **algebraically closed field**.

$\mathcal{F} \models ACF_0 \ \Leftrightarrow \mathcal{F}$ acf of **characteristic 0** $\ \Leftrightarrow Th(\mathcal{F}) = Th(\mathbb{C})$

$\mathcal{F} \models ACF_p \ \Leftrightarrow \mathcal{F}$ acf of **characteristic p** $\ \Leftrightarrow \ Th(\mathcal{F}) = Th(\mathbb{F}_p^{alg})$

For each σ L-sentence either $ACF_k \models \sigma$ or $ACF_k \models \neg\sigma$

Algebraically closed fields

Corollary. Let σ be an L-sentence, $L = \{+, \times, 0, 1\}$. TFAE:

(i) σ is true in the complex field.

(ii) σ is true in some acf of characteristic 0.

(iii) σ is true in every acf of characteristic 0.

(iv) There is an m such that for all $p > m$, σ is true in all acf of characteristic p.

(v) There are arbitrarily large p such that σ is true in some acf of characteristic p.

(iii) $\Rightarrow$ (iv): By compactness, $\exists\, T_0 \subseteq ACF_0$ finite s.t. $T_0 \models \sigma$.

(v) $\Rightarrow$ (i): If $\mathbb{C} \not\models \sigma$, then $\mathbb{C} \models \neg\sigma$, contradiction.

Let $f\colon \mathbb{C}^n \to \mathbb{C}^n$ be injective & polynomial. Then f is surj.

Claim. Every injective polynomial map $f\colon (\mathbb{F}_p^{alg})^n \to (\mathbb{F}_p^{alg})^n$ is surjective.

Proof of the Claim.

If not, let $\bar{a} \in (\mathbb{F}_p^{alg})^s$ be the coefficients of f and let $\bar{b} \in (\mathbb{F}_p^{alg})^n$ not in the range of f. Let K be the subfield of $\mathbb{F}_p^{alg}$ generated by $\bar{a}, \bar{b}$. Then the restriction of f to K^n is an injective but not surjective polynomial map $K^n \to K^n$. But $\mathbb{F}_p^{alg} = \cup \mathbb{F}_{p^n}$ is locally finite, so K is finite, contradiction. □

Let $f\colon \mathbb{C}^n \to \mathbb{C}^n$ be a counterexample and let d be the largest degree of the coordinate functions of f. Let Φ_d be the sentence saying "every injective polynomial function in n variables with n coordinate functions with degree at most d is surjective".

$\mathbb{F}_p^{alg} \models \Phi_d$ for all p by the Claim. So $\mathbb{C} \models \Phi_d$, contradiction.

Every field F has an algebraic closure

Let $L = \{+, \cdot, 0, 1\} \cup \{\underline{c} : c \in F\}$ and T the L-theory given by:

(1) the axioms of fields;

(2) $ED(\mathcal{F})$: axioms of the ring operations on the elements of F;

(3) for each non-zero polynomial $p \in F[x]$, an axiom saying that p splits.

Given a finite $T_0 \subset T$, there are finitely many polynomials from (3), so we can find a finite extension of F that is a model of T_0.

By compactness, T has a model $\mathcal{A}$.

$F^{\mathcal{A}} := \{\underline{c}^{\mathcal{A}} : c \in F\}$ is a field isomorphic to F.

$\overline{F} := \{a \in A : a \text{ is algebraic over } F^{\mathcal{A}}\}$ is algebraically closed.

The algebraic closure of F is unique

Let E, K be algebraic closures of F.

Set $L = \{+, \cdot, 0, 1\} \cup \{\underline{c} : c \in E\} \cup \{\underline{d} : d \in K\}$.

Let T be the L-theory given by:

(1) the axioms of fields;

(2) axioms of the ring operations on the elements of E and K.

Given $T_0 \subset T$, only finitely many elements of E and K appear, and there is a finite field extension of F that models T_0.

By compactness, there is a model $\mathcal{A}$ of T.

$E^{\mathcal{A}} := \{\underline{c}^{\mathcal{A}} : c \in E\}$ is isomorphic to E.

$K^{\mathcal{A}} := \{\underline{d}^{\mathcal{A}} : d \in K\}$ is isomorphic to K.

$E^{\mathcal{A}}$ and $K^{\mathcal{A}}$ are isomorphic and agree on F.

Exercises 2.

(1) Prove the compactness theorem is equivalent to the compactness of the topological space of satisfiable maximal L-theories (see previous slide)

(2) Prove the compactness theorem is equivalent to the compactness of the quotient of L-structure by elementarily equivalence (see previous slide)

(3) Show that the following are not first-order axiomatizable. That is, there is no first-order theory T whose models are

(i) the finite sets (or finite groups, or finite fields, etc.),

(ii) the connected graphs,

(iii) the torsion groups

Some references

(1) Lou van den Dries' **Logic Notes**

(2) David Marker's article in the **The Princeton Companion to Mathematics**, Princeton University Press (2008), IV.23 *Logic and Model Theory*, pg 635–646.

(3) David Marker's book: **Model Theory: An Introduction**, Springer GTM 217 (2002).

Let X be the set of satisfiable maximal L-theories.

Set $T_\varphi = \{T \in X : \varphi \in T\}$ for any L-sentence φ.

This is a basis for a topology τ on X.

> Every f.s. L-theory is satisfiable $\iff$ (X, τ) is compact

$\boxed{\Rightarrow}$ $\mathcal{U} = \{T_\varphi : \varphi \in S\}$ open covering $\qquad \bigcap_{\varphi \in S} T_\varphi^c = \emptyset = \bigcap_{\varphi \in S} T_{\neg\varphi}$

$\Sigma = \{\neg\varphi \mid \varphi \in S\}$ unsatisfiable

$\exists$ finite $\Sigma_0 \subseteq \Sigma$ unsatisfiable $\qquad \bigcap_{\varphi \in \Sigma_0} T_\varphi = \emptyset$

$\{T_{\neg\varphi} : \varphi \in \Sigma_0\} \subseteq \mathcal{U}$ finite subcovering so (X, τ) is compact.

Let X be the set of satisfiable maximal L-theories.

Set $T_\varphi = \{T \in X : \varphi \in T\}$ for any L-sentence φ.

This is a basis for a topology τ on X.

If $\mathcal{A} \models T$ max
$T = Th(\mathcal{A})$

Every f.s. L-theory is satisfiable $\iff$ (X, τ) is compact

$\Leftarrow$ Σ f.s. L-theory

$\mathcal{C} = \{ T_{\neg\varphi}^c : \varphi \in \Sigma \}$ closed set with the finite intersection property:

$\bigcap_{i=1}^n T_{\neg\varphi_i}^c = \bigcap_{i=1}^n T_{\varphi_i} = T_{\varphi_1 \wedge \dots \wedge \varphi_n} \neq \emptyset$ because Σ f.s. (X,τ) compact $\Rightarrow$

$\bigcap_{\varphi \in \Sigma} T_{\neg\varphi}^c \neq \emptyset \neq \bigcap_{\varphi \in \Sigma} T_\varphi$ Σ has a model

Let X be the set of satisfiable maximal L-theories.

Set $T_\varphi = \{T \in X : \varphi \in T\}$ for any L-sentence φ.

This is a basis for a topology τ on X.

Every f.s. L-theory is satisfiable $\iff$ (X, τ) is compact

Equivalently, let S be the set of L-structures and set $Y = S/\sim$

$$\mathcal{A} \sim \mathcal{B} \Leftrightarrow Th(\mathcal{A}) = Th(\mathcal{B}) \qquad \text{Set } \mathcal{A}^* = [\mathcal{A}] \in Y$$

Set $S_\varphi = \{\mathcal{A}^* \in Y : \mathcal{A} \models \varphi\}$ for any L-sentence φ.

This is a basis for a topology τ on Y.

Every f.s. L-theory is satisfiable $\iff$ (Y, τ) is compact

Nonstandard analysis

Let $\mathcal{R}$ be an ordered field. Then the following are equivalent:

- $Th(\mathcal{R}) = Th(\mathbb{R}) = RCF$ ($\mathcal{R}$ is a **real closed field**)
- Every positive element in $\mathcal{R}$ is a square & every polynomial of odd degree has a root in $\mathcal{R}$.
- $\mathcal{R}(\sqrt{-1})$ is algebraically closed.

K acf char 0 $R \subseteq K$ max rcf $R(\sqrt{-1}) = K$ $\mathbb{C} = \mathbb{R}(\sqrt{-1})$

$(K, \oplus, \otimes)$ is definable in $(R, <, +, \times)$ $K = R^2$

So algebraic varieties, groups over K are definable in a rcf R

Nonstandard analysis

Let $\mathcal{R}$ be an ordered field. Then the following are equivalent:

- $Th(\mathcal{R}) = Th(\mathbb{R}) = RCF$ ($\mathcal{R}$ is a **real closed field**)
- Every positive element in $\mathcal{R}$ is a square & every polynomial of odd degree has a root in $\mathcal{R}$.
- $\mathcal{R}(\sqrt{-1})$ is algebraically closed.

> There are non-Archimedean real closed fields.

Let $L = \{<, +, \times, 0, 1, c\}$, $T = Th(\mathbb{R}) \cup \{0 < c < \frac{1}{n} : n \in \mathbb{N}\}$

For any finite $T_0 \subseteq T$, $\mathbb{R} \models T_0$. By compactness, T has a model.

Let $\mathcal{R}$ be a model of T. Does $\{\frac{1}{n}\}_{n\in\mathbb{N}} \to c^{\mathcal{R}}$? NO :

$c < 2c = c + c < \frac{1}{n} \iff c < \frac{1}{2n}$, contradiction.

$\{\frac{1}{n}\}_{n\in\mathbb{N}}$ does not converge in $\mathcal{R}$

Nonstandard analysis

Let $\mathcal{R}$ be an ordered field. Then the following are equivalent:

- $Th(\mathcal{R}) = Th(\mathbb{R}) = RCF$ ($\mathcal{R}$ is a **real closed field**)
- Every positive element in $\mathcal{R}$ is a square & every polynomial of odd degree has a root in $\mathcal{R}$.
- $\mathcal{R}(\sqrt{-1})$ is algebraically closed.

There are non-Archimedean real closed fields.

Let $L = \{<, +, \times, 0, 1, c\}$, $T = Th(\mathbb{R}) \cup \{0 < c < \frac{1}{n} : n \in \mathbb{N}\}$

For any finite $T_0 \subseteq T$, $\mathbb{R} \models T_0$. By compactness, T has a model.

Let $\mathcal{R}$ be a model of T. Does $\{\frac{1}{n}\}_{n\in\mathbb{N}} \to c^{\mathcal{R}}$?

There are real closed fields where no sequence converges, unless it is eventually constant.

A first proof in Number Theory: The Division Algorithm

Let $a, b \in \mathbb{Z}, b \neq 0$. Then there are $q, r \in \mathbb{Z}$, $0 \leqslant r < |b|$ such that

$$a = qb + r$$

Proof.

Let $S = \{a + kb : k \in \mathbb{Z} \;\&\; a + kb \geqslant 0\}$. $S \neq \emptyset$.

Let $r = \min S$, $r = a + k_0 b$. Set $q = -k_0$

If $b > 0$ and $r \geqslant b$, then $r - b = a + (k_0 - 1)b \geqslant 0$

If $b < 0$ and $r \geqslant -b$, then $r + b = a + (k_0 + 1)b \geqslant 0$

Either way, $r \neq \min S$, contradiction. So $r < |b|$ $\square$

Peano Arithmetic (PA)

Let $L = \{<, +, \times, 0, 1\}$. PA is the L-theory with axioms:

$+$ & $\times$ are commutative, associative, with identities 0 & 1

$<$ is a linear order that agrees with $+$ & $\times$

$\forall xy\, (x < y \;\leftrightarrow\; \exists z\, (z > 0 \wedge x + z = y))$

$\forall x\, (x \geqslant 0 \wedge\; (x > 0 \rightarrow x \geqslant 1))$

$(\varphi(0) \wedge \forall x\, (\varphi(x) \rightarrow \varphi(x+1))) \rightarrow \forall x\, \varphi(x)$

for any $\varphi(x)$

Peano Arithmetic (PA)

Let $L = \{<, +, \times, 0, 1\}$. PA is the L-theory with axioms:

$+$ & $\times$ are commutative, associative, with identities 0 & 1

$<$ is a linear order that agrees with $+$ & $\times$

$\forall xy\,(x < y \;\leftrightarrow\; \exists z\,(z > 0 \wedge x + z = y))$

$\forall x\,(x \geqslant 0 \wedge \;(x > 0 \rightarrow x \geqslant 1))$

$(\varphi(0) \wedge \forall x\,(\varphi(x) \rightarrow \varphi(x+1))) \rightarrow \forall x\,\varphi(x)$

for any $\varphi(x)$

What about negative numbers?

The Integers

Define $\mathbb{Z} = (\mathbb{N} \times \mathbb{N})/\sim$ where

$$(a, b) \sim (c, d) \quad \Leftrightarrow \quad a + d = b + c$$

We can think of $[(a, b)]$ as $a - b$. The following are well-defined:

$$[(a, b)] \oplus [(c, d)] = [(a + c, b + d)]$$

$$[(a, b)] \otimes [(c, d)] = [(ac + bd, ad + bc)]$$

$$[(a, b)] \prec [(c, d)] \quad \Leftrightarrow \quad a + d < b + c$$

$$\overline{0} = [(a, a)]$$

$$\overline{1} = [(a + 1, a)]$$

$(\mathbb{N}, <, +, \times, 0, 1)$ embeds into $(\mathbb{Z}, \prec, \oplus, \otimes, \overline{0}, \overline{1})$ through the map

$$k \;\mapsto\; [(a + k, a)]$$

Nonstandard models of Arithmetic

$(\mathbb{N}, <, +, \times, 0, 1)$ is called **the standard model** of PA.

Any other model of PA is called **nonstandard**.

Theorem. There are nonstandard models of PA.

Proof.

Let $L = \{<, +, \times, 0, 1, c\}$, c a constant symbol.

Let $T = \text{PA} \cup \{c > n : n \in \mathbb{N}\}$ and $T_0 \subseteq T$ finite.

Then $(\mathbb{N}, <, +, \times, 0, 1, m + 1)$ is a model of T_0,

where m is the largest of the $c > n$ axioms in T_0.

By compactness, T has a model $\mathcal{A}$.

$c^{\mathcal{A}}$ is a nonstandard element (and so are $c^{\mathcal{A}} + 1$, etc.) $\square$

Is there a countable nonstandard model? Yes!

How small are the models we can find?

Theorem. (Löwenheim-Skolem $\downarrow$) Let T be a satisfiable L-theory and C be the set of constants in L. Then there is a model of T with cardinality $|C| + \aleph_0$. In particular, if C is at most countable, then T has a countable model.

Proof.

By Henkin's proof of the compactness theorem, there is a model of T of cardinality $|L^*| = |L| + \aleph_0$. □

How big are the models we can find?

Theorem. (Löwenheim-Skolem $\uparrow$) Let T be a satisfiable L-theory and C be the set of constants in L. If T has an infinite model, then there is a model of T with cardinality λ, for each infinite $\lambda > |C|$.

Proof.

Let $\mathcal{A}$ be an infinite model of T and I be a set, $|I| = \lambda > |C|$.

Set $L' = L \cup \{c_i : i \in I\}$ and $T' = T \cup \{c_i \neq c_j : i, j \in I, i \neq j\}$

For any finite $T_0 \subset T'$, $\mathcal{A}$ (infinite) can be made a model of T_0.

By compactness, T' has a model (with cardinality at least λ).

By Henkin's proof, T' (and T) has a model with cardinality λ. $\square$

Limitations of first-order axiomatization

Finite groups, fields, graphs etc. are not first-order axiomatizable.

If a theory has only finite models, their size is bounded.

Proof.

Let T be an L-theory whose models are all finite. Suppose, by a contradiction, that for each $n \in \mathbb{N}$, T has a model $\mathcal{A}_n$, $|A_n| > n$.

Let $L' = L \cup \{c_n : n \in \mathbb{N}\}$, c_n constant symbols. Let $T' = T \cup \{c_i \neq c_j : i \neq j\}$. If $T_0 \subset T$, $|T_0| = n$, then $\mathcal{A}_n \models T_0$.

By compactness, T' has a model $\mathcal{A}$, A is infinite and $\mathcal{A} \models T$. Contradiction. □

Similarly for *torsion groups* or *connected graphs*.

Limitations of first-order axiomatization

Finite groups, fields, graphs etc. are not first-order axiomatizable.

> If a theory has only finite models, their size is bounded.

Proof.

Let T be an L-theory whose models are all finite. Suppose, by a contradiction, that for each $n \in \mathbb{N}$, T has a model $\mathcal{A}_n$, $|A_n| > n$.

Let $L' = L \cup \{c_n : n \in \mathbb{N}\}$, c_n constant symbols. Let $T' = T \cup \{c_i \neq c_j : i \neq j\}$. If $T_0 \subset T$, $|T_0| = n$, then $\mathcal{A}_n \models T_0$.

By compactness, T' has a model $\mathcal{A}$, A is infinite and $\mathcal{A} \models T$. Contradiction. □

Similarly for *torsion groups* or *connected graphs*.

Why we put up with the limitations of first-order logic?

Because of the Compactness Theorem!

Twin prime conjecture

(Polignac, 1849) There are infinitely many primes p such that $p+2$ is also prime. [p & $p+2$ are called **twin primes**].

TPC Holds $\iff$ There is a pair of nonstandard twin primes in a nonstandard model $\mathcal{A} \models Th(\mathbb{N})$

Proof. let $\psi(x)$: "x is a prime" $\qquad \varphi(x) \equiv \psi(x) \wedge \psi(x+2)$

If TPC is TRUE, let $L = \{<, +, \times, 0, 1, c\}$ &

$T = Th(\mathbb{N}) \cup \{\varphi(c) \wedge c > n\}_{n \in \mathbb{N}}$. If $T_0 \subseteq T$ is finite, then $\mathbb{N} \models T_0$

By compactness, T has a model $\mathcal{A}$. $c^{\mathcal{A}}, c^{\mathcal{A}}+2$ are nonstandard twin primes.

If TPC is FALSE let $n_1 = 3$, $n_2 = 5$, $n_3 = 11$, ... n_r so

$\mathbb{N} \models \forall x \left[\varphi(x) \to \bigvee_{i=1}^{r} x = n_i\right]$ so if $\mathcal{A} \models Th(\mathbb{N})$ then

$\mathcal{A}$ has no nonstandard twin primes.

Reading: Model Theory and Number Theory

- Thomas Scanlon, **Diophantine Geometry from Model Theory** (2001).
- Kobi Peterzil & Sergei Starchenko, **Tame complex analysis and o-minimality** (ICM 2010)
- Jonathan Pila, **O-minimality and Diophantine Geometry** (ICM 2014)