

Gödel's Completeness Theorem for First-Order Predicate Calculus

Contents

1	Statement of the Problem	1
2	Preliminaries	1
2.1	Structures and Truth	1
2.2	Consistency	2
3	Extending the Language: Henkin Constants	2
3.1	Why We Need Witnesses	2
3.2	Iterative Extension of the Language	2
4	Lindenbaum's Lemma: Extension to a Complete Set	5
5	Constructing the Term Model	6
5.1	The Carrier	6
5.2	The Nature of the Term Model	6
5.3	Interpreting the Symbols	6
6	The Truth Lemma	7
7	Completing the Proof	8
8	Example: The Ax–Grothendieck Theorem	9

1 Statement of the Problem

Fix a first-order language $\mathcal{L}$ with a *finite signature*: a finite set of predicate symbols $P_1, \dots, P_k$, function symbols $f_1, \dots, f_m$, and constants $a_1, \dots, a_n$. Let $\mathcal{L}^0$ denote the set of all *closed* formulas (sentences) of the language $\mathcal{L}$.

Theorem 1 (Gödel, 1930). *Let $\Gamma \subseteq \mathcal{L}^0$ and $\varphi \in \mathcal{L}^0$. Then*

$$\Gamma \models \varphi \iff \Gamma \vdash \varphi.$$

The direction ($\Leftarrow$) — *soundness* — is proved by a standard induction on derivations, and we omit it. The substantial content of the theorem is the direction ($\Rightarrow$), *completeness*, to which this text is devoted.

Contrapositively, completeness is equivalent to the following statement:

Lemma 2 (Model Existence). *Every consistent set $\Gamma \subseteq \mathcal{L}^0$ is satisfiable, i.e., there exists an $\mathcal{L}$ -structure $\mathfrak{M}$ such that $\mathfrak{M} \models \varphi$ for all $\varphi \in \Gamma$.*

The rest of this text is devoted to the proof of this lemma.

2 Preliminaries

2.1 Structures and Truth

Definition 3. An $\mathcal{L}$ -structure $\mathfrak{M} = \langle M, \{P_i^{\mathfrak{M}}\}, \{f_j^{\mathfrak{M}}\}, \{a_l^{\mathfrak{M}}\} \rangle$ consists of:

- a nonempty set M — the *carrier*;
- an interpretation of each predicate symbol P_i as a relation $P_i^{\mathfrak{M}} \subseteq M^{\text{ar}(P_i)}$;
- an interpretation of each function symbol f_j as a function $f_j^{\mathfrak{M}}: M^{\text{ar}(f_j)} \rightarrow M$;
- an interpretation of each constant a_l as an element $a_l^{\mathfrak{M}} \in M$.

The truth of a closed formula φ in a structure $\mathfrak{M}$ (notation: $\mathfrak{M} \models \varphi$) is defined by the standard recursion on the construction of φ :

- $\mathfrak{M} \models P_i(t_1, \dots, t_k)$ if $\langle t_1^{\mathfrak{M}}, \dots, t_k^{\mathfrak{M}} \rangle \in P_i^{\mathfrak{M}}$;
- $\mathfrak{M} \models \neg\psi$ if $\mathfrak{M} \not\models \psi$;
- $\mathfrak{M} \models \psi_1 \wedge \psi_2$ if $\mathfrak{M} \models \psi_1$ and $\mathfrak{M} \models \psi_2$;
- $\mathfrak{M} \models \exists x \psi(x)$ if there exists $a \in M$ such that $\mathfrak{M} \models \psi(a)$ (where a is treated as a new constant).

2.2 Consistency

A set Γ is called *consistent* if $\Gamma \not\models \perp$, that is, no contradiction can be derived from it.

3 Extending the Language: Henkin Constants

3.1 Why We Need Witnesses

Suppose the formula $\exists x \psi(x)$ is true in a model. Then the model must contain a *specific element* realizing ψ . In a syntactic construction of the model there is no such element — we build the carrier out of terms of the language — so we must decide in advance to “promise” a name for every such element.

3.2 Iterative Extension of the Language

Since adding new constants generates new closed formulas (including new existential ones), the extension of the language is carried out *iteratively*.

We build an increasing chain of languages:

$$\mathcal{L} = \mathcal{L}_0 \subseteq \mathcal{L}_1 \subseteq \mathcal{L}_2 \subseteq \dots$$

Step $n \rightarrow n+1$. Suppose $\mathcal{L}_n$ has already been built. Fix an enumeration of all the existential sentences of the language $\mathcal{L}_n$:

$$\exists x \psi_0^{(n)}(x), \quad \exists x \psi_1^{(n)}(x), \quad \dots$$

Remark 4 (Witness Registry). Among the formulas $\exists x \psi_k^{(n)}(x)$ there may be some for which a witness has already been assigned at an earlier step $i < n$ (for instance, the formula $\exists x \psi(x)$ may literally recur because it entered the language $\mathcal{L}_n$ via several different paths of construction). To avoid wastefully duplicating constants for one and the same formula, we keep a *registry* $W: \{\text{ex. formulas}\} \rightarrow \{\text{constants}\}$: before introducing a new constant for $\exists x \psi_k^{(n)}(x)$, we check whether $W(\exists x \psi_k^{(n)}(x))$ is already defined; if so, we use the existing constant; if not, we introduce a fresh one and record it in W .

This technical refinement does not affect the correctness of the further construction: without the registry (that is, with duplicate constants for recurring formulas) the proof of Lemma 7 and of the truth lemma goes through unchanged, since a “superfluous” constant is simply an additional, formally independent element of the future carrier. The registry only makes the model more economical. It is important not to confuse this with *replacing* an already-introduced constant — the constants $c_{n,k}$ are once fixed and remain unchanged thereafter.

For each such sentence $\exists x \psi_k^{(n)}(x)$ that has no witness in the registry W , we introduce a *fresh* constant $c_{n,k}$, not occurring in any $\mathcal{L}_i$, $i \leq n$. Set

$$\mathcal{L}_{n+1} = \mathcal{L}_n \cup \{c_{n,k} : k \in \mathbb{N}\}.$$

The extended language and the extended axiom set. Define

$$\mathcal{L}_\omega = \bigcup_{n \in \mathbb{N}} \mathcal{L}_n.$$

Lemma 5 (Closure of $\mathcal{L}_\omega$ under Henkin Extension). *For every closed formula of the form $\exists x \psi(x)$ of the language $\mathcal{L}_\omega$, a witnessing constant for it already exists in $\mathcal{L}_\omega$.*

Proof. The formula $\exists x \psi(x)$ is finite and therefore uses only finitely many symbols of the signature. Since $\mathcal{L}_\omega = \bigcup_n \mathcal{L}_n$ is the union of an *increasing* chain of languages, all the symbols occurring in the formula appear by some finite step; hence the formula itself belongs entirely to some $\mathcal{L}_n$. But then $\exists x \psi(x)$ occurred in the enumeration of the existential sentences of the language $\mathcal{L}_n$ at the step where $\mathcal{L}_{n+1}$ was built (see Remark 4), and a constant $c_{n,k}$ was introduced for it (or an existing one reused) together with the Henkin axiom $\exists x \psi(x) \rightarrow \psi(c_{n,k})$. This axiom lies in $H \subseteq \mathcal{L}_\omega^0$. $\square$

Remark 6. It is exactly this lemma that explains why ω steps of extension suffice and no transfinite iteration is needed: any existential formula, *however long it may be*, is still finite, and so it “gets caught” at some finite step n and receives a witness. A step $\omega + 1$ could only be needed in a situation where formulas of infinite length appeared in the language — but a first-order language admits no such thing.

With each pair (n, k) we associate a *Henkin axiom*:

$$\exists x \psi_k^{(n)}(x) \rightarrow \psi_k^{(n)}(c_{n,k}).$$

Denote by H the set of all such axioms and set

$$\Gamma^* = \Gamma \cup H.$$

Lemma 7. Γ^* is consistent in the language $\mathcal{L}_\omega$.

Proof. It suffices to show that adding *one* Henkin axiom to a consistent set preserves consistency; then the statement for all of H follows by induction on the number of axioms actually used in a hypothetical derivation of $\perp$ (the derivation is finite, so there are only finitely many such axioms, and each contains its own fresh constant, not occurring in the others — this is guaranteed by the construction).

Base step. Let Σ be a consistent set of formulas, and let $H \equiv \exists x \psi(x) \rightarrow \psi(c)$ be a Henkin axiom, where c is a fresh constant occurring neither in Σ nor in ψ . We show that $\Sigma \cup \{H\}$ is consistent.

Suppose otherwise:

$$\Sigma \cup \{\exists x \psi(x) \rightarrow \psi(c)\} \vdash \perp.$$

Step 1: the deduction theorem. Moving the Henkin axiom out of the hypotheses, we get

$$\Sigma \vdash (\exists x \psi(x) \rightarrow \psi(c)) \rightarrow \perp,$$

which, by the laws of propositional logic (negation introduction, ex falso quodlibet), is equivalent to a proof of the negation of the implication itself:

$$\Sigma \vdash \exists x \psi(x) \wedge \neg \psi(c).$$

This yields *two independent* derivations from Σ :

$$(i) \quad \Sigma \vdash \exists x \psi(x), \quad (ii) \quad \Sigma \vdash \neg \psi(c).$$

Step 2: replacing the constant by a variable. Consider derivation (ii). Since c occurs nowhere in Σ , nor in ψ (except at the explicitly written substitution site), its presence in the derivation tree is purely nominal: replacing c everywhere in this tree by a fresh variable y , bound by no quantifier, again yields a correct derivation tree (the axioms and rules of logic apply to y exactly as they do to c , since nowhere was it used that this is specifically a constant). Hence

$$\Sigma \vdash \neg \psi(y).$$

Step 3: generalization and contradiction. The variable y occurs free in no formula of Σ (it was chosen fresh), so the rule of universal generalization is legitimately applicable to the derivation $\neg \psi(y)$:

$$\Sigma \vdash \forall y \neg \psi(y).$$

By the laws governing quantifiers, $\forall y \neg \psi(y)$ is logically equivalent to $\neg \exists y \psi(y)$, and, up to renaming the bound variable, to the formula $\neg \exists x \psi(x)$. Hence

$$(iii) \quad \Sigma \vdash \neg \exists x \psi(x).$$

Conclusion. Comparing (i) and (iii): Σ proves both $\exists x \psi(x)$ and $\neg \exists x \psi(x)$ at once, that is, $\Sigma \vdash \perp$ — contradicting the consistency of Σ . Note that in this argument the fresh

constant c of the original theory Σ nowhere appears again: the contradiction is obtained strictly within Σ .

From one axiom to an arbitrary finite set. The base step eliminates exactly one added Henkin axiom. To extend the result to all of H , one must proceed carefully: the derivation $\Gamma \cup H \vdash \perp$ is finite, so it actually uses only a finite list of axioms $H_1, \dots, H_m \in H$, but they *cannot* be eliminated all at once by a single application of the deduction theorem.

Remark 8 (why the “naive” route does not work). Moving the whole list out of the hypotheses at once, we get

$$\Gamma \vdash \neg(H_1 \wedge \dots \wedge H_m) \equiv \Gamma \vdash \neg H_1 \vee \dots \vee \neg H_m.$$

From the provability of a disjunction in classical logic it does *not* follow that at least one disjunct is separately provable — unlike the semantic “truth of a disjunction”, at the level of formal derivability this is not a property of $\vdash$. So the argument cannot be split into the branches $\Gamma \vdash \neg H_i$: the axioms must be eliminated *sequentially*, not all at once.

Sequential elimination is possible thanks to the following observation: the list $H_1, \dots, H_m$ can always be *ordered* so that the constant of axiom H_m is fresh not only for Γ , but also for all the other axioms $H_1, \dots, H_{m-1}$. This follows directly from the construction: constants are introduced level by level, $n = 0, 1, 2, \dots$, and within one level — for different existential formulas, with each constant $c_{n,k}$ used only in *its own* Henkin axiom and never occurring in the formulas on the basis of which earlier constants were introduced. It suffices to order $H_1, \dots, H_m$ by nondecreasing level of their constants (and arbitrarily within a level).

Set $\Sigma_0 = \Gamma$ and $\Sigma_j = \Gamma \cup \{H_1, \dots, H_j\}$ for $j = 1, \dots, m$. By construction, the constant of axiom H_j does not occur in Σ_{j-1} . Applying the base step to the pair (Σ_{j-1}, H_j) , we get: if $\Sigma_{j-1} \cup \{H_j\} = \Sigma_j$ is inconsistent, then so is Σ_{j-1} . Applying this argument sequentially for $j = m, m-1, \dots, 1$ (“peeling off” axioms from the end of the list), we conclude: if $\Sigma_m = \Gamma \cup \{H_1, \dots, H_m\}$ is inconsistent, then so is $\Sigma_0 = \Gamma$.

Thus, if $\Gamma \cup H \vdash \perp$, then $\Gamma \vdash \perp$ — contrary to the assumption that Γ is consistent. Hence $\Gamma \cup H = \Gamma^*$ is consistent. A Henkin axiom may physically occur in a derivation tree of a contradiction, but, as the base step applied one axiom at a time shows, it is structurally always eliminable: the contradiction in which it is used collapses into a shorter contradiction, ultimately obtained directly from Γ . $\square$

4 Lindenbaum’s Lemma: Extension to a Complete Set

Definition 9. A set of sentences $\Delta \subseteq \mathcal{L}_\omega^0$ is called *complete and consistent* (a *CC-set*), if:

- (i) Δ is consistent;
- (ii) for every $\varphi \in \mathcal{L}_\omega^0$ exactly one of $\varphi, \neg\varphi$ belongs to Δ .

Lemma 10 (Lindenbaum). *Every consistent $\Gamma^* \subseteq \mathcal{L}_\omega^0$ can be extended to a CC-set $\Delta \supseteq \Gamma^*$.*

Proof. The language $\mathcal{L}_\omega$ is countable (a union of a countable chain of countable languages), so all the sentences of $\mathcal{L}_\omega^0$ can be enumerated: $\varphi_0, \varphi_1, \varphi_2, \dots$

We build an increasing chain:

$$\Gamma_0 = \Gamma^*, \quad \Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{\varphi_n\} & \text{if } \Gamma_n \cup \{\varphi_n\} \text{ is consistent,} \\ \Gamma_n & \text{otherwise.} \end{cases}$$

Each step is well defined. By the law of excluded middle, either $\Gamma_n \cup \{\varphi_n\}$ is consistent or it is not — there is no third option. (Note: here we do not require the check of consistency to be *decidable* — its *existence* as a dichotomy suffices.)

Set $\Delta = \bigcup_n \Gamma_n$.

Δ is consistent. Suppose $\Delta \vdash \perp$. Then this derivation is finite and uses only formulas from some Γ_N — a contradiction.

Δ is complete. Let $\varphi = \varphi_n$. If $\Gamma_n \cup \{\varphi_n\}$ is consistent, then $\varphi_n \in \Gamma_{n+1} \subseteq \Delta$. Otherwise $\Gamma_n \cup \{\varphi_n\} \vdash \perp$, so $\Gamma_n \vdash \neg\varphi_n$, whence $\neg\varphi_n \in \Delta$. Thus Δ contains exactly one of $\varphi_n, \neg\varphi_n$. $\square$

Remark 11. In the countable case the construction is fully explicit (no transfinite recursion and no axiom of choice). For uncountable languages one would need Zorn’s lemma, but we do not use it here.

Remark 12 (Nonconstructivity of Δ and Multiplicity of Models). The object Δ exists by virtue of classical principles (the law of excluded middle and countable iteration), but in general it is *non-effective*: we cannot algorithmically determine exactly which formulas ended up in it, unless they are explicitly derivable from Γ^* . At each step we make a nonconstructive choice between φ_n and $\neg\varphi_n$ — the existence of such a choice is guaranteed by the law of excluded middle, but the specific outcome is not available to us.

A consequence of this is that different “completions” of Γ to a CC-set produce, in general, *non-isomorphic* models. This is precisely why a consistent theory typically has many different models. This is not a defect of the construction but a reflection of reality: Gödel’s theorem of completeness proves the *existence* of a model, not its uniqueness or constructive presentability.

5 Constructing the Term Model

From now on Δ is a fixed CC-set containing Γ^* .

5.1 The Carrier

Let T be the set of all *closed terms* of the language $\mathcal{L}_\omega$ (terms with no free variables). Define on T the relation:

$$s \sim t \stackrel{\text{def}}{\iff} (s = t) \in \Delta.$$

Lemma 13. $\sim$ is an equivalence relation.

Proof. Reflexivity: $(t = t) \in \Delta$, since $\vdash t = t$. Symmetry and transitivity follow similarly from the fact that Δ contains all the axioms of equality. $\square$

The carrier of the model:

$$M = T/\sim, \quad [t] = \{s \in T : s \sim t\}.$$

5.2 The Nature of the Term Model

The key observation: the carrier $M = T/\sim$ consists of *terms*, not of some abstract objects. So the interpretation of the symbols of the language is not “a computation in a pre-given structure”, but a *definition by syntax*.

In other words: we do not know, and need not know, “what” the Henkin constant $c_{n,k}$ is as an object. It is an element of the model $[c_{n,k}]$ by definition. The question “does $P(c_{n,k})$ hold?” has an answer: “yes, if $P(c_{n,k}) \in \Delta$; no, if $\neg P(c_{n,k}) \in \Delta$ ” — and exactly one of these holds, since Δ is complete. No further “knowledge” about $c_{n,k}$ is required or assumed.

Similarly, the value of a function $f^{\mathfrak{M}}([t_1], \dots, [t_r])$ is simply the class of the term $[f(t_1, \dots, t_r)]$, that is, the application of the function symbol at the level of term syntax. Semantics is entirely reduced to syntax via membership in Δ .

5.3 Interpreting the Symbols

Before specifying the interpretation, we must fix what underlies the formal system: we assume that the predicate calculus with equality includes (besides reflexivity, symmetry, and transitivity) the *congruence axiom scheme* for every function symbol f_j of arity r and every predicate symbol P_i of arity r :

$$\begin{aligned} \forall \vec{x} \forall \vec{y} \left(\bigwedge_{i=1}^r x_i = y_i \rightarrow f_j(\vec{x}) = f_j(\vec{y}) \right), \\ \forall \vec{x} \forall \vec{y} \left(\bigwedge_{i=1}^r x_i = y_i \rightarrow (P_i(\vec{x}) \leftrightarrow P_i(\vec{y})) \right). \end{aligned}$$

This is a standard assumption (Enderton, Mendelson, and others): without it, equality would be merely an arbitrary equivalence relation, not compatible with the rest of the signature, and the definitions below would be ill-defined. Since these axioms are purely logical, they are derivable from an empty set of premises, and hence belong to Δ (as does every theorem of the calculus, in any complete consistent set).

- **Constants.** $a_i^{\mathfrak{M}} = [a_i]$ for every constant $a_i \in \mathcal{L}$ (and likewise for the Henkin constants $c_{n,k}$).
- **Function symbols.**

$$f_j^{\mathfrak{M}}([t_1], \dots, [t_r]) := [f_j(t_1, \dots, t_r)].$$

- **Predicate symbols.**

$$P_i^{\mathfrak{M}}([t_1], \dots, [t_r]) = \top \iff P_i(t_1, \dots, t_r) \in \Delta.$$

Lemma 14 (Well-Definedness). *Both definitions are independent of the choice of representatives of the equivalence classes.*

Proof. Let $s_i \sim t_i$ for all $i = 1, \dots, r$, that is, $(s_i = t_i) \in \Delta$.

Functions. The congruence axiom for f_j is a logical theorem, hence lies in Δ . Applying it (via universal instantiation and modus ponens, using $(s_i = t_i) \in \Delta$ for each i), we obtain $(f_j(s_1, \dots, s_r) = f_j(t_1, \dots, t_r)) \in \Delta$, that is, $f_j(s_1, \dots, s_r) \sim f_j(t_1, \dots, t_r)$, and hence $[f_j(s_1, \dots, s_r)] = [f_j(t_1, \dots, t_r)]$.

Predicates. Similarly, from the congruence axiom for P_i and $(s_i = t_i) \in \Delta$ we obtain $(P_i(s_1, \dots, s_r) \leftrightarrow P_i(t_1, \dots, t_r)) \in \Delta$. Since Δ is complete and closed under derivation, it follows that $P_i(s_1, \dots, s_r) \in \Delta \iff P_i(t_1, \dots, t_r) \in \Delta$, that is, the value of $P_i^{\mathfrak{M}}$ indeed does not depend on the choice of representatives. $\square$

Remark 15. This is not “set-theoretic magic” — well-definedness essentially uses the fact that the congruence axioms explicitly belong to the logical basis of the predicate calculus with equality. If equality were an arbitrary congruent relation, given without this scheme of axioms, the definitions of $f^{\mathfrak{M}}$ and $P^{\mathfrak{M}}$ would need separate justification, or would simply be ill-defined.

6 The Truth Lemma

Lemma 16 (Truth Lemma). *For every closed $\varphi \in \mathcal{L}_\omega^0$:*

$$\mathfrak{M} \models \varphi \iff \varphi \in \Delta.$$

Proof. Induction on the complexity of φ .

Atomic case $\varphi = P_i(t_1, \dots, t_r)$. By the definition of $P_i^{\mathfrak{M}}$: $\mathfrak{M} \models P_i(t_1, \dots, t_r) \iff P_i(t_1, \dots, t_r) \in \Delta$.

Case $\varphi = \neg\psi$. By completeness of Δ : $\neg\psi \in \Delta \iff \psi \notin \Delta$. By the induction hypothesis: $\psi \notin \Delta \iff \mathfrak{M} \not\models \psi$. Altogether: $\neg\psi \in \Delta \iff \mathfrak{M} \models \neg\psi$. ✓

Case $\varphi = \psi_1 \wedge \psi_2$. $\psi_1 \wedge \psi_2 \in \Delta$ iff $\psi_1 \in \Delta$ and $\psi_2 \in \Delta$ (since Δ is closed under derivation in both directions and contains the axioms for conjunction), which by induction is equivalent to $\mathfrak{M} \models \psi_1 \wedge \psi_2$. ✓

Key case $\varphi = \exists x \psi(x)$.

($\Rightarrow$) Suppose $\mathfrak{M} \models \exists x \psi(x)$. Then there exists $[t] \in M$ such that $\mathfrak{M} \models \psi(t)$. By induction $\psi(t) \in \Delta$, whence $\vdash \exists x \psi(x)$, and therefore $\exists x \psi(x) \in \Delta$.

($\Leftarrow$) Suppose $\exists x \psi(x) \in \Delta$. By construction there is a Henkin axiom $\exists x \psi(x) \rightarrow \psi(c)$, which lies in $\Gamma^* \subseteq \Delta$. Modus ponens: $\psi(c) \in \Delta$. By the induction hypothesis $\mathfrak{M} \models \psi(c)$, and hence $\mathfrak{M} \models \exists x \psi(x)$. ✓ $\square$

7 Completing the Proof

Before completing the proof, we shall need a general fact about structures: the truth of a formula depends only on the interpretation of those symbols that actually occur in it, not on the whole signature of the language in which the structure is given.

Lemma 17 (Coincidence of Interpretations). *Let $\mathcal{L}' \subseteq \mathcal{L}''$ be languages (the signature $\mathcal{L}'$ is part of the signature $\mathcal{L}''$), and let $\mathfrak{M}'' = \langle M, I'' \rangle$ be an $\mathcal{L}''$ -structure, where I'' is its interpretation function for the symbols. Define the reduct of $\mathfrak{M}''$ to the signature $\mathcal{L}'$ as*

$$\mathfrak{M}''|_{\mathcal{L}'} = \langle M, I' \rangle, \quad \text{where } I' = I''|_{\text{symbols of } \mathcal{L}'}$$

(the same carrier M , and the interpretation I' is the restriction of I'' to the symbols of $\mathcal{L}'$ only; the interpretations of symbols in $\mathcal{L}'' \setminus \mathcal{L}'$ are simply discarded).

Then for any formula φ of the language $\mathcal{L}'$ (that is, containing no symbols outside $\mathcal{L}'$):

$$\mathfrak{M}'' \models \varphi \iff \mathfrak{M}''|_{\mathcal{L}'} \models \varphi.$$

Proof. Induction on the construction of the formula φ . At the atomic step φ has the form $P(t_1, \dots, t_r)$ or $s = t$ for a symbol P and terms of $\mathcal{L}'$; truth is computed via $I''(P)$ or via $I'(P) = I''(P)$ — the same relation, since $P \in \mathcal{L}'$. Similarly for terms: the values $t^{\mathfrak{M}''}$ and $t^{\mathfrak{M}''|_{\mathcal{L}'}}$ coincide, since all the function symbols and constants occurring in t lie in $\mathcal{L}'$ and are therefore interpreted identically in I'' and in I' . The induction steps for $\neg, \wedge, \exists$ carry the definition of truth verbatim from $\mathfrak{M}''$ to $\mathfrak{M}''|_{\mathcal{L}'}$ and back, since the definition of truth on these connectives does not mention the signature explicitly. $\square$

Remark 18. Informally: restricting the signature is a purely structural operation of trimming away unused interpretations, changing neither the carrier M , nor the behavior of the structure on formulas that do not use the discarded symbols.

Proof of Lemma 2. Let $\Gamma \subseteq \mathcal{L}^0$ be consistent.

Step 1. Extending the language. We build the chain $\mathcal{L}_0 \subseteq \mathcal{L}_1 \subseteq \dots$ and $\mathcal{L}_\omega = \bigcup_n \mathcal{L}_n$ with the Henkin constants, and form $\Gamma^* = \Gamma \cup H$. By Lemma 7, Γ^* is consistent.

Step 2. Extending via Lindenbaum. By Lemma 10 we extend Γ^* to a CC-set $\Delta \subseteq \mathcal{L}_\omega^0$.

Step 3. The term model. We build an $\mathcal{L}_\omega$ -structure $\mathfrak{M}$ with carrier $T/\sim$ and the interpretation of symbols as described in Section 5.

Step 4. Applying the truth lemma. By Lemma 16, for every $\varphi \in \Delta$ we have $\mathfrak{M} \models \varphi$. Since $\Gamma \subseteq \Gamma^* \subseteq \Delta$, we get $\mathfrak{M} \models \Gamma$.

Step 5. Reducing to the original language. We apply Lemma 17 to the languages $\mathcal{L} \subseteq \mathcal{L}_\omega$ and the structure $\mathfrak{M}$:

- (a) the original theory Γ is formulated in the language $\mathcal{L}$;
- (b) we built the extended language $\mathcal{L}_\omega \supseteq \mathcal{L}$, by adding the Henkin constants;
- (c) the term structure $\mathfrak{M}$ is given in the language $\mathcal{L}_\omega$, and by the previous step $\mathfrak{M} \models \Gamma$;
- (d) all the formulas of Γ are written in the language $\mathcal{L}$ (they contain no Henkin constants), so we take the reduct $\mathfrak{M}|_{\mathcal{L}}$ — the same structure with carrier M , but with the interpretation of the symbols $c_{n,k}$ forgotten;
- (e) by Lemma 17, applied to each formula $\varphi \in \Gamma$, we obtain $\mathfrak{M}|_{\mathcal{L}} \models \Gamma$.

□

Corollary 19 (Gödel’s Completeness Theorem). $\Gamma \models \varphi \implies \Gamma \vdash \varphi$.

Proof. If $\Gamma \not\models \varphi$, then $\Gamma \cup \{\neg\varphi\}$ is consistent. By Lemma 2 there exists a model $\mathfrak{M}$ such that $\mathfrak{M} \models \Gamma$ and $\mathfrak{M} \models \neg\varphi$, that is, $\mathfrak{M} \not\models \varphi$. Hence $\Gamma \not\models \varphi$. □

8 Example: The Ax–Grothendieck Theorem

One of the most striking applications of model theory is the proof of the following fact from algebraic geometry.

Theorem 20 (Ax–Grothendieck). *Every injective polynomial map $f: \mathbb{C}^n \rightarrow \mathbb{C}^n$ is surjective.*

Remarkably, the proof itself requires no topology, no analysis, and no explicit construction of a preimage — only model theory.

Proof (Ax, 1968). **Step 1: formalization.** The statement “every injective polynomial map $\mathbb{F}^n \rightarrow \mathbb{F}^n$ is surjective”, for a fixed dimension n and fixed polynomial degrees $(d_1, \dots, d_n)$, is expressed by a single first-order formula in the language of rings:

$$\sigma_{n,\vec{d}} \equiv \forall \vec{a} \left[\left(\forall \vec{x} \forall \vec{y} (f_{\vec{a}}(\vec{x}) = f_{\vec{a}}(\vec{y}) \rightarrow \vec{x} = \vec{y}) \right) \rightarrow \forall \vec{z} \exists \vec{x} f_{\vec{a}}(\vec{x}) = \vec{z} \right],$$

where $f_{\vec{a}}$ denotes the map with the coefficient tuple $\vec{a}$. The full statement of the theorem for $\mathbb{C}$ is $\mathbb{C} \models \sigma_{n,\vec{d}}$ for all n and $\vec{d}$.

Step 2: the case of finite fields. If $\mathbb{F}$ is finite, then $\mathbb{F}^n$ is finite. An injection of a finite set into itself is surjective, for cardinality reasons. Hence, $\mathbb{F} \models \sigma_{n,\vec{d}}$ for every finite field $\mathbb{F}$ and all $n, \vec{d}$.

Step 3: the theory ACF_0 and Vaught's test. Recall that the theory of *algebraically closed fields of characteristic zero* (ACF_0) is described by the axioms of an algebraically closed field together with the axiom scheme $\neg(1 + \dots + 1) = 0$ (one for each prime p). This theory is *complete*: by *Vaught's test*, a theory of countable signature that has no finite models and is κ -categorical for some infinite κ is complete. ACF_0 is categorical in every uncountable cardinality (all algebraically closed fields of characteristic zero of one uncountable cardinality are isomorphic), whence completeness.

In particular, $\mathbb{C}$ is the unique (up to isomorphism) uncountable model of ACF_0 of cardinality $2^{\aleph_0}$, and $\text{ACF}_0 \models \varphi$ is equivalent to $\mathbb{C} \models \varphi$ for any sentence φ .

Step 4: transfer via the completeness theorem. Let a specific $\sigma_{n,\vec{d}}$ be given.

Suppose that $\mathbb{C} \not\models \sigma_{n,\vec{d}}$. Since ACF_0 is complete and $\mathbb{C} \models \text{ACF}_0$, this would mean $\text{ACF}_0 \models \neg\sigma_{n,\vec{d}}$. By the completeness theorem (Theorem 1) there exists a formal derivation $\text{ACF}_0 \vdash \neg\sigma_{n,\vec{d}}$. This derivation is finite, so it uses only finitely many axioms of ACF_0 — in particular, only finitely many axioms $\neg p \cdot 1 = 0$, say for primes $p \leq p_0$.

Hence $\neg\sigma_{n,\vec{d}}$ is already derivable from

$$\text{ACF}_{>p_0} := \text{ACF} \cup \{\neg p \cdot 1 = 0 : p \leq p_0\},$$

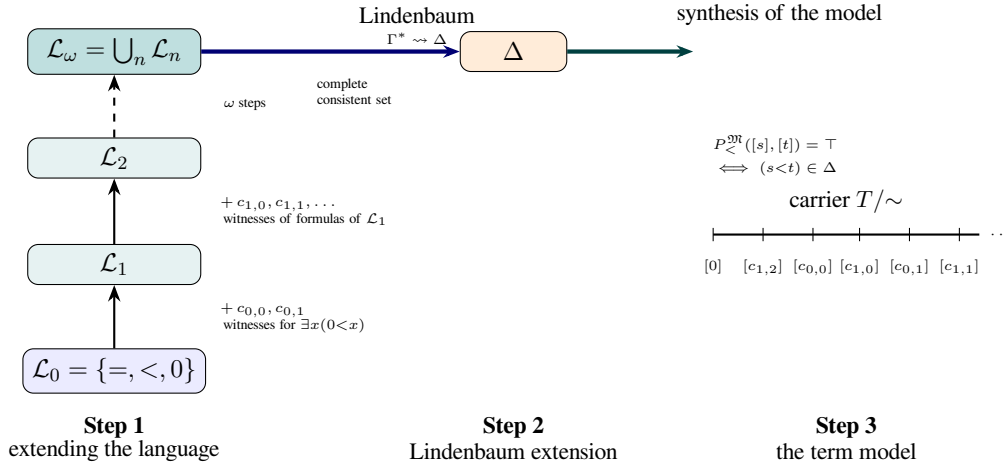
and therefore $\text{ACF}_{>p_0} \models \neg\sigma_{n,\vec{d}}$. But for any prime $p > p_0$, the algebraic closure $\overline{\mathbb{F}_p}$ is a model of $\text{ACF}_{>p_0}$, and hence $\overline{\mathbb{F}_p} \models \neg\sigma_{n,\vec{d}}$. At the same time $\overline{\mathbb{F}_p}$ contains the finite subfields $\mathbb{F}_{p^k}$, and restricting the argument of Step 2 gives $\overline{\mathbb{F}_p} \models \sigma_{n,\vec{d}}$ — a contradiction.

Hence $\mathbb{C} \models \sigma_{n,\vec{d}}$. □

Remark 21 (methodological significance). The proof exhibits a fundamental feature of the model-theoretic method: from the completeness of ACF_0 and the completeness theorem we conclude that there exists a purely algebraic formal derivation of the Ax–Grothendieck theorem from the axioms of ring theory, even though this derivation is nowhere exhibited and would in practice be extraordinarily complicated. Thus model theory establishes the *existence* of a proof without constructing it explicitly — and this is precisely why such a method is called a *transfer* or a *paraphrase* argument: a fact available in one class of structures is carried over to another through semantic mediation.

Illustration: The Construction on a Running Example

The following diagram brings together the three steps of the construction for the language $\mathcal{L} = \{=, <, 0\}$ from Section 7: the growth of the language by levels, the passage to a complete consistent set, and the resulting carrier of the term model, with the order that has already emerged on it.



Summary of the Construction

Step	What We Do
1	Extend the language: add the Henkin constants $c_{n,k}$ iteratively (with the witness registry, Remark 4)
2	Extend Γ with the Henkin axioms — obtaining Γ^* , consistent by Lemma 7
3	Use the Lindenbaum algorithm to build a CC-set $\Delta \supseteq \Gamma^*$
4	The carrier of the model is the set of equivalence classes of closed terms of $\mathcal{L}_\omega$
5	Define the predicates and functions via Δ ; well-definedness is given by Lemma 14 (the congruence axioms)
6	The truth lemma (induction on complexity) gives $\mathfrak{M} \models \Delta$
7	The reduct $\mathfrak{M} _{\mathcal{L}}$ (Lemma 17) is a model of the original Γ