MATHEMATICAL LOGIC

ANUSH TSERUNYAN

These lecture notes introduce the main ideas and basic results of mathematical logic from a fairly modern prospective, providing a number of applications to other fields of mathematics such as algebra, algebraic geometry, and combinatorics. They consist of two parts: basic model theory and basic recursion theory.

Basic model theory. Model theory is a study of mathematical structures, examples of which include groups, rings, fields, graphs, and partial orders. We will first abstractly study structures and definability, theories, models and categoricity, as well as formal proofs, and this will culminate in proofs of the Gödel Completeness and Compactness Theorems—two of the most useful tools of logic.

We will discuss applications of the Compactness theorem in combinatorics, deriving finitary analogues of the infinitary combinatorial statements such as the infinite Ramsey theorem, van der Waerden's or Szemerédi's theorems, graph colorings, etc.

As another application of the Compactness theorem, we briefly discuss nonstandard extensions of the structure of natural numbers and study some of their properties.

On the algebraic side, we apply the developed techniques to algebraically closed fields, which yields a rigorous proof of the Lefschetz Principle (a first-order sentence is true in the field of complex numbers if and only if it is true in all algebraically closed fields of sufficiently large characteristic) and an amusingly slick proof of Ax's theorem (if a polynomial function $\mathbb{C}^n \to \mathbb{C}^n$ is injective, then it is surjective).

Finally, we will study quantifier elimination and model-completeness, and, as a quick application, give a transparent proof of Hilbert's Nullstellensatz.

Basic recursion theory. At the beginning of the 20^{th} century mathematics experienced a crisis due to the discovery of certain paradoxes (e.g. Russell's paradox) in previous attempts to formalize abstract notions of sets and functions. To put analysis on a firm foundation, similar to the axiomatic foundation for geometry, Hilbert proposed a program aimed at a direct consistency proof of analysis. This would involve a system of axioms that is consistent, meaning free of internal contradictions, and complete, meaning rich enough to prove all true statements. But the search for such a system was doomed to fail: Gödel proved in the early 1930s that any system of axioms that can be listed by some "computable process", and subsumes Peano arithmetic, is either incomplete or inconsistent. This is the Gödel Incompleteness theorem. To prove this theorem, we begin with a robust definition of "computable process" (algorithm), followed by a rather short investigation of computable functions and sets. The investigation will be short because we will quickly discover that many interesting functions and sets are not computable, as radiantly illustrated by the Gödel Incompleteness theorem and Church's theorem on undecidability of first-order logic.

Credits. These notes owe a great deal to [Mos08] and [vdD10]; in fact, some parts are almost literally copied from them. I also used my handwritten lecture notes from Matthias Aschenbrenner's model theory course taught at UCLA, as well as [Mar02] and [End01].

Contents

1. First order logic: the semantic aspect

Like any other field of mathematics, mathematical logic starts with a pile of definitions, the importance and use of which will become apparent as we go. Right now, our position is analogous to that of an instructor of geometry who has to define the concept of a differential manifold from scratch without assuming knowledge of point set topology and differentiability. So one has to patiently make his way through the definitions keeping in mind that the end goal is worth it. Let the story begin...

1.A. **Structures**

Every mathematician recognizes a mathematical structure as such when he sees it. Here are some.

Examples 1.1.

(a) A *graph* is a pair $\boldsymbol{G} = (G, E)$, where $G \neq \emptyset$ is the set of nodes and E is a binary relation on G, i.e. $E \subseteq G^2$.

(b) A *partial ordering* is a pair $\boldsymbol{P} = (P, \leq)$, where P is a set and $\leq$ is a binary relation on it satisfying the following conditions:
   (i) (Reflexivity) $\forall x \in P$, $x \leq x$,
   (ii) (Antisymmetry) $\forall x, y \in P$, if $x \leq y$ and $y \leq x$, then $x = y$,
   (iii) (Transitivity) $\forall x, y, z \in P$, if $x \leq y$ and $y \leq z$, then $x \leq z$.

(c) A *group* is a quadruple $\boldsymbol{\Gamma} = (\Gamma, 1, \cdot, ()^{-1})$, where Γ is a set, 1 is a fixed element of G (a constant) and $\cdot, ()^{-1}$ are binary and unary operations on G, respectively, such that the following conditions hold:
   (i) (Associativity) $\forall x, y, z \in G$, $x \cdot (y \cdot z) = (x \cdot y) \cdot z$,
   (ii) (Identity) $\forall x \in G$, $1 \cdot x = x \cdot 1 = x$,
   (iii) (Inverse) $\forall x \in G$, $x \cdot x^{-1} = x^{-1} \cdot x = 1$.

(d) An *ordered field* is a 6-tuple $\boldsymbol{F} = (F, 0, 1, +, \cdot, <)$, where F is a set, $0, 1$ are some fixed elements of F, $+$ and $\cdot$ are binary operations, and $<$ is a binary relation on F such that certain conditions are satisfied (too many to list here).

What is common between these examples? Well, they all have an underlying set together with either relations, operations or constant elements (or all of the above as in Example 1.1(d)) defined on it. Let's formalize this and give an abstract definition of a *mathematical structure*.

Definition 1.2. A *structure* is a quadruple $\boldsymbol{S} = (S, \mathcal{C}, \mathcal{F}, \mathcal{R})$, where S is a set, $\mathcal{C}$ is a set of elements from S (constants), $\mathcal{F}$ is a set of operations on S (i.e. each element of F is a function from S^n to S for some $n \geq 1$) and $\mathcal{R}$ is a set of relations on S (i.e. each element of $\mathcal{R}$ is a subset of S^n for some $n \geq 1$).

Although this definition covers all of the examples above, it is a bit awkward to use when defining a class of structures that have the same format, i.e. the same number of constants, functions, and relations of the same arity. It gets even worse when the structures in that class must also satisfy certain axioms. For example, when defining the class of groups, we not only have to demand that, in those structures, $|\mathcal{C}| = 1$, $|\mathcal{F}| = 2$, $\mathcal{R} = \emptyset$, and one of the operations in $\mathcal{F}$ is binary and the other is unary, but we also have to require that conditions (i)-(iii) of Example 1.1(c) hold. To write these conditions down, we need a coherent system of naming the constants, functions and relations in these structures, i.e. we have to specify that 1 refers to the unique element in $\mathcal{C}$ and $\cdot$ refers to the unique binary element in $\mathcal{F}$. So why don't we first fix a set of names (like $\{1, \cdot, ()^{-1}\}$) and then include their correspondence with the actual constants, functions and relations in the definition of a structure? In fact, that is exactly what we will do.

Definition 1.3. A *signature* is a quadruple

$$\sigma = (\mathcal{C}, \mathcal{F}, \mathcal{R}, \mathfrak{a}),$$

where $\mathcal{C}, \mathcal{F}, \mathcal{R}$ are pairwise disjoint sets (of symbols), which we refer to as the *sets of constant*, *function*, and *relation symbols*, respectively, and

$$\mathfrak{a} : \mathcal{F} \cup \mathcal{R} \to \mathbb{N}^+,$$

which we refer to as the *arity function* and call $P \in \mathcal{F} \cup \mathcal{R}$ an *n-ary* symbol if $\mathfrak{a}(P) = n$. We put $\mathrm{Const}(\sigma) := \mathcal{C}$, $\mathrm{Func}(\sigma) := \mathcal{F}$, $\mathrm{Rel}(\sigma) := \mathcal{R}$, and $\mathfrak{a}_\sigma := \mathfrak{a}$ (although we write $\mathfrak{a}$ is σ is clear from the context). Furthermore, for each $n \in \mathbb{N}^+$, we denote by $\mathrm{Func}_n(\sigma)$ and $\mathrm{Rel}_n(\sigma)$ the sets of n-ary symbols in $\mathrm{Func}(\sigma)$ and $\mathrm{Rel}(\sigma)$, respectively.

The sets $\mathrm{Const}(\sigma), \mathrm{Func}(\sigma), \mathrm{Rel}(\sigma)$ should be thought of as *names* for constant elements, functions (operations), and relations, and not the actual constant elements, functions, and relations themselves! N.b., any of these sets can be empty.

Examples 1.4.

(a) The signature for graphs is
$$\sigma_{\mathrm{gr}} = (\emptyset, \emptyset, \{E\}, (E \mapsto 2)),$$
However, this is too formal and hard to read, so in order to avoid headache (think of a signature for ordered fields!) we simply write
$$\sigma_{\mathrm{gr}} = (E),$$
and then specify that E is a binary relation symbol.

(b) The signature for monoids is
$$\sigma_{\mathrm{mon}} = (1, \cdot),$$
where $\cdot$ is a binary function symbol and 1 is a constant symbol.

(c) The signature for groups is
$$\sigma_{\mathrm{mon}} = (1, \cdot, ()^{-1}),$$
where $\cdot$ and $()^{-1}$ are binary and unary function symbols, respectively, and 1 is a constant symbol.

(d) The signature for rings is
$$\sigma_{\mathrm{ring}} = (0, 1, +, -, \cdot),$$
where $+, -, \cdot$ are binary function symbols and $0, 1$ are constant symbols.

(e) The signature for arithmetic is
$$\sigma_{\mathrm{arthm}} = (0, S, +, \cdot),$$
where 0 is a constant symbol, S is a unary function symbol (S stands for "successor"), and $+, \cdot$ are binary function symbols.

(f) The signature for sets is
$$\sigma_{\mathrm{set}} = (\in),$$
where $\in$ is a binary relation symbol.

Although in this examples the signatures are finite, it is not required by the definition.
Now we are ready to define a structure in a given signature $\sigma = (\mathcal{C}, \mathcal{R}, \mathcal{F})$.

Definition 1.5. A *σ-structure* is a pair $\boldsymbol{S} = (S, \mathfrak{i})$, where S is a set and $\mathfrak{i}$ is a map (correspondence) that assigns

- to each $c \in \mathrm{Const}(\sigma)$, an element $\mathfrak{i}(c)$ of S,
- to each $f \in \mathrm{Func}(\sigma)$ an operation $\mathfrak{i}(f) : S^{\mathfrak{a}(f)} \to S$,
- to each $R \in \mathrm{Rel}(\sigma)$, a relation $\mathfrak{i}(R) \subseteq S^{\mathfrak{a}(R)}$.

We call S the universe of the structure $\boldsymbol{S}$. The choice of the letter $\mathfrak{i}$ is because we think of $\mathfrak{i}$ as the *interpretation* of the symbols of σ in the structure $\boldsymbol{S}$. To simplify the notation, we write $q^{\boldsymbol{S}}$ instead of $\mathfrak{i}(q)$, for all symbols q in σ, so instead of $(S, \mathfrak{i})$, we write
$$\boldsymbol{S} = (S, \left\{c^{\boldsymbol{S}}\right\}_{c \in \mathcal{C}}, \left\{R^{\boldsymbol{S}}\right\}_{R \in \mathcal{R}}, \left\{f^{\boldsymbol{S}}\right\}_{f \in \mathcal{F}}).$$

For finite signatures, we use an even simpler notation as in the following examples.

Examples 1.6.

(a) A complete graph on n vertices is a σ_{gr}-structure
$$\boldsymbol{K_n} = (V, E^{\boldsymbol{K_n}}),$$
where V is a set of n elements (vertices) and $E^{\boldsymbol{K_n}} = V^2 \setminus \mathrm{Id}_V$, $\mathrm{Id}_V := \{(v, v) : v \in V\}$.

(b) $\mathbb{Z}$, as a group, is a σ_{gp}-structure
$$\boldsymbol{Z} = (\mathbb{Z}, 1^{\boldsymbol{Z}}, \cdot^{\boldsymbol{Z}}, \left(()^{-1}\right)^{\boldsymbol{Z}}),$$
where $1^{\boldsymbol{Z}}$ is $0 \in \mathbb{Z}$, $\cdot^{\boldsymbol{Z}}$ is the usual addition operation $(a, b) \mapsto a + b$, and $\left(()^{-1}\right)^{\boldsymbol{Z}}$ is the usual additive inverse operation $a \mapsto -a$.

(c) $\mathbb{R}$, as a field, is a σ_{ring}-structure:

$$\boldsymbol{R} = (\mathbb{R}, 0^{\boldsymbol{R}}, 1^{\boldsymbol{R}}, +^{\boldsymbol{R}}, -^{\boldsymbol{R}}, \cdot^{\boldsymbol{R}}),$$

where all of the symbols are interpreted in the usual way.

(d) Here is a useless σ_{ring}-structure:

$$\boldsymbol{R}_{\text{crazy}} = (\mathbb{R}, 0^{\boldsymbol{R}_{\text{crazy}}}, 1^{\boldsymbol{R}_{\text{crazy}}}, +^{\boldsymbol{R}_{\text{crazy}}}, -^{\boldsymbol{R}_{\text{crazy}}}, \cdot^{\boldsymbol{R}_{\text{crazy}}}),$$

where $0^{\boldsymbol{R}_{\text{crazy}}}, 1^{\boldsymbol{R}_{\text{crazy}}}$ are equal to π, $+^{\boldsymbol{R}_{\text{crazy}}}$ is the $\sin(x+y)$ function, $-^{\boldsymbol{R}_{\text{crazy}}}$ is the $x+y$ function and $\cdot^{\boldsymbol{R}_{\text{crazy}}}$ is the $x+4y$ function. Clearly $\boldsymbol{R}_{\text{crazy}}$ is far from being a ring although it is a structure in the signature of rings.

(e) The structure of natural numbers is a σ_{arthm}-structure and it will be the central object of this course:

$$\boldsymbol{N} = (\mathbb{N}, 0^{\boldsymbol{N}}, S^{\boldsymbol{N}}, +^{\boldsymbol{N}}, \cdot^{\boldsymbol{N}}),$$

where $0^{\boldsymbol{N}}, +^{\boldsymbol{N}}, \cdot^{\boldsymbol{N}}$ are defined in the usual way, and $S^{\boldsymbol{N}}$ is the successor operation (i.e. the unary function of adding 1).

Since it is annoying to keep writing $\boldsymbol{S}$ in the superscript to denote the interpretation of symbols of σ in a σ-structure $\boldsymbol{S}$, we will omit it if the interpretation is the usual/expected one (as suggested by the notation), as long as it is clear that we mean the interpretations rather than the symbols. For example, we will write $\boldsymbol{R} = (\mathbb{R}, 0, 1, +, -, \cdot)$ instead of $\boldsymbol{R} = (\mathbb{R}, 0^{\boldsymbol{R}}, 1^{\boldsymbol{R}}, +^{\boldsymbol{R}}, -^{\boldsymbol{R}}, \cdot^{\boldsymbol{R}})$ if it is the structure in Example 1.6(c), but we won't use this shorthand notation with anything like Example 1.6(d).

In algebra, one of the first things you learn after the definition of a group is the definition of a subgroup, homomorphism and isomorphism. We do the same with arbitrary structures.

Notation 1.7. For any set S and $\vec{a} \in S^n$, let $|\vec{a}|$ denote the length of $\vec{a}$, i.e. $|\vec{a}| := n$. For $i < n$, we refer to the i^{th} coordinate of $\vec{a}$ by writing $(\vec{a})_i$; we also simply write a_i if there is no ambiguity, so $\vec{a} = (a_0, \dots, a_{n-1})$. For a function h defined on S, we write $h(\vec{a})$ for $\big(h(a_0), \dots, h(a_{n-1})\big)$.

Definition 1.8. For σ-structures $\boldsymbol{A}, \boldsymbol{B}$, we say that $\boldsymbol{A}$ is a *substructure* of $\boldsymbol{B}$, written $\boldsymbol{A} \subseteq \boldsymbol{B}$, if $A \subseteq B$ and the interpretations of σ by $\boldsymbol{A}$ and $\boldsymbol{B}$ coincide on A, more precisely:

- $c^{\boldsymbol{A}} = c^{\boldsymbol{B}}$, for each $c \in \text{Const}(\sigma)$;
- $f^{\boldsymbol{A}} = f^{\boldsymbol{B}}|_{A^{\mathfrak{a}(f)}}$, for each $f \in \text{Func}(\sigma)$;
- $R^{\boldsymbol{A}} = R^{\boldsymbol{B}} \cap A^{\mathfrak{a}(R)}$, for each $R \in \text{Rel}(\sigma)$.

For example, $(\mathbb{N}, 0, +)$, $(3\mathbb{Z}, 0, +)$, and $(-6\mathbb{N}, 0, +)$, as well as $(\{0\}, 0, +)$, are substructures of $(\mathbb{Z}, 0, +)$ and the reader is invited to characterize all substructures of $(\mathbb{Z}, 0, +)$. Note that even though $(\mathbb{Z}, 0, +)$ is a group, it's substructure may not be. Similarly, $(\mathbb{R}, 0, 1, +, -, \cdot)$ is a field its substructure $(\mathbb{Z}, 0, 1, +, -, \cdot)$ is not.

For a σ-structure $\boldsymbol{B}$ and $A \subseteq B$, we say that A is *a universe of a substructure* of $\boldsymbol{B}$ if there is a substructure $\boldsymbol{A} \subseteq \boldsymbol{B}$ whose universe is A. For example, if σ only has relation symbols, then any subset $A \subseteq B$ is a universe of a substructure. In particular, if (V, E) is a graph and $U \subseteq V$, then $(U, E \cap U^2)$, i.e. the induced subgraph on U, is a substructure of (V, E). However, note that being a subgraph is not the same as being a substructure of a graph: indeed, a subgraph of a graph (V, E) can be missing some edges between vertices it contains even though these edges may be present in E and this kind of subgraph isn't a substructure of (V, E).

More generally, for a σ-structure $\boldsymbol{B}$ and $A \subseteq B$, we say that A *contains the constants of* $\boldsymbol{B}$ if $c^{\boldsymbol{B}} \in A$ for each $c \in \text{Const}(\sigma)$. We also say that A is *closed under the functions of* $\boldsymbol{B}$ if $f^{\boldsymbol{B}}\big(A^{\mathfrak{a}(f)}\big) \subseteq A$ for each $f \in \text{Func}(\sigma)$. Here a useful characterization of when a subset is a universe of a substructure, whose proof is immediate from the definition of a substructure and is left to the reader.

Lemma 1.9. *Let $\boldsymbol{B}$ be a σ-structure and $A \subseteq B$.*

(1.9.a) *There is at most one substructure of $\boldsymbol{B}$ with universe A.*

(1.9.b) *A is the universe of a substructure of $\boldsymbol{B}$ if and only if A contains the constants of $\boldsymbol{B}$ and is closed under the functions of $\boldsymbol{B}$.*

Proposition 1.10. *Any (finite or infinite, even uncountable) intersection of substructures of the same structure is again a substructure.*

Proof. By (1.9.b), it is enough to check that the intersection still contains all constants and is closed under all functions, which is immediate. □

Let $\boldsymbol{B}$ be a σ-structure and $S \subseteq B$. In the light of Proposition 1.10, we define the *substructure generated by* S as the smallest substructure $\boldsymbol{A}$ containing S, namely: the intersection of all substructures of $\boldsymbol{B}$ that contain S; we denote this substructure by $\langle S\rangle_{\boldsymbol{B}}$. We have the following constructive characterization of the universe of $\langle S\rangle_{\boldsymbol{B}}$.

Proposition 1.11. *Let $\boldsymbol{B}$ be a σ-structure and $S \subseteq B$.*

(1.11.a) *The underlying set of $\langle S\rangle_{\boldsymbol{B}}$ is $S_\infty := \bigcup_{n\in\mathbb{N}} S_n$, where $S_0 := S \cup \left\{c^{\boldsymbol{B}} : c \in \mathrm{Const}(\sigma)\right\}$ and*

$$S_{n+1} = S_n \cup \bigcup_{f\in \mathrm{Func}(\sigma)} f^{\boldsymbol{B}}\left(S_n^{\mathfrak{a}(f)}\right).$$

(1.11.b) $|\langle S\rangle_{\boldsymbol{B}}| \le |S\times\mathbb{N}\times\sigma| = \max\left\{|S|, \aleph_0, |\sigma|\right\}$, *where* $\aleph_0 := |\mathbb{N}|$.

Proof. The fact that $\langle S\rangle_{\boldsymbol{B}} \supseteq S_\infty$ is due to (1.9.b). To obtain the converse it is enough to show, again by (1.9.b), that S_∞ is closed under the functions of $\boldsymbol{B}$. This is immediate because each function of $\boldsymbol{B}$ is *finitely based* (i.e. takes finitely-many inputs) and for any vector $\vec{a} \in S_\infty^n$, there is some $m \in \mathbb{N}$ such that $\vec{a} \in S_m^n$. □

For example, the substructure of $\boldsymbol{R} = (\mathbb{R}, 0, 1, +, \cdot)$ generated by $\emptyset$ is $(\mathbb{N}, 0, 1, +, \cdot)$ (why?).

Definition 1.12. Let $\boldsymbol{A}, \boldsymbol{B}$ be σ-structures. A function $h : A \to B$ is called a *σ-homomorphism* (or just *homomorphism*) if h respects the interpretation of σ, more precisely:

- $h(c^{\boldsymbol{A}}) = c^{\boldsymbol{B}}$, for each $c \in \mathrm{Const}(\sigma)$;
- $h\left(f^{\boldsymbol{A}}(\vec{a})\right) = f^{\boldsymbol{B}}\left(h(\vec{a})\right)$, for each $f \in \mathrm{Func}(\sigma)$ and $\vec{a} \in A^{\mathfrak{a}(f)}$;
- $R^{\boldsymbol{A}}(\vec{a}) \Rightarrow R^{\boldsymbol{B}}\left(h(\vec{a})\right)$, for each $R \in \mathrm{Rel}(\sigma)$ and $\vec{a} \in A^{\mathfrak{a}(R)}$.

Denote this by $h : \boldsymbol{A} \to \boldsymbol{B}$.

Note that in this definition, we only require $\Rightarrow$ for relations. To realize why, think about when R is the equality relation: require the reverse implication $\Leftarrow$ to also hold, would be equivalent to requiring all homomorphisms to be injection. Another justification of this asymmetry is the fact that if we look at the graphs of functions $f^{\boldsymbol{A}}$ and $f^{\boldsymbol{B}}$ as $\left(\mathfrak{a}(f)+1\right)$-ary relations $R_f^{\boldsymbol{A}}$ and $R_f^{\boldsymbol{B}}$, then, putting $b = f^{\boldsymbol{A}}(\vec{a})$, the condition $h\left(f^{\boldsymbol{A}}(\vec{a})\right) = f^{\boldsymbol{B}}\left(h(\vec{a})\right)$ is equivalent to $R_f^{\boldsymbol{A}}(\vec{a}, b) \Rightarrow R_f^{\boldsymbol{B}}\left(h(\vec{a}), h(b)\right)$.

Corollary 1.13. *For any $h : \boldsymbol{A} \to \boldsymbol{B}$, $h(A)$ is a universe of a substructure of $\boldsymbol{B}$.*

Proof. The following immediately follows from the definition of homomorphism and (1.9.b). □

Definition 1.14. Let $\boldsymbol{A}, \boldsymbol{B}$ be σ-structures. A function $h : A \to B$ is called a *σ-isomorphism* (or just isomorphism) if h is bijective and both h and h^{-1} are σ-homomorphisms; in this case we write $h : \boldsymbol{A} \xrightarrow{\sim} \boldsymbol{B}$. The structures $\boldsymbol{A}, \boldsymbol{B}$ are called isomorphic if there is an isomorphism between them; denote this by $\boldsymbol{A} \cong \boldsymbol{B}$.

Definition 1.15. Let $\boldsymbol{A}, \boldsymbol{B}$ be σ-structures and $h : \boldsymbol{A} \to \boldsymbol{B}$. Recalling that $h(A)$ is the universe of a substructure $\boldsymbol{B}' \subseteq \boldsymbol{B}$, we call h a *σ-embedding* (or just *embedding*) if h is an isomorphism between $\boldsymbol{A}$ and $\boldsymbol{B}'$. We denote this by $h : \boldsymbol{A} \hookrightarrow \boldsymbol{B}$.

Observation 1.16. *A σ-homomorphism $h : \boldsymbol{A} \to \boldsymbol{B}$ is a σ-embedding if and only if it is injective and $R^{\boldsymbol{A}}(\vec{a}) \Leftrightarrow R^{\boldsymbol{B}}(h(\vec{a}))$ for all $R \in \mathrm{Rel}(\sigma)$ and $\vec{a} \in A^{\mathfrak{a}(R)}$.*

Note that if $\boldsymbol{A} \subseteq \boldsymbol{B}$ then the inclusion map is an embedding. This wouldn't be true if in the definition of substructure we had $\Rightarrow$ for relations instead of $\Leftrightarrow$.

Sometimes in algebra we consider the universe of a ring as an abelian group under addition, in other words, we "forget" the multiplication operation. We make this precise here.

Definition 1.17. Let σ, σ' be signatures with $\sigma \subseteq \sigma'$, let $\boldsymbol{A}$ be a σ-structure and $\boldsymbol{B}$ be a σ'-structure. We say that $\boldsymbol{A}$ is a *reduct* of $\boldsymbol{B}$ (or $\boldsymbol{B}$ an *expansion* of $\boldsymbol{A}$), written $\boldsymbol{A} = \boldsymbol{B}|_\sigma$, if $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same underlying set and the same interpretations of the symbols of σ.

For example, $(\mathbb{R}, 0, +)$ is a reduct of $(\mathbb{R}, 0, 1, +, \cdot)$, which in its turn is a reduct of $(\mathbb{R}, 0, 1, +, -, \cdot, <)$.

1.B. **Language and interpretation**

Now we have to define the language of First Order Logic (FOL) that will allow us to express statements about τ-structures, like axioms (i)-(iii) in Example 1.1(c). Although the definitions below are very natural, they are somewhat annoying to write and even to read. The readers are advised to try to come up with the definitions themselves before (instead of?) reading.

Let σ denote a signature for the rest of the section.

Definition 1.18. The *alphabet* $\mathbb{FOL}(\sigma)$ of the first order language in the signature σ consists of the symbols in σ and the following additional symbols:

- logical symbols $\doteq\ \neg\ \wedge\ \vee\ \to\ \forall\ \exists$
- punctuation symbols , ( )
- symbols for variables $v_0, v_1, v_2, \dots$

The symbols $\forall$ and $\exists$ are called *quantifiers*. Below, finite sequences of symbols from $\mathbb{FOL}(\sigma)$ are referred to as *words in* $\mathbb{FOL}(\sigma)$ or $\mathbb{FOL}(\sigma)$-words.

Remark 1.19. We use $\doteq$ in $\mathbb{FOL}(\sigma)$ instead of the regular symbol $=$ for equality to avoid ambiguity and confusion with the usual equality. More precisely, for $\mathbb{FOL}(\sigma)$-words w_1, w_2, we would write $w_1 = w_2$ to mean that these words are equal (are literally the same sequences of symbols); this unambiguously reads as a *statement* (for us, humans) about the words w_1 and w_2 and is not confused with the $\mathbb{FOL}(\sigma)$-*word* $w_1 \doteq w_2$.

Definition 1.20. A *σ-term* (or a *term in* $\mathbb{FOL}(\sigma)$) is a word formed via the following recursive rules:

(1.20.i) each $c \in \mathrm{Const}(\sigma)$ is a term;

(1.20.ii) each variable is a term;

(1.20.iii) if $t_1, \dots, t_n$ are terms and $f \in \mathrm{Func}_n(\sigma)$, then $f(t_1, \dots, t_n)$ is a term.

We let $\mathrm{Terms}(\sigma)$ denote the set of all σ-terms.

Examples 1.21.

(a) $(v_0 \cdot 1)^{-}1 \cdot v_3$ is a term in $\mathbb{FOL}(\sigma_{\mathrm{gp}})$. Note that the way this term is written is technically incorrect, we should have written $\cdot((\,)^{-1}(\cdot(v_0, 1)), v_3)$, but the latter is almost impossible to read, so we will keep abusing notation and write the former way.

(b) $S(0 + v_2) + S(S(S(v_2)))$ is a term in $\mathbb{FOL}(\sigma_{\mathrm{arthm}})$. For each $n \in \mathbb{N}$ (think $n := 7$) and σ_{arthm}-term t, we use abbreviation $S^n(t)$ for $\underbrace{S(S(\dots S(t)\dots))}_{n \text{ times}}$.

(c) Variables $v_0, v_1, \dots$ are the only terms in $\mathbb{FOL}(\sigma_{\mathrm{gr}})$.

We also casually use letters different than $v_0, v_1, \dots$ to denote variables, e.g. v, u, x, y, z.

We will naturally interpret terms in a structure as functions, but before giving a precise definition, we recall that one interprets a polynomial p over, say, $\mathbb{R}$ as function on $\mathbb{R}^n$, where the inputs are values from $\mathbb{R}$ assigned to the variables that appear in p. However, one can also introduce dummy variables that don't appear p and include them as part of the input yielding a function defined on a higher power of $\mathbb{R}$. For example, the polynomial $p := x^2 + xy - 3$ viewed as a polynomial in (x, y) is interpreted as a function on $\mathbb{R}^2$, but we can also view p as a polynomial in (x, y, z), in which case it would be interpreted as a function on $\mathbb{R}^3$. To make this distinction clear, we write $p(x, y)$ in the former and $p(x, y, z)$ in the latter cases.

Definition 1.22. Let t be a σ-term and let $\vec{v} := (v_{k_0}, v_{k_1}, \dots, v_{k_{n-1}})$ be a vector of variables of $\mathbb{FOL}(\sigma)$, so $(\vec{v})_i = v_{k_i}$. We call the word $t[\vec{v}]$ an *extended σ-term* if $\vec{v}$ includes all of the variables that appear in t. We let $\mathrm{ExtTerms}(\sigma)$ denote the set of all extended σ-terms.

Definition 1.23. Let $\boldsymbol{A}$ be a σ-structure and $t[\vec{v}]$ be an extended σ-term. We define the *interpretation* of $t[\vec{v}]$ in $\boldsymbol{A}$ as a function $t^{\boldsymbol{A}}[\vec{v}] : A^{|\vec{v}|} \to A$ by induction on the construction of t as follows: for $\vec{a} \in A^{|\vec{v}|}$,

(1.23.i) if $t = c$, where $c \in \mathrm{Const}(\sigma)$, then $t^{\boldsymbol{A}}[\vec{v}](\vec{a}) = c^{\boldsymbol{A}}$;

(1.23.ii) if $t = (\vec{v})_i$, then $t^{\boldsymbol{A}}[\vec{v}](\vec{a}) = (\vec{a})_i$;

(1.23.iii) if $t = f(t_1, \dots, t_n)$, where $t_1, \dots, t_n$ are terms and $f \in \mathrm{Func}_n(\sigma)$, then

$$t^{\boldsymbol{A}}[\vec{v}](\vec{a}) = f^{\boldsymbol{A}}\Big(t_1^{\boldsymbol{A}}[\vec{v}](\vec{a}), \dots, t_n^{\boldsymbol{A}}[\vec{v}](\vec{a})\Big).$$

So one should think of $t[\vec{v}]$ as a name of the function $t^{\boldsymbol{A}}[\vec{v}] : A^{|\vec{v}|} \to A$. Note that if $t = v_1$, then $t[v_1]$ is interpreted as a unary function, while $t[v_1, v_2]$ as a binary function (although it does not depend on v_2).

Example 1.24. Let $\boldsymbol{A} := (\mathbb{Z}, 1^{\boldsymbol{A}}, \cdot^{\boldsymbol{A}})$ and $\boldsymbol{B} := (\mathbb{Z}, 1^{\boldsymbol{B}}, \cdot^{\boldsymbol{B}})$, where $1^{\boldsymbol{A}} := 1 \in \mathbb{Z}$ and $\cdot^{\boldsymbol{A}}$ is the usual multiplication, whereas $1^{\boldsymbol{B}} := 0 \in \mathbb{Z}$ and $\cdot^{\boldsymbol{B}}$ is the usual addition. Letting $t := (v_1 \cdot 1) \cdot v_3$, we see that $t^{\boldsymbol{A}}[v_1, v_3](2, 3) = 6$, whereas $t^{\boldsymbol{B}}[v_1, v_3](2, 3) = 5$. Also, $t^{\boldsymbol{A}}[v_1, v_3]$ is the map $(a, b) \mapsto ab$ from $\mathbb{Z}^2$ to $\mathbb{Z}$, whereas $t^{\boldsymbol{A}}[v_1, v_2, v_3]$ is the map $(a, c, b) \mapsto ab$ from $\mathbb{Z}^3$ to $\mathbb{Z}$.

Definition 1.25. A *σ-formula* (or a *formula in* $\mathbb{FOL}(\sigma)$) is a word formed via the following recursive rules:

(1.25.i) if s, t are terms, then $s \doteq t$ is a formula;
(1.25.ii) if $t_1, \dots, t_n$ are terms and $R \in \mathrm{Rel}_n(\sigma)$, then $R(t_1, \dots, t_n)$ is a formula;
(1.25.iii) if φ and ψ are formulas, then $\neg(\varphi)$, $(\varphi) \wedge (\psi)$, $(\varphi) \vee (\psi)$, $(\varphi) \to (\psi)$ are formulas;
(1.25.iv) if φ is a formula and v a variable symbol, then $\forall v \varphi$, $\exists v \varphi$ are formulas.

Let $\mathrm{Formulas}(\sigma)$ denote the set of all σ-formulas.

The formulas in (1.25.i) and (1.25.ii) are called *atomic*. Also, if a formula is formed without using (1.25.iv), it is called *quantifier free* (or *q.f.* for short).

According to Definition 1.25, $\Big(\forall x(x \doteq y)\Big) \wedge \Big(\neg(x \doteq z)\Big)$ is a formula (in any signature), although, as one may guess, the third occurrence of x has nothing to do with its first two occurrences, where x is used as the variable of the quantifier $\forall$. The use of x as the variable for the quantifier is a bad idea because it makes reading of the formula hard and confusing—imagine writing $x \int_0^1 x dx$ instead of $x \int_0^1 t dt$ in a calculus course! Thus, we make a convention to not use such bad notation.

Convention 1.26. We say that the variable v is *quantified* in the formula φ if in its construction rule (1.25.iv) was used with the variable v, i.e. φ contains a subformula of the form $\forall v \psi$ or $\exists v \psi$. We make the convention that each variable v can be used with a quantifier only once, i.e. a subword of the form $\mathrm{Q} v \psi$ occurs at most once, where Q is either $\forall$ or $\exists$, and if it does appear, then v is not allowed to appear elsewhere other than in ψ.

This convention disqualifies words like $\Big(\forall x(x \doteq y)\Big) \wedge \Big(\neg(x \doteq z)\Big)$ as formulas; one should write $\Big(\forall t(t \doteq y)\Big) \wedge \Big(\neg(x \doteq z)\Big)$ instead.

A variable v is *free* in a formula φ if it occurs in φ and is not quantified. A formula without free variables is called a *sentence*. Note that all statements (theorems, conjectures, etc.) in mathematics are sentences (in the signature of set theory).

Example 1.27. In the formula $\Big(\forall t(t \doteq y)\Big) \wedge \Big(\neg(x \doteq z)\Big)$, t is quantified, whereas x, y, z are free. The formula $\forall x \forall y \forall z \Big(\Big(\forall t(t \doteq y)\Big) \wedge \Big(\neg(x \doteq z)\Big)\Big)$ is a sentence.

Below we will abbreviate $\neg(t_1 \doteq t_2)$ by $t_1 \neq t_2$. Furthermore, for a vector of variables $\vec{v} := (v_{k_0}, v_{k_1}, \dots, v_{k_{n-1}})$, we write $\forall \vec{v}$ to mean $\forall v_{k_0} \forall v_{k_1} \dots \forall v_{k_{n-1}}$ and similarly for $\exists$.

We will naturally interpret formulas in a structure as relations. Just like with terms, the arity of this relation will depend on the number of additional dummy variables involved in the interpretation.

Definition 1.28. Let φ be a σ-formula and let $\vec{v}$ of variables of $\mathbb{FOL}(\sigma)$. We call the word $\varphi[\vec{v}]$ an *extended σ-formula* if $\vec{v}$ includes all of the free variables of φ and does not contain any variable that is quantified in φ. Let $\mathrm{ExtFormulas}(\sigma)$ denote the set of all extended σ-formulas.

Proposition 1.29. $|\mathrm{ExtFormulas}(\sigma)| = |\mathrm{Formulas}(\sigma)| = \max\{\aleph_0, \sigma\}$.

Proof. Follows by simple cardinal arithmetic. The main fact used is that for infinite sets A, B, $|A \times B| = \max\{|A|, |B|\}$. The proof of this fact uses the Axiom of Choice. □

Notation 1.30. For any set A, we put $A^0 := \{\emptyset\}$, so A^0 has exactly one element and its definition has nothing to do with A. This allows thinking of the interpretations of constant symbols as 0-ary functions $A^0 \to A$. Also, a 0-ary relation R on A is just a subset of $A^0 = \{\emptyset\}$, so it is either true (i.e. $R = A^0$) or false (i.e. $R = \emptyset$).

Definition 1.31. Let $\boldsymbol{A}$ be a σ-structure and $\varphi[\vec{v}]$ an extended σ-formula. We define the *interpretation* of $\varphi[\vec{v}]$ in $\boldsymbol{A}$ as a $|\vec{v}|$-ary relation $\varphi^{\boldsymbol{A}}[\vec{v}]$ on A by induction on the construction of φ as follows: for $\vec{a} \in A^{|\vec{v}|}$,

(1.31.i) if $\varphi = t_1 \doteq t_2$, then $\varphi^{\boldsymbol{A}}[\vec{v}](\vec{a})$ holds if and only if $t_1^{\boldsymbol{A}}[\vec{v}](\vec{a}) = t_2^{\boldsymbol{A}}[\vec{v}](\vec{a})$;

(1.31.ii) if $\varphi = R(t_1, \dots, t_k)$, then $\varphi^{\boldsymbol{A}}[\vec{v}](\vec{a})$ holds if and only if $R^{\boldsymbol{A}}(t_1^{\boldsymbol{A}}[\vec{v}](\vec{a}), \dots, t_k^{\boldsymbol{A}}[\vec{v}](\vec{a}))$ holds;

(1.31.iii) if $\varphi = \neg\psi$, then $\varphi^{\boldsymbol{A}}[\vec{v}](\vec{a})$ holds if and only if $\psi^{\boldsymbol{A}}[\vec{v}](\vec{a})$ fails;

(1.31.iv) if $\varphi = \psi \wedge \theta$, then $\varphi[\vec{v}](\vec{a})$ holds if and only if $\psi^{\boldsymbol{A}}[\vec{v}](\vec{a})$ and $\theta^{\boldsymbol{A}}[\vec{v}](\vec{a})$ both hold;

(1.31.v) if $\varphi = \psi \vee \theta$, then $\varphi[\vec{v}](\vec{a})$ holds if and only if $\psi^{\boldsymbol{A}}[\vec{v}](\vec{a})$ or $\theta^{\boldsymbol{A}}[\vec{v}](\vec{a})$ holds;

(1.31.vi) if $\varphi = \forall u \psi(\vec{v}, u)$, in particular u is not in $\vec{v}$ and $\psi[\vec{v}, u]$ is an extended formula, then $\varphi[\vec{v}](\vec{a})$ holds if and only if for each $b \in A$, $\psi^{\boldsymbol{A}}[\vec{v}, u](\vec{a}, b)$ holds;

(1.31.vii) if $\varphi = \exists u \psi(\vec{v}, u)$, in particular u is not in $\vec{v}$ and $\psi[\vec{v}, u]$ is an extended formula, then $\varphi[\vec{v}](\vec{a})$ holds if and only if there exists $b \in A$ such that $\psi^{\boldsymbol{A}}[\vec{v}, u](\vec{a}, b)$ holds.

We also say that $\boldsymbol{A}$ *satisfies* (or *models*) $\varphi[\vec{v}](\vec{a})$, written $\boldsymbol{A} \models \varphi[\vec{v}](\vec{a})$, to mean that $\varphi^{\boldsymbol{A}}[\vec{v}](\vec{a})$ holds.

Note that the above definition applies when φ is a sentence and $\vec{v} = \emptyset$. In this case, $\boldsymbol{A} \models \varphi$ is a 0-ary relation on $\boldsymbol{A}$ (see Notation 1.30), so it is true or false and we read $\boldsymbol{A} \models \varphi$ as "φ is *true* (or *holds*) in $\boldsymbol{A}$".

Note that some of the logical symbols we use are redundant: we could restrict to using only $\neg, \vee, \exists$ and the rest would be expressible in terms of these. So what we usually do is the following: we use all of the symbols when it is convenient, but in our inductive proofs we only take care of the cases with $\neg, \vee, \exists$ or $\neg, \wedge, \forall$ or other equivalent combinations.

Convention 1.32. Instead of writing $t[\vec{v}](\vec{a})$ and $\varphi[\vec{v}](\vec{a})$, we simply write $t(\vec{a})$ and $\varphi(\vec{a})$ whenever it causes no ambiguity or confusion. Furthermore, we will drop the word "extended" and simply call $t[\vec{v}]$ and $\varphi[\vec{v}]$ a term and a formula. Lastly, we write $t(\vec{v})$ and $\varphi(\vec{v})$ for an extended term $t[\vec{v}]$ and an extended formula $\varphi[\vec{v}]$ as the use of round brackets is more familiar from, say, algebra; e.g. for a polynomial p in variables x, y, we commonly write $p(x, y, z)$ to view it as a polynomial in x, y, z.

Examples 1.33.

(a) For a σ_{arthm}-formula $\varphi := S(S(0)) \doteq v_0$, $\boldsymbol{N} := (\mathbb{N}, 0, S, +, \cdot) \models \varphi[v_0](2)$. We may also simply write $\boldsymbol{N} \models S^2(0) \doteq 2$.

(b) In this example we define a σ_{arthm}-sentence expressing the Goldbach conjecture. Let:
- $x \leq y$ stand for $\exists z(z + x \doteq y)$; similarly, for $\geq$; one also easily defines $<$ and $>$;
- $x|y$ (divides) stand for $\exists z(y \doteq z \cdot x)$;
- $\dot{n}$ stand for $S^n(0)$, for each $n \in \mathbb{N}$; for example, $\dot{0}$ and $\dot{2}$ stand for 0 and $S(S(0))$, respectively;
- $\mathbf{Even}(x)$ stand for $\dot{2}|x$;
- $\mathbf{Prime}(x)$ stand for $\forall y\big(y|x \to (y \doteq 1 \vee y \doteq x)\big)$;
- **GoldbachConj** stand for $\forall x\Big(x > 2 \wedge \mathbf{Even}(x) \to \exists(y, z)(\mathbf{Prime}(y) \wedge \mathbf{Prime}(z) \wedge x \doteq y + z)\Big)$.

As of Nov 28, 2017, we still don't know whether $\boldsymbol{N} \models$ **GoldbachConj** or not.

(c) Let $\boldsymbol{N}_{\exp} := (\mathbb{N}, 0, S, +, \cdot, \exp)$, where $0, S, +, \cdot$ are interpreted as usual and exp is the binary exponentiation function: $\exp(n, m) := n^m$ for nonzero n and $\exp(0, m) := 0$. Thanks to A. Wiles, we now know that $\boldsymbol{N}_{\exp} \models \forall n \forall x \forall y \forall z[(n \geq 3 \wedge \exp(x, n) + \exp(y, n) \doteq \exp(z, n)) \to (x \doteq 0 \vee y \doteq 0)]$, where $n \geq \dot{3}$ stands for $n \neq \dot{0} \wedge n \neq \dot{1} \wedge n \neq \dot{2}$.

(d) $\boldsymbol{R} \models \exists y(a \doteq y \cdot y)$ holds for all non-negative $a \in \mathbb{R}$.

Lemma 1.34. *Let $\boldsymbol{A}, \boldsymbol{B}$ be two σ-structures. If $h : \boldsymbol{A} \to \boldsymbol{B}$ is a homomorphism, then for any term $t(\vec{v})$ and $\vec{a} \in A^{|\vec{v}|}$,*

$$h(t^{\boldsymbol{A}}(\vec{a})) = t^{\boldsymbol{B}}(h(\vec{a})).$$

Proof. We prove by induction on the construction (length) of t.

- If $t = c$ for $c \in \mathrm{Const}(\sigma)$, then $t^{\mathbf{A}}(\vec{a}) = c^{\mathbf{A}}$ and hence we have
$$h(t^{\mathbf{A}}(\vec{a})) = h(c^{\mathbf{A}}) = c^{\mathbf{B}} = t^{\mathbf{B}}(h(\vec{a}))$$
because h is a homomorphism.
- If $t = (\vec{v})_i$, then $t^{\mathbf{A}}(\vec{a}) = (\vec{a})_i$ and hence we have
$$h(t^{\mathbf{A}}(\vec{a})) = h((\vec{a})_i) = t^{\mathbf{B}}(h(\vec{a})).$$
- If $t = f(t_1, \dots, t_n)$ for $f \in \mathrm{Func}_n(\sigma)$, then
$$\begin{aligned} h\Big(t^{\mathbf{A}}(\vec{a})\Big) &= h\Big(f^{\mathbf{A}}\Big(t_1^{\mathbf{A}}(\vec{a}), \dots, t_n^{\mathbf{A}}(\vec{a})\Big)\Big) \\ \Big[h \text{ is a homomorphism}\Big] \quad &= f^{\mathbf{B}}\Big(h\Big(t_1^{\mathbf{A}}(\vec{a})\Big), \dots, h\Big(t_n^{\mathbf{A}}(\vec{a})\Big)\Big) \\ \Big[\text{by induction}\Big] \quad &= f^{\mathbf{B}}\Big(t_1^{\mathbf{B}}\Big(h(\vec{a})\Big), \dots, t_n^{\mathbf{B}}\Big(h(\vec{a})\Big)\Big) \\ &= t^{\mathbf{B}}\Big(h(\vec{a})\Big). \end{aligned}$$
□

Proposition 1.35. *Let $\mathbf{A}, \mathbf{B}$ be two σ-structures. If $h : \mathbf{A} \to \mathbf{B}$ is an isomorphism, then for any formula $\varphi(\vec{v})$ and $\vec{a} \in A^{|\vec{v}|}$,*
$$\mathbf{A} \models \varphi(\vec{a}) \iff \mathbf{B} \models \varphi(h(\vec{a})).$$

Proof. We prove by induction on the construction of φ. For the induction step, it is enough to consider only the following cases: $\varphi = \neg\psi$, $\varphi = \psi_1 \wedge \psi_2$, and $\varphi = \exists v \psi$.

- If $\varphi = t_1 \doteq t_2$, then
$$\begin{aligned} \mathbf{A} \models \varphi(\vec{a}) &\iff t_1^{\mathbf{A}}(\vec{a}) = t_2^{\mathbf{A}}(\vec{a}) \\ \Big[h \text{ is injective}\Big] \quad &\iff h\Big(t_1^{\mathbf{A}}(\vec{a})\Big) = h\Big(t_2^{\mathbf{A}}(\vec{a})\Big) \\ \Big[\text{by Lemma 1.34}\Big] \quad &\iff t_1^{\mathbf{B}}\Big(h(\vec{a})\Big) = t_2^{\mathbf{B}}\Big(h(\vec{a})\Big) \\ &\iff \mathbf{B} \models \varphi\Big(h(\vec{a})\Big). \end{aligned}$$
- If $\varphi = R(t_1, \dots, t_n)$, then the calculation is similar to the previous case, but does not use the injectivity of h.
- If $\varphi = \neg\psi$, then
$$\begin{aligned} \mathbf{A} \models \varphi(\vec{a}) &\iff \mathbf{A} \nvDash \psi(\vec{a}) \\ \Big[\text{by induction}\Big] \quad &\iff \mathbf{B} \nvDash \psi(\vec{a}) \\ &\iff \mathbf{B} \models \varphi\Big(h(\vec{a})\Big). \end{aligned}$$
- If $\varphi = \psi_1 \wedge \psi_2$, then the calculation is similar to the previous case.
- If $\varphi = \exists v \psi$, then
$$\begin{aligned} \mathbf{A} \models \varphi(\vec{a}) &\iff \exists a' \in A, \mathbf{A} \models \psi(\vec{a}, a') \\ \Big[\text{by induction}\Big] \quad &\iff \exists a' \in A, \mathbf{B} \models \psi\Big(h(\vec{a}), h(a')\Big) \\ \Big[\text{use surjectivity of } h \text{ for } \Longleftarrow\Big] \quad &\iff \exists b' \in B, \mathbf{B} \models \psi\Big(h(\vec{a}), b'\Big) \\ &\iff \mathbf{B} \models \varphi\Big(h(\vec{a})\Big). \end{aligned}$$
□

Proposition 1.36. *If a σ-structure $\mathbf{A}$ is a reduct of a σ'-structure $\mathbf{A}'$, then for every σ-formula $\varphi(\vec{v})$ and $\vec{a} \in A^n$,*
$$\mathbf{A} \models \varphi(\vec{a}) \iff \mathbf{A}' \models \varphi(\vec{a}).$$

Proof. Trivial induction on formulas and possibly also terms. □

1.C. **Definability**

A central notion in the study of a structure is that of a *definable set*. As the name suggests, these are all sets whose definition can be written in the first-order language.

Definition 1.37. Let $\boldsymbol{A}$ be a σ-structure and $P \subseteq A$. A set $S \subseteq M^n$ is called *P-definable* (or *definable from P*) in $\boldsymbol{A}$ if there is a σ-formula $\varphi(\vec{x},\vec{y})$, with $|\vec{y}| = n$, and $\vec{p} \in P^{|\vec{x}|}$ such that $S = \{\vec{a} \in A^n : \boldsymbol{A} \models \varphi(\vec{p},\vec{a})\}$.

If $P = \emptyset$, we say that S is *0-definable*, and if $P = A$, we say that S is *definable*. We say that an element $\vec{b} \in A^n$ is *P-definable* if the singleton $\{\vec{b}\}$ is P-definable. For a set $D \subseteq A$, a function $f : D^n \to A$ is called *P-definable* if its graph $\mathrm{Graph}(f) := \{(\vec{a}, b) \in D^n \times A : f(\vec{a}) = b\}$ is P-definable.

We denote by $\mathcal{D}^{\boldsymbol{A}}_n(P)$ the collection of all P-definable subsets of A^n and we let $\mathcal{D}^{\boldsymbol{A}}(P)$ denote their union over $n \in \mathbb{N}^+$.

Observation 1.38. *The cardinality of $\mathcal{D}^{\boldsymbol{A}}(P)$ is at most that of* $|\mathrm{ExtFormulas}(\sigma) \times P^{<\mathbb{N}}| \leq \max\{\aleph_0, \sigma, |P|\}$.

To show definability of a subset, one has to come up with a definition for it. Sometimes this requires deep knowledge about the structure, c.f. Example 1.40(c) below. Proving nondefinability of sets is often much harder and requires advanced tools such as the Compactness and the Completeness theorems, and quantifier elimination. However, here is a cheap yet very useful tool provided by Proposition 1.35.

For a function $h : A \to A$ and a set $S \subseteq A$, we say that h fixes S *pointwise* (resp. *setwise*) if $h(s) = s$ for any $s \in S$ (resp. $h(S) \subseteq S$).

Lemma 1.39. *Let $\boldsymbol{A}$ be a σ-structure and $P \subseteq A$. Any automorphism of $\boldsymbol{A}$ that pointwise fixes P, setwise fixes every P-definable set.*

Proof. Immediately follows from *Proposition* 1.35. □

Examples 1.40.

(a) In $\boldsymbol{R} := (\mathbb{R}, 0, 1, +, \cdot)$, the set of positive numbers is 0-definable by the formula $\varphi_{>0}(x) := x \neq 0 \wedge \exists y(x \doteq y^2)$, where y^2 is an abbreviation for $y \cdot y$. Using this, one can define the binary relation $<\subseteq \mathbb{R}^2$ by the formula $\varphi_<(x,y) := \varphi_{>0}(y - x)$ (0-definable). Thus $\boldsymbol{R}$ and $\boldsymbol{R}_< := (\mathbb{R}, 0, 1, +, \cdot, <)$ have the same definable sets.

(b) In $(\mathbb{Q}, <)$ the set of positive rationals is not definable; this follows from Lemma 1.39 as $q \mapsto q + 1$ is an automorphism.

(c) In $(\mathbb{Z}, +, \cdot)$, the set $\mathbb{N}$ is definable because by Lagrange's Four Square theorem that states that any natural number is a sum of squares of four integers.

(d) The definable sets of $\boldsymbol{N} := (\mathbb{N}, 0, S, +, \cdot)$ are called *arithmetical*. It is easy to see that a set is definable in $\boldsymbol{N}$ if and only if it is 0-definable. We will show later on in the course that all computer-recognizable (programmable) sets are arithmetical, but these occupy only a tiny corner of the class of arithmetical sets.

(e) In $\boldsymbol{C} := (\mathbb{C}, 0, 1, +, \cdot)$, the set $\{\sqrt{2}, -\sqrt{2}\}$ is 0-definable by $\varphi(z) := z^2 \doteq \dot{2}$, where z^2 and $\dot{2}$ are abbreviations for $z \cdot z$ and $1 + 1$, respectively. However, $\sqrt{2}$ itself isn't 0-definable! This follows from the fact that $\boldsymbol{C}$ admits *quantifier elimination* (as we will see later), so the only definable sets are varieties (i.e. those defined by polynomials) and Boolean combinations thereof.

(f) In any graph $\boldsymbol{G} := (V, E)$, the set

$$\{(u, v) \in V^2 : \text{the edge-distance between } u \text{ and } v \text{ is } \leq 2\}$$

is 0-definable by the formula

$$\varphi(x, y) := xEy \vee \exists z(xEz \wedge zEy).$$

Similarly, one can show that for any $n \geq 1$, the set

$$\{(u, v) \in V^2 : \text{the edge-distance between } u \text{ and } v \text{ is } \leq n\}$$

is 0-definable. However it turns out that the set

$$\left\{(u,v) \in V^2 : u \text{ and } v \text{ are connected}\right\}$$

is not even definable in some (actually most) graphs; e.g. for a graph consisting of two disconnected infinite paths. We will prove this later on in the course using the Compactness theorem.

To understand definable sets better, let's examine their structure from a set theoretic/geometric point of view. Clearly all relations of the structure, as well as the equality relation (i.e. the diagonal), are definable. What set theoretic/geometric operations keep definable sets definable? Here we answer this question.

Notation 1.41. For any set $S \subseteq A \times B$ and $a \in A$, we write $S_a := \{b \in B : (a,b) \in S\}$ and refer to this set as the *fiber of S over a*.

Definition 1.42. For a set A and a collection $\mathcal{S} \subseteq \bigcup_{n\geq 1} \mathscr{P}(A^n)$, put $\mathcal{S}_n := \mathcal{S} \cap \mathscr{P}(A^n)$. For a set $P \subseteq A$, call $\mathcal{S}$ *P-constructively closed* if:

(1.42.i) *Boolean algebra*: Each $\mathcal{S}_n$ is a Boolean algebra, i.e. contains $\emptyset$ and A^n and is closed under complements and (finite) unions.

(1.42.ii) *Symmetry*: Each $\mathcal{S}_n$ is symmetric, i.e. closed under any permutation of coordinates.

(1.42.iii) *Projections*: The projection onto the first n coordinates of any set in $\mathcal{S}_{n+1}$ is in $\mathcal{S}_n$, i.e. $\forall S \in \mathcal{S}_{n+1}$, $\mathrm{proj}_{(0,1,\dots,n-1)}(S) \in \mathcal{S}_n$.

(1.42.iv) *Lifts*: The Cartesian product of A with any set in $\mathcal{S}_n$ is in $\mathcal{S}_{n+1}$, i.e. $S \in \mathcal{S}_n \Rightarrow A \times S \in \mathcal{S}_{n+1}$.

(1.42.v) *P-fibers*: The fiber over any element $p \in P$ of any set in $\mathcal{S}_{1+n}$ is in $\mathcal{S}_n$, i.e. $S \in \mathcal{S}_{1+n} \Rightarrow S_p \in \mathcal{S}_n$.

Proposition 1.43. *For any σ-structure $\boldsymbol{A}$ and $P \subseteq A$, the collection $\mathcal{D}^{\boldsymbol{A}}(P)$ of P-definable subsets of $\boldsymbol{A}$ is the smallest P-constructively closed collection containing the constant singletons $\left\{c^{\boldsymbol{A}}\right\}$, the graphs $\mathrm{Graph}(f^{\boldsymbol{A}})$ of the functions, and the relations $=\subseteq A^2$ and $R^{\boldsymbol{A}}$, for all $c, f, R \in \sigma$.*

Proof. Exercise. □

In particular, the set $\mathcal{D}^{\boldsymbol{A}}_n(P)$ is a (Boolean) algebra, i.e. it is closed under finite unions and complements and contains $\emptyset$ and A^n. Let $\mathcal{T}^{\boldsymbol{A}}_n(P)$ denote the topology on A^n generated by $\mathcal{D}^{\boldsymbol{A}}_n(P)$; considering this topology will prove to be very useful later. Because $\mathcal{D}^{\boldsymbol{A}}_n(P)$ is closed under finite intersections, in it is actually a basis for the topology. Furthermore, because it is closed under complements, each set in it is both open and closed, i.e. *clopen*. Thus, the topology $\mathcal{T}^{\boldsymbol{A}}_n(P)$ is *zero-dimensional*, i.e. admits a basis of clopen sets. However, this topology might not be Hausdorff as Example 1.40(e) shows. Lastly, whether or not it is compact is tied to a model-theoretic property called *saturation*, which we will learn later.

1.D. **Theories, models, and axiomatization**

Given a signature σ, a set of σ-sentences is called a *σ-theory*. The sentences in a theory T are often referred to as *axioms*.

Definition 1.44. We say that a **nonempty** σ-structure $\boldsymbol{M}$ *satisfies* (or *models*) a σ-theory T, written $\boldsymbol{M} \models T$, if $\boldsymbol{M} \models \varphi$, for every $\varphi \in T$. Equivalently, we also say that $\boldsymbol{M}$ is a *σ-model* (or just *model*) of T.

Notation 1.45. For a σ-theory T, let $\mathcal{M}_\sigma(T)$ denote the class[1] of its σ-models, i.e. nonempty σ-structures that satisfy it.

Definition 1.46. For a class $\mathcal{C}$ of σ-structures, a σ-theory T is called an *axiomatization* of $\mathcal{C}$ (or we say that T *axiomatizes* $\mathcal{C}$) if $\mathcal{M}_\sigma(T) = \mathcal{C}$. A class $\mathcal{C}$ is called (resp. *finitely*) *axiomatizable* if it admits a (resp. finite) axiomatization. A σ-theory S is called an *axiomatization* of T if it is an axiomatization of $\mathcal{M}_\sigma(T)$, and we call T *finitely axiomatizable* if it admits a finite axiomatization.

Here are examples of axiomatizations for various classes of structures.

Examples 1.47.

[1] $\mathcal{M}_\sigma(T)$ is closed under σ-isomorphism, so it is "too large to be a set"—it is a *proper class*. Besides, if σ has an infinite model, then it has a model of each cardinality, see Theorem 3.15.

(a) For any signature σ and any $n \in \mathbb{N}^+$ (think $n := 7$), the class $\mathcal{C}_{\leq n}$ of all σ-structures with at most n elements is axiomatized by the axiom

$$\varphi_{\leq n} := \exists x_1 \dots \exists x_n \forall y \bigvee_{i=1}^{n} y \doteq x_i.$$

On the other hand, the class $\mathcal{C}_{\geq n}$ of all σ-structures with at least n elements is axiomatized by the axiom

$$\varphi_{\geq n} := \exists x_1 \dots \exists x_n \bigwedge_{1 \leq i < j \leq n} x_i \doteq x_j.$$

Thus, the class $\mathcal{C}_{=n}$ of all σ-structures with exactly n elements is axiomatized by the axiom $\varphi_{=n} := \varphi_{\leq n} \wedge \varphi_{\geq n}$.

(b) For any signature σ, the class $\mathcal{C}_\infty$ of all infinite σ-structures is axiomatized by the theory

$$T_\infty := \{\varphi_{\geq n} : n \in \mathbb{N}^+\}.$$

We will see later that T_∞ is not finitely axiomatizable (in any signature σ) because the Compactness theorem implies that any finitely axiomatizable theory actually contains a finite axiomatization, which is clearly not the case for T_∞.

(c) Graphs (undirected with no loops): Recalling the signature $\sigma_{\text{gr}} := (E)$, the class of undirected graphs with no loops is axiomatized by the theory GRAPHS consisting of the following axioms:
(i) (Undirected) $\forall x \forall y (xEy \to yEx)$,
(ii) (No loops) $\forall x (\neg xEx)$.
In particular, this class is finitely axiomatizable. Below, by a *graph*, we mean an undericted graph with no loops.

(d) Bipartite graphs: A graph $G := (V, E)$ is called *bipartite* if V admits a partition $V_1 \sqcup V_2$ such that there are no edges between two vertices that are both in V_1 or both in V_2. Note that this definition asks for an existence of subsets V_1, V_2 of V, which is not first-order expressible. However, by basic graph theory, being bipartite is equivalent to not containing any cycle of odd length, which is axiomatized by the theory

$$T_{\text{bp}} := \{\psi_{2n} : n \in \mathbb{N}^+\},$$

where ψ_k states the nonexistence of a cycle of length exactly k (write this down explicitly).

(e) Partial orderings: Letting $\sigma_{\text{po}} := (\leq)$, the class of partial orderings (σ_{po}-structures) is axiomatized by the theory PO consisting of the following axioms:
(PO1) (Reflexivity) $\forall x (x \leq x)$.
(PO2) (Antisymmetry) $\forall x \forall y (x \leq y \wedge y \leq x \to x \doteq y)$,
(PO3) (Transitivity) $\forall x \forall y \forall z (x \leq y \wedge y \leq z \to x \leq z)$.

(f) Groups: Recalling the signature $\sigma_{\text{gp}} := \left(1, \cdot, ()^{-1}\right)$, the class of groups is axiomatized by the theory GROUPS consisting of the following axioms:
(G1) (Associativity) $\forall x \forall y \forall z [x \cdot (y \cdot z) \doteq (x \cdot y) \cdot z]$,
(G2) (Identity) $\forall x (1 \cdot x \doteq x \cdot 1 \doteq x)$,
(G3) (Inverse) $\forall x (x \cdot x^{-1} \doteq 1 \doteq x^{-1} \cdot x)$.
The reader is invited to find an axiomatization for the class of $(\cdot)$-structures that are groups under $\cdot$.

(g) Rings and fields: Similarly, one defines the theory RINGS of rings in the signature $\sigma_{\text{ring}} := (0, 1, +, -, \cdot)$ (too many axioms to write, but still finitely-many), and then the theory FIELDS of fields is defined as RINGS together with the following three axioms:
(F1) (Nonzero) $0 \neq 1$,
(F2) (Commutativity) $\forall x \forall y [x \cdot y \doteq y \cdot x]$,
(F3) (Multiplicative inverse) $\forall x \exists y [xy \doteq yx \doteq 1]$,

(h) Characteristic p fields: Here is a σ_{ring} axiomatizing the class of fields of characteristic p, for a prime number p:

$$\text{FIELDS}_p := \text{FIELDS} \cup \left\{ \underbrace{1+1+\ldots+1}_{p} \doteq 0 \right\}.$$

(i) Characteristic 0 fields:

$$\text{FIELDS}_0 := \text{FIELDS} \cup \left\{ \underbrace{1+1+\ldots+1}_{p} \neq 0 : p \text{ prime} \right\}.$$

(j) Algebraically closed fields: The following σ_{ring}-theory axiomatizes the class of algebraically closed fields:

$$\text{ACF} := \text{FIELDS} \cup \left\{ \forall a_0 \forall a_1 \ldots \forall a_n \exists r [a_n r^n + a_{n-1} r^{n-1} + \ldots + a_1 r + a_0 \doteq 0] : n \in \mathbb{N} \right\}.$$

(k) Algebraically closed fields of fixed characteristic: Letting n be either 0 or prime, the following is an axiomatization for a class of algebraically closed fields of characteristic n:

$$\text{ACF}_n := \text{ACF} \cup \text{FIELDS}_n.$$

As we see, many interesting classes of structures admit a (first-order) axiomatization. However, we will show later on in the course that many other very interesting classes of structures are not axiomatizable, e.g. connected graphs, disconnected graphs, cyclic groups, torsion groups, nontorsion groups, etc.

Given a σ-structure $\boldsymbol{A}$, we put $\text{Th}(\boldsymbol{A}) := \{\varphi : \varphi \text{ is a } \sigma\text{-sentence and } \boldsymbol{A} \models \varphi\}$. It can often be very hard to tell whether a given σ-sentence is in $\text{Th}(\boldsymbol{A})$. For example, for the structure $\boldsymbol{N} := (\mathbb{N}, 0, S, +, \cdot)$ of natural numbers, we still don't know whether the sentence expressing Goldbach's conjecture belongs to $\text{Th}(\boldsymbol{N})$. Thus, it is desirable to find a simpler axiomatization for $\text{Th}(\boldsymbol{A})$ for a structure $\boldsymbol{A}$ of interest. The following is Peano's attempt to do so for $\boldsymbol{N}$.

Example 1.48. The theory PA of arithmetic, called Peano Arithmetic, in the signature $\sigma_{\text{arthm}} := (0, S, +, \cdot)$, consists of the following (infinitely-many) axioms:

(PA1) $\forall x[\neg S(x) \doteq 0]$
(PA2) $\forall x \forall y[S(x) \doteq S(y) \to x \doteq y]$
(PA3) $\forall x[x + 0 \doteq x]$
(PA4) $\forall x \forall y[S(x + y) \doteq x + S(y)]$
(PA5) $\forall x[x \cdot 0 \doteq 0]$
(PA6) $\forall x \forall y[x \cdot S(y) \doteq x \cdot y + x]$
(PA7) (Axiom schema of induction) For each σ_{arthm}-formula $\varphi(x, \vec{y})$, where x is a variable and $\vec{y}$ is a vector of variables, the following is an axiom:

$$\Big(\varphi(0, \vec{y}) \wedge \forall x \Big[\varphi(x, \vec{y}) \to \varphi(S(x), \vec{y})\Big]\Big) \to \forall x \varphi(x, \vec{y}).$$

It should be emphasized that PA is an infinite theory: (PA7) is not one axiom, rather it is a collection of infinitely-many axioms, one for each extended formula $\varphi(x, \vec{y})$, so we call it an *axiom schema*.

Clearly, $\boldsymbol{N} \models \text{PA}$, where $\boldsymbol{N} := (\mathbb{N}, 0, S, +, \cdot)$. However, as we will see later on, it is a consequence of Gödel's Incompleteness theorem that PA doesn't axiomatize $\text{Th}(\boldsymbol{N})$.

We end this section with perhaps the most important theory in mathematics.

Example 1.49. The Zermelo-Fraenkel set theory, ZFC, is a theory in the signature $\sigma_{\text{set}} := (\in)$, in which all of the mathematics is derived. Its list of axiom schemas (again, infinitely-many axioms) is a little too long to be listed here, so it is enough to mention that they express some basic facts about sets such as existence of unions, definable subsets, an infinite set, etc.

1.E. **Semantic versions of implication, consistency, and completeness**

Definition 1.50. We say that a σ-theory T *satisfies* a σ-sentence φ, written $T \models \varphi$, if every model of T satisfies φ, i.e. $\forall \boldsymbol{M} \models T (\boldsymbol{M} \models \varphi)$. Equivalently, we say that T *semantically implies* φ.

Examples 1.51.

(a) We know from group theory that GROUPS $\models \forall x \forall y \forall y' (yx \doteq e \doteq xy' \to y \doteq y')$.

(b) One can easily show that for any $n \geq 0$ and p prime,

$$\text{FIELDS}_p \models \underbrace{1 + 1 + \ldots + 1}_{n} \doteq 0$$

if and only if p divides n.

(c) It is also easy to see that for all $n \geq 1$, $\text{FIELDS}_0 \models \underbrace{1 + 1 + \ldots + 1}_{n} \neq 0$.

Definition 1.52. A σ-theory T is said to be

- *satisfiable* (or *semantically consistent*) if it has a model.
- *semantically σ-complete* (or just *semantically complete* if σ is understood) if for every σ-sentence φ, $T \models \varphi$ or $T \models \neg\varphi$. Call a σ-theory $\bar{T} \supseteq T$ a *semantic completion* if it is satisfiable and semantically complete.

Let $\top$ denote the sentence $\forall x (x \doteq x)$ and set $\bot := \neg\top$. The following explains the term *semantically consistent*.

Observation 1.53. *For a σ-theory T, the following are equivalent:*

(1) *T is satisfiable.*
(2) *$T \nvdash \bot$.*
(3) *$T \nvdash \varphi$ for some σ-sentence φ.*

Definition 1.54. Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be σ-structures. We say that $\boldsymbol{A}$ and $\boldsymbol{B}$ are *elementarily equivalent*[2], written $\boldsymbol{A} \equiv \boldsymbol{B}$, if $\text{Th}(\boldsymbol{A}) = \text{Th}(\boldsymbol{B})$.

By Proposition 1.35, isomorphic structures are elementarily equivalent. However, the converse is false! For example, it is a homework problem to show that $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ are elementarily equivalent, but they clearly cannot be isomorphic, having different cardinalities.

The following is a convenient rephrasing of semantic completeness in terms of elementary equivalence.

Proposition 1.55 (Semantic completeness, rephrased)**.** *A σ-theory T is semantically complete if and only if for any $\boldsymbol{A}, \boldsymbol{B} \models T$, $\boldsymbol{A} \equiv \boldsymbol{B}$.*

Proof. Left as an exercise. □

In the light of this, it is easy to see that most of the examples:theories given above are semantically incomplete; indeed, for example, the theory GROUPS is semantically incomplete because there is a group that has an element of order 3, hence satisfies the sentence $\exists x (x \cdot x \cdot x \doteq e)$, and there is a group that does not. However, we will show later that the theory ACF_n, for n either prime or 0, is semantically complete, making model theory highly applicable in algebraic geometry.

Definition 1.56. A σ-theory T is said to be *maximally σ-complete* (or just *maximally complete* if σ is understood) if for every σ-sentence φ, $\varphi \in T$ or $\neg\varphi \in T$ (both hold if and only if T is not satisfiable). Call a σ-theory $\bar{T} \supseteq T$ a *maximal completion* of T if it is maximally complete.

Observation 1.57.

(1.57.a) *For any σ-structure $\boldsymbol{M}$, $\text{Th}(\boldsymbol{M})$ is satisfiable and maximally complete.*
(1.57.b) *Thus, every satisfiable theory admits a satisfiable maximal completion.*

[2]The author does not understand the choice of the term *elementary*; if she had to choose, it would perhaps be *first-order equivalent*.

1.F. **Elementarity**

For σ-structures $\boldsymbol{A} \subseteq \boldsymbol{B}$, it is an interesting question as to which formulas $\boldsymbol{A}$ and $\boldsymbol{B}$ agree on. The following is all we can say for general $\boldsymbol{A} \subseteq \boldsymbol{B}$.

Proposition 1.58. *Substructures agree on quantifier free formulas; more precisely, for σ-structures $\boldsymbol{A} \subseteq \boldsymbol{B}$, any quantifier free σ-formula φ and $\vec{a} \in A^n$, we have*

$$\boldsymbol{A} \models \varphi(\vec{a}) \iff \boldsymbol{B} \models \varphi(\vec{a}).$$

Proof. Easy induction on the construction of φ. The step only consists of the cases $\varphi = \neg\varphi$ and $\varphi = \varphi \wedge \psi$, whereas the base cases (i.e. when φ is atomic) follow from Lemma 1.34 and the fact that the inclusion map $A \subseteq B$ is a homomorphism. □

However, the calculations of a structure and a substructure of the validity of formulas with quantifiers may differ. Typically, a formula of the form $\exists x \varphi(x)$ may be valid in the bigger structure but may not be in the substructure simply because the objects for which φ holds (which we refer to as *witnesses* to $\exists x \varphi(x)$) may all be outside of the universe of the substructure. For example, in the signature $\sigma_{\text{mon}} := (1, \cdot)$, a substructure of a group may not be a subgroup because not all elements might have inverses in the substructure. Even if it was a subgroup, it might disagree with the ambient group about the truth of statements like "being abelian" or "a particular element commutes with everybody" (they may be true in the subgroup, but false in the ambient group). The best we can say is the following.

Definition 1.59. A σ-formula is called *universal* (resp. *existential*) if it is of the form $\forall x_1 \forall x_2 ... \forall x_n \psi$ (resp. $\exists x_1 \exists x_2 ... \exists x_n \psi$), where ψ is quantifier free and $n \geq 0$; in particular, quantifier free formulas are both universal and existential.

Proposition 1.60. *Let $\boldsymbol{A}, \boldsymbol{B}$ be σ-structures with $\boldsymbol{A} \subseteq \boldsymbol{B}$ and let $\varphi(\vec{v})$ be a σ-formula. For any $\vec{a} \in A^{|\vec{v}|}$,*

(1.60.a) *if φ is universal, then $\boldsymbol{B} \models \varphi(\vec{a}) \implies \boldsymbol{A} \models \varphi(\vec{a})$;*
(1.60.b) *if φ is existential, then $\boldsymbol{A} \models \varphi(\vec{a}) \implies \boldsymbol{B} \models \varphi(\vec{a})$.*

Proof. Exercise. □

The following definitions isolate those substructures which agree with the ambient structure on *all* of the statements about the elements of the substructure.

Definition 1.61. Let $\boldsymbol{A}, \boldsymbol{B}$ be σ-structures.

- A homomorphism $h : \boldsymbol{A} \to \boldsymbol{B}$ is called an *elementary embedding*, written $h : \boldsymbol{A} \hookrightarrow_e \boldsymbol{B}$, if for all formulas $\varphi(\vec{v})$ and tuples $\vec{a} \in A^{|\vec{v}|}$,
$$\boldsymbol{A} \models \varphi(\vec{a}) \iff \boldsymbol{B} \models \varphi(h(\vec{a})).$$
- We say that $\boldsymbol{A}$ *elementarily embeds* into $\boldsymbol{B}$, written $\boldsymbol{A} \hookrightarrow_e \boldsymbol{B}$, if there is $h : \boldsymbol{A} \hookrightarrow_e \boldsymbol{B}$.
- We call a substructure $\boldsymbol{A} \subseteq \boldsymbol{B}$ *elementary*, written $\boldsymbol{A} \preceq \boldsymbol{B}$, if the inclusion map $\boldsymbol{A} \subseteq \boldsymbol{B}$ is an elementary embedding.

Observation 1.62. *If $\boldsymbol{A} \preceq \boldsymbol{B}$, then $\boldsymbol{A} \equiv \boldsymbol{B}$.*

Caution 1.63. The converse of Observation 1.62 is false; in fact, there are graphs $\boldsymbol{H} \subseteq \boldsymbol{G}$ such that $\boldsymbol{H} \cong \boldsymbol{G}$ (in partcular, $\boldsymbol{H} \equiv \boldsymbol{G}$), but $\boldsymbol{H} \not\preceq \boldsymbol{G}$. The reader is invited to construct such graphs.

What does it take for a substructure to be elementary? Recall that for a subset of a structure to be a universe of a substructure, it has to contain all the constants and be closed under all the functions of the structure, see (1.9.b). For a substructure to be elementary, a stronger closure condition is needed: it has to contain a witness to every formula of the form $\exists x \varphi(x)$. This is stronger than just asking to contain constants and values of functions. The following makes all this precise.

Proposition 1.64 (Tarski–Vaught test)**.** *Let $\boldsymbol{A}$ be a substructure of a σ-structure $\boldsymbol{B}$. $\boldsymbol{A}$ is an elementary substructure of $\boldsymbol{B}$ if and only if for every formula $\varphi(\vec{x}, y)$ and $\vec{a} \in A^{|\vec{x}|}$,*

$$\boldsymbol{B} \models \exists y \varphi(\vec{a}, y) \iff \exists a' \in A \text{ such that } \boldsymbol{B} \models \varphi(\vec{a}, a').$$

Proof. $\Rightarrow$: Supposing $\boldsymbol{A} \prec \boldsymbol{B}$, we check the Tarski–Vaught condition:

$$\begin{aligned} \boldsymbol{B} \models \varphi(\vec{a}) &\iff \boldsymbol{B} \models \exists y \varphi(\vec{a}, y) \\ \Big[\text{elementarity}\Big] &\iff \boldsymbol{A} \models \exists y \varphi(\vec{a}, y) \\ \Big[\text{definition of } \models\Big] &\iff \exists a' \in A \text{ such that } \boldsymbol{A} \models \varphi(\vec{a}, a') \\ \Big[\text{elementarity}\Big] &\iff \exists a' \in A \text{ such that } \boldsymbol{B} \models \varphi(\vec{a}, a'). \end{aligned}$$

$\Leftarrow$: Suppose the Tarski–Vaught condition holds and show by induction on the construction of formulas that for every σ-formula φ and $\vec{a} \in A^{|\vec{x}|}$, we have

$$\boldsymbol{A} \models \varphi(\vec{a}) \iff \boldsymbol{B} \models \varphi(\vec{a}).$$

Proposition 1.58 takes care of the atomic formulas and the cases $\varphi = \neg\psi$ and $\varphi = \psi_0 \wedge \psi_2$ are straightforward, so we only consider the case $\varphi(\vec{x}) = \exists y \psi(\vec{x}, y)$. Fix $\vec{a} \in A^{|\vec{x}|}$ and check:

$$\begin{aligned} \boldsymbol{B} \models \varphi(\vec{a}) &\iff \boldsymbol{B} \models \exists y \psi(\vec{a}, y) \\ \Big[\text{Tarski–Vaught condition}\Big] &\iff \exists a' \in A \text{ such that } \boldsymbol{B} \models \psi(\vec{a}, a') \\ \Big[\text{induction}\Big] &\iff \exists a' \in A \text{ such that } \boldsymbol{A} \models \psi(\vec{a}, a') \\ \Big[\text{definition of } \models\Big] &\iff \boldsymbol{A} \models \exists y \psi(\vec{a}, y) \iff \boldsymbol{A} \models \varphi(\vec{a}). \end{aligned}$$

□

Given a σ-structure $\boldsymbol{B}$ and $S \subseteq B$, we could define a substructure generated by S as the smallest substructure containing S, which exists because intersection of substructures is still a substructure. However, intersection of elementary substructures may not be elementary (again, there are simple examples with graphs), so we cannot define "the elementary substructure generated by S".

Recall the procedure described in Proposition 1.11 of constructing the substructure generated by S, where we add the constants of $\boldsymbol{B}$ to S and then close under all the functions of $\boldsymbol{B}$ by iteratively adding the values of these functions $\aleph_0$-many times. According to the Tarski–Vaught test, to be a universe of an *elementary* substructure, a set has to also contain a witness to ever formula that claim existence of an object and is true in $\boldsymbol{B}$. So in our iterative procedure, at every step, we have to additionally throw in a witness to every such formula—that's all. In fact, only throwing witnesses will also add the values of functions because for every $f \in \mathrm{Func}(\sigma)$ and $\vec{a} \in B^{\mathfrak{a}(f)}$, the value $f^{\boldsymbol{B}}(\vec{a})$ is the (unique) witness to the formula $\exists y f(\vec{a}) \doteq y$. Same is true for constant symbols.

To "throw in a witness to every such formula", we need to make a choice of a witness for every such formula.

Definition 1.65. Let $\boldsymbol{B}$ be a σ-structure and $\varphi(\vec{x}, y)$ be an extended σ-formula. A *Skolem function for* $\varphi(\vec{x}, y)$ is a partial function[3] $f_{\varphi(\vec{x},y)} : B^{|\vec{x}|} \rightharpoonup B$ such that, for each $\vec{b} \in B^{|\vec{x}|}$, if $\boldsymbol{B} \models \exists y \varphi(\vec{b}, y)$, then $f_{\varphi(\vec{x},y)}(\vec{b})$ witnesses this, i.e. $\vec{b} \in \mathrm{dom}(f_{\varphi(\vec{x},y)})$ and $\boldsymbol{B} \models \varphi\Big(\vec{b}, f_{\varphi(\vec{x},y)}(\vec{b})\Big)$.

Remark 1.66. Skolem functions exist due to the Axiom of Choice, so, in general, they are not definable.

The following theorem summarizes the above discussion.

Theorem 1.67 (Downward Löwenheim–Skolem)**.** *Let $\boldsymbol{B}$ be a σ-structure and $S \subseteq B$. There exists $\boldsymbol{A} \preceq \boldsymbol{B}$ with $A \supseteq S$ such that $|A| \leq \max(|S|, |\sigma|, \aleph_0)$.*

Proof. Using the Axiom of Choice, we let $f_{\varphi(\vec{x},y)}$ be a Skolem function for each extended σ-formula $\varphi(\vec{x}, y)$. As in Proposition 1.11, we recursively construct an increasing sequence $(S_n)_{n \in \mathbb{N}}$ of subsets of B as follows: put $S_0 := S$, and assuming S_n is defined, let

$$S_{n+1} := S_n \cup \bigcup_{\varphi(\vec{x},y)} f_{\varphi(\vec{x},y)}\Big(S_n^{|\vec{x}|}\Big),$$

[3] A partial function $f : X \rightharpoonup Y$ is a function whose domain $\mathrm{dom}(f)$ is a (possibly empty) subset of X.

where $\varphi(\vec{x},y)$ ranges over all extended σ-formulas and $f_{\varphi(\vec{x},y)}\left(S_n^{|\vec{x}|}\right)$ means $f_{\varphi(\vec{x},y)}\left(S_n^{|\vec{x}|}\cap \mathrm{dom}(f_{\varphi(\vec{x},y)})\right)$. Letting $A:=\bigcup_{n\in\mathbb{N}} S_n$ and it is now straightforward to check that A is a universe of a substructure, which passes the Tarski–Vaught test and is thus elementary. □

Remark 1.68. It is possible that the same formula has multiple witnesses and the Axiom of Choice makes a choice of one of them. Depending on this (highly noncanonical) choice, the resulting substructures may be different, and this is why there is no notion of "*the* elementary substructure generated by S".

Theorem 1.69 (Weak Downward Löwenheim–Skolem)**.** *Any satisfiable σ-theory T has a model of cardinality at most* $\max|\sigma|,\aleph_0$.

Proof. Let $\boldsymbol{M}\models T$ and apply the Downward Löwenheim–Skolem theorem to $S:=\emptyset$. □

1.G. **The Skolem "paradox"**

The Weak Downward Löwenheim–Skolem theorem has the following at first striking consequence: if ZFC is satisfiable (which we really hope it is), then it has a countable model. This may seem strange because this countable model $\boldsymbol{M}$ satisfies the sentence that there is an uncountable set since Cantor's theorem that the reals are uncountable is true in $\boldsymbol{M}$. Does this imply that ZFC is not satisfiable? Of course not and here are the two reasons why (second being the main reason).

(1) Replacing the universe of M with $\mathbb{N}$, we may assume that $M=\mathbb{N}$, so $\epsilon^{\boldsymbol{M}}$ is just a binary relation on $\mathbb{N}$, i.e. a subset of $\mathbb{N}^2$. So what if somehow $\boldsymbol{M}$ satisfies the statement that reads as "there is an uncountable set"? It is just some statement about this binary relation $\epsilon^{\boldsymbol{M}}$ and it does not imply anything about the actual sets and the cardinality of M.
(2) Even if M was a set of sets and $\epsilon^{\boldsymbol{M}}$ was the true $\in$, then the countability of M would simply imply that $\boldsymbol{M}$'s version of the real numbers, $\mathbb{R}^{\boldsymbol{M}}$, is indeed countable (for us), i.e. there is a bijection $f:\mathbb{R}^{\boldsymbol{M}}\to\mathbb{N}$. This bijection is a set, namely a subset of $\mathbb{R}^{\boldsymbol{M}}\times\mathbb{N}$, but it may not be an element of M—the latter doesn't contain *all* sets, only countably-many of them. In fact, since $\boldsymbol{M}$ satisfies the statement "$\mathbb{R}^{\boldsymbol{M}}$ is uncountable", we conclude that $f\notin M$ for sure! In other words, M does not "see" the countability of $\mathbb{R}^{\boldsymbol{M}}$ and thus thinks that $\mathbb{R}^{\boldsymbol{M}}$ is uncountable. It's like how people thought the world was endless before they discovered it was round since all they could see was the ocean up to the line of the horizon and for all they knew it continued forever. The only difference is that we eventually obtained the knowledge that Earth is round and finite, while $\boldsymbol{M}$ never will.

2. First order logic: the syntactic aspect

So far, we have been dealing with the semantic (model-theoretic) aspect of FOL, i.e. structures/models, satisfiability, definability, etc. In this section we turn to the syntactic aspect, namely proof systems and formal proofs.

We fix a signature σ for this section and everything below is assumed to be in this signature.

2.A. **The axioms and the rule of inference of** $\mathbb{FOL}(\sigma)$

Here we define the set $\mathrm{Axioms}(\sigma)$ of the *default axioms* of, as well as a *rule of inference* of $\mathbb{FOL}(\sigma)$. These axioms would be satisfied by every σ-structure and the rule of inference would preserve satisfiability.

Unlike the definition of a σ-theory, $\mathrm{Axioms}(\sigma)$ includes formulas with free variables. This is necessary as in the course of a proof, even if our goal is to prove a sentence, we often make quantified variables free. For example, when proving

$$\forall f:[0,1]\to\mathbb{R},\ f \text{ is continuous} \Rightarrow f \text{ is bounded}, \tag{$*$}$$

we start the proof by letting the variable f denote a function and this variable stays free until the end of the proof, where we *generalize* by saying "but f is arbitrary, so $(*)$ is true".

We need the following technical definition in order to state the axioms that involve variables.

Definition 2.1. Let φ be a formula and t be a term. We say that *t is free for v in φ* (or *t is OK to be plugged-in for v in φ*) if neither v nor any variable in t is quantified in φ. If t is free for v in φ, we define $\varphi(t/v)$ to be the formula obtained from φ by replacing all occurrences of v by t.

Convention 2.2. Below, whenever we write $\varphi(t/v)$, it is assumed that t is free for v in φ.

Convention 2.3. From now on, we treat $\varphi \vee \psi; \varphi \wedge \psi; \exists v \varphi$ as abbreviations for $\neg\varphi \to \psi; \neg(\varphi \to \neg\psi); \neg\forall v \neg\varphi$.

We use an axiom schema to describe an infinite set of axioms that have the same form, so each of the schemas below, defines an infinite set. The set Axioms(σ) consists of the union of the sets of axioms defined via the following axiom schemas.

Propositional axioms. Letting φ, ψ, χ range over all σ-formulas, the following are in Axioms(σ).

(1) If-true-then-implied: $\varphi \to (\psi \to \varphi)$.
(2) Implication-is-transitive: $(\varphi \to \psi) \to \Big[(\varphi \to (\psi \to \chi)) \to (\varphi \to \chi)\Big]$.
(3) Proof-by-contradiction: $(\neg\varphi \to \psi) \to \Big[(\neg\varphi \to \neg\psi) \to \varphi\Big]$.

Quantifier axioms. Letting

- φ, ψ, χ range over all σ-formulas,
- v range over all variable symbols,
- t range over all σ-terms that are OK to be plugged in for v in φ,
- u range over all variable symbols that are OK to be plugged in for v in φ,

the following are in Axioms(σ).

(4) Instantiation: $\forall v \varphi \to \varphi(t/v)$.
(5) Generalization: $\varphi \to \forall u \varphi(u/v)$.

Remark 2.4. What we really want to write here is $\varphi \to \forall v \varphi(v)$, but if v is free φ (which is exactly when this axiom is most likely to be used), $\varphi \to \forall v \varphi(v)$ is not a formula according to Convention 1.26.

Equality axioms. Letting

- n range over $\mathbb{N}^+$,
- u, v, w range over all variable symbols,
- f range over $\text{Func}_n(\sigma)$,
- R range over $\text{Rel}_n(\sigma)$,
- $\vec{x}, \vec{y}$ range over all vectors of variables of length n,

the following are in Axioms(σ).

(6) Equality-is-equivalence:
  (6.a) Equality-is-reflexive: $u \doteq u$.
  (6.b) Equality-is-symmetric: $u \doteq v \to v \doteq u$.
  (6.c) Equality-is-transitive: $(u \doteq v \wedge v \doteq w) \to u \doteq w$.
(7) Functions-respect-equality: $(\vec{x} \doteq \vec{y}) \to \Big(f(\vec{x}) \doteq f(\vec{y})\Big)$.
(8) Relations-respect-equality: $(\vec{x} \doteq \vec{y}) \to \Big(R(\vec{x}) \to R(\vec{y})\Big)$.

We now state the only *rule of inference* we need to derive new statements from axioms.

Rule of inference. For any σ-formulas φ, ψ, we say that ψ is *obtained from* φ *and* $\varphi \to \psi$ *via Modus Ponens*.

(9) A cartoon of Modus Ponens: $\boxed{\varphi, \varphi \to \psi} \overset{\text{MP}}{\rightsquigarrow} \boxed{\psi}$.

Definition 2.5. Let $\varphi(\vec{v})$ be an extended σ-formula, so $\forall \vec{v} \varphi$ is a sentence. We say that φ is *satisfied/true* in a σ-structure $\boldsymbol{A}$, written $\boldsymbol{A} \models \varphi$, if $\boldsymbol{A} \models \forall \vec{v} \varphi$.

The following better be true.

Observation 2.6. *Every σ-structure $\boldsymbol{A}$ satisfies each σ-formula in* Axioms(σ) *and Modus Ponens preserves the satisfiability of σ-formulas in $\boldsymbol{A}$.*

2.B. **Formal proofs**

Definition 2.7. Let T be a set of σ-formulas (not necessarily sentences) and let φ be a σ-formula. A *proof*[4] of φ from T is a finite sequence $\varphi_1, \varphi_2, \dots \varphi_n$ of σ-formulas such that $\varphi_n = \varphi$ and for each $i = 1, \dots, n$,

- either: $\varphi_i \in \text{Axioms}(\sigma) \cup T$,
- or: φ_i follows from the previous φ_j-s by Modus Ponens, i.e. for some $j, k < i$ (not necessarily $j < k$), $\varphi_k = \varphi_j \to \varphi_i$; in this case, we say that φ_i is obtained by Modus Ponens from φ_j, φ_k.

We say that T *proves* φ, written $T \vdash \varphi$, if there exists a proof of φ from T. When $T = \emptyset$, we just write $\vdash \varphi$.

The following better be true.

Proposition 2.8 (Soundness)**.** *Our proof system is* sound*, i.e. for any σ-theory T, if $T \vdash \varphi$ then $T \models \varphi$.*

Proof. This follows by induction on the length of the formal proof of φ and Observation 2.6. □

The converse is also true and it is exactly the content of Gödel's Completeness Theorem 3.2, which is *much* harder to prove, just wait.

The following illustrates formal proofs and how tedious (even hard) it can be to find formal proofs of statements that are "obviously" true.

Proposition 2.9 (Basic provable facts)**.** *Let χ, θ be a σ-formulas.*

(2.9.a) *Self-implication:* $\vdash \theta \to \theta$.

(2.9.b) *Everything-implies-an-axiom:* $\chi \vdash \theta \to \chi$.

Proof. (2.9.a) Here is a (very) formal proof:

(i) $(\theta \to (\theta \to \theta)) \to [(\theta \to ((\theta \to \theta) \to \theta)) \to (\theta \to \theta)]$ [Axiom (2) for $\varphi := \chi := \theta$ and $\psi := (\theta \to \theta)$],

(ii) $\theta \to (\theta \to \theta)$ [Axiom (1) for $\varphi := \psi := \theta$],

(iii) $(\theta \to ((\theta \to \theta) \to \theta)) \to (\theta \to \theta)$ [Modus Ponens (i), (ii)],

(iv) $\theta \to ((\theta \to \theta) \to \theta)$ [Axiom (1) for $\varphi := \theta$ and $\psi := (\theta \to \theta)$],

(v) $\theta \to \theta$ [Modus Ponens (iii), (iv)].

(2.9.b) By the if-true-then-implied axiom (1), $\vdash \chi \to (\theta \to \chi)$. We also trivially have $\chi \vdash \chi$, so an application of Modus Ponens finishes the proof. □

The following lemma makes coming up with proofs much easier.

Lemma 2.10 (Deduction theorem)**.** *For a set T of σ-formulas and σ-formulas χ, φ,*

$$T, \chi \vdash \varphi \text{ if and only if } T \vdash \chi \to \varphi.$$

Proof. $\Leftarrow$: Follows by an application of Modus Ponens.
$\Rightarrow$: Letting $\varphi_1, \dots, \varphi_n$ with $\varphi_n = \varphi$ be a proof of φ from $T \cup \{\chi\}$, we show that $T \vdash \chi \to \varphi$ by induction on n.
Case 1: $\varphi \in \text{Axioms}(\sigma) \cup T$. Then $T \vdash \chi \to \varphi$ by (2.9.b).
Case 2: $\varphi = \chi$. Then $T \vdash \chi \to \varphi$ by (2.9.a).
Case 3: $\varphi = \varphi_n$ is obtained by Modus Ponens from φ_i and φ_j for some $i, j < n$. Then $\varphi_j = \varphi_i \to \varphi$. By the induction hypothesis, $T \vdash \chi \to \varphi_i$ and $T \vdash \chi \to (\varphi_i \to \varphi)$. By Axiom (2),

$$T \vdash (\chi \to \varphi_i) \to \Big[(\chi \to (\varphi_i \to \varphi)) \to (\chi \to \varphi)\Big]$$

so applying Modus Ponens twice gives $T \vdash \chi \to \varphi$. □

The Deduction theorem makes life much easier in proving that a formula is provable from T. The following illustrates this.

Proposition 2.11 (Further provable facts)**.** *For any σ-formulas φ, ψ, variable symbol v, and a σ-term t that is OK to be plugged in for v in φ, the following σ-formulas are provable from the empty theory.*

(2.11.a) *Double-negation-elimination:* $\neg\neg\varphi \to \varphi$.

(2.11.b) *Double-negation-introduction:* $\varphi \to \neg\neg\varphi$.

(2.11.c) *If-false-then-implies:* $\neg\varphi \to (\varphi \to \psi)$.

[4]The author thanks Itay Neeman for explaining how to apply this concept when doing mathematics.

(2.11.d) *Forward-contrapositive:* $(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$.
(2.11.e) *Truth:* $\top := \forall v(v \doteq v)$.
(2.11.f) *Contradiction-implies-everything:* $(\neg\varphi \wedge \varphi) \to \psi$.
(2.11.g) *Falsity-implies-everything:* $\bot \to \psi$, *where* $\bot := \neg\top$.
(2.11.h) *Witness-implies-existence:* $\varphi(t/v) \to \exists v\varphi$.

Proof. (2.11.a) The Deduction theorem reduces to proving $\neg\neg\varphi \vdash \varphi$. The following instance of the proof-by-contradiction axiom (3) is our driving statement:

$$\vdash (\neg\varphi \to \neg\varphi) \to \Big[(\neg\varphi \to \neg\neg\varphi) \to \varphi\Big]. \tag{2.12}$$

(2.9.a) gives $\vdash \neg\varphi \to \neg\varphi$ and (2.9.b) gives $\neg\neg\varphi \vdash \neg\varphi \to \neg\neg\varphi$, so a couple of Modus Ponenses applied to (2.12) yield $\neg\neg \vdash \varphi$.

(2.11.b) Very similar to the proof of double-negation-elimination (2.11.a) and is left to the reader.

(2.11.c) The Deduction theorem reduces to proving $\neg\varphi, \varphi \vdash \psi$. By the proof-by-contradiction (3) axiom,

$$\vdash (\neg\psi \to \varphi) \to \Big[(\neg\psi \to \neg\varphi) \to \psi\Big]$$

and by everything-implies-an-axiom (2.9.b) $\varphi \vdash \neg\psi \to \varphi$ and $\neg\varphi \vdash \neg\psi \to \neg\varphi$, so a couple of Modus Ponenses finish the proof.

(2.11.d) Again, the Deduction theorem reduces to proving $\varphi \to \psi, \neg\psi \vdash \neg\varphi$. The driving statement is again given by the proof-by-contradiction (3) axiom:

$$\vdash (\neg\neg\varphi \to \psi) \to \Big[(\neg\neg\varphi \to \neg\psi) \to \neg\varphi\Big]. \tag{2.13}$$

Using the implication-is-transitive axiom (2), the facts $\vdash \neg\neg\varphi \to \varphi$ (by (2.11.a)) and $\vdash \varphi \to \psi$ (by (2.9.a)), a couple of Modus Ponenses yield $\varphi \to \psi \vdash \neg\neg\varphi \to \psi$. (2.9.b) gives $\neg\psi \vdash \neg\neg\varphi \to \neg\psi$, so a couple of Modus Ponenses applied to (2.13) yields $\varphi \to \psi, \neg\psi \vdash \neg\varphi$.

(2.11.e) Recalling that $\top$ stands for $\vdash \forall v(v \doteq v)$, this follows from the equality-is-reflexive (6.a) and generalization axioms (5) topped with an application of Modus Ponens.

(2.11.f) $\varphi \wedge \neg\varphi$ is the abbreviation of $\neg(\neg\varphi \to \neg\varphi)$, so, denoting $\chi := (\neg\varphi \to \neg\varphi)$, the driving statement is the following instance of the proof-by-contradiction axiom (3):

$$\vdash (\neg\psi \to \chi) \to \Big[(\neg\psi \to \neg\chi) \to \psi\Big].$$

The rest of the proof is left to the reader.

(2.11.g) Recall that $\bot$ stands for $\neg\top$ and, by (2.11.e), we have $\vdash \top$. It remains to put these two together and apply contradiction-implies-everything (2.11.f).

(2.11.h) Recall that $\exists v\varphi$ stands for $\neg\forall v\neg\varphi$, so, using forward-contrapositive (2.11.d) and double-negation-introduction (2.11.b), our task reduces to proving $\vdash \forall v\neg\varphi \to \neg\varphi(t/v)$, which is an instance of the instantiation axiom (4). □

Lemma 2.14 (Constant Substitution)**.** *Let c be a symbol that is not in σ and let v be free in a σ-formula φ. For a set T of σ-formulas,*

$$T \vdash \varphi(c/v) \text{ if and only if } T \vdash \varphi,$$

where in the first statement c is treated as a constant symbol and T is viewed as a $(\sigma \sqcup \{c\})$-theory.

Proof. The direction $\Leftarrow$ follows by applications of the generalization (5) and instantiation (4) axioms, followed by Modus Ponens. The reverse implication is a straightforward induction on the length of the formal proof. This comes down to proving that if $\varphi(c/v)$ is an axiom, then so is φ, which is only worth checking for the axioms that explicitly deal with variables, i.e. (4)–(8). □

2.C. **Syntactic versions of consistency and completeness**

In this subsection, we define analogues of the notions defined in Subsection 1.E using $\vdash$ instead of $\models$.

Definition 2.15. A σ-theory T is said to be

- (syntactically) *consistent* if there is no σ-sentence φ such that $T \vdash \varphi \wedge \neg\varphi$;
- (syntactically) *σ-complete* (or just complete if σ is understood) if for any σ-sentence φ, $T \vdash \varphi$ or $T \vdash \neg\varphi$.

Note that a satisfiable theory is consistent by the Soundness of the proof system. We also have the analogue of Proposition 2.16 for $\vdash$.

Proposition 2.16. *For a σ-theory T, the following are equivalent:*

(1) *T is syntactically consistent.*
(2) *$T \vdash \bot$.*
(3) *$T \vdash \varphi$ for some σ-sentence φ.*

Proof. Follows from (2.11.f) and (2.11.g). □

Proposition 2.17 (Compactness for $\vdash$)**.** *A σ-theory T is consistent if and only if every finite subset of T is consistent.*

Proof. A proof of $\bot$ is finite, so it uses only a finite subset of T. □

Lemma 2.18 (Modifying consistent theories)**.** *Let T be a σ-theory.*

(2.18.a) *For any σ-sentence χ, $T \cup \{\chi\}$ is inconsistent if and only if $T \vdash \neg\chi$.*
(2.18.b) *If T is consistent, then for any σ-sentence χ, at least one of $T \cup \{\chi\}$ and $T \cup \{\neg\chi\}$ is consistent.*
(2.18.c) *For any extended σ-formula $\varphi(v)$, if $T \cup \{\exists v \varphi(v)\}$ is consistent and c is a constant symbol that does not appear in $T \cup \{\exists v \varphi\}$, then $T \cup \{\varphi(c/v)\}$ is consistent.*

Proof. (2.18.a) The right-to-left direction is immediate so we show the other direction. Assume $T \cup \{\chi\}$ is inconsistent and hence $T, \chi \vdash \bot$. By the Deduction theorem, $T \vdash \chi \to \bot$, and hence, by forward-contrapositive (2.11.d), double-negation-elimination (2.11.a), and Modus Ponens, we get $T \vdash \top \to \neg\chi$. But $\vdash \top$ by (2.11.e), so applying Modus Ponens again yields $T \vdash \neg\chi$.

(2.18.b) We prove the contrapositive. Assume both $T \cup \{\chi\}$ and $T \cup \{\neg\chi\}$ are inconsistent. Then by (2.18.a), $T \vdash \neg\chi$ and $T \vdash \neg\neg\chi$. Thus $T \vdash \chi \wedge \neg\chi$ and hence T is inconsistent.

(2.18.c) Here too, we prove the contrapositive. Assume $T \cup \{\varphi(c/v)\}$ is inconsistent. Then by (2.18.a), $T \vdash \neg\chi(c/v)$. By the Constant Substitution Lemma 2.14, $T \vdash \neg\varphi(v)$, and by the generalization axiom (5), $T \vdash \forall v \neg\varphi(v)$, so $T \vdash \neg\exists v \varphi(v)$. Thus, by (2.18.a) again, $T \cup \{\exists v \varphi(v)\}$ is inconsistent. □

The compactness statement of Proposition 2.17 is actually equivalent to the fact that the following topological space is compact: let $\mathcal{T}$ be the set of all consistent maximally complete theories and take the topology generated by the sets of the form $\langle\varphi\rangle := \{T \in \mathcal{T} : T \vdash \varphi\}$, where φ ranges over all σ-sentences. Although, it is not as immediate as with semantic consistency, the analogue of (1.57.b) is also true for syntactic consistency:

Lemma 2.19. *Any syntactically consistent σ-theory T has a consistent maximal completion, i.e. there exists a (nonunique, in general) syntactically consistent maximally complete σ-theory $\overline{T} \supseteq T$.*

Proof. We give two proofs: one for countable σ and one for arbitrary σ; the first one is a (seemingly) more hands on construction and the readers not familiar with Zorn's lemma may find it more helpful.

Case: σ is countable. In this case there are only countably-many σ-formulas, so we can enumerate all σ-sentences $(\chi_n)_{n\in\mathbb{N}}$. Put $T_0 := T$, and recursively construct an increasing sequence $(T_n)_{n\in\mathbb{N}}$ of consistent theories as follows. Assuming that T_n is defined and is consistent, put $T_{n+1} := T_n \cup \{\chi_n\}$ if $T_n \nvdash \neg\chi_n$, and put $T_{n+1} := T_n \cup \{\neg\chi_n\}$, otherwise. It follows from (2.18.a) that T_{n+1} is consistent. Putting $\overline{T} := \bigcup_n T_n$, note that $\overline{T} \supseteq T$ and $\overline{T}$ is consistent: indeed, if it was inconsistent, then, by 2.17, some finite subset $F \subseteq \overline{T}$ would be inconsistent, but this F would be trapped in some T_n, i.e. $F \subseteq T_n$, making T_n inconsistent, which is a contradiction. Lastly, it is immediate from the construction that $\overline{T}$ is maximally complete.

Case: σ is arbitrary. By 2.17, inconsistent theories have inconsistent finite subsets (i.e. inconsistency has finite base), so arbitrary increasing unions of consistent theories are consistent. Thus, by Zorn's lemma, there is a $\subseteq$-maximal consistent theory $\overline{T} \supseteq T$ and it remains to show that it is maximally complete. Indeed, for any σ-sentence χ, one of $\overline{T} \cup \{\chi\}$ or $\overline{T} \cup \{\neg\chi\}$ is consistent by (2.18.b) of Lemma 2.18, so, by $\subseteq$-maximality, $\overline{T}$ must already contain χ or $\neg\chi$. □

3. Completeness of the proof system and its consequences

The soundness of the proof system 2.8 says that if we have a "syntactical certificate" that something is true (i.e. is a syntactic consequence of T), then it is indeed true (in every model of T). What about the converse: is the validity of φ in every model of T witnessed by an actual formal proof from the axioms of T? If the answer to this question was no, mathematicians would appear in a pretty rough shape since it would be possible that some (first order) statement was true in every model of T (e.g. Hilbert's Nullstellensatz for algebraically closed fields), but we would have no (first order) way of proving it. Fortunately, the answer is YES and that is the content of the Completeness Theorem to which we devote this section.

3.A. **Syntactic-semantic duality, completeness and compactness**

We have already defined some syntactic and semantic notions for a theory T, and, in this subsection, we draw analogies between them. Finally, we state the Completeness theorem, which in my opinion should have been called the Syntactic-Semantic Duality theorem. It is called Completeness because it shows that the proof system defined in the previous section is "complete" in the sense that the axioms that we threw in, together with the rule of inference, are enough to prove any statement that is true in every model of T (i.e. semantically implied by T).

The following table compares the notions we have defined.

Table 3.1. Syntactic-semantic duality

Notions	**Syntactic** (Proof-theoretic)	**Semantic** (Model-theoretic)
Consistency	$T \nvdash \bot$	$T \nvDash \bot$, i.e. T is satisfiable
Implication	$T \vdash \varphi$	$T \models \varphi$
Completeness	$\forall \varphi$, $T \vdash \varphi$ or $T \vdash \neg\varphi$	$\forall \varphi$, $T \models \varphi$ or $T \models \neg\varphi$
Compactness	$T \vdash \varphi \implies \exists$ finite $T_0 \subseteq T$, $T_0 \vdash \varphi$	$T \models \varphi \implies \exists$ finite $T_0 \subseteq T$, $T_0 \models \varphi$

Although the statements in each row are clearly analogous, there is no immediate reason to think that they may be equivalent. For example, it is not clear at all whether the semantic version of the compactness statement is true. This is why one should appreciate the following.

Theorem 3.2 (Completeness of FOL; Gödel, 1929)**.** *Every consistent σ-theory T is satisfiable. In fact, it has a model of cardinality at most* $\max\{|\sigma|, \aleph_0\}$.

Remark 3.3 (silly). The completeness of FOL should NOT be confused with the completeness of a theory; these are two completely different notions, they just use the same adjective (unfortunate terminology). I put this remark here because I have had students ask me whether Gödel's Completeness theorem contradicts his Incompleteness theorem. The first one means Completeness of the FOL proof system, whereas the second means Incompleteness of the theory PA.

Before proceeding with a proof of this theorem, let us mention a couple of very important immediate corollaries.

Corollary 3.4 (Syntactic-semantic duality)**.** *For a σ-theory T and a σ-sentence φ,*

$$T \vdash \varphi \iff T \models \varphi.$$

In particular, the statements in each row of the above table are equivalent.

Proof. We only prove that $T \models \varphi$ implies $T \vdash \varphi$ since the rest easily follows from it. We show the contrapositive. Suppose $T \nvdash \varphi$, in particular T is consistent (inconsistent theories prove everything). Moreover, $T \cup \{\neg\varphi\}$ is consistent by (2.18.a), so the Completeness theorem gives a model $\boldsymbol{M} \models T \cup \{\neg\varphi\}$, and hence, $T \nvDash \varphi$. □

Remark 3.5. If one somehow manages to prove a first-order statement φ about all models of T using non-first-order methods, the syntactic-semantic duality implies that there is a first-order proof of φ from T and using external methods was unnecessary.

A theory is called *finitely satisfiable* if every finite subset of it is satisfiable. Rephrasing the semantic version of the compactness statement above, we get (probably) the most useful theorem of logic:

Theorem 3.6 (Compactness)**.** *If a σ-theory T is finitely satisfiable, then it is satisfiable. In fact, it has a model of cardinality at most* $\max\{|\sigma|, \aleph_0\}$.

Proof. Because T is finitely satisfiable, every finite subset of it is consistent. Hence T is consistent and the Completeness theorem applies. □

The Compactness theorem has a wide range of applications and we will mention some of them in the upcoming lectures.

3.B. **Henkin's proof of Gödel's Completeness Theorem**

In this subsection we give a proof of Gödel's Completeness theorem that is due to Henkin.

We start with a consistent theory T in a signature σ and our goal is to build a model for it. To appreciate the difficulty of this task, think of the following particular case: given a set of (first order) conditions together with the field axioms, how hard would it be to construct a field satisfying those conditions? In this example at least, our knowledge of algebra may help finding or constructing such a field, but to build a model for T, it's not even clear where to start.

The first question we need to address is what underlying set we should take for our future model. In general, the objects in the underlying sets of different structures are of different nature; for example, the objects in the group $GL_n(\mathbb{R})$ are matrices, whereas those in the group S_n are permutations. But of course, we can always take isomorphic copies of these structures whose underlying sets are build of the same "material", such as names or symbols. More precisely, given a σ-structure $\boldsymbol{A} := (A, \sigma)$, we can give a name c_a to each element $a \in A$, obtaining a new underlying set $C_A := \{c_a : a \in A\}$ and a σ-structure $\boldsymbol{A}' := (C_A, \sigma)$ isomorphic to $\boldsymbol{A}$, but the objects in the underlying set of $\boldsymbol{A}'$ are just names (i.e. symbols). It's like taking $GL_n(\mathbb{R})$ and replacing the matrices with their pictures (JPEG images if you will).

We can use this idea of naming the elements of a given structure even further. Given $\mathrm{Th}(\boldsymbol{A})$, we usually cannot recover the structure $\boldsymbol{A}$ even if we know the underlying set A. However, we can upgrade our signature σ by adding names for elements of A and then the theory of the natural expansion of $\boldsymbol{A}$ to a structure in this upgraded signature would completely determine the structure $\boldsymbol{A}$.

Definition 3.7. For a σ-structure $\boldsymbol{A}$, define a new signature

$$\sigma_A := \sigma \cup \{c_a : a \in A\},$$

where the c_a are treated as (distinct) constant symbols in σ_A. Let $\boldsymbol{A}' := (\boldsymbol{A}, A)$ denote the expansion of $\boldsymbol{A}$ to a σ_A-structure, where the constant symbols c_a are interpreted as one expects: $c_a^{\boldsymbol{A}'} := a$, for every $a \in A$. Call this structure $\boldsymbol{A}'$ the *natural σ_A-expansion* of $\boldsymbol{A}$. Call $\mathrm{Th}(\boldsymbol{A}')$ the *elementary diagram* of $\boldsymbol{A}$, and denote it by $\mathrm{ElDiag}(\boldsymbol{A})$. Also, denote by $\mathrm{Diag}(\boldsymbol{A})$ the set of all quantifier free sentences in $\mathrm{Th}(\boldsymbol{A}')$ and call it the *diagram* of $\boldsymbol{A}$.

Now the structure $\boldsymbol{A}' := (A, \sigma_A)$ is such that every element in its underlying set A has a name (i.e. a corresponding constant symbol) in the signature σ_A, so $\mathrm{Th}(\boldsymbol{A}')$ will tell us exactly how the constant symbols, function symbols and relation symbols in σ are interpreted in $\boldsymbol{A}$; for example, if $a_1, a_2, a_3 \in A$ and $f^{\boldsymbol{A}'}(a_1, a_2) = a_3$, then $\mathrm{Th}(\boldsymbol{A}')$ would contain the σ_A-sentence $f(c_{a_1}, c_{a_2}) = c_{a_3}$; for groups this would correspond to the multiplication table. Moreover, $\mathrm{ElDiag}(\boldsymbol{A})$ also includes quantified statements about the elements of A. In particular, $\boldsymbol{A} \models \exists v \varphi$ if and only if $\exists v \varphi \in \mathrm{ElDiag}(\boldsymbol{A})$. Furthermore, the latter holds if and only if $\varphi(c/v) \in \mathrm{ElDiag}(\boldsymbol{A})$, for some constant symbol $c \in \sigma_A$. We refer to this c as a *Henkin witness* below.

Why is this useful for us in proving the Completeness theorem? Well, we are to build a model of T, so we have to define interpretations of the symbols in σ so they agree with T. Therefore, it would be really nice if T could tell us exactly how to define those interpretations because if we do exactly as T says, then we would naturally end up with a σ-structure modeling the quantifier free sentences of T. It would be even better, if T could tell us which formulas of the form $\exists v \varphi$ our future model should satisfy. In other words, we would like our T to "look like" an elementary diagram of some σ-structure, so we can take that σ-structure as our model. The following definition makes all this precise.

Definition 3.8. For a signature σ, a σ-theory H and a σ-formula $\exists v\varphi$, we say that H admits a *Henkin witness* for $\exists v\varphi$ if $H \vdash \varphi(c/v)$ for some $c \in \mathrm{Const}(\sigma)$. A σ-theory H is called a *σ-Henkin theory* (or just a Henkin theory) if H is consistent, maximally complete, and admits a Henkin witness for every σ-formula of the form $\exists v\varphi$ that it proves.

Observation 3.9. *For any σ-structure $\boldsymbol{A}$, $\mathrm{ElDiag}(\boldsymbol{A})$ is a σ_A-Henkin theory.*

Note that the existence of a σ-Henkin theory implies that σ has at least one constant symbol. Our initial signature σ may not contain enough constants to be used as Henkin witnesses, so we artificially create them and throw them into σ; more precisely, we define increasing sequence (σ_n) of signatures as follows: take $\sigma_0 := \sigma$ and, for each $n \in \mathbb{N}$, let

$$\sigma_{n+1} := \sigma_n \cup \left\{c_{\exists v\varphi} : \varphi \in \mathrm{Formulas}(\sigma_n)\right\},$$

where symbols in the last set are treated as constant symbols in σ_{n+1}. Lastly, let $\overline{\sigma} := \Uparrow_n \sigma_n$.

Lemma 3.10 (Constructing a Henkin theory)**.** *Any consistent σ-theory T admits a $\overline{\sigma}$-Henkin extension $H \supseteq T$.*

Proof. The proof is very similar to that of the Downward Löwenheim–Skolem theorem 1.67, but instead of structures we build theories and instead of adding Tarski–Vaught witnesses we add Henkin witnesses. We define an increasing sequence (T_n) of extensions of T, where each T_n is a consistent maximally complete σ_n-theory. Let T_0 be a consistent σ_0-completion of T and suppose that T_n is defined. Putting

$$H_n := \left\{\varphi(c_{\exists v\varphi}/v) : \exists v\varphi \in T_n\right\},$$

it follows from (2.18.c) (or rather its proof) that $T_n \cup H_n$ is still consistent because for each $\varphi(c_{\exists v\varphi}/v) \in H_n \setminus T_n$, $\exists v\varphi \in T_n$ and the symbol $c_{\exists v\varphi}$ does not appear in T_n. Let T_{n+1} be a consistent σ_{n+1}-completion of $T_n \cup H_n$. Finally taking $H := \bigcup_{n\in\mathbb{N}} T_n$, it is easy to verify that H is a consistent maximally $\overline{\sigma}$-complete extension of T. Because every $\overline{\sigma}$-sentence $\exists v\varphi$ is actually a σ_n-sentece for some $n \in \mathbb{N}$ (once again, formulas are finite), it follows by our construction that H contains $\varphi(c_{\exists v\varphi}/v)$ whenever it contains $\exists v\varphi$, so H is a $\overline{\sigma}$-Henkin theory. □

Having constructed a $\overline{\sigma}$-Henkin theory H, we now construct a model of H, i.e. a $\overline{\sigma}$-structure satisfying H and then take its reduct to the signature σ (i.e. forget the names of Henkin witnesses). Thus, for the rest of the proof, we let H be a τ-Henkin theory, for some signature τ, and we build a τ-structure that models H.

Lemma 3.11 (Henkin theories calculate terms)**.** *Let H a τ-Henkin theory. For any τ-term t with no variables, $t \doteq c \in H$ for some $c \in \mathrm{Const}(\tau)$.*

Proof. We aim at getting such $c \in \tau$ as a Henkin witness to $\exists v\varphi(v)$, where $\varphi(v) := t \doteq v$, so it is enough to show that $H \vdash \exists v\varphi$. But, by the equality-is-equivalence (6), generalization (5), and instantiation (4) axioms, $H \vdash t \doteq t$, and the latter sentence is precisely $\varphi(t/v)$. Thus, $H \vdash \varphi(t/v)$, so using witness-implies-existence (2.11.h), we get $H \vdash \exists v\varphi$. □

Lemma 3.12 (Constructing a model for a Henkin theory)**.** *Any τ-Henkin theory H has a model.*

Proof. As our first attempt, we take $A := \mathrm{Const}(\tau)$ as the universe of our future model $\boldsymbol{A}$ with the following interpretations:

$$\begin{aligned}
c^{\boldsymbol{A}} &:= c, && \text{for each } c \in \mathrm{Const}(\tau);\\
f^{\boldsymbol{A}}(\vec{a}) = b &:\Leftrightarrow f(\vec{a}) \doteq b \in H, && \text{for each } f \in \mathrm{Func}(\tau),\ \vec{a} \in A^{|\mathfrak{a}(f)|}, \text{ and } b \in A;\\
R^{\boldsymbol{A}}(\vec{a}) &:\Leftrightarrow R(\vec{a}) \in H, && \text{for each } R \in \mathrm{Rel}(\tau) \text{ and } \vec{a} \in A^{|\mathfrak{a}(R)|}.
\end{aligned}$$

This construction almost works except that it may well be that $a \doteq b \in H$, for distinct $a, b \in A$. Because of this, $\boldsymbol{A}$ is not even a τ-structure since the interpretations of the function symbols of τ (second clause above) are not well defined, but this is a secondary issue and would be fixed once the first is fixed. We fix the first issue by modding out A by the equivalence relation $\sim_H$ on A defined by:

$$a \sim_H b :\Leftrightarrow a \doteq a \in H.$$

It follows from the equality-is-equivalence axioms and the completeness of H that $\sim_H$ is indeed an equivalence relation.

Put $M := A/\sim$, so $M = \{[a]_H : a \in A\}$, where $[a]_H$ denotes the $\sim_H$-equivalence class of a; for a vector $\vec{a} := (a_1, a_2, \dots a_n) \in A^n$, we also write $[\vec{a}]_H$ to mean $\big([a_1]_H, [a_2]_H, \dots, [a_n]_H\big)$. We define a τ-structure $\boldsymbol{M}$ with universe M and the following interpretations:

$$\begin{array}{rcll} c^{\boldsymbol{M}} & := & [c]_H, & \text{for each } c \in \mathrm{Const}(\tau); \\ f^{\boldsymbol{M}}([\vec{a}]_H) = [b]_H & :\Leftrightarrow & f(\vec{a}) \doteq b \in H, & \text{for each } f \in \mathrm{Func}(\tau),\ \vec{a} \in A^{|\mathfrak{a}(f)|}, \text{ and } b \in A; \\ R^{\boldsymbol{A}}([\vec{a}]_H) & :\Leftrightarrow & R(\vec{a}) \in H, & \text{for each } R \in \mathrm{Func}(\tau) \text{ and } \vec{a} \in A^{|\mathfrak{a}(R)|}. \end{array}$$

Claim 1. $\boldsymbol{M}$ is well-defined.

Proof of Claim. One has to prove that the definitions of $R^{\boldsymbol{M}}$ and $f^{\boldsymbol{M}}$ do not depend on the choice of the representatives of the equivalence classes, but this immediately follows from the functions-respect-equality (7) and relations-respect-equality (8). Moreover, for each $f \in \mathrm{Func}(\tau)$, one has to verify that for all $\vec{a} \in A^{\mathfrak{a}(f)}$, there *does exist* $b \in A$ with $f(\vec{a}) \doteq b \in H$, but this is just an instance of Lemma 3.11. ⊠

Claim 2. For every τ-term t with no variables and $b \in A$, $t^{\boldsymbol{M}} = [b]_H$ if and only if $t \doteq b \in H$.

Proof of Claim. We induct on the recursive construction of t. The case of t being a variable is excluded, so the only base case is $t = c$ for $c \in \mathrm{Const}(\tau)$. But then, by definition,

$$c^{\boldsymbol{M}} = [b]_H \Leftrightarrow c \doteq b \in H.$$

Now assume that $t = f(t_1, t_2, \dots, t_n)$ for some $f \in \mathrm{Func}_n(\tau)$ and $n \in \mathbb{N}$. Let $\vec{a} := (a_1, a_2, \dots, a_n) \in A^n$ be such that $t_i^{\boldsymbol{M}} = [a_i]$ for all i. By induction, we have that $t_i \doteq a_i \in H$, so by the functions-respect-equality (7) axiom, we also have that $H \vdash f(t_1, t_2, \dots, t_n) = f(\vec{a})$. But then, by definition,

$$f^{\boldsymbol{M}}([\vec{a}]_H) = [b]_H \Leftrightarrow f(\vec{a}) \doteq b \in H \Leftrightarrow H \vdash f(t_1, t_2, \dots, t_n) \doteq b,$$

where the last equivalence is due to the functions-respect-equality (6) axiom. ⊠

Claim 3. $\boldsymbol{M} \models H$.

Proof of Claim. We show that for every extended τ-formula $\theta(\vec{x})$ and $\vec{a} \in A^{|\vec{x}|}$,

$$\boldsymbol{M} \models \theta\big([\vec{a}]_H\big) \iff \theta(\vec{a}/\vec{x}) \in H,$$

by induction on the length of θ. The base case of equality is handled by the previous claim, and the base case of a relation symbol follows from the same claim and the definition of $\boldsymbol{M}$ using the equality-is-equivalence (6) and relations-respect-equality (8) axioms.

For the step of induction, we consider the cases with connectives $\neg$ and $\to$ and the quantifier $\exists$. The case of $\neg$ follows easily from the induction hypothesis and the consistency and maximal completeness of H. For the case $\theta = \varphi \to \psi$, observe:

$$\begin{aligned} \boldsymbol{M} \models \theta\big([\vec{a}]_H\big) &\iff \boldsymbol{M} \models \neg\varphi\big([\vec{a}]_H\big) \text{ or } \boldsymbol{M} \models \psi\big([\vec{a}]_H\big) \\ \left[\begin{array}{l}\text{induction hypothesis, uses}\\ \text{that } \neg\varphi \text{ is shorter than } \theta\end{array}\right] &\iff \neg\varphi(\vec{a}/\vec{x}) \in H \text{ or } \psi(\vec{a}/\vec{x}) \in H \\ \left[\begin{array}{l}\text{by the if-true-then-implied (1) axiom}\\ \text{and if-false-then-implies (2.11.c)}\end{array}\right] &\iff \varphi \to \psi \in H. \end{aligned}$$

We now handle the remaining case of $\theta = \exists v \varphi$ as follows:

$$\begin{aligned} \boldsymbol{M} \models \theta\big([\vec{a}]_H\big) &\iff \text{there is } [b]_H \in M \text{ such that } \boldsymbol{M} \models \varphi\big([\vec{a}]_H, [b]_H\big) \\ \Big[\text{induction hypothesis}\Big] &\iff \text{there is } b \in A \text{ such that } \varphi(\vec{a}/\vec{x}, b/v) \in H \\ \left[\begin{array}{l}\Longrightarrow \text{ is by witness-implies-existence}\\ \text{(2.11.h) and } \Longleftarrow \text{ is because } H \text{ admits}\\ \text{a Henkin witness for } \exists v \varphi(\vec{a}/\vec{x})\end{array}\right] &\iff \exists v \varphi(\vec{a}/\vec{x}, v) \in H. \end{aligned}$$

⊠

The last claim finishes the proof of the lemma. □

Proof of the Completeness Theorem 3.2 (Henkin, 1949). By Lemma 3.10, there is a $\overline{\sigma}$-Henkin theory $H \supseteq T$. Now applying Lemma 3.12 to $\overline{\sigma}$ and H, we get a model $\boldsymbol{M}$ of H of cardinality at most $|\overline{\sigma}|$ and hence at most $\kappa := \max\{|\sigma|, \aleph_0\}$. Finally, we take the reduct of $\boldsymbol{M}$ to the signature σ. □

From now on, we will not differentiate between the syntactic and semantic notions in Table 3.1.

3.C. **Upward Löwenheim–Skolem theorem**

One of the numerous consequences of the Compactness theorem is the following general statement about cardinalities of models.

Theorem 3.13 (Weak Upward Löwenheim–Skolem)**.** *For a σ-theory T the following are equivalent:*

(1) *T has an infinite model.*
(2) *For every $n \in \mathbb{N}$, T has a model of cardinality at least n.*
(3) *T has a model of any cardinality $\kappa \geq \max\{|\sigma|, \aleph_0\}$.*

Proof. We prove the only nontrivial direction (2)⇒(3). Put $\overline{\sigma} := \sigma \cup \{c_\alpha\}_{\alpha<\kappa}$, where c_α are constant symbols that are not in σ. The theory

$$\overline{T} := T \cup \left\{c_\alpha \neq c_\beta : \alpha \neq \beta, \alpha, \beta < \kappa\right\}$$

is finitely satisfiable by (2). Thus, by the Compactness theorem, $\overline{T}$ has a model $\boldsymbol{M}$ of cardinality at most κ since $|\overline{\sigma}| = \kappa \geq \aleph_0$. On the other hand, $|M| \geq \kappa$ since $c_\alpha^{\boldsymbol{M}} \neq c_\beta^{\boldsymbol{M}}$ for distinct $\alpha, \beta < \kappa$. Thus $|M| = \kappa$. □

This theorem implies for example that PA has uncountable models!

Recall that the Downward Löwenheim–Skolem theorem gives us an elementary substructure $\boldsymbol{A}$ of a given σ-structure $\boldsymbol{B}$ of any cardinality $\kappa \leq |B|$ as long as $\kappa \geq \max\{|\sigma|, \aleph_0\}$. We would like to also get an upward version of this, i.e. start with a σ-structure $\boldsymbol{A}$ and get an elementary extension $\boldsymbol{B} \geq \boldsymbol{A}$ of any cardinality $\geq \max\{|A|, |\sigma|, \aleph_0\}$. To achieve this, we may consider applying the previous theorem to $\mathrm{Th}(\boldsymbol{A})$. However, this would only give us a structure $\boldsymbol{B}$ that is elementarily equivalent to $\boldsymbol{B}$, i.e. $\boldsymbol{A} \equiv \boldsymbol{B}$, whereas we want $\boldsymbol{A} \hookrightarrow_e \boldsymbol{B}$. So instead, we apply the previous theorem to the elementary diagram $\mathrm{ElDiag}(\boldsymbol{A})$ of $\boldsymbol{A}$ (see Definition 3.7), and the following lemma tells us why.

Lemma 3.14. *For σ-structures $\boldsymbol{A}, \boldsymbol{B}$, if an expansion $\boldsymbol{B}'$ of $\boldsymbol{B}$ is a model of $\mathrm{ElDiag}(\boldsymbol{A})$, then $\boldsymbol{A} \hookrightarrow_e \boldsymbol{B}$. In particular, there is an isomorphic copy of $\boldsymbol{B}$ containing $\boldsymbol{A}$ as an elementary substructure.*

Proof. Let $h : A \to B$ be the map $a \mapsto c_a^{\boldsymbol{B}'}$, i.e. h maps an element $a \in A$ to the interpretation in $\boldsymbol{B}'$ of the corresponding constant symbol c_a. It is straightforward to check that f is an elementary embedding. □

Theorem 3.15 (Upward Löwenheim–Skolem)**.** *Any infinite σ-structure $\boldsymbol{A}$ has an elementary extension of any cardinality $\kappa \geq \max\{|A|, |\sigma|, \aleph_0\}$; more precisely, there is a σ-structure $\boldsymbol{B}$ such that $|B| = \kappa$ and $\boldsymbol{A} \preceq \boldsymbol{B}$.*

Proof. By the weak upward Löwenheim–Skolem, get a model $\boldsymbol{B}$ of $\mathrm{ElDiag}(\boldsymbol{A})$ of cardinality κ and apply the previous lemma. □

3.D. **Nonstandard models of arithmetic**

A *nonstandard model of Peano Arithmetic* is any model of PA that is not isomorphic to the *standard* model $\boldsymbol{N} := (\mathbb{N}, 0, S, +, \cdot)$. As mentioned above, PA has uncountable models and hence they are nonstandard. In this subsection, we construct a countable nonstandard model of PA using the Compactness theorem.

For the rest of the subsection we work in the signature $\sigma_{\mathrm{arthm}} := (0, S, +, \cdot)$.

For each $n \in \mathbb{N}$, recursively define a σ_{arthm}-term $\Delta(n)$ by setting $\Delta(0) := 0$ and $\Delta(n+1) := S\big(\Delta(n)\big)$. Note that for every $n \in \mathbb{N}$, $\boldsymbol{N} \models \Delta(n) \doteq n$ and hence $\mathbb{N} = \left\{\Delta(n)^{\boldsymbol{N}} : n \in \mathbb{N}\right\}$.

Proposition 3.16. *The theory $\mathrm{Th}(\boldsymbol{N})$, and hence also PA, admits a countable nonstandard model.*

Proof. Let w be a new constant symbol not in σ_{arthm} and consider the signature $\sigma := \sigma_{\mathrm{arthm}} \cup \{w\}$. Put

$$T := \mathrm{Th}(\boldsymbol{N}) \cup \{w \neq \Delta(n) : n \in \mathbb{N}\}.$$

T is finitely satisfiable because for any finite $T_0 \subseteq T$, letting n be the maximum number with $w \neq \Delta(n) \in T_0$, the expansion of $\boldsymbol{N}$ to a σ-structure with w being interpreted as $n+1$ satisfies T_0. Thus, by the Compactness theorem, T has a countable model $\boldsymbol{M}$.

To see that this $\boldsymbol{M}$ is nonstandard, assume for contradiction that there is an isomorphism $h : \boldsymbol{N} \to \boldsymbol{M}$. Since $h(\Delta(n)^{\boldsymbol{N}}) = \Delta(n)^{\boldsymbol{M}}$, $h(\mathbb{N}) = \left\{\Delta(n)^{\boldsymbol{M}} : n \in \mathbb{N}\right\}$. But then $w^{\boldsymbol{M}} \notin h(\mathbb{N})$, hence h is not surjective, a contradiction. □

3.E. **From finite to infinite and back**

The Compactness theorem provides a transfer principle between finitary and infinitary statements, and we discuss both directions here.

3.E.1. *From finite to infinite.* The following instance of Weak Upward Löwenheim-Skolem Theorem 3.13 is a simple of example of transfer from finite to infinite.

Corollary 3.17. *If a σ-theory T has arbitrarily large finite models, then it has an infinite model.*

Given that some property P holds for all finite subsets of a given structure, we can often conclude via the Compactness theorem that P holds for the entire structure. For example, if every finite subgraph of a graph is k-colorable, then the entire graph is k-colorable. Similarly, if every finite subgraph is contained in a larger subgraph that admits a perfect matching, then the entire graph admits a perfect matching. In both of these examples, P asserts existence of some object O (relations or functions) satisfying a certain property Q. Knowing that for each finite subset F of the structure, such an object O_F exists, one might hope to write the underlying set A our structure as a directed union of finite subsets and build a corresponding directed sequence of objects O_F (e.g. partial k-colorings) that coheres with $\subseteq$ on these finite sets; then, taking the "union" of these objects O_F, one would obtain an object O that is global, i.e. works for the entire structure. The Compactness theorem asserts the existence of such a coherent sequence, but in our proofs, we don't even have to think about it. Here is an example.

Corollary 3.18. *If every finite subgraph of a graph $\boldsymbol{G} \coloneqq (V, E)$ is 3-colorable, then the entire graph $\boldsymbol{G}$ is 3-colorable.*

Proof. Extend the signature $\sigma_{\mathrm{gr}} \coloneqq (E)$ to $\overline{\sigma}$ by adding a set C_V of names (constant symbols) for elements of V as well as unary relation symbols R_0, R_1, R_2 corresponding to the 3 colors. Let T be the $\overline{\sigma}$-theory that includes $\mathrm{ElDiag}(\boldsymbol{G})$ as well as finitely-many sentences asserting that R_0, R_1, R_2 form a partition, which is a 3-coloring. Any finite $T_0 \subseteq T$ only mentions finitely-many constants from C_V, so T_0 is satisfiable by the induced subgraph of $\boldsymbol{G}$. Thus, the Compactness theorem gives a model $\overline{\boldsymbol{G}}$ of T, which is 3-colorable and (elementarily) embeds $\boldsymbol{G}$, so $\boldsymbol{G}$ is also 3-colorable. □

Observe that in this and many other examples, the Compactness theorem switches (non-first-order) quantifiers $\forall\exists$ to $\exists\forall$. Indeed, we are given that *for all* finite subsets F *there is* a certain object R_F that "works" for F, and what we get is that *there is* a certain object R that "works" *for all* F at once.

3.E.2. *From infinite to finite.* In arithmetic combinatorics and Ramsey theory, it often happens that one proves an infinitary theorem (e.g. theorems of Ramsey, van der Waerden, Szemerédi, etc.) by infinitary means (i.e. idealistic tools, without keeping track of ε's and bounding errors) and then deduces its finitary version via a so-called *compactness-and-contradiction* argument. The latter uses the fact that product of finite topological spaces is compact by Tychonoff's theorem. Here we give an example of such a proof using the Compactness theorem rather than a compactness-and-contradiction argument. Our example will be the deduction of the Finite Ramsey theorem from its famous Infinite counterpart.

For a set V and $d \geq 1$, let $[V]^d$ denote the set of d-element subsets of V. (Think of $[V]^2$ as the set of edges of the undirected complete graph on V.) For $k \geq 0$, put $\bar{k} \coloneqq \{0, 1, \dots, k-1\}$. A *$k$-coloring* of $[V]^d$ is just a function $\chi : [V]^d \to \bar{k}$. A set $E \subseteq [\mathbb{N}]^d$ is said to be *χ-monochromatic* if all elements of E have the same color, i.e. $\chi|_E$ is constant. A vertex-set $A \subseteq \mathbb{N}$ is called *χ-monochromatic* if $[A]^d$ is χ-monochromatic.

Theorem 3.19 (Infinite Ramsey)**.** *For any 2-coloring χ of $[\mathbb{N}]^2$, $\mathbb{N}$ has an infinite χ-monochromatic subset.*

Proof. For $a \in \mathbb{N}$ and $A \subseteq \mathbb{N}$, put $[a, A] \coloneqq \big\{\{a, a'\} : a' \in A \setminus \{a\}\big\}$. We inductively define sets a decreasing sequence (A_n) of infinite subsets of $\mathbb{N}$ such that $[a_n, A_{n+1}]$ is χ-monochromatic, where $a_n \coloneqq \min A_n$. Putting $A_0 \coloneqq \mathbb{N}$, suppose that A_n is defined and infinite. Hence, the set $[a_n, A_n]$ is infinite and 2-colored by χ. By the Infinite Pigeonhole Principle, there is an infinite χ-monochromatic subset $E_n \subseteq [a_n, A_n]$. Let $A_{n+1} \subseteq A_n$ be the set defined by $E_n = [a_n, A_{n+1}]$.

Let $A \coloneqq \{a_n : n \in \mathbb{N}\}$ and define a coloring $\chi' : A \to \bar{2}$ by coloring a_n with the common color of $[a_n, A_{n+1}]$. By the Infinite Pigeonhole Principle, again, there is an infinite χ'-monochromatic subset $B \subseteq A$, whose common χ'-color is, say, 0. We claim that B is also χ-monochromatic with common χ-color of $[B]^2$ also being 0: indeed, for any $a_n, a_m \in B$ with $n < m$, $a_m \in A_{n+1}$, so $\chi\big(\{a_n, a_m\}\big) = \chi'(a_n) = 0$. □

Example 3.20. The Infinite Ramsey theorem can be used to show that every sequence $(x_n)_{n\in\mathbb{N}}$ of reals has a monotone subsequence. Indeed, color a pair $n < m$ blue if $x_n < x_m$, and red, otherwise.

We now derive the Finite Ramsey theorem from this using the Compactness theorem. The original combinatorial proof is much messier (look it up).

Theorem 3.21 (Finite Ramsey)**.** *For every $m \geq 1$, there exists $n \geq m$ such that for any 2-coloring χ of $[\bar{n}]^2$, there is a χ-monochromatic subset of $\bar{n}$ of size m.*

Proof. Let σ be the signature containing constant symbols c_n, for every $n \in \mathbb{N}$, and a binary relation symbol R. Think of R as a symbol for coloring: the color of $\{x, y\}$ is 1 if $R(x, y)$ holds, and it is 0, otherwise. Fix $m \geq 1$, and for each $n \geq m$, let φ_n be a σ-sentence expressing that $c_0, c_1, \ldots, c_{n-1}$ are pairwise distinct and the set $\{c_0, c_1, \ldots, c_{n-1}\}$ does not have a monochromatic subset of cardinality m (there are only finitely-many such subsets, so the nonexistence of a monochromatic one is expressed by a very big, yet finite, conjunction).

Now suppose towards a contradiction that for any $n \geq m$, there is a 2-coloring of $[\bar{n}]^2$ such that $\bar{n}$ has no monochromatic subsets of cardinality m. Thus, the theory $T := \{\varphi_n : n \in \mathbb{N}\}$ is finitely satisfiable, and hence, has a model $\boldsymbol{M}$. Let $N := \left\{c_n^{\boldsymbol{M}} : n \in \mathbb{N}\right\}$. By the Infinite Ramsey theorem, N has an infinite monochromatic subset A, i.e. either for all distinct $a, a' \in A$, $R^{\boldsymbol{M}}(a, a')$, or for all distinct $a, a' \in A$, $\neg R^{\boldsymbol{M}}(a, a')$. Let n be large enough so that $A \cap \left\{c_i^{\boldsymbol{M}} : i < n\right\}$ has at least m elements. Then it is clear that $\boldsymbol{M} \nvDash \varphi_n$, a contradiction. □

3.F. **Nonaxiomatizable classes**

One can use the Compactness theorem to show that many interesting classes of structures are not axiomatizable.

3.F.1. *Bounded classes.*

Proposition 3.22. *Let $\mathcal{C}$ be a class of σ-structures. If the cardinalities of the structures in $\mathcal{C}$ are bounded (i.e. bounded by some, possibly infinite, cardinal κ), then $\mathcal{C}$ is not axiomatizable, unless all structures in $\mathcal{C}$ have at most n elements, for some fixed $n \in \mathbb{N}$.*

Proof. Follows from the Weak Upward Löwenheim–Skolem Theorem 3.13. □

Example 3.23. *Cyclic groups*. By the last proposition, the class of cyclic groups is not axiomatizable.

3.F.2. *Infinite conjunctions.*

Notation 3.24. For a set T of σ-formulas and a vector $\vec{x}$ of variables, we write $T(\vec{x})$ for T to imply, in addition, that the free variables of each formula $\varphi \in T$ are among $\vec{x}$ (i.e. $\varphi(\vec{x})$ is an extended formula). In this case, for a vector $\vec{c}$ of constant symbols with $|\vec{c}| = |\vec{x}|$, we put

$$T(\vec{c}/\vec{x}) := \{\varphi(\vec{c}/\vec{x}) : \varphi \in T\}.$$

We also put

$$\begin{aligned} \exists\vec{x}T &:= \{\exists\vec{x}\varphi : \varphi \in T\} \\ \forall\vec{x}T &:= \{\forall\vec{x}\varphi : \varphi \in T\} \\ \neg T &:= \{\neg\varphi : \varphi \in T\}. \end{aligned}$$

Lastly, if T is finite, we put

- $\bigvee T := \bigvee_{\varphi\in T} \varphi$,
- $\bigwedge T := \bigwedge_{\varphi\in T} \varphi$.

(Unlike the first three definitions, the latter two denote σ-formulas.)

Proposition 3.25. *Let $\mathcal{C}$ be a class of σ-structures defined as follows: for some σ-theory T_0 and set $T(\vec{x})$ of σ-formulas, we have that for every σ-structure $\boldsymbol{A}$,*

$$\boldsymbol{A} \in \mathcal{C} \iff \boldsymbol{A} \models T_0 \text{ and for each } \vec{a} \in A^{|\vec{x}|} \text{ there is } \varphi \in T \text{ such that } \boldsymbol{A} \models \varphi(\vec{a}). \tag{3.26}$$

Then $\mathcal{C}$ is not axiomatizable, unless for some finite $T' \subset T$, the theory

$$T_0 \cup \left\{\forall \vec{x}\left(\bigvee T'\right)\right\}$$

axiomatizes $\mathcal{C}$.

Proof. Suppose for contradiction that there is an axiomatization S of $\mathcal{C}$. Enhance the signature by adding a vector $\vec{c}$ of new constant symbols of length $|\vec{x}|$ note that the theory

$$S' := S \cup \neg T(\vec{c}/\vec{x})$$

is finitely satisfiable; indeed, otherwise, by (2.18.a) of Lemma 2.18, for some finite $T' \subseteq T$, $S \vdash \bigvee T'(\vec{c}/\vec{x})$, and hence $S \vdash \forall \vec{x}(\bigvee T')$ by Constant Substitution Lemma 2.14 and the generalization axiom (5). Since every model of $T' \cup \{\forall \vec{x}(\bigvee T')\}$ is in $\mathcal{C}$, it follows that $T_0 \cup \{\forall \vec{x}(\bigvee T')\}$ axiomatizes $\mathcal{C}$, contrary to our assumption. Thus, S' is finitely satisfiable.

But then the Compactness theorem yields a model $\mathcal{M} \models S'$, which must be in $\mathcal{C}$ even though it violates (3.26), a contradiction. □

Examples 3.27.

(a) *Nonbipartite graphs.* Let $T := \{\varphi_{2k+1} : k \in \mathbb{N}\}$, where φ_n expresses that there is a cycle of length n. Clearly, for a graph $\boldsymbol{G} := (V, E)$,

$$\boldsymbol{G} \text{ is nonbipartite} \iff \boldsymbol{G} \models \text{GRAPHS and there is } \varphi \in T \text{ with } \boldsymbol{G} \models \varphi.$$

Thus, the hypothesis of Proposition 3.25 is met, so this class is not axiomatizable.

(b) *Connected graphs.* Let $T(x, y) := \{\varphi_n(x, y) : n \in \mathbb{N}\}$, where $\varphi_n(x, y)$ expresses that there is a path between x and y of length at most n. Clearly, for a graph $\boldsymbol{G} := (V, E)$,

$$\boldsymbol{G} \text{ is connected} \iff \boldsymbol{G} \models \text{GRAPHS and for each } u, v \in V \text{ there is } \varphi \in T \text{ with } \boldsymbol{G} \models \varphi_n(u, v).$$

Thus, the hypothesis of Proposition 3.25 is met, so this class is not axiomatizable.

3.F.3. *Infinite disjunctions.* For this type of nonaxiomatizable classes, we need an observation referred to as *Exists Elimination*.

Observation 3.28 (Exists Elimination)**.** *For an extended σ-formula $\varphi(\vec{x})$, a vector $\vec{c}$ of length $|\vec{x}|$ of constant symbols that are not in σ, and a σ-formula ψ,*

$$\varphi(\vec{c}/\vec{x}) \models \psi \text{ if and only if } \exists \vec{x} \varphi \models \psi,$$

where on the left side, the theory $\{\varphi(\vec{c}/\vec{x})\}$ is a $\overline{\sigma} := \sigma \cup \{\vec{c}\}$-theory.

Proposition 3.29. *Let $\mathcal{C}$ be a class of σ-structures defined as follows: for some set $T(\vec{x})$ of σ-formulas, we have that for every σ-structure $\boldsymbol{A}$,*

$$\boldsymbol{A} \in \mathcal{C} \iff \text{there exists } \vec{a} \in A^{|\vec{x}|} \text{ such that for all } \varphi \in T, \boldsymbol{A} \models \varphi(\vec{a}). \tag{3.30}$$

Then, for any σ-sentence χ, if every structure $\boldsymbol{A} \in \mathcal{C}$ satisfies χ, then there is finite $T' \subseteq T$ with $\exists \vec{x}(\bigwedge T') \models \chi$. In particular, $\mathcal{C}$ is not axiomatizable, unless the theory

$$\left\{\exists \vec{x}\left(\bigwedge T'\right) : T' \subseteq T \text{ finite}\right\}$$

axiomatizes $\mathcal{C}$.

Proof. The last statement follows from the first by applying it to each sentence χ of a hypothetical axiomatization S of $\mathcal{C}$.

To prove the first statement, let χ as in the hypothesis and enhance the signature by adding a vector $\vec{c}$ of new constant symbols of length $|\vec{x}|$. Note that, by (3.30), the σ-reducts of models of $T(\vec{c}/\vec{x})$ are in $\mathcal{C}$, so in particular, they all satisfy χ, and hence $T(\vec{c}/\vec{x}) \models \chi$. By the Compactness theorem, there is a finite $T' \subseteq T$ with $T'(\vec{c}/\vec{x}) \models \chi$, equivalently, $\bigwedge T'(\vec{c}/\vec{x}) \models \chi$, so $\exists \vec{x}(\bigwedge T') \models \chi$, by Exists Elimination 3.28. □

Example 3.31. *Disconnected graphs.* Let $T(x, y) := \{\varphi_n(x, y) : n \in \mathbb{N}\}$, where $\varphi_n(x, y)$ expresses that there is no path between x and y of length n. Clearly, for a graph $\boldsymbol{G} := (V, E)$,

$$\boldsymbol{G} \text{ is disconnected} \iff \text{there are } u, v \in V \text{ such that for all } \varphi \in \text{GRAPHS} \cup T, \ \boldsymbol{G} \models \varphi(u, v).$$

Thus, the hypothesis of Proposition 3.29 is met, so this class is not axiomatizable.

4. Complete theories

As mentioned above, it is easy to see that every consistent theory has a (consistent) completion. So why don't we only consider complete theories and not have to deal with the issues that come with incomplete theories? For example, why don't we just work with $\mathrm{Th}(\boldsymbol{N})$ instead of PA? The problem is that it is hard (in a very precise sense) to check whether a given statement is an axiom of $\mathrm{Th}(\boldsymbol{N})$ or not. For example, is the Twin Prime Conjecture in $\mathrm{Th}(\boldsymbol{N})$? We wish we knew. The whole point of mathematics is to derive complicated statements from "easy-to-verify" axioms. We will see in the next section that a good rigorous approximation of "easy-to-verify" is that we can write a computer program that checks whether a given sentence is an axiom or not. For example, all of the theories in Eq. (1.47) satisfy this criterion.

Now the question is: having defined some reasonable theory, like ACF_p, is it complete? In other words, are these axioms enough to capture the first-order essence of say algebraically closed fields of characteristic p? In this section we develop a sufficient condition for verifying completeness, using which we show that ACF_p is complete.

4.A. The Łoś–Vaught test

Definition 4.1. Let κ be a cardinal. A σ-theory T is called *κ-categorical* if any two models of T of cardinality κ are isomorphic. We say that T is *uncountably categorical* if it is κ-categorical for some uncountable cardinal κ.

Examples 4.2.

(a) The theory $\mathrm{VEC}_{\mathbb{Q}}$ of vector spaces over $\mathbb{Q}$ is uncountably categorical; in fact, it is κ-categorical, for every uncountable cardinal κ.

Proof. This is by virtue of the fact that every vector space has a basis and to construct an isomorphism between vector spaces it is enough to find a bijection between their bases. Details to be added. □

(b) Let DLO be the theory of dense linear orderings without endpoints, i.e. DLO comprises of the following axioms in the signature $\sigma := (<)$:
   (i) Antireflexivity: $\forall x(x \not< x)$
   (ii) Transitivity: $\forall x \forall y \forall z \Big[(x < y \wedge y < z) \to x < z\Big]$
   (iii) Linearity: $\forall x \forall y \Big[(x \neq y \wedge x \not< y) \to y < x\Big]$
   (iv) Density: $\forall x \forall y \Big[x < y \to \exists z (x < z < y)\Big]$
   (v) No endpoints: $\forall x \exists y \exists z (y < x < z)$
   It is not hard to show that DLO is $\aleph_0$-categorical and hence $(\mathbb{Q}, <)$ is the only (up to isomorphism) countable dense linear ordering without end points. We leave proving this as an exercise.

(c) For a finite σ-structure $\boldsymbol{A}$, $\mathrm{Th}(\boldsymbol{A})$ is *absolutely categorical*, i.e. any two models $\boldsymbol{B}, \boldsymbol{B}' \models \mathrm{Th}(\boldsymbol{A})$ are isomorphic. It is an exercise to show that the theory of any finite structure is absolutely categorical.

(d) We will see shortly that an argument similar to that for vector spaces shows that ACF_p is κ-categorical as well (for every uncountable cardinal κ).

Proposition 4.3 (Łoś–Vaught test)**.** *Let T be a σ-theory that does not have finite models. If T is κ-categorical for some $\kappa \geq \max\{|\sigma|, \aleph_0\}$, then T is complete.*

Proof. Let $\boldsymbol{A}, \boldsymbol{B} \models T$. We need to show that $\boldsymbol{A} \equiv \boldsymbol{B}$, by Proposition 1.55. Since $\boldsymbol{A}$ and $\boldsymbol{B}$ are infinite, we can apply the Weak Upward Löwenheim–Skolem theorem 3.13 and get $\boldsymbol{A}' \models \mathrm{Th}(\boldsymbol{A})$ and $\boldsymbol{B}' \models \mathrm{Th}(\boldsymbol{B})$ such that $|A'| = \kappa = |B'|$. Because T is κ-categorical, $\boldsymbol{A}' \cong \boldsymbol{B}'$ and hence $\boldsymbol{A}' \equiv \boldsymbol{B}'$. Thus, $\boldsymbol{A} \equiv \boldsymbol{A}' \equiv \boldsymbol{B}' \equiv \boldsymbol{B}$. □

This immediately gives that the theories $\mathrm{VEC}_{\mathbb{Q}}$ and DLO are complete. One cannot help mentioning here the following very important theorem that started the modern model theory:

Theorem (Morley, 1965)**.** *Let T be a theory in a countable signature σ. If T is uncountably categorical, then it is κ-categorical for every uncountable cardinal κ.*

Thus, it is not a coincidence that the theory of vector spaces is κ-categorical for all uncountable cardinals κ. The proof of this theorem is far outside the realm of this course, but it is worth mentioning that the most important ingredient of it is showing that if a structure is such that all of its definable sets are either finite or cofinite (complement is finite), then it admits a "basis" similar to the vector space basis, and so one can use the same argument as for vector spaces to construct isomorphisms.

One has to also mention the following long standing open problem that, although being model-theoretic in nature, has been best understood (but not completely solved) in the context of descriptive set theory. Let's examine the possible infinite cardinalities for the number of countable nonisomorphic models of a given σ-theory T. It is not hard to see that there are at most $2^{|\max|\sigma|,\aleph_0|}$-many σ-structures with universe $\mathbb{N}$, so when σ is countable, there are at most continuum-many σ-structures. The following conjecture is trivially true when the Continuum Hypothesis (CH) holds, but it is still open when CH fails.

Vaught's Conjecture 4.4. *Let σ be a countable signature and T be a complete σ-theory having infinite models. If T has uncountably-many nonisomorphic countable models, then it has continuum-many nonisomorphic countable models.*

4.B. **Algebraically closed fields and the Lefschetz principle**

We now aim at satisfying the conditions of the Łoś–Vaught test for ACF_p.

The proof of the following is similar to that of the theory of vector spaces being uncountably categorical, and can be safely omitted by the reader not comfortable with field theory.

Definition 4.5. Let $F \subseteq K$ be fields and let $B \subseteq K$. We denote by $F(B)$ the subfield of K generated by B. Call $B \subseteq K$ *algebraically independent over F* if for any finite subset $B_0 \subseteq B$ and $b \in B \setminus B_0$, b is not algebraic over $F(B_0)$, i.e. it is not a root of a polynomial over $F(B_0)$. Call B a transcendence basis for K over F if it is $\subseteq$-maximal algebraically independent over F.

Lemma 4.6. *Let $F \subseteq K$ be fields. If F is countable and K is uncountable, then any transcendence basis B of K over F has cardinality $|K|$.*

Proof. By the maximality of B, each element of K is a root of a polynomial over $F(B)$, so $|K| = \aleph_0 \cdot |F(B)^{<\mathbb{N}}|$. By the countability of F, $|F(B)| = \aleph_0 \cdot |F| \cdot |B| = \aleph_0 \cdot |B|$, and hence $|F(B)^{<\mathbb{N}}| = \aleph_0 \cdot |B|$, so $|K| = \aleph_0 \cdot |B|$. This implies that B is uncountable because otherwise K would be countable. Hence $\aleph_0 \cdot |B| = |B|$, so $|K| = |B|$. □

Proposition 4.7. *For p prime or 0, ACF_p is κ-categorical for every uncountable cardinal κ.*

Proof. Let $\boldsymbol{K}_1, \boldsymbol{K}_2 \models \mathrm{ACF}_p$ with $|K_1| = |K_2| = \kappa$. For $i = 1, 2$, let F_i be the prime field of $\boldsymbol{K}_i$, i.e. the substructure of $\boldsymbol{K}_i$ generated by $\emptyset$. (If $p = 0$, then F_i is a copy of $\mathbb{Q}$; otherwise it is a copy of $\mathbb{Z}/p\mathbb{Z}$.) Since F_1 and F_2 are clearly isomorphic (as rings), we can assume without loss of generality that $F_1 = F_2 =: F$. Let B_i be a transcendence basis over F in $\boldsymbol{K}_i$, so K_i is equal to the algebraic closure $\overline{F(B_i)}$ of $F(B_i)$ in K_i. By Lemma 4.6, $|B_i| = |K_i| = \kappa$, so there is a bijection $f : B_1 \xrightarrow{\sim} B_2$ (yay!). This f uniquely extends to an isomorphism from $F(B_1)$ to $F(B_2)$, which in its turn extends (not necessarily uniquely) to an isomorphism of $K_1 = \overline{F(B_1)}$ onto $K_2 = \overline{F(B_2)}$. □

Lemma 4.8. *Every algebraically closed field is infinite.*

Proof. For any finite field $F := \{a_1, \dots, a_n\}$, the polynomial $(x - a_1)(x - a_2)\dots(x - a_n) + 1$ does not have a root in F. Thus F is not algebraically closed. □

Corollary 4.9. *ACF_p is complete, for any prime p and for $p = 0$.*

Proof. Follows from the Łoś–Vaught test 4.3, Proposition 4.7, and Lemma 4.8 put together. □

The following was once just a principle (a belief) in algebraic geometry, but it was later on formalized and turned into a theorem by A. Robinson:

Theorem 4.10 (Lefschetz Principle)**.** *Let $\boldsymbol{C} := (\mathbb{C}, 0, 1, +, -, \cdot)$. For a σ_{ring}-sentence φ the following are equivalent:*

(1) $\boldsymbol{C} \models \varphi$.
(2) $\boldsymbol{K} \models \varphi$, *for some* $\boldsymbol{K} \models \mathrm{ACF}_0$.

(3) $\mathrm{ACF}_0 \models \varphi$.
(4) *For sufficiently large primes* p, $\mathrm{ACF}_p \models \varphi$.
(5) *For infinitely-many primes* p, *there is* $\boldsymbol{K} \models \mathrm{ACF}_p$ *such that* $\boldsymbol{K} \models \varphi$.

Proof. (1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3): Follows from the completeness of ACF_0.
(3) $\Rightarrow$ (4): By the Compactness theorem, there is a finite $T \subseteq \mathrm{ACF}_0$ such that $T \models \varphi$. But then, by the definitions of ACF_0 and ACF_p, for sufficiently large prime p, $\mathrm{ACF}_p \models T$, so $\mathrm{ACF}_p \models \varphi$.

(4) $\Rightarrow$ (5): Trivial.

(5) $\Rightarrow$ (3): We prove the contrapositive: assume (3) fails. But then $\mathrm{ACF}_0 \models \neg\varphi$ and hence, by (3) $\Rightarrow$ (4), for sufficiently large primes p, $\mathrm{ACF}_p \models \neg\varphi$. Therefore (5) is false. □

Corollary 4.11 (Ax's theorem)**.** *Let* $f : \mathbb{C}^n \to \mathbb{C}^n$ *be a polynomial map, i.e.* $f = (f_1, \ldots, f_n)$, *where each* $f_i(z_1, \ldots, z_n)$ *is a polynomial in* $z_1, \ldots, z_n$ *with coefficients in* $\mathbb{C}$. *If* f *is injective then it is surjective.*

Proof (Robinson). For fixed n and fixed degree $d := \max_i \{\deg(f_i)\}$, the statement is first-order expressible by a σ_{ring}-sentence $\varphi_{n,d}$, and hence, instead of proving it for the field $\mathbb{C}$, by the Lefschetz principle, it is enough to prove $\varphi_{n,d}$ for the algebraic closure $\overline{\mathbb{F}}_p$ of $\mathbb{F}_p := \mathbb{Z}/p\mathbb{Z}$, for all primes p. So, fix a polynomial map $f : \overline{\mathbb{F}}_p^n \to \overline{\mathbb{F}}_p^n$ of degree d.

It is not hard to check that $\overline{\mathbb{F}}_p$ is an increasing union $\Uparrow_{k\in\mathbb{N}} F_k$ of finite fields F_k, where $F_0 := \mathbb{F}_p$.[5] Thus, letting $k_0 \geq 0$ be large enough so that all of the coefficients involved in the definition of f are in F_{k_0}, we can write:

$$\overline{\mathbb{F}}_p^n = \bigcup_{k\geq k_0} F_k^n.$$

But then, because F_k is a field and the definition of f only uses field operations and elements of F_k, F_k^n is closed under f, i.e. $f(F_k^n) \subseteq F_k^n$, for all $k \geq k_0$. Because f is injective, the Pigeonhole Principle (yay!) gives $f(F_k^n) = F_k^n$, so

$$f(\overline{\mathbb{F}}_p^n) = f\left(\bigcup_{k\geq k_0} F_k^n\right) = \bigcup_{k\geq k_0} f(F_k^n) = \bigcup_{k\geq k_0} F_k^n = \overline{\mathbb{F}}_p^n. \qquad \square$$

4.C. **Reducts of arithmetic**

PA was constructed as an attempt to build a "computationally recognizable" axiomatization of $\mathrm{Th}(\boldsymbol{N})$, where "computationally recognizable" means that there is a computational procedure (i.e. a computer program) for recognizing whether a given sentence is an axiom (we will make this more in the next section). However, as we shall see, Gödel's Incompleteness theorem states that PA is incomplete, so it does not axiomatize $\mathrm{Th}(\boldsymbol{N})$. In fact, there is no "computationally recognizable" axiomatization for $\mathrm{Th}(\boldsymbol{N})$, i.e. any subtheory $T \subseteq \mathrm{Th}(\boldsymbol{N})$ is either incomplete or "computationally unrecognizable".

What about reducts of $\boldsymbol{N}$? Does the theory of $(\mathbb{N}, 0, S)$ or even of $(\mathbb{N}, 0, S, +)$ admit a "computationally recognizable" axiomatization? In other words, where is the border of recognizability? It turns out that unlike $\boldsymbol{N}$, the theories of $(\mathbb{N}, 0, S)$ and $(\mathbb{N}, 0, S, +)$ admit "computationally recognizable" axiomatizations, and this is what we will focus on in this subsection.

We start with $\boldsymbol{N}_S := (\mathbb{N}, 0, S)$. Let $\sigma_S := (0, S)$. Here is our first (and last) attempt of axiomatizing $\mathrm{Th}(\boldsymbol{N}_S)$. Let theory T_S consist of the following axioms:

(S1) Zero has no predecessor: $\forall x(S(x) \neq 0)$.
(S2) The successor function is one-to-one: $\forall x \forall y(S(x) \doteq S(y) \to x \doteq y)$.
(S3) Any nonzero number is a successor of something: $\forall x(x \neq 0 \to \exists y(x \doteq S(y)))$.
(S4) Axiom schema: For each $n \in \mathbb{N}$, there are no n-cycles: $\forall x(S^n(x) \neq x)$, where $S^n(*) := \underbrace{S(S(\ldots S(*)\ldots))}_{n \text{ times}}$.

[5] Indeed, a finite field F has only finitely many polynomials over it because $x^{n|F^*|} = 1$ for each $n \in \mathbb{N}$, where $F^* := F \setminus \{0\}$. Thus, throwing in a root for each results in a finite field again.

Note that (S4) is an axiom schema, i.e. it contains an axiom for every $n \in \mathbb{N}$; in particular, T_S is infinite.

It is clear that any model $\boldsymbol{M}$ of T_S has a standard part $\overline{\mathbb{N}} := \left\{\Delta(n)^{\boldsymbol{M}} : n \in \mathbb{N}\right\}$, where $\Delta(n) := S^n(0)$. Define a binary relation $\sim$ on M as follows: for all $a, b \in M$,

$$a \sim b \iff \text{if for some } n \in \mathbb{N}, \boldsymbol{M} \models S^n(a) \doteq b \text{ or } \boldsymbol{M} \models S^n(b) \doteq a.$$

If a is standard, i.e. $a \in \overline{\mathbb{N}}$, then the equivalence class $[a]$ of a is exactly $\overline{\mathbb{N}}$. If $a \in M$ is nonstandard, then $[a]$ does not have a least element (why?) and hence looks like a $\mathbb{Z}$-chain:

$$\ldots \to * \to a \to S^{\boldsymbol{M}}(a) \to S^{\boldsymbol{M}}(S^{\boldsymbol{M}}(a)) \to \ldots$$

Thus $\boldsymbol{M}$ is a union of $\overline{\mathbb{N}}$ and a bunch of $\mathbb{Z}$-chains. Let $\mathcal{Z}_{\boldsymbol{M}}$ denote the set of $\mathbb{Z}$-chains in $\boldsymbol{M}$ and put $\zeta_{\boldsymbol{M}} := |\mathcal{Z}_{\boldsymbol{M}}|$. Then $|M| = |\mathbb{N}| + \zeta_{\boldsymbol{M}} \cdot |\mathbb{Z}|$ and hence, by basic cardinal arithmetic, M has cardinality $\zeta_{\boldsymbol{M}}$ unless $\zeta_{\boldsymbol{M}}$ is finite, i.e. $|M| = \max\{\zeta_{\boldsymbol{M}}, \aleph_0\}$. In particular, if M is uncountable, then $|M| = \zeta_{\boldsymbol{M}}$.

Proposition 4.12. *T_S is κ-categorical, for every uncountable cardinal κ.*

Proof. Let $\boldsymbol{A}, \boldsymbol{B} \models T_S$ with $|A| = |B| = \kappa$. By above, $\zeta_{\boldsymbol{A}} = |A| = \kappa = |B| = \zeta_{\boldsymbol{B}}$. Thus, there is a bijection $f : \mathcal{Z}_{\boldsymbol{A}} \to \mathcal{Z}_{\boldsymbol{B}}$. Now the standard parts of $\boldsymbol{A}$ and $\boldsymbol{B}$ are clearly isomorphic. Moreover, any $\mathbb{Z}$-chain $C \in \mathcal{Z}_{\boldsymbol{A}}$ is isomorphic to $f(C)$ because any two $\mathbb{Z}$-chains are clearly isomorphic. Thus, combining all these individual isomorphisms together, we get an isomorphism from $\boldsymbol{A}$ to $\boldsymbol{B}$. □

From this and the Łoś–Vaught test, we get

Corollary 4.13. *T_S is complete.*

Turning to $\boldsymbol{N}_+ := (\mathbb{N}, 0, S, +)$, we let $\sigma_+ := (0, S, +)$ and T_+ be the theory consisting of all of the axioms of PA except for the ones involving multiplication, so it is clear that T_+ is "computationally recognizable". The proof of the following theorem will be omitted for now since it uses the technique of quantifier elimination, which will be discussed later.

Theorem 4.14 (Presburger, 1929)**.** *T_+ is complete.*

Thus, as we shall see, the "computational unrecognizability" phenomenon starts with $\boldsymbol{N} := (\mathbb{N}, 0, S, +, \cdot)$.

5. Incomplete theories

We start with an informal definition, which we will formalize later on.

Definition 5.1 (Informal)**.** A σ-theory T is called *recursive* if there is a computer program such that given a σ-sentence φ, it returns YES if $\varphi \in T$, and NO otherwise.

We saw in the previous section that the theories of $(\mathbb{N}, 0, S)$ and $(\mathbb{N}, 0, S, +)$ admit recursive axiomatizations. However, the situation changes once we add multiplication because it enables prime numbers and makes it possible to encode tuples of natural numbers into a single number, and we have the following groundbreaking theorem:

Theorem 5.2 (Incompleteness; Gödel, 1931)**.** *Any recursive theory $T \subseteq \mathrm{Th}(\boldsymbol{N})$ is incomplete. In particular,* PA *is incomplete.*

This section is devoted to the proof of several versions of this theorem and some of its consequences, as well as making the definition of *recursive* precise.

5.A. Sketch of proof of the Incompleteness theorem

> Gödel's theorem is like the late paintings of Claude Monet. It is easy to perceive, but from a certain distance. A close look reveals only fastidious details that one perhaps does not want to know.[6]
>
> *Jean-Yves Girard*

[6]Thanks to Anton Bernshteyn for suggesting to include this quote.

There are infinitely-many proofs of this theorem, but mainly, they split into two groups depending on what they use: self-reference or diagonalization. We will give rigorous proofs of each kind later on. However, mainly for historical reasons, in this subsection we sketch the idea of Gödel's original proof, which uses self-reference. We shall give a more rigorous version of this proof later after developing the basics of recursion theory.

Definition 5.3 (Informal)**.** A function $f : \mathbb{N}^k \to \mathbb{N}$ is called *recursive* if there is a computer program such that given $\vec{a} \in \mathbb{N}^k$ as input, it outputs $f(\vec{a})$. A set/relation $A \subseteq \mathbb{N}^k$ is called recursive if its characteristic function is recursive.

First thing one shows is that recursive functions are arithmetical, i.e. definable (and hence 0-definable) in $\boldsymbol{N} := (\mathbb{N}, 0, S, +, \cdot)$. Thus, any function we can write a computer program for is first-order expressible in the signature of arithmetic.

Let σ be a finite signature, whose symbols are $s_0, \ldots, s_n$. Recalling Convention 2.3, we enumerate the symbols of $\mathbb{FOL}(\sigma)$ as follows:

$$s_0\ s_1\ \ldots\ s_n\ \doteq\ \neg\ \to\ \forall\ ,\ (\ )\ v_0,\ v_1,\ v_2,\ \ldots \tag{5.4}$$

and call the index of a symbol its *code*. For example, the code of s_0 is 0, the code of $\doteq$ is $n+1$ and the code of v_i is $n+8+i$. Using prime numbers and the fact that prime number factorization is unique, we can encode a tuple $(n_1, \ldots, n_k)$ of natural numbers into a single natural number $\langle n_1, \ldots, n_k \rangle := p_1^{n_1+1} \cdot \ldots \cdot p_k^{n_k+1}$, so we can encode formulas since they are just tuples of symbols of $\mathbb{FOL}(\sigma)$. In fact, we can make sure that the coding and decoding operations are recursive (think of computer programs that would do this).

Thus, for a word w in the alphabet $\mathbb{FOL}(\sigma)$, let $\ulcorner w \urcorner$ denote its code. It is now not hard to see that a σ-theory T is recursive if and only if the set $\ulcorner T \urcorner$ of codes of its axioms is recursive (as a subset of $\mathbb{N}$).

Now let σ be the signature of arithmetic, i.e. $\sigma := \sigma_{\text{arthm}}$, and thus we have the above coding since σ_{arthm} is finite. For every $n \in \mathbb{N}$, put $\Delta(n) := S^n(0)$. It is tedious but straightforward to show that there is a recursive function $\text{Sub}_0 : \mathbb{N}^2 \to \mathbb{N}$ such that for any σ_{arthm}-formula φ in which v_0 is not quantified, and for any $m \in \mathbb{N}$,

$$\text{Sub}_0(\ulcorner \varphi \urcorner, m) = \ulcorner \varphi(\Delta(m)/v_0) \urcorner.$$

In English: this function takes m and the code of φ, and returns the code of the formula obtained from φ by replacing all occurrences of v_0 by the term $\Delta(m)$.

As mentioned above, all recursive functions are arithmetical. Hence, there is a σ_{arthm}-formula $\mathbf{Sub}_0(x, y, z)$ such that for all $a, b, c \in \mathbb{N}$,

$$\text{Sub}_0(a, b) = c \iff \boldsymbol{N} \models \mathbf{Sub}_0(a, b, c).$$

Without loss of generality, we can assume v_0 is not quantified in $\mathbf{Sub}_0(x, y, z)$.

Lemma 5.5 (Fixed point for $\boldsymbol{N}$)**.** *For each σ_{arthm}-formula $\varphi(v)$ there is a σ_{arthm}-sentence θ such that*

$$\boldsymbol{N} \models \theta \leftrightarrow \varphi(\ulcorner \theta \urcorner).$$

Proof. Put $\psi(v_0) := \exists z \Big(\mathbf{Sub}_0(v_0, v_0, z) \wedge \varphi(z) \Big)$ and $m := \ulcorner \psi(v_0) \urcorner$. Now we feed $\psi(v_0)$ its own code by letting $\theta := \psi\Big(\Delta(m)/v_0\Big)$, and thus $\text{Sub}_0(m, m) = \ulcorner \psi\Big(\Delta(m)/v_0\Big) \urcorner = \ulcorner \theta \urcorner$. Watch the magic happen:

$$\begin{aligned}
\boldsymbol{N} \models \theta &\iff \boldsymbol{N} \models \psi(m) \\
&\iff \boldsymbol{N} \models \exists z \Big(\mathbf{Sub}_0(m, m, z) \wedge \varphi(z) \Big) \\
&\iff \text{there exists } b \in \mathbb{N} \text{ such that } b = \text{Sub}_0(m, m) \text{ and } \boldsymbol{N} \models \varphi(b) \\
&\iff \boldsymbol{N} \models \varphi(\ulcorner \theta \urcorner).
\end{aligned}$$

If you feel cheated, join the club. □

This lemma says that every unary arithmetical relation $\varphi(v)$ asserts of (the code of) some sentence θ exactly what θ asserts about $\boldsymbol{N}$. It enables self-reference in the language of arithmetic, using which we can express the Liar Paradox (i.e. Cantor's diagonalization method), which is what lies at the heart of the proof of the Incompleteness theorem.

As an immediate corollary we get the following result that is actually stronger than the Gödel's Incompleteness theorem:

Theorem 5.6 (Tarski, 1939)**.** *$Th(\boldsymbol{N})$ is not arithmetical, i.e. the set $\ulcorner \mathrm{Th}(\boldsymbol{N})\urcorner := \{\ulcorner\varphi\urcorner : \varphi \in \mathrm{Th}(\boldsymbol{N})\}$ is not definable in $\boldsymbol{N}$.*

Proof. Left as a homework problem. □

Because formal proofs are just finite sequences of formulas, we can code them using the operation of coding n-tuples. Given a recursive σ_{arthm}-theory T, it is straightforward to check that the following relation is recursive: for $a, b \in \mathbb{N}$,

$$\mathrm{Proof}_T(a,b) :\Leftrightarrow a \text{ is a code of a } \sigma_{\mathrm{arthm}}\text{-formula } \varphi \text{ and } b \text{ is a code of a proof of } \varphi \text{ from } T.$$

To write a program for this, one has to check the definition of the formal proof, i.e. that every formula in the finite sequence coded by e is either an axiom of $\mathbb{FOL}(\sigma_{\mathrm{arthm}})$, or belongs to T (this is where we need T to be recursive), or can be obtained from the previous formulas in the sequence by applying Modus Ponens.

As before, since all recursive functions are arithmetical, there is a σ_{arthm}-formula $\mathbf{Proof}_T(x,y)$ such that for all $a, b \in \mathbb{N}$,

$$\mathrm{Proof}_T(a,b) \iff \boldsymbol{N} \models \mathbf{Proof}_T(a,b).$$

Given this, we have a σ_{arthm}-formula defining the relation of provability in $\boldsymbol{N}$:

$$\mathbf{Provable}_T(x) := \exists y \mathbf{Proof}_T(x,y),$$

and hence, for any σ-formula φ,

$$\varphi \text{ is provable in } T \iff \boldsymbol{N} \models \mathbf{Provable}_T(\ulcorner\varphi\urcorner).$$

Proof of the Incompleteness theorem 5.2. We let $T \subseteq \mathrm{Th}(\boldsymbol{N})$ be recursive and show that it is incomplete by finding a sentence that $\boldsymbol{N}$ satisfies but T does not prove.

Applying the Fixed Point lemma to

$$\varphi(v) := \neg\mathbf{Provable}_T(v),$$

we get a σ_{arthm}-sentence γ_T such that

$$\boldsymbol{N} \models \gamma_T \leftrightarrow \neg\mathbf{Provable}_T(\ulcorner\gamma_T\urcorner).$$

The *Gödel sentence* γ_T says about itself that it is not provable in T (just like in the Liar Paradox, the liar says "I am a liar"). Because $T \subseteq \mathrm{Th}(\boldsymbol{N})$, we have

$$\begin{aligned}
T \vdash \gamma_T &\implies \boldsymbol{N} \models \gamma_T \\
&\iff \boldsymbol{N} \models \neg\mathbf{Provable}_T(\ulcorner\gamma_T\urcorner) \\
&\iff \text{for all } b \in \mathbb{N}, \boldsymbol{N} \models \neg\mathbf{Proof}_T(\ulcorner\gamma_T\urcorner, b) \\
&\iff \text{for all } b \in \mathbb{N}, b \text{ is not a code of a proof of } \gamma_T \\
&\iff T \nvdash \gamma_T,
\end{aligned}$$

and thus, $T \nvdash \gamma_T$. But this means that $\boldsymbol{N} \models \neg\mathbf{Provable}_T(\ulcorner\gamma_T\urcorner)$, so $\boldsymbol{N} \models \gamma_T$, demonstrating the incompleteness of T. □

Here is another proof of the Incompleteness theorem that is shorter but nonconstructive:

Another proof of the Incompleteness theorem 5.2. If T was recursive and complete, then the formula

$$\mathbf{Provable}_T(x)$$

would define the set $\ulcorner \mathrm{Th}(\boldsymbol{N})\urcorner$ in $\boldsymbol{N}$ because, by the completeness of T, a sentence φ is provable from T if and only if $\ulcorner\varphi\urcorner \in \ulcorner \mathrm{Th}(\boldsymbol{N})\urcorner$. Thus $\ulcorner \mathrm{Th}(\boldsymbol{N})\urcorner$ would be arithmetical, contradicting Tarski's Theorem 5.6. □

5.B. **Quine: a program that prints its own code**

A more down to earth version of the Fixed Point lemma is a computer program that prints its own code, commonly referred to as a quine[7]. In this subsection, we will write such a program using informal pseudo-code in the hope of obtaining a better (hands-on) understanding of how the self-reference is implemented via the substitution function.

To write a quine, we will use a pseudo-code, whose syntax resembles that of the programming language C. In our pseudo-code, $:=$ is the command that assigns a value to a variable. A key point is that any programming language has the means of distinguishing when a symbol stands for a variable and when it is just a symbol with no content: x is a variable, whereas 'x' is just the symbol x with no content. Similarly, $x := 7$ is a programming code that assigns value 7 to the variable x, whereas "$x := 7$" is just a sequence/string of symbols with no content.

To get a quine, we can just mimic the proof of the Fixed Point lemma above: first write a program PrintSub(x, c, y) that takes as input strings (i.e. sequences of symbol) x, y and a symbol (character) c, and prints the result of substitution in x of y for c, i.e. it iterates through x and every time it encounters the symbol given in c, it replaces with the string y. Then we take the diagonal of this function: PrintDiagSub$(x) :=$ PrintSub$(x, 'x', x)$. This program now takes a string x as input and in the content of x replaces every occurrence of the *symbol* x with the *content* of the variable x (which is a string of symbols). It remains to feed the program PrintDiagSub(x) its own code: Quine() := PrintDiagSub(the code of PrintDiagSub(x)).

Below, the sequence/string of symbols written in double-quotes are interpreted by the programming language as just that string of symbols and not as programming language commands.

We start by writing a program without input that assigns the variable x some string (e.g. "math$x :=$ is $:=$ $x$fun$x :=$ ") using the command $x :=$ "math$x :=$ is $:=$ $x$fun$x :=$ ", and then, it iterates through the content of x and prints every symbol in it; however, whenever it encounters the pattern "$x :=$ ", it, in addition, prints the opening quote symbol '"', then the *content* of the variable x, then the closing quote symbol '"'.

```
NotYetQuine()
  {
     x := "mathx := is := xfunx := ";
     for(i := 0; i < length(x); i := i + 1)
     {
       Print(x[i]);
       if(i ≥ 1 ∧ x[i − 1] = 'x' ∧ x[i] = ':=')
       {
         Print('"');
         Print(x);
         Print('"');
       }
     }
  }
```

This is not quite a quine yet and we leave it as an exercise to determine what this program actually prints. Now, we'll get an actual quine by replacing the string "math$x :=$ is $:=$ $x$fun$x :=$ " with the code above, but with the string "math$x :=$ is $:=$ $x$fun$x :=$ " removed from the code.

[7]This is in the honor of philosopher Willard Van Orman Quine, who studied self-reference and is the author of the Quine paradox: "Yields falsehood when preceded by its quotation" yields falsehood when preceded by its quotation.

```
Quine()
  {
      x := "Quine()
            {
                        x := ;
                        for(i := 0; i < length(x); i := i + 1)
                        {
                          Print(x[i]);
                          if(i ≥ 1 ∧ x[i − 1] = 'x' ∧ x[i] = ':=')
                          {
                            Print('"');
                            Print(x);
                            Print('"');
                          }
                        }
            }"
        for(i := 0; i < length(x); i := i + 1)
        {
          Print(x[i]);
          if(i ≥ 1 ∧ x[i − 1] = 'x' ∧ x[i] = ':=')
          {
            Print('"');
            Print(x);
            Print('"');
          }
        }
    }
```

This program will print exactly its own code, character-by-character, up to the spacing/formatting (which is there only to increase readability).

For the rest of the section, we will be occupied with making the notion of recursive precise and developing tools for proving a stronger version of Gödel's Incompleteness theorem that applies not only to subtheories of $\mathrm{Th}(\boldsymbol{N})$, but also to theories (in an arbitrary finite signature σ), which have PA "encoded" in them; for example, $\mathrm{PA}\cup\{\neg\gamma_{\mathrm{PA}}\}$ and ZFC.

5.C. **A quick introduction to recursion theory**

In this subsection we give a model (of computation) to capture intuitive notions such as algorithm, computable functions, etc. It is a general belief, known as the Church–Turing thesis, that this model captures the mentioned notions pretty well. One evidence of it is that it is very robust in the sense that all other seemingly different models of computation that people had defined turned out to be equivalent.

Definition 5.7. For a relation $R\subseteq\mathbb{N}^{k+1}$ and $\vec{a}\in\mathbb{N}^k$, let $\mu_x(R(\vec{a},x))$ be the least $x\in\mathbb{N}$ for which $R(\vec{a},x)$ holds, if such x exists; otherwise, $\mu_x(R(\vec{a},x))$ is undefined and we write $\mu_x(R(\vec{a},x)):=\bot$. This operation applied to R is called the *search* (or *minimalization*) operation. We call the search operation is *successful* at $\vec{a}\in\mathbb{N}^k$, $\mu_x(R(\vec{a},x))$ is defined. We also say that the search operation is *successful* if it is successful at every $\vec{a}\in\mathbb{N}^k$.

For example, $\mu_x(x^2>7)=3$.

Definition 5.8. A function $f:\mathbb{N}^k\to\mathbb{N}$ is called *recursive* (or *computable*) if it is one of the basic functions in (R1) or is obtained from the latter by finitely-many applications of the operations of composition (R2) and successful search (R3):

(R1)
- Addition: $(x,y)\mapsto x+y:\mathbb{N}^2\to\mathbb{N}$
- Multiplication: $(x,y)\mapsto x\cdot y:\mathbb{N}^2\to\mathbb{N}$

- Order: $\mathbf{1}_{\leq} : \mathbb{N}^2 \to \mathbb{N}$, where $\mathbf{1}_{\leq}$ is the indicator function of $\leq$ (i.e. $\mathbf{1}_{\leq}(x,y) = 1$ if $x \leq y$, and 0, otherwise)
- Projection functions: $P_i^k : \mathbb{N}^k \to \mathbb{N}$ given by $P_i^k(x_1,\dots,x_n) := x_i$, for each $k \in \mathbb{N}$ and $i \in \{1,\dots,k\}$.

(R2) Composition: For $g : \mathbb{N}^m \to \mathbb{N}$ and $h_1,\dots,h_m : \mathbb{N}^k \to \mathbb{N}$, their *composition* is the function $f := g(h_1,\dots,h_m) : \mathbb{N}^k \to \mathbb{N}$.

(R3) Successful search: For $g : \mathbb{N}^{k+1} \to \mathbb{N}$, a function $f : \mathbb{N}^k \rightharpoonup \mathbb{N}$ is said to be obtained from g via *successful search* if $f(\vec{a}) = \mu_x(g(\vec{a},x) = 0)$ for each $\vec{a} \in \mathbb{N}^k$.

A relation $R \subseteq \mathbb{N}^k$ is said to be *recursive* if its indicator function $\mathbf{1}_R : \mathbb{N}^k \to \mathbb{N}$ is recursive.

Although the class of recursive functions is obtained by closing the set of functions in (R1) under operations (R2) and (R3), it is closed under many other operations. The most important among these is:

(R4) Primitive recursion: Let $g : \mathbb{N}^k \to \mathbb{N}$ and $h : \mathbb{N}^k \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. We say that $f : \mathbb{N}^k \times \mathbb{N} \to \mathbb{N}$ is defined by *primitive recursion* from g,h if for all $\vec{a} \in \mathbb{N}^k$ and $n \in \mathbb{N}$,

$$f(\vec{a},0) = g(\vec{a})$$
$$f(\vec{a},n+1) = h(\vec{a},n,f(\vec{a},n))$$

This is often included in the definition of recursive functions. However, we prefer showing that it is a consequence of the definition rather than including it in the latter since keeping the definition minimalistic makes it easier to prove that the class of recursive functions is contained in other classes of functions (less cases to consider). We now develop some tools, which we will use to show that recursive functions are closed under primitive recursion.

The following proposition provides examples of recursive functions and further closure properties, which are used below without mention.

Lemma 5.9.

(a) *The relations* $\geq, =$ *are recursive.*

(b) *Constant functions* $C_m^k : \mathbb{N}^k \to \mathbb{N}$ *are recursive, where* $C_i^k(\vec{a}) := m$*, for all* $\vec{a} \in \mathbb{N}^k$*.*

(c) *The successor function* $S : \mathbb{N} \to \mathbb{N}$ *is recursive.*

(d) *The set of recursive relations is an algebra, i.e. it is closed under complements and finite unions/intersections.*

(e) *Successful search for any recursive relation: Let* $R \subseteq \mathbb{N}^{k+1}$ *be recursive such that for all* $\vec{a} \in \mathbb{N}^k$ *there exists* $x \in \mathbb{N}$ *with* $(\vec{a},x) \in R$*. Then the function* $f : \mathbb{N}^k \to \mathbb{N}$ *given by*

$$f(\vec{a}) = \mu_x R(\vec{a},x)$$

is recursive.

(f) *Definition by cases: Let* $R_1,\dots,R_m \subseteq \mathbb{N}^k$ *be recursive such that for each* $\vec{a} \in \mathbb{N}^k$ *exactly one of* $R_1(\vec{a}),\dots,R_m(\vec{a})$ *holds, and suppose that* $g_1,\dots,g_m : \mathbb{N}^k \to \mathbb{N}$ *are recursive. Then* $g : \mathbb{N}^k \to \mathbb{N}$ *given by*

$$g(\vec{a}) = \begin{cases} g_1(\vec{a}) & \text{if } R_1(\vec{a}) \\ \vdots & \vdots \\ g_m(\vec{a}) & \text{if } R_k(\vec{a}) \end{cases}$$

is recursive.

Proof. For (a), observe that $\mathbf{1}_{\geq}(x,y) = \mathbf{1}_{\leq}(P_2^2(x,y),P_1^2(x,y))$ and $\mathbf{1}_{=}(x,y) = \mathbf{1}_{\leq}(x,y) \cdot \mathbf{1}_{\geq}(x,y)$.

We prove (b) by induction on m. For $m = 0$, observe that $C_0^k(\vec{a}) = \mu_x(P_{k+1}^{k+1}(\vec{a},x) = 0)$. Assume C_m^k is recursive and note that

$$C_{m+1}^k(\vec{a}) = \mu_x(C_m^k(\vec{a}) < x) = \mu_x(\mathbf{1}_{\geq}(C_m^{k+1}(\vec{a},x),P_{k+1}^{k+1}(\vec{a},x)) = 0).$$

For (c), just note that $S(a) = a + C_1^1(a)$.

For (d), observe that $\neg P(\vec{a}) \Leftrightarrow \mathbf{1}_P(\vec{a}) = C_0^k(\vec{a})$ and $\mathbf{1}_{P\wedge Q}(\vec{a}) = \mathbf{1}_P(\vec{a}) \cdot \mathbf{1}_Q(\vec{a})$. Thus $\neg P$ and $P \wedge Q$ are recursive if so are P and Q.

For (e), note that $f(\vec{a}) = \mu_x(\mathbf{1}_{\neg R}(\vec{a},x) = 0)$.

Part (f) is left to the reader. □

By definition, a relation is recursive if and only if its indicator function is recursive. The following is a converse to this, providing a convenient way of verifying recursiveness of functions.

Proposition 5.10 (Graph property)**.** *A function $f : \mathbb{N}^k \to \mathbb{N}$ is recursive if and only if its graph $R \subseteq \mathbb{N}^{k+1}$ is recursive.*

Proof. For each $\vec{a} \in N^k$ and $b \in \mathbb{N}$,

$$R(\vec{a}, b) \iff f(\vec{a}) = b \iff f\Big(P_1^{k+1}(\vec{a}, b), \dots, P_k^{k+1}(\vec{a}, b)\Big) = P_{k+1}^{k+1}(\vec{a}, b),$$

so R is recursive if f is. Conversely, for each $\vec{a} \in N^k$, $f(\vec{a}) = \mu_x R(\vec{a}, x)$, so f is recursive if R is. □

We will see below that the class of recursive relation is not closed under quantification over $\mathbb{N}$. However, the following shows that it is closed under bounded quantification and we use it below without mention.

Proposition 5.11 (Bounded quantification)**.** *The class of recursive relations is closed under bounded quantification, i.e. if a relation $R \subseteq \mathbb{N}^{k+1}$ is recursive, then the following relations are also recursive: for each $\vec{a} \in \mathbb{N}^k$, $b \in \mathbb{N}$,*

$$P(\vec{a}, b) :\Leftrightarrow (\exists x < b)\, R(\vec{a}, x),$$
$$Q(\vec{a}, b) :\Leftrightarrow (\forall x < b)\, R(\vec{a}, x).$$

Proof. The second statement follows from the first by taking negations. For the first, observe that for each $\vec{a} \in \mathbb{N}^k, b \in \mathbb{N}$, there is an $x \in \mathbb{N}$ such that $R(\vec{a}, x) \vee x \geq b$, so the search for such an x is successful and hence the following relation is recursive:

$$P(\vec{a}, b) \iff \mu_x\Big(R(\vec{a}, x) \vee x \geq b\Big) < b.$$ □

This gives another batch of examples of recursive functions.

Lemma 5.12.

(a) *The function $\dot{-} : \mathbb{N}^2 \to \mathbb{N}$ defined by $n \dot{-} m := \max\{n - m, 0\}$ is recursive.*

(b) *The remainder function* Rem $: \mathbb{N}^2 \to \mathbb{N}$, *defined by $(a, b) \mapsto$ the remainder of a when divided by b, is recursive.*

(c) *The function* Pair $: \mathbb{N}^2 \to \mathbb{N}$ *defined by*

$$(x, y) \to \frac{(x + y)(x + y + 1)}{2} + x$$

is a recursive bijection.

(d) *The functions* Left, Right $: \mathbb{N} \to \mathbb{N}$ *defined by*

$$\text{Pair}(x, y) = z \iff \text{Left}(z) = x \wedge \text{Right}(z) = y$$

are recursive.

Proof. We leave parts (a),(b) and (c) to the reader. For (d), observe that $\text{Left}(z) = \mu_x(\exists y_{<z+1} \text{Pair}(x, y) = z)$ and similarly for Right. □

Towards our goal of showing closure under primitive recursion, we need to understand how to check the calculation done by via primitive recursion; in other words, we need a "proof/certificate" that to verify the calculations ourselves.

Proposition 5.13 (Dedekind's analysis of recursion)**.** *If $f : \mathbb{N}^{k+1} \to \mathbb{N}$ is defined by primitive recursion from g, h as in (R4), then for all $\vec{a} \in \mathbb{N}^k$, $n \in \mathbb{N}$ and $m \in \mathbb{N}$,*

$$\begin{aligned} f(\vec{a}, n) = m \iff \exists \vec{b} \in \mathbb{N}^{<\mathbb{N}} \text{ such that } |\vec{b}| = n + 1 \\ \text{and } \vec{b}(0) = g(\vec{a}) \\ \text{and for each } i < n, \vec{b}(i + 1) = h(\vec{a}, n, \vec{b}(i)) \\ \text{and } \vec{b}(n) = m. \end{aligned}$$

Proof. Obvious. □

To be able to express the right hand side of Dedekind's analysis of recursion, we need to be able to recursively encode and decode tuples of natural numbers of arbitrary length into single natural numbers. We do it using the most basic result in number theory.

Chinese Remainder Theorem 5.14. *Let $d_0,\dots,d_{n-1}$ be pairwise coprime and put $d = d_0 \cdot d_1 \cdot \ldots \cdot d_{n-1}$. Then the map*

$$h : \mathbb{Z}/d\mathbb{Z} \to \mathbb{Z}/d_0\mathbb{Z} \times \ldots \times \mathbb{Z}/d_{n-1}\mathbb{Z}$$

defined by

$$[a]_d \mapsto ([a]_{d_0},\dots,[a]_{d_{n-1}})$$

is a well-defined group isomorphism.

Proof. That h is well-defined follows from the fact that every d_i divides d, and that h is a homomorphism follows from the fact that the remainder function respects addition. Since the groups on the left and right of the homomorphism have the same number of elements, by the Pigeonhole Principle, we only have to show that h is injective. To this end, assume that $h([a]_d) = 0$. Thus every d_i divides a and hence d divides a because d_i are pairwise coprime. Therefore, $[a]_d = 0$ and hence $\ker(h)$ is trivial. □

Lemma 5.15 (Gödel's β-function)**.** *The function $\beta : \mathbb{N}^2 \to \mathbb{N}$ defined by*

$$\beta(w,i) := \mathrm{Rem}\Big(\mathrm{Left}(w), 1 + (i+1)\mathrm{Right}(w)\Big)$$

is recursive and has the property that for every sequence $(w_0,\dots,w_{n-1})$, there exists $w \in \mathbb{N}$ such that for all $i < n$,

$$\beta(w,i) = w_i.$$

Proof. The fact that β is recursive follows from 5.12, so we prove the second statement. We let

$$s := \max\{n, w_0, w_1, \dots, w_{n-1}\}$$

and $b := s!$, and verify that

$$d_0 := 1 + (0+1)\cdot b, d_1 := 1 + (1+1)\cdot b, \dots, d_{n-1} := 1 + n \cdot b$$

are pairwise coprime as follows: if a prime p divides $1 + (i+1)\cdot b$ and $1 + (j+1)\cdot b$, for $i < j$, then it divides their difference $(j-i)\cdot b = (j-i)\cdot s!$. Since $j - i < n \leq s$, p must divide $s! = b$, contradicting that p divides $1 + (i+1)\cdot b$.

By the Chinese Remainder theorem, there is $a < d_0 \cdot \ldots \cdot d_{n-1}$ such that $\mathrm{Rem}(a, d_i) = w_i$. Thus setting $w := \mathrm{Pair}(a,b)$, we get

$$w_i = \mathrm{Rem}(a,d_i) = \mathrm{Rem}\Big(\mathrm{Left}(w), 1 + (i+1)\mathrm{Right}(w)\Big) = \beta(w,i).$$ □

Using Gödel's β-function, we define the following encoding/decoding tuples functions, which are clearly recursive.

(5.16.i) For each $k \in \mathbb{N}$ (think $k = 7$), we encode $\mathbb{N}^k$ into $\mathbb{N}$ via the function $\langle\cdot\rangle_k : \mathbb{N}^k \to \mathbb{N}$ defined by

$$\langle a_0,\dots,a_{k-1}\rangle_k := \mu_x\left(\beta(x,0) = k \wedge \bigwedge_{i=1}^{k} \beta(x,i) = a_{i-1}\right).$$

Note that we record the length of the turple $(a_0,\dots,a_{k-1})$ in the 0^{th} coordinate to be able to recover the tuple unambiguously. This also ensures that the images of the functions $\langle\cdot\rangle_k$ are disjoint for distinct k. Below, we often omit writing the subscript k in $\langle a_0,\dots,a_{k-1}\rangle_k$ and simply write $\langle a_0,\dots,a_{k-1}\rangle$ because there is no ambiguity; however, we do write the subscript whenever we wish to emphasize that the $\langle\cdot\rangle_k$ are different functions (with different domains) for distinct k.

(5.16.ii) Thus, if $a \in \mathbb{N}$ is in the image of $\langle\cdot\rangle_k$ for some $k \in \mathbb{N}$, we can recover this k as the 0^{th} coordinate of the sequence that a encodes. In other words, we define the *length* function $\mathrm{lh} : \mathbb{N} \to \mathbb{N}$ by $\mathrm{lh}(a) := \beta(a,0)$.

(5.16.iii) For each $i \in \mathbb{N}$ (think $i = 4$), if $a \in \mathbb{N}$ is in the image of $\langle\cdot\rangle_k$ for some $k \in \mathbb{N}$, we can recover its i^{th} coordinate as the $(i+1)^{\text{th}}$ coordinate of the sequence encoded by a. In other words, we define the *i^{th} coordinate* function $(\cdot)_i : \mathbb{N} \to \mathbb{N}$ by $(a)_i := \beta(a, i+1)$. We clearly have $(\langle a_0,\dots,a_{k-1}\rangle)_i = a_i$ for each $k \in \mathbb{N}$ and each $(a_0,\dots,a_{k-1}) \in \mathbb{N}^k$.

(5.16.iv) We define the initial *segment function* $\mathrm{InitSeg}:\mathbb{N}^2\to\mathbb{N}$ by

$$\mathrm{InitSeg}(a,i):=\mu_x\Big(\mathrm{lh}(x)=i\wedge(\forall j<i)\,[(x)_j=(a)_j]\Big).$$

Thus, $\mathrm{InitSeg}(\langle a_0,\dots,a_n\rangle,i)=\langle a_0,\dots,a_{i-1}\rangle$.

(5.16.v) We define the *concatenation* function $*:\mathbb{N}^2\to\mathbb{N}$ by

$$a*b:=\mu_x\Big(\mathrm{lh}(x)=\mathrm{lh}(a)+\mathrm{lh}(b)\wedge\Big(\forall i<\mathrm{lh}(a)\Big)\Big[(x)_i=(a)_i\Big]\wedge\Big(\forall i<\mathrm{lh}(b)\Big)\Big[(x)_{\mathrm{lh}(a)+i}=(b)_i\Big]\Big).$$

Thus, $\langle a_0,\dots,a_{n-1}\rangle*\langle b_0,\dots,b_{m-1}\rangle=\langle a_0,\dots,a_{n-1},b_0,\dots,b_{m-1}\rangle$.

Proposition 5.17. *Recursive functions are closed under the operation of primitive recursion, i.e. if g,h,f are as in (R4) and g,h are recursive, then f too is recursive.*

Proof. We implement Dedekind's analysis of recursion as follows. Define an auxiliary function $\tilde f:\mathbb{N}^{k+1}\to\mathbb{N}$ by

$$\tilde f(\vec a,n)=\mu_x(\mathrm{lh}(x)=n+1\wedge(x)_0=g(\vec a)\wedge(\forall i<n)(x)_{i+1}=h(\vec a,i,(x)_i)),$$

and note that $f(\vec a,n)=(\tilde f(\vec a,n))_n$. Since $\tilde f$ is clearly recursive, f is also recursive. □

Primitive recursion enables us to show that any function that admits a recursive definition is recursive. E.g. $n\to 2^n$ is recursive because

$$\begin{cases} 2^0 &= 1\\ 2^{n+1} &= 2\cdot 2^n\end{cases}.$$

We now define a nice subclass of recursive functions, namely that of *primitive recursive* functions, which is still rich enough to contain most of the functions that can be implemented as computer programs. In fact, most of the recursive functions mentioned so far are actually primitive recursive.

Definition 5.18. The class of *primitive recursive functions* is the smallest class containing the successor function $S:\mathbb{N}\to\mathbb{N}$, the constant functions $C^k_m:\mathbb{N}^k\to\mathbb{N}$, $k,m\in\mathbb{N}$ and the projection functions $P^k_i(x_1,\dots,x_k)=x_i$, $i\le k\in\mathbb{N}$, and is closed under composition (R2) and primitive recursion (R4). A relation $R\subseteq\mathbb{N}^k$ is called *primitive recursive* if its characteristic function $\mathbf{1}_R:\mathbb{N}^k\to\mathbb{N}$ is primitive recursive.

Lemma 5.19. *The following functions are primitive recursive.*

(a) *Addition* $+:\mathbb{N}^2\to\mathbb{N}$ *and multiplication* $\cdot:\mathbb{N}^2\to\mathbb{N}$.

(b) *Predecessor:* $PD(n):=\begin{cases}0 & \text{if } n=0\\ n-1 & \text{otherwise}\end{cases}$.

(c) *Insured subtraction:* $n\dot- m:=\max\{n-m,0\}$.

(d) *Inverse-bit:* $\overline{\mathrm{bit}}(n):=\begin{cases}0 & \text{if } n>0\\ 1 & \text{otherwise.}\end{cases}$

(e) *Equality:* $\mathbf{1}_=(n,m):=1$ *if* $n=m$ *and* $\mathbf{1}_=(n,m):=0$ *otherwise.*

(f) *Less than or equal to:* $\mathbf{1}_\le(n,m):=1$ *if* $n\le m$ *and* $\mathbf{1}_\le(n,m):=0$ *otherwise.*

Proof. Use primitive recursion to define

- addition from the projection and successor functions,
- multiplication from the constant 0 and addition functions,
- predecessor from the constant 0 function and the projection function $\mathrm{proj}^2_1(n,y):=n$,
- insured subtraction from the constant 0 and the predecessor functions.

For equality, observe that $\mathbf{1}_=(n,m)=\overline{\mathrm{bit}}\Big((n\dot- m)+(m\dot- n)\Big)$, so for less than or equal to, we have

$$n\le m \text{ if and only if } (n\dot- m)=0.$$ □

It is easy to check that Lemma 5.9 holds with *recursive* replaced by *primitive recursive*.

The following makes it easy to verify that Lemma 5.12 and 5.15 also hold with "recursive" replaced by "primitive recursive".

Lemma 5.20 (Bounded search)**.** *Let $R \subseteq \mathbb{N}^{n+1}$ be a recursive relation. Then the function $f : \mathbb{N}^{n+1} \to \mathbb{N}$ defined by $f(\vec{a}, y) := \mu_{x<y} R(\vec{a}, x)$ is primitive recursive, where*

$$\mu_{x<y} R(\vec{a}, x) = \begin{cases} \mu_x R(\vec{a}, x) & \text{if } (\exists x < y) R(\vec{a}, x) \\ y & \text{otherwise.} \end{cases}$$

Proof. We define $f(\vec{a}, y)$ by primitive recursion as follows: let $f(\vec{a}, 0) := 0$ and

$$f(\vec{a}, y+1) := \begin{cases} f(\vec{a}, y) & \text{if } f(\vec{a}, y) < y \\ y & \text{if } f(\vec{a}, y) = y \wedge R(\vec{a}, y) \\ y+1 & \text{otherwise.} \end{cases}$$

□

The proof of 5.15 yields a primitive recursive function $B : \mathbb{N} \to \mathbb{N}$, defined by $B(s) = \prod_{i<s}(1 + (1+i)s!)$, such that for every $n \in \mathbb{N}$ and $\vec{a} \in \mathbb{N}^n$,

▌ whenever $N \geq \max\{n, a_0, \ldots, a_{n-1}\}$, there is $a < B(N)$ such that $\beta(a, i) = a_i$ for each $i < n$.

Using this together with 5.20 one can easily show that all of the encoding/decoding functions in (5.16) are primitive recursive.

The following lemma allows recursive definitions using all previously computed values of a function as opposed to only the last computed value.

Lemma 5.21 (Complete primitive recursion)**.** *For $f : \mathbb{N}^{n+1} \to \mathbb{N}$, let*

$$\bar{f}(\vec{a}, n) := \langle f(\vec{a}, 0), \ldots, f(\vec{a}, n-1) \rangle.$$

Then:

(a) *f is primitive recursive if and only if $\bar{f}$ is primitive recursive.*

(b) *If $g : \mathbb{N}^{k+1} \to \mathbb{N}$ is primitive recursive, then so is $f : \mathbb{N}^{k+1} \to \mathbb{N}$ defined by $f(\vec{a}, n) := g(\vec{a}, \bar{f}(\vec{a}, n))$.*

Proof. We prove part (a), leaving (b) to the reader.
⇐: Put $f(\vec{a}, n) := (\bar{f}(\vec{a}, n+1))_n$.
⇒: We define $\bar{f}(\vec{a}, n)$ by primitive recursion as follows:

$$\begin{cases} \bar{f}(\vec{a}, 0) & := & <> \\ \bar{f}(\vec{a}, n+1) & := & \bar{f}(\vec{a}, n) * \langle f(\vec{a}, n) \rangle \end{cases}.$$

□

One may ask if there are any recursive functions that are not primitive recursive. The answer is YES (of course) and the reason is that primitive recursive functions can only realize bounded search, whereas recursive ones can realize a potentially unbounded successful search. Thus, to construct an example of a recursive but not primitively recursive function, one needs to come up with one involving successful search, but the bound for the search can not be computed from the input in primitively recursive fashion. We do so using the general method of *parameterization-and-diagonalization*.

Antidiagonalization 5.22 (Cantor)**.** *Let X be a set and $\Upsilon : X \times X \to \mathbb{N}$ be a function. The* antidiagonal $\mathrm{AntiDiag}_\Upsilon : X \to \mathbb{N}$ *of Υ defined by*

$$\mathrm{AntiDiag}_\Upsilon(x) := \begin{cases} 0 & \text{if } \Upsilon(x,x) \neq 0 \\ 1 & \text{otherwise} \end{cases}$$

is not among the fibers of Υ, i.e. the set $\left\{\Upsilon_p \in \mathbb{N}^X : p \in X\right\}$, where $\Upsilon_p := \Upsilon(p, \cdot) : X \to \mathbb{N}$.

Proof. For any $p \in X$, if $\mathrm{AntiDiag}_\Upsilon = \Upsilon_p$, then $\mathrm{AntiDiag}_\Upsilon(p) = 0$ if and only if $0 \neq \Upsilon(p,p) =: \Upsilon_p(p) = \mathrm{AntiDiag}_\Upsilon(p)$, a contradiction. □

Definition 5.23. For sets X, Y, P and a class $\mathcal{C}$ of functions $X \to Y$, we call a function $\Upsilon : P \times X \to Y$ a *P-parameterization* of $\mathcal{C}$ if for each $f : X \to Y$, $f \in \mathcal{C}$ if and only if there is $p \in P$ with $\Upsilon_p = f$.

Corollary 5.24. *Any $\mathbb{N}$-parameterization of the class of all recursive (resp. primitive recursive) functions $\mathbb{N} \to \mathbb{N}$ is itself not recursive (resp. not primitive recursive).*

Proof. For any such parameterization Υ, its antidiagonal $\mathrm{AntiDiag}_\Upsilon$ is has a primitive recursive definition from Υ, so if Υ is recursive (resp. primitive recursive), then such is $\mathrm{AntiDiag}_\Upsilon$, so it has to show up as one of the fibers of Υ, contradicting Eq. (5.22). □

Proposition 5.25. *There exists a recursive $\mathbb{N}$-parameterization $\Upsilon : \mathbb{N}\times\mathbb{N}\to\mathbb{N}$ for the class of all primitive recursive functions $\mathbb{N}\to\mathbb{N}$. In particular, this Υ is an example of a recursive but not primitive recursive function.*

Proof. Outlined in a homework problem. □

A similar proof also shows that there is no recursive $\mathbb{N}$-parameterization of the class of all recursive functions $\mathbb{N}\to\mathbb{N}$. Similarly, thinking of recursive functions as programs, the set of their codes is not recursive, i.e. there is no recursive binary relation R such that for any unary recursive relation Q there is n such that for all x,

$$Q(x) \iff R(n,x).$$

This is known as the undecidability of the *halting problem*.

Here is a more concrete and important example of a recursive function that is not primitive recursive:

Definition 5.26. *Ackermann function* is the function $A:\mathbb{N}^2\to\mathbb{N}$ inductively defined as follows:

$$\begin{cases} A(0,x) &= x+1\\ A(n+1,0) &= A(n,1)\\ A(n+1,x+1) &= A(n,A(n+1,x)) \end{cases}.$$

The proof that this function is recursive but not primitive recursive is left as a homework problem together with the proof that the graph of this function is primitive recursive. The last fact shows that the graph property (Proposition 5.10) does not hold for primitive recursive functions.

5.D. **Representability in a theory**

In the sketch of the proof of the Incompleteness theorem above, we used the fact that recursive functions are arithmetical, i.e. definable in $\boldsymbol{N}$. Thus the proof only applied to theories that $\mathbb{N}$ satisfies. If we want to prove incompleteness for other theories, like $\mathrm{PA}\cup\{\neg\gamma_{\mathrm{PA}}\}$, we have to develop a notion of definability inside a theory rather than a structure. This is what the following definition is supposed to capture.

Definition 5.27 (Representability)**.** Let T be a σ_{arthm}-theory in the signature σ_{arthm} of arithmetic.

- We say that a relation $R\subseteq\mathbb{N}^n$ is *representable in T* if there is a formula $\varphi(\vec{x})$ such that for all $\vec{a}\in\mathbb{N}^n$,
$$R(\vec{a}) \implies T\models\varphi(\Delta(\vec{a})) \text{ and } \neg R(\vec{a}) \implies T\models\neg\varphi(\Delta(\vec{a})),$$
where $\Delta(\vec{a})=(\Delta(a_1),\dots,\Delta(a_n))$. Such φ is said to represent the relation R in T.
- We say that a function $f:\mathbb{N}^n\to\mathbb{N}$ is *representable in T (by a formula)* if there is a formula $\varphi(\vec{x},y)$ such that for all $\vec{a}\in\mathbb{N}^n$,
$$T\models\forall y\Big[\varphi(\Delta(\vec{a}),y)\leftrightarrow y=\Delta(f(\vec{a}))\Big].$$
Such φ is said to *represent* the function f in T.
- A function $f:\mathbb{N}^n\to\mathbb{N}$ is said to be *representable in T by a term* if there is a σ_{arthm}-term $t(\vec{x})$ such that for all $\vec{a}\in\mathbb{N}^n$,
$$T\models t(\Delta(\vec{a}))=\Delta(f(\vec{a})).$$
Such t is said to represent f in T.

Proposition 5.28. *Let T be a σ_{arthm}-theory and $f:\mathbb{N}^n\to\mathbb{N}$.*

(a) *If f is representable in T by a term, then it is also representable in T by a formula.*

(b) *Suppose that for any distinct $m,k\in\mathbb{N}$, $T\models\Delta(m)\neq\Delta(k)$. Then, if f is representable in T by a formula, then the graph of f is representable in T by the same formula.*

Proof. For part (a), letting $t(\vec{x})$ be a term representing f, it is straightforward to check that the formula $\varphi(\vec{x},y):=t(\vec{x})\doteq y$ represents f.

As for (b), let $\varphi(\vec{x},y)$ be a formula representing f in T and fix arbitrary $\vec{a}\in\mathbb{N}^n$ and $b\in\mathbb{N}$. By instantiating $y:=\Delta(b)$, we get

$$T\models\varphi(\Delta(\vec{a},\Delta(b)))\leftrightarrow\Delta(b)=\Delta(f(\vec{a})).$$

Thus, it is clear that if $f(\vec{a}) = b$ then $T \models \varphi(\Delta(\vec{a}), \Delta(b))$, and if $f(\vec{a}) \neq b$ then the additional hypothesis on T guarantees that $T \models \neg\varphi(\Delta(\vec{a}), \Delta(b))$. □

The following shows that we could have defined representability of relations using that of functions (not the other way around).

Proposition 5.29. *If T is a σ_{arthm}-theory such that $T \models \Delta(1) \neq 0$ and $R \subseteq \mathbb{N}^n$, then*

R is representable in T if and only if $\mathbf{1}_R$ is representable in T.

Proof. $\Rightarrow$: Let $\varphi(\vec{x})$ represent R in T and put

$$\psi(\vec{x}, y) \doteq \Big[\varphi(\vec{x}) \wedge y \doteq \Delta(1)\Big] \vee \Big[\neg\varphi(\vec{x}) \wedge y \doteq 0\Big].$$

We show that $\psi(\vec{x}, y)$ represents $\mathbf{1}_R$ in T. Fix $\vec{a} \in \mathbb{N}^n$ and consider cases as to whether $R(\vec{a})$ holds.

Assume $R(\vec{a})$ holds, so $T \models \varphi(\Delta(\vec{a}))$, $\mathbf{1}_R(\vec{a}) = 1$, and we have to show

$$T \models \forall y\Big[\psi(\Delta(\vec{a}), y) \leftrightarrow y = \Delta(1)\Big].$$

Fixing a model $\boldsymbol{M} \models T$ and an arbitrary $y \in M$, we see that, since $\boldsymbol{M} \models \varphi(\Delta(\vec{a}))$,

$$\boldsymbol{M} \models \psi(\Delta(\vec{a}), y) \iff \boldsymbol{M} \models \Big[\varphi(\Delta(\vec{a})) \wedge y = \Delta(1)\Big] \iff \boldsymbol{M} \models y = \Delta(1).$$

A similar argument handles the case $\neg R(\vec{a})$.

$\Leftarrow$: Let $\varphi(\vec{x}, y)$ represent $\mathbf{1}_R$ and put $\psi(\vec{x}) \doteq \varphi(\vec{x}, \Delta(1))$. We show that $\psi(\vec{x})$ represents R in T. For every $\vec{a} \in \mathbb{N}^n$, instantiating $y := \Delta(1)$ in the definition of representability, we get

$$T \models \varphi(\Delta(\vec{a}), \Delta(1)) \leftrightarrow \Delta(1) = \Delta(\mathbf{1}_R(\vec{a})).$$

Thus, it is clear that if $R(\vec{a})$ holds then $T \models \varphi(\Delta(\vec{a}), \Delta(1))$, and if $R(\vec{a})$ fails then $T \models \Delta(1) \neq 0$ guarantees that $T \models \neg\varphi(\Delta(\vec{a}), \Delta(1))$. □

Proposition 5.30. *All recursive functions and relations are representable in* PA.

Proof. By Proposition 5.29, it is enough to show for functions.

Because the standard part of any model of PA is isomorphic to $\boldsymbol{N}$, the terms $t_+(x, y) := x + y$, $t_\cdot(x, y) := x \cdot y$ and $t_i^{(n)}(x_1, \ldots, x_n) := x_i$ represent, respectively, the addition, multiplication and the projection functions. For the same reason, the formula $x \leq y := \exists z(z + x = y)$ represents the relation $\leq$, and hence $\mathbf{1}_\leq$ is representable as well by 5.29. It remains to show that representability is closed under composition (R2) and safe search (R3).

For (R2) , assume that $\varphi(\vec{x}, y)$ represents the function $g : \mathbb{N}^k \to \mathbb{N}$ and $\psi_i(\vec{v}, u)$ represent the functions $h_i : \mathbb{N}^n \to \mathbb{N}$, where $\vec{x}$ is an k-vector and $\vec{v}$ is a n-vector. We show that

$$\theta(\vec{v}, y) \doteq \exists \vec{x} \bigwedge_{i=1}^{k} \psi_i(\vec{v}, x_i) \wedge \varphi(\vec{x}, y)$$

represents $f := g(h_1, \ldots, h_k)$. Fix $\vec{a} \in \mathbb{N}^n$ and let $d = f(\vec{a})$. We have to show that

$$\text{PA} \models \forall y\Big[\theta(\Delta(\vec{a}), y) \leftrightarrow y \doteq \Delta(d)\Big].$$

Let $b_i := h_i(\vec{a})$ and put $\vec{b} := (b_1, \ldots, b_k)$. Then $f(\vec{a}) = g(\vec{b}) = d$. Therefore,

$$\text{PA} \models \forall y\Big[\varphi(\vec{b}, y) \leftrightarrow y \doteq \Delta(d)\Big] \text{ and } \text{PA} \models \forall z\Big[\psi_i(\Delta(\vec{a}), z) \leftrightarrow z \doteq \Delta(b_i)\Big], \text{ for } i = 1, \ldots, k.$$

Thus, arguing in models gives the desired statement.

For (R3), let $\varphi(\vec{x}, y, z)$ represent the function $g : \mathbb{N}^{n+1} \to \mathbb{N}$, where $\vec{x}$ is an n-vector and g is such that for all $\vec{a} \in \mathbb{N}^n$ there is $b \in \mathbb{N}$ with $g(\vec{a}, b) = 0$. We show that

$$\psi(\vec{x}, z) \doteq \varphi(\vec{x}, z, 0) \wedge \forall u\Big[u < z \to \neg\varphi(\vec{x}, u, 0)\Big]$$

represents $f(\vec{a}) := \mu_z(g(\vec{a}, z) = 0)$. Fix $\vec{a} \in \mathbb{N}^n$ and let $b := f(\vec{a})$, so $g(\vec{a}, b) = 0$ and b is the least such natural number. We have to show that

$$\text{PA} \models \forall z\Big[\psi(\Delta(\vec{a}), z) \leftrightarrow z \doteq \Delta(b)\Big].$$

Fix $\boldsymbol{M} \models \mathrm{PA}$ and an element $\zeta \in M$. We know that if $\zeta = \Delta(b)$, then, because φ represents g in PA, $\boldsymbol{M} \models \varphi(\Delta(\vec{a}), \Delta(b), 0)$ and $\boldsymbol{M} \models \neg\varphi(\Delta(\vec{a}), \Delta(b'), 0)$ for each $b' < b$. Because $\mathrm{PA} \models \forall u\Big(u < \Delta(b) \to \bigvee_{b'<b} u \doteq \Delta(b')\Big)$, it follows that $\boldsymbol{M} \models \psi(\Delta(\vec{a}), \Delta(b))$ and hence $\boldsymbol{M} \models \psi(\Delta(\vec{a}), \zeta)$. Conversely, if $\boldsymbol{M} \models \psi(\Delta(\vec{a}), \zeta)$, then $\boldsymbol{M} \models \varphi(\Delta(\vec{a}), \zeta)$ and $\boldsymbol{M} \models \forall u\Big(u < \zeta \to \neg\varphi(\Delta(\vec{a}), \zeta, 0)\Big)$. But $\boldsymbol{M} \models \varphi(\Delta(\vec{a}), \Delta(b))$, so $\boldsymbol{M} \models \zeta \leq \Delta(b)$, so ζ is a standard element of M and hence it must be $\Delta(b)$. □

In a later subsection, we will also prove the converse of this proposition, so representability in PA actually characterizes recursive functions.

Remark 5.31. The proof of Proposition 5.30 goes through with PA replaced by any σ_{arthm}-theory T that is strong enough to prove all of the q.f. statements about the standard part $\{\Delta(n) : n \in \mathbb{N}\}$ in the signature of arithmetic extended with $<$. More particularly, all we need is that $T \vdash \Delta(n) \neq \Delta(m)$ for distinct $n, m \in \mathbb{N}$ and

$$T \models \forall x\Big(x < \Delta(m) \to \bigvee_{n<m} x \doteq \Delta(n)\Big).$$

5.E. **Gödel coding**

Here we describe a coding of formulas and proofs, and all functions necessary to prove the Fixed Point lemma and the Incompleteness theorem.

For the rest of the section, let σ be a finite signature.

- We list the symbols of $\mathbb{FOL}(\sigma)$ as it is done in (5.4) and for each symbol in this list, we let $\mathrm{SN}(s)$ denote its index.
- For a σ-term t, define its *Gödel code* $\ulcorner t \urcorner$ as follows

$$\ulcorner t \urcorner = \begin{cases} \langle \mathrm{SN}(s) \rangle & \text{if } t = s \text{ is a variable or a constant symbol} \\ \langle \mathrm{SN}(f), \ulcorner t_1 \urcorner, \dots, \ulcorner t_n \urcorner \rangle & \text{if } f \text{ is an } n\text{-ary function symbol and } t = f(t_1, \dots, t_n). \end{cases}$$

  Note that for a variable or a constant symbol s, $\ulcorner s \urcorner$ may not be equal to $\mathrm{SN}(s)$.
- For a σ-formula φ, define its *Gödel code* $\ulcorner \varphi \urcorner$ as follows

$$\ulcorner \varphi \urcorner = \begin{cases} \langle \mathrm{SN}(\doteq), \ulcorner t_1 \urcorner, \ulcorner t_2 \urcorner \rangle & \text{if } \varphi = (t_1 \doteq t_2) \\ \langle \mathrm{SN}(R), \ulcorner t_1 \urcorner, \dots, \ulcorner t_n \urcorner \rangle & \text{if } R \text{ is an } n\text{-ary relation symbol and } \varphi \doteq R(t_1, \dots, t_n) \\ \langle \mathrm{SN}(\neg), \ulcorner \psi \urcorner \rangle & \text{if } \varphi = \neg\psi \\ \langle \mathrm{SN}(\to), \ulcorner \psi_1 \urcorner, \ulcorner \psi_2 \urcorner \rangle & \text{if } \varphi = \psi_1 \to \psi_2 \\ \langle \mathrm{SN}(\forall), \ulcorner v \urcorner, \ulcorner \psi \urcorner \rangle & \text{if } \varphi = \forall v \psi. \end{cases}$$

Lemma 5.32. *The following subsets of* $\mathbb{N}$ *are primitive recursive:*

(5.32.a) $\mathrm{Variable} := \{\ulcorner x \urcorner : x \text{ is a variable}\}$
(5.32.b) $\mathrm{Term} := \{\ulcorner t \urcorner : t \text{ is a } \sigma\text{-term}\}$
(5.32.c) $\mathrm{Formula} := \{\ulcorner \varphi \urcorner : \varphi \text{ is a } \sigma\text{-formula}\}$

Proof. In all proofs we use complete primitive recursion Lemma 5.21.
(5.32.a) $a \in \mathrm{Variable}$ if and only if $\mathrm{lh}(a) = 1$ and $(a)_0$ is even.
(5.32.b) $\mathrm{Term}(a)$ if and only if $\mathrm{Variable}(a)$ or a is a code for a constant symbol or $(a)_0$ is a code for an n-ary functions symbol with $n = \mathrm{lh}(a) - 1$ and $\forall i < n, \mathrm{Term}((a)_{i+1})$.
(5.32.c) is left to the reader. It gets messy if one wants to also check our convention about quantified variables. □

Lemma 5.33. *There is a primitive recursive function* $\mathrm{Sub} : \mathbb{N}^3 \to \mathbb{N}$ *such that for any* σ*-formula* φ*, variable* v *and* σ*-term* t *that is OK to be plugged in for* v *in* φ*,*

$$\mathrm{Sub}(\ulcorner \varphi \urcorner, \mathrm{SN}(v), \ulcorner t \urcorner) = \ulcorner \varphi(t/v) \urcorner.$$

Proof. Define $\mathrm{Sub}(a,m,k) :=$

$$\begin{cases} k & \text{if } \mathrm{Variable}(a) \text{ and } (a)_0 = m \\ \left\langle (a)_0, \mathrm{Sub}((a)_1,m,k), \dots, \mathrm{Sub}((a)_{\mathrm{lh}(a)-1},m,k) \right\rangle & \text{if } \mathrm{lh}(a) > 0 \text{ and } (a)_0 \neq \mathrm{SN}(\forall) \\ \langle (a)_0, (a)_1, \mathrm{Sub}((a)_2,m,k)\rangle & \text{if } \mathrm{lh}(a) > 0 \text{ and } (a)_0 = \mathrm{SN}(\forall) \text{ and } (a)_1 \neq m \\ a & \text{otherwise.} \end{cases}$$

This is clearly primitive recursive (using complete recursion). □

Lemma 5.34. *The following relations are primitive recursive:*

(5.34.a) $\mathrm{FreeVar} := \{(\ulcorner\varphi\urcorner, \mathrm{SN}(v)) : v \text{ occurs free in } \varphi\} \subseteq \mathbb{N}^2$
(5.34.b) $\mathrm{OKtoSub} := \{(\ulcorner\varphi\urcorner, \ulcorner t\urcorner, \mathrm{SN}(v)) : t \text{ is OK to be plugged in for variable } v \text{ in } \varphi\} \subseteq \mathbb{N}^3$
(5.34.c) $\mathrm{Sentence} := \{\ulcorner\varphi\urcorner : \varphi \text{ is a sentence}\} \subseteq \mathbb{N}$
(5.34.d) $\mathrm{Axiom} := \{\ulcorner\varphi\urcorner : \varphi \text{ is an axiom of } \mathbb{FOL}(\sigma)\} \subseteq \mathbb{N}$
(5.34.e) $\mathrm{MP} := \{(\ulcorner\varphi\urcorner, \ulcorner\varphi \to \psi\urcorner, \psi) : \varphi, \psi \text{ are } \sigma\text{-formulas}\} \subseteq \mathbb{N}^3$

where φ, t v range over formulas, terms and variables of $\mathbb{FOL}(\sigma)$.

Proof. This is an easy but tedious programming exercise. For example: for all $a \in \mathbb{N}$,

$$\mathrm{Sentence}(a) \iff \mathrm{Formula}(a) \text{ and } \forall i_{<a} \neg \mathrm{FreeVar}(a,i).$$

The readers are invited to check the rest of the relations themselves if they feel like programming. □

Definition 5.35. For a σ-theory T, define binary relations $\mathrm{Proof}_T, \mathrm{Refute}_T \subseteq \mathbb{N}^2$ by

$$\mathrm{Proof}_T := \{(\langle \ulcorner\varphi_1\urcorner, \dots, \ulcorner\varphi_n\urcorner\rangle, \ulcorner\varphi\urcorner) : (\varphi_1, \dots, \varphi_n) \text{ is a proof of } \varphi \text{ from } T\},$$
$$\mathrm{Refute}_T := \{(\langle \ulcorner\varphi_1\urcorner, \dots, \ulcorner\varphi_n\urcorner\rangle, \ulcorner\varphi\urcorner) : (\varphi_1, \dots, \varphi_n) \text{ is a proof of } \neg\varphi \text{ from } T\},$$

where φ_i and φ vary over σ-formulas.

For a σ-theory T, put $\ulcorner T\urcorner := \{\ulcorner\varphi\urcorner : \varphi \in T\}$. We say that T is *recursive* (resp. *primitive recursive*, *arithmetical*) if such is $\ulcorner T\urcorner$.

Lemma 5.36. *If a σ-theory T is recursive (resp. primitive recursive, arithmetical), then such is* Proof_T.

Proof. This is because for all $a \in \mathbb{N}$, $\mathrm{Proof}_T(a,b)$ if and only if $\mathrm{lh}(a) > 0$ and $(a)_{\mathrm{lh}(a)-1} = b$ and for every $k < \mathrm{lh}(a)$ either $(a)_k \in \mathrm{Axiom}$ or $(a)_k \in \ulcorner T\urcorner$ or $(\exists i < k)\,(\exists j < k)\,\mathrm{MP}((a)_i, (a)_j, (a)_k)$. □

5.F. **Robinson's system** Q

Now we describe a finite subtheory of $\mathrm{Th}(\boldsymbol{N})$, namely Robinson's[8] system Q, which is much weaker than PA, but still strong enough to represent recursive functions, c.f. Remark 5.31. The advantage of it over PA is that it is finite, and we will use this later in proving that the empty σ_{arthm}-theory is undecidable. However, this subsection can be safely skipped by readers, who are willing to accept that we can represent all recursive functions in some finite subtheory of $\mathrm{Th}(\boldsymbol{N})$.

Definition 5.37 (Robinson's system Q)**.** The following are the axioms of Q:

(Q1) $\forall x[\neg S(x) \doteq 0]$,
(Q2) $\forall x \forall y[S(x) \doteq S(y) \to x \doteq y]$,
(Q3) $\forall x[x + 0 \doteq x]$,
(Q4) $\forall x \forall y[S(x+y) \doteq x + S(y)]$,
(Q5) $\forall x[x \cdot 0 \doteq 0]$,
(Q6) $\forall x \forall y[x \cdot S(y) \doteq x \cdot y + x]$,
(Q7) $\forall x(x \neq 0 \to \exists y(x \doteq S(y)))$.

[8]This is due to Raphael Robinson and not Abraham or Julia Robinsons as I falsely thought.

So the difference between PA and Q is that the induction schema of PA is replaced by a single axiom stating that every nonzero element has a predecessor (which is clearly provable in PA). This theory is pretty weak: for example, it does not prove the associativity/commutativity of the addition/multiplication. However, every model of Q has a standard part:

Proposition 5.38.

(a) *For any model $\boldsymbol{M}$ of* Q*, there is a unique homomorphism $f : \boldsymbol{N} \to \boldsymbol{M}$. In fact, this f is a σ_{arthm}-embedding and hence we can view $\boldsymbol{N}$ as a substructure of $\boldsymbol{M}$.*

(b) *For any quantifier free formula $\varphi(\vec{x})$ and $\vec{a} \in \mathbb{N}^k$,*

$$\boldsymbol{N} \models \varphi(\vec{a}) \iff \mathrm{Q} \vdash \varphi(\Delta(\vec{a})),$$

where $\Delta(\vec{a}) := (\Delta(a_1), \dots, \Delta(a_k))$.

Proof. Part (b) follows from (a) since for $\boldsymbol{M} \models \mathrm{Q}$, $\boldsymbol{N} \subseteq \boldsymbol{M}$ and hence

$$\boldsymbol{N} \models \varphi(\vec{a}) \iff \boldsymbol{M} \models \varphi(\Delta(\vec{a})),$$

because φ is quantifier free. Because $\boldsymbol{M}$ was an arbitrary model of Q, we are done by the Completeness theorem.

As for part (a), the proof is exactly the same as for models of PA. The uniqueness is clear because we f has to preserve 0 and S and thus $f(\Delta(n)^{\boldsymbol{N}}) = \Delta(n)^{\boldsymbol{M}}$. This function is injective because $S^{\boldsymbol{M}}$ is injective and $0^{\boldsymbol{M}}$ does not have a predecessor. It remains to show that f preserves $+$ and $\cdot$. We show that $f(n+m) = f(n)+f(m)$ by induction on m, and we leave the case of $\cdot$ to the reader. For $m = 0$, this follows from axiom (Q3). Now assume $f(n+m) = f(n)+f(m)$. Then $f(n+S(m)) = f(S(n+m)) = S(f(n+m)) = S(f(n)+f(m)) = f(n)+S(f(m)) = f(n)+f(S(m))$, where we used the facts that f respects S and that $\boldsymbol{M}$ satisfies axiom (Q4). □

Let $x \le y$ and $x < y$ abbreviate the formulas $\exists z(z + x \doteq y)$ and $x \neq y \wedge \exists z(z + x \doteq y)$, respectively. Keep in mind that $z + x$ may not be equal to $x + z$ in a model of Q. Since the statement $x \le y$ is not quantifier free, it does not follow from the previous lemma that a model of Q and $\boldsymbol{N}$ have to agree on the ordering of natural numbers (the standard part of $\boldsymbol{M}$). However, it turns out to still be true:

Lemma 5.39 (Q preserves the ordering on $\mathbb{N}$)**.** *For all $n, m \in \mathbb{N}$,*

(a) $\mathrm{Q} \vdash x \le \Delta(n) \to \bigvee_{i=0}^{n} x = \Delta(i)$;

(b) $n \le m \iff \mathrm{Q} \vdash \Delta(n) \le \Delta(m)$;

(c) $\neg n \le m \iff \mathrm{Q} \vdash \neg\Delta(n) \le \Delta(m)$;

(d) $\mathrm{Q} \vdash x \le \Delta(n) \vee \Delta(n+1) \le x$;

(e) $\mathrm{Q} \vdash x \le \Delta(n) \vee \Delta(n) < x$.

Proof. For part (b), the right-to-left direction follows immediately from (a). As for the other direction, if $n \le m$, then let $k = m - n$ and thus $\boldsymbol{N} \models \Delta(k) + \Delta(n) = \Delta(m)$. By (b) of 5.38, $\mathrm{Q} \vdash \Delta(k) + \Delta(n) = \Delta(m)$ and thus $\mathrm{Q} \vdash \Delta(n) \le \Delta(m)$.

For (e), first consider $n = 0$. Then by (Q3), $\mathrm{Q} \vdash 0 \le x$, so the desired statement follows from the definition of the formula $y < z$. Now let $n \neq 0$ and hence $n = m + 1$. By (d), $\mathrm{Q} \vdash x \le \Delta(m) \vee \Delta(n) \le x$. Thus, arguing in Q and using (a), either $x = \Delta(k)$ for some $k < n$, or $x = \Delta(n)$, or $\Delta(n) \ge x$. Hence, again using (a) and the definition of the formula $y < z$, we get that either $x \le \Delta(n)$ or $\Delta(n) < x$.

We leave the proofs of (c) and (d) to the reader, and we prove (a) by induction on n. Let $\boldsymbol{M} \models \mathrm{Q}$. For $n = 0$, assume $a \in M$ and $\boldsymbol{M} \models a \le 0$. Thus, there is $b \in M$ such that $\boldsymbol{M} \models b + a \doteq 0$. Now if $a \neq 0^{\boldsymbol{M}}$, then a has a predecessor, i.e. for some $c \in \boldsymbol{M}$, $\boldsymbol{M} \models a \doteq S(c)$ and thus $\boldsymbol{M} \models b + S(c) \doteq 0$. Arguing inside $\boldsymbol{M}$, $0 = b + S(c) = S(b + c)$, which contradicts the fact that 0 is not a successor. Thus $a = 0$.

Now assume the statement is true for n and assume $\boldsymbol{M} \models a \le \Delta(n+1)$. Hence there is $b \in M$ such that $b + a = \Delta(n+1)$ (arguing inside $\boldsymbol{M}$). Now if $a = 0$, we are done. Otherwise, it has a predecessor $c \in M$ and thus $S(b + c) = b + S(c) = \Delta(n+1)$. By injectivity of S, we get $b + c = \Delta(n)$ and hence $c \le \Delta(n)$. By the induction hypothesis, c is equal to one of $\Delta(i)$ for $i = 0, \dots, n$ and thus a is equal to one of $\Delta(j)$ for $j = 1, \dots, n+1$. □

Proposition 5.40. *All recursive functions and relations are representable in* Q*.*

Proof. The proof is word-by-word the same as for Proposition 5.30 because we have proven above that the properties of PA used in that proof also hold for Q: namely, the required properties are (b) of 5.38 and (a,c,d) of 5.39. □

In a later subsection, we will also prove the converse of this proposition, so representability in Q actually characterizes recursive functions.

End of the revised part. The texts above and below might not be consistent.

5.G. **The First Incompleteness Theorem (Rosser's form)**

Define a function $\mathrm{Sub}_0 : \mathbb{N}^2 \to \mathbb{N}$ by $\mathrm{Sub}_0(a,n) = \mathrm{Sub}(a, \mathrm{SN}(v_0), \Delta(n))$. It is clear that Sub_0 is primitive recursive since such is Sub.

For a σ_{arthm}-formula θ, put $[\theta] := \Delta(\ulcorner\theta\urcorner)$.

Lemma 5.41 (Fixed point for Q)**.** *For every σ_{arthm}-formula $\varphi(v)$, there is a σ_{arthm}-sentence θ such that*

$$\mathrm{Q} \models \theta \leftrightarrow \varphi([\theta]).$$

Proof. Let $\boldsymbol{Sub}_0(x,y,z)$ be a σ_{arthm}-formula representing Sub_0 in Q. We can assume without loss of generality that the variable v_0 does not appear in $\boldsymbol{Sub}_0$ and φ. Put

$$\psi(v_0) \doteq \exists z(\boldsymbol{Sub}_0(v_0, v_0, z) \wedge \varphi(z)),$$

and let $m = \ulcorner\psi\urcorner$. Put $\theta \doteq \psi(\Delta(m))$. Then $\mathrm{Sub}_0(m,m) = \ulcorner\psi(\Delta(m))\urcorner = \ulcorner\theta\urcorner$ and hence, by the definition of representability,

$$\mathrm{Q} \models \boldsymbol{Sub}_0(\Delta(m), \Delta(m), z) \leftrightarrow z = [\theta]. \tag{i}$$

In particular,

$$\mathrm{Q} \models \boldsymbol{Sub}_0(\Delta(m), \Delta(m), [\theta]). \tag{ii}$$

Therefore, we have

$$\begin{aligned} \mathrm{Q} \models \theta &\iff \mathrm{Q} \models \psi(\Delta(m)) \\ &\iff \mathrm{Q} \models \exists z(\boldsymbol{Sub}_0(\Delta(m), \Delta(m), z) \wedge \varphi(z)) \\ [\Longrightarrow \text{ is because of (i)}] &\iff \mathrm{Q} \models \boldsymbol{Sub}_0(\Delta(m), \Delta(m), [\theta]) \wedge \varphi([\theta]) \\ [\Longleftarrow \text{ is because of (ii)}] &\iff \mathrm{Q} \models \varphi([\theta]). \end{aligned}$$

□

Now we are ready to prove the Incompleteness theorem for all σ_{arthm}-theories $T \supseteq \mathrm{Q}$. However, we would like to prove a slightly stronger version that applies to theories in signatures other than σ_{arthm} that are rich enough to encode Q in them. We make this precise in the following definition.

Definition 5.42. Let T_1, T_2 be theories in finite signatures σ_1, σ_2, respectively. An *interpretation* of T_1 in T_2 is a map π from the set of σ_1-sentences to the set of σ_2-sentences such that

(i) $T_1 \models \theta \implies T_2 \models \pi(\theta)$,
(ii) $T_2 \models \pi(\neg\theta) \leftrightarrow \neg\pi(\theta)$,
(iii) $T_2 \models \pi(\varphi \wedge \psi) \leftrightarrow \pi(\varphi) \wedge \pi(\psi)$,
(iv) there is a primitive recursive function $\pi^* : \mathbb{N} \to \mathbb{N}$ such that $\pi^*(\ulcorner\theta\urcorner) = \ulcorner\pi(\theta)\urcorner$,

where θ, φ, ψ range over σ_1-sentences, and in the last equality, $\ulcorner\ \urcorner$ denotes the coding function of $\mathbb{FOL}(\sigma_1)$ on the left and of $\mathbb{FOL}(\sigma_2)$ on the right.

If there is an interpretation of T_1 in T_2, we say that T_2 interprets T_1. For example, ZFC interprets Q. Also, if $T_1 \subseteq T_2$, then by taking the identity function as π^*, we see that T_2 interprets T_1.

Below let σ be a finite signature.

Lemma 5.43. *Let T be a (resp. primitive) recursive σ-theory that interprets* Q *and let π be an interpretation of* Q *in T. Then the following relations are (resp. primitive) recursive:*

$$\begin{aligned} \mathrm{Proof}_{\pi,T}(a,b) &\iff b \text{ is an } \mathbb{FOL}(\sigma_{\mathrm{arthm}})\text{-code of a } \sigma_{\mathrm{arthm}}\text{-sentence } \varphi \text{ and} \\ &\quad\ a \text{ is an } \mathbb{FOL}(\sigma)\text{-code of a proof of } \pi(\varphi) \text{ from } T, \\ \mathrm{Refute}_{\pi,T}(a,b) &\iff b \text{ is an } \mathbb{FOL}(\sigma_{\mathrm{arthm}})\text{-code of a } \sigma_{\mathrm{arthm}}\text{-sentence } \varphi \text{ and} \\ &\quad\ a \text{ is an } \mathbb{FOL}(\sigma)\text{-code of a proof of } \pi(\neg\varphi) \text{ from } T. \end{aligned}$$

Proof. Observe that

$$\begin{array}{rcl} \mathrm{Proof}_{\pi,T}(a,b) & \iff & \mathrm{Sentence}_{\sigma_{\mathrm{arthm}}}(b) \text{ and } \mathrm{Proof}_T(a,\pi^*(b)), \\ \mathrm{Refute}_{\pi,T}(a,b) & \iff & \mathrm{Sentence}_{\sigma_{\mathrm{arthm}}}(b) \text{ and } \mathrm{Proof}_T(a,\pi^*(\langle \mathrm{SN}(\neg), b\rangle)). \end{array}$$

□

First Incompleteness Theorem 5.44 (Rosser's form)**.** *Any consistent recursive σ-theory that interprets* Q *is incomplete.*

Let us contemplate about the proof a bit before we present it. In the proof of the Incompleteness theorem for $T \subseteq \mathrm{Th}(\boldsymbol{N})$, we constructed a sentence γ that basically expressed the Liar Paradox: it said about itself that it is not provable. Let us try to use the same idea here: let π be an interpretation of Q in T and let $\boldsymbol{Proof}_{\pi,T}(x,y)$ be a σ_{arthm}-formula representing $\mathrm{Proof}_{\pi,T}$ in Q. Then by the Fixed Point lemma for Q, we get a σ_{arthm}-sentence γ such that

$$\mathrm{Q} \models \gamma \leftrightarrow \forall x \neg \boldsymbol{Proof}_{\pi,T}(x,[\gamma]). \tag{$*$}$$

It is true that $T \nvdash \pi(\gamma)$ since otherwise there will be a code $a \in \mathbb{N}$ of a proof of $\pi(\gamma)$ from T and hence $\mathrm{Q} \models \boldsymbol{Proof}_{\pi,T}(\Delta(a),[\gamma])$. But then by $(*)$, $\mathrm{Q} \models \neg\gamma$ and thus $T \models \pi(\neg\gamma)$, so $T \models \neg\pi(\gamma)$, contradicting the consistency of T.

However, we don't get any contradiction if we assume $T \not\models \neg\pi(\gamma)$. Indeed, assuming the latter, the consistency of T implies that $T \nvdash \pi(\gamma)$ and hence there is no natural number that is a code of a proof of $\pi(\gamma)$ from T, i.e. $\neg\mathrm{Proof}_{\pi,T}(a,\ulcorner\gamma\urcorner)$, for all $a \in \mathbb{N}$. Then, for every $a \in \mathbb{N}$, $\mathrm{Q} \models \neg\boldsymbol{Proof}_{\pi,T}(\Delta(a),[\gamma])$. Unfortunately, this does NOT imply that $\mathrm{Q} \models \forall x \neg\boldsymbol{Proof}_{\pi,T}(x,[\gamma])$ because there may well be a model $\boldsymbol{M}$ of Q with a nonstandard element $w \in M \setminus \mathbb{N}$ such that $\boldsymbol{M} \models \boldsymbol{Proof}_{\pi,T}(w,[\gamma])$ and there is no contradiction here.

So, the Liar Paradox doesn't work here and Rosser's trick is to use the idea of the following joke[9]:

An an economist and his friend stumble upon a \$100 bill lying on the street. The friend says ``Hey, look, theres a \$100 bill on the sidewalk'' and bends over to pick it up, but the economist stops him, saying ``Don't bother because that's impossible --- if it were really a \$100 bill, someone would have picked it up by now.''

Rosser's proof of the Incompleteness First Incompleteness Theorem 5.44. Let π be an interpretation of Q in T, and let $\boldsymbol{Proof}_{\pi,T}(x,y)$ and $\boldsymbol{Refute}_{\pi,T}(x,y)$ be σ_{arthm}-formulas representing $\mathrm{Proof}_{\pi,T}$ and $\mathrm{Refute}_{\pi,T}$ in Q. Then by the Fixed Point lemma for Q, we get a σ_{arthm}-sentence ρ such that

$$\mathrm{Q} \models \rho \leftrightarrow \forall x(\boldsymbol{Proof}_{\pi,T}(x,[\rho]) \to (\exists u < x)\boldsymbol{Refute}_{\pi,T}(u,x)). \tag{1}$$

The *Rosser sentence* ρ expresses the unprovability of its translation in T in a round-about way: it asserts

For every proof of myself, there is a shorter proof of my negation.

We show that neither $T \nvdash \pi(\rho)$ nor $T \nvdash \neg\pi(\rho)$.

Case 1: suppose $T \nvdash \pi(\rho)$. Then there is a code $m \in \mathbb{N}$ of a proof of $\pi(\rho)$ from T and hence

$$\mathrm{Q} \nvdash \boldsymbol{Proof}_{\pi,T}(\Delta(m),[\rho]). \tag{2}$$

Because T is consistent, $T \nvdash \neg\pi(\rho)$, and hence, by the definition of interpretation, $T \nvdash \pi(\neg\rho)$. Thus $\forall k \in \mathbb{N}$, $\neg\mathrm{Refute}_{\pi,T}(k,\ulcorner\rho\urcorner)$ and hence $\mathrm{Q} \nvdash \neg\boldsymbol{Refute}_{\pi,T}(\Delta(k),[\rho])$; in particular, this is true for all $k < m$. Therefore, by (a) of Lemma 5.39,

$$\mathrm{Q} \nvdash (\forall u < \Delta(m))\neg\boldsymbol{Refute}_{\pi,T}(u,[\rho]). \tag{3}$$

From (2) and (3), we get

$$\mathrm{Q} \nvdash \exists x(\boldsymbol{Proof}_{\pi,T}(x,[\rho]) \wedge (\forall u < x)\neg\boldsymbol{Refute}_{\pi,T}(u,x)),$$

which implies $\mathrm{Q} \nvdash \neg\rho$ by (1). Therefore, $T \nvdash \pi(\neg\rho)$ and hence $T \nvdash \neg\pi\rho$, contradicting the consistency of T.

Case 2: suppose $T \nvdash \neg\pi(\rho)$. Thus $T \nvdash \pi(\neg\rho)$, so there is a code $k \in \mathbb{N}$ of a proof of $\pi(\neg\rho)$ from T. Hence $\mathrm{Refute}_{\pi,T}(k,\ulcorner\rho\urcorner)$ holds and by representability in Q,

$$\mathrm{Q} \nvdash \boldsymbol{Refute}_{\pi,T}(\Delta(k),[\rho]). \tag{4}$$

[9]The author has heard this joke from Itay Neeman in the context of searching for an apartment to rent in LA.

Also, for any $n \in \mathbb{N}$, $\neg \mathrm{Proof}_{\pi,T}(n, \ulcorner \rho \urcorner)$ holds by the consistency of T, and thus

$$\mathrm{Q} \vdash \neg \boldsymbol{Proof}_{\pi,T}(\Delta(n), [\rho]). \tag{5}$$

We argue in models, so fix $\boldsymbol{M} \models \mathrm{Q}$. By (e) of Lemma 5.39, for every $a \in M$, $a \leq \Delta(k)$ or $\Delta(k) < a$. In the first case, by (a) of Lemma 5.39, we get that $a = \Delta(n)$ for some $n \leq k$, and thus $\boldsymbol{M} \models \neg \boldsymbol{Proof}_{\pi,T}(a, [\rho])$, by (5). In the second case, i.e. if $\Delta(k) < a$,

$$\boldsymbol{M} \models (\exists u < a) \boldsymbol{Refute}_{\pi,T}(u, [\rho]),$$

by (4). Therefore, for all $a \in M$,

$$\boldsymbol{M} \models \boldsymbol{Proof}_{\pi,T}(a, [\rho]) \to (\exists u < a) \boldsymbol{Refute}_{\pi,T}(u, [\rho]).$$

Thus

$$\mathrm{Q} \vdash \forall x (\boldsymbol{Proof}_{\pi,T}(x, [\rho]) \to (\exists u < x) \boldsymbol{Refute}_{\pi,T}(u, x)),$$

and hence $\mathrm{Q} \vdash \rho$, by (1). But then $T \vdash \pi(\rho)$, contradicting the consistency of T. □

5.H. **The Second Incompleteness Theorem and Löb's theorem**

Let σ be a finite signature and let T be a recursive σ-theory. Recall (see Definition 5.35) that the relations $\mathrm{Proof}_T, \mathrm{Refute}_T \subseteq \mathbb{N}^2$ are recursive. Let $\boldsymbol{Proof}_T(x, y)$ and $\boldsymbol{Refute}_T(x, y)$ be σ_{arthm}-formulas representing them in Q, and put $\boldsymbol{Provable}_T(y) \doteq \exists x \boldsymbol{Proof}_T(x, y)$. Also recall that by $\bot$ we denote the sentence $\exists x (x \neq x)$.

Definition 5.45. For T as above, we define a σ_{arthm}-sentence that expresses the consistency of T as follows:

$$\boldsymbol{Con}_T \doteq \neg \boldsymbol{Provable}_T([\bot]).$$

Lemma 5.46. *Let T be a recursive σ-theory interpreting* PA *and let π be an interpretation. Also, let ρ_T be the Rosser sentence for T as in the proof of 5.44 above. Then* $\mathrm{PA} \vdash \boldsymbol{Con}_T \to \rho_T$.

Proof. We claim that Rosser's proof of the First Incompleteness theorem can be carried out in PA. It would take too long to actually prove this, but the main point is the following: Rosser's proof is completely syntactic, i.e. playing with formal proofs (we only used models and the Completeness theorem because we were too lazy to do formal proofs, but in principle we could have constructed all necessary formal proofs). Syntactic arguments such as the proofs of the Fixed Point lemma or Deduction theorem can be expressed and carried through PA because all they use is induction, which PA has.

Thus, in particular PA proves that if T is consistent then $T \nvdash \pi(\rho_T)$:

$$\mathrm{PA} \vdash \boldsymbol{Con}_T \to \forall x \neg \boldsymbol{Proof}_{\pi,T}(x, [\rho_T]).$$

On the other hand, it follows from the definition of ρ_T that

$$\mathrm{PA} \vdash \forall x \neg \boldsymbol{Proof}_{\pi,T}(x, [\rho_T]) \to \rho_T.$$

Therefore, $\mathrm{PA} \vdash \boldsymbol{Con}_T \to \rho_T$. □

From this we immediately get yet another foundational theorem by Gödel:

Second Incompleteness Theorem 5.47. *Let T be a recursive σ-theory interpreting* PA *and let π be an interpretation. Then $T \nvdash \pi(\boldsymbol{Con}_T)$, i.e. T cannot prove its own consistency.*

Proof. By the previous lemma and the fact that π is an interpretation of PA in T, we get

$$T \vdash \pi(\boldsymbol{Con}_T) \to \pi(\rho_T).$$

Thus, if $T \vdash \pi(\boldsymbol{Con}_T)$ then $T \vdash \pi(\rho_T)$, which is a contradiction. □

Lemma 5.48. *Let σ be a finite signature and T a recursive σ-theory. For any σ-sentences φ, θ, the following statements are provable in* PA*:*

(a) *The Deduction theorem:* $\boldsymbol{Provable}_{T \cup \{\theta\}}([\varphi]) \leftrightarrow \boldsymbol{Provable}_T([\theta \to \varphi])$.
(b) *Proof by contradiction:* $\boldsymbol{Provable}_T([\neg\theta \to \bot]) \leftrightarrow \boldsymbol{Provable}_T([\theta])$.
(c) *Lemma about consistency:* $\boldsymbol{Con}_{T \cup \{\neg\theta\}} \leftrightarrow \neg \boldsymbol{Provable}_T(\theta)$.

Proof. For parts (a) and (b), one has to note that the proofs of the corresponding theorems can be formalized in PA since all they use is syntactic arguments and induction. As for (c), it follows from (a) and (b) and we leave this as an exercise. □

Because $\boldsymbol{N}$ is a model of PA, we know that whatever PA proves is true about the natural numbers, in other words, for every σ_{arthm}-sentence θ,

$$\boldsymbol{N} \models \boldsymbol{Provable}_{\text{PA}}(\ulcorner\theta\urcorner) \to \theta.$$

Does PA know this? That is: does it prove $\boldsymbol{Provable}_{\text{PA}}(\ulcorner\theta\urcorner) \to \theta$ for all θ? Here is the answer:

Theorem 5.49 (Löb, 1955)**.** *For every σ_{arthm}-sentence θ,* PA *does not prove $\boldsymbol{Provable}_{\text{PA}}(\ulcorner\theta\urcorner) \to \theta$ unless it proves θ itself, i.e.*

$$\text{PA} \vdash \boldsymbol{Provable}_{\text{PA}}(\ulcorner\theta\urcorner) \to \theta \iff \text{PA} \vdash \theta.$$

Proof. We prove the left-to-right direction since the other one is trivial. Assume for contradiction that $\text{PA} \vdash \boldsymbol{Provable}_{\text{PA}}(\ulcorner\theta\urcorner) \to \theta$ yet $\text{PA} \nvdash \theta$. Thus the theory $S := \text{PA} \cup \{\neg\theta\}$ is consistent. By contrapositive, $\text{PA} \vdash \neg\theta \to \neg\boldsymbol{Provable}_{\text{PA}}(\ulcorner\theta\urcorner)$ and hence,

$$S \vdash \neg\boldsymbol{Provable}_{\text{PA}}(\ulcorner\theta\urcorner). \tag{$*$}$$

By (c) of Lemma 5.48, we have

$$\text{PA} \vdash \boldsymbol{Con}_S \leftrightarrow \neg\boldsymbol{Provable}_{\text{PA}}(\theta),$$

thus also

$$S \vdash \boldsymbol{Con}_S \leftrightarrow \neg\boldsymbol{Provable}_{\text{PA}}(\theta),$$

so, by Modus Ponens with $(*)$, we get $S \vdash \boldsymbol{Con}_S$, contradicting the Second Incompleteness theorem. □

6. Undecidable theories

Fix a finite signature σ.

Definition 6.1. For a σ-theory T, let $\text{Thm}(T)$ denote the set of its theorems, i.e. $\text{Thm}(T) := \{\varphi : T \vdash \varphi\} \subseteq \mathbb{N}$, where φ ranges over all σ-sentences. If $\ulcorner\text{Thm}(T)\urcorner$ is recursive, T is called decidable.

After various incompleteness results, we are now convinced that sufficiently rich recursive theories T such as PA or ZFC are incomplete. But maybe we can still write a program that for a given sentence φ decides whether it is a theorem of T or not? More precisely, is T decidable? (If the answer was yes for example for ZFC, mathematicians would be unemployed and the world would be an uninteresting place to live in.) This section is devoted to answering this question.

6.A. Σ^0_1 sets and Kleene's theorem

Below, let Γ be set of subsets of various finite powers of $\mathbb{N}$; e.g., $\Gamma = \mathcal{R}$, where $\mathcal{R}$ is the sets of all recursive sets, more precisely,

$$\mathcal{R} := \left\{A \subseteq \mathbb{N}^k : \ A \text{ recursive}, k \in \mathbb{N}\right\}.$$

Notation 6.2. For each $k \in \mathbb{N}$, put $\Gamma(\mathbb{N}^k) := \left\{A \subseteq \mathbb{N}^k : A \in \Gamma\right\}$, so $\Gamma = \bigcup_{k\in\mathbb{N}} \Gamma(\mathbb{N}^k)$. Also, put

$$\begin{aligned}
\neg\Gamma &:= \left\{\mathbb{N}^k \setminus A : A \in \Gamma(\mathbb{N}^k), k \in \mathbb{N}\right\} \\
\exists^{\mathbb{N}}\Gamma &:= \left\{\text{proj}_{k+1}(R) : R \in \Gamma(\mathbb{N}^{k+1}), k \in \mathbb{N}\right\} \\
\forall^{\mathbb{N}}\Gamma &:= \neg\exists^{\mathbb{N}}\neg\Gamma.
\end{aligned}$$

Definition 6.3. A set (relation) $A \subseteq \mathbb{N}^k$ is called Σ^0_1 if for some recursive relation $R \subseteq \mathbb{N}^{k+1}$, we have for all $\vec{a} \in \mathbb{N}^k$,

$$\vec{a} \in A \iff \exists y R(\vec{a}, y).$$

In other words, Σ^0_1 sets are exactly the projections of recursive sets. We also denote by Σ^0_1 the collection of all Σ^0_1 sets, i.e. $\Sigma^0_1 := \exists^{\mathbb{N}}\mathcal{R}$. Finally, put $\Pi^0_1 := \neg\Sigma^0_1$ and $\Delta^0_1 := \Sigma^0_1 \cap \Pi^0_1$.

Note that $\Pi^0_1 = \forall^{\mathbb{N}}\mathcal{R}$, and here are some closure properties of Σ^0_1 and Π^0_1:

Lemma 6.4. (a) *Σ^0_1 is closed under finite unions/intersections and projections, i.e. if $P, Q \subseteq \mathbb{N}^k$, $R \subseteq \mathbb{N}^{k+1}$ are Σ^0_1, then so are*

$$P \vee Q,\ P \wedge Q,\ \exists z R(\cdot, z).$$

Hence, Π^0_1 is closed under finite unions/intersections and co-projections, i.e. if $P, Q \subseteq \mathbb{N}^k$, $R \subseteq \mathbb{N}^{k+1}$ are Π^0_1, then so are

$$P \vee Q,\ P \wedge Q,\ \forall z R(\cdot, z).$$

(b) *Σ^0_1 is closed under recursive preimages, i.e. if $f : \mathbb{N}^k \to \mathbb{N}$ is recursive and $A \subseteq \mathbb{N}$ is Σ^0_1, then the relation $B = f^{-1}(A)$ is Σ^0_1. Same is true for Π^0_1.*

Proof. We leave (a) as a homework exercise, and we prove (b). Let $R \subseteq \mathbb{N}^2$ be a recursive relation such that for all $n \in \mathbb{N}$, $n \in A \iff \exists m R(n, m)$. But then the relation $Q \subseteq \mathbb{N}^{k+1}$ defined by

$$(\vec{a}, m) \in Q \iff R(f(\vec{a}), m)$$

is recursive and hence the relation

$$\vec{a} \in B \iff \exists m Q(\vec{a}, m)$$

is Σ^0_1. The statement about Π^0_1 follows from that about Σ^0_1 and the fact that preimages commute with complements. □

Lemma 6.5. *For a σ-theory T, if T is recursive, then $\ulcorner \mathrm{Thm}(T) \urcorner$ is Σ^0_1.*

Proof. If T is recursive, then so is the relation $\mathrm{Proof}_T \subseteq \mathbb{N}^2$ defined in the previous subsection. But then for all $a \in \mathbb{N}$

$$a \in \ulcorner \mathrm{Thm}(T) \urcorner \iff \exists x \mathrm{Proof}_T(x, a).$$ □

Let Π^0_1 denote the set of complements of Σ^0_1 relations, i.e. $\Pi^0_1 = \left\{\neg R : R \in \Sigma^0_1\right\}$, and let $\Delta^0_1 := \Sigma^0_1 \cap \Pi^0_1$. Also, let Recursive denote the set of recursive relations.

Lemma 6.6 (Kleene's theorem)**.** $\Delta^0_1 = \mathrm{Recursive}$.

Proof. $\supseteq$: It is clear that Recursive $\subseteq \Sigma^0_1$ (why?) and since Recursive is closed under complements, Recursive $\subseteq \Delta^0_1$.
$\subseteq$: Let $R \subseteq \mathbb{N}^k$ be a Δ^0_1 relation. Hence, there are recursive relations $P, Q \subseteq \mathbb{N}^{k+1}$ such that $\forall \vec{a} \in \mathbb{N}^k$

$$\vec{a} \in R \iff \exists x P(\vec{a}, x), \qquad \vec{a} \in \neg R \iff \exists x Q(\vec{a}, x).$$

But then the function $f : \mathbb{N}^k \to \mathbb{N}$ defined by $f(\vec{a}) = \mu_x(P \vee Q(\vec{a}, x))$ is recursive and hence so is R since $\vec{a} \in R \iff f(\vec{a}) \in P$. □

From this we immediately get the following decidability result:

Proposition 6.7. *Every complete recursive σ-theory T is decidable.*

Proof. Using the fact that for every σ-sentence φ, $\varphi \notin \mathrm{Thm}(T) \iff \neg\varphi \in \mathrm{Thm}(T)$, we get that for every $a \in \mathbb{N}$,

$$a \notin \ulcorner \mathrm{Thm}(T) \urcorner \iff a \notin \mathrm{Sentence}_\sigma \text{ or } \langle \mathrm{SN}(\neg), a \rangle \in \ulcorner \mathrm{Thm}(T) \urcorner.$$

By Lemma 6.5, $\ulcorner \mathrm{Thm}(T) \urcorner$ is Σ^0_1. Because $\neg\mathrm{Sentence}_\sigma$ is recursive (hence Σ^0_1) and Σ^0_1 is closed under recursive preimages and finite unions (6.4), the right hand side is Σ^0_1 and thus so is the completement of $\ulcorner \mathrm{Thm}(T) \urcorner$. Therefore, $\ulcorner \mathrm{Thm}(T) \urcorner$ is Δ^0_1 and hence is recursive (by Kleene's theorem). □

As a corollary, we get that ACF_p, $p = 0$ or prime, and the theory of vector spaces over a countable field[10] are decidable.

Another corollary of Kleene's theorem is a strengthening of Proposition 5.10.

Corollary 6.8. *For a function $f : \mathbb{N}^k \to \mathbb{N}$, the following are equivalent:*

[10] As it is written, 6.7 applies only to finite signatures and if a countable field F is not finite, the signature σ_F of the theory of vector spaces over F is infinite. However, we can still assign codes to symbols in σ_F so that we can decode all the information about the symbol from its code in a primitive recursive way. Thus everything proven above applies to σ_F as well.

(1) f *is recursive.*
(2) $\mathrm{Graph}(f) \subseteq \mathbb{N}^{k+1}$ *is recursive.*
(3) $\mathrm{Graph}(f) \subseteq \mathbb{N}^{k+1}$ *is* Σ^0_1.

Proof. By Proposition 5.10, it is enough to show (3) $\Rightarrow$ (2), for which, by Kleene's theorem, it is enough to show that if $\mathrm{Graph}(f)$ is Σ^0_1 then it is also Π^0_1. Indeed, for any $\vec{a} \in \mathbb{N}^k$ and $b \in \mathbb{N}$,

$$(\vec{a}, b) \in \mathrm{Graph}(f) \Leftrightarrow \forall y \in \mathbb{N}\Big[(\vec{a}, y) \notin \mathrm{Graph}(f) \vee y = b\Big].$$

It remains to note that because $\mathrm{Graph}(f)$ is Σ^0_1, the relation $(\vec{a}, y) \notin \mathrm{Graph}(f)$ is Π^0_1, so the expression on the right defines a Π^0_1 relation. □

6.B. **Universal Σ^0_1 relation and Church's theorem**

For any sets A, B, any relation $R \subseteq A \times B$, and $a \in A$, put $R(a) := \{b \in B : (a, b) \in R\}$. In this subsection we construct a Σ^0_1 relation $R \subseteq \mathbb{N}^2$ that is universal for recursive relations, i.e. any recursive relation $P \subseteq \mathbb{N}$ is of the form $P = R(a)$, for some $a \in \mathbb{N}$. Using this we prove that any consistent theory interpreting Q is undecidable. We start by proving the converse of 5.30.

Proposition 6.9. *Let T be a recursive consistent σ_{arthm}-theory. Then any relation $R \subseteq \mathbb{N}^k$ representable in T is recursive. In particular, any function $f : \mathbb{N}^k \to \mathbb{N}$ representable in T is recursive.*

Proof. The statement about functions follows from that about relations because if f is representable, then such is its graph ((b) of Proposition 5.28), therefore, by the first statement, the graph is recursive, and hence such is f (Proposition 5.10).

Let $R \subseteq \mathbb{N}^k$ be representable in T by a formula $\varphi(\vec{x})$. By the definition of representability and because T is consistent, for all $\vec{a} \in \mathbb{N}^k$, we have

$$\vec{a} \in R \iff T \vdash \varphi(\Delta(\vec{a})) \iff \ulcorner \varphi(\Delta(\vec{a})) \urcorner \in \ulcorner \mathrm{Thm}(T) \urcorner.$$

Lemma 6.5, $\mathrm{Thm}(T)$ is Σ^0_1 and the function $s : \mathbb{N}^k \to \mathbb{N}$ defined by $\vec{a} \to \ulcorner \varphi(\Delta(\vec{a})) \urcorner$ is clearly primitive recursive. Hence, the right hand side is Σ^0_1 by (b) of Lemma 6.4.

Because the definition of representability is symmetric for R and $\neg R$, we have that $\neg R$ is also representable (by $\neg\varphi$) and hence, by what we have already proven, $\neg R$ is Σ^0_1. Therefore, by Kleene's theorem, R is recursive. □

This, together with Propositions 5.30 and 5.40, gives the following characterization of recursive functions.

Corollary 6.10. *A function $f : \mathbb{N}^k \to \mathbb{N}$ is recursive if and only if it is representable in* PA *if and only if it is representable in* Q.

This allows us to construct a relation that enumerates all recursive subsets of $\mathbb{N}$ as follows:

Definition 6.11. Recall the primitive recursive function $\mathrm{Sub}_0(a, n)$ that has the property that for every σ_{arthm}-formula φ,

$$\mathrm{Sub}_0(\ulcorner \varphi \urcorner, n) = \ulcorner \varphi(\Delta(n)/v_0) \urcorner.$$

For a σ-theory T that interprets Q by π, define a relation $U_T \subseteq \mathbb{N}^2$ by

$$U_{\pi,T}(a, n) \iff \pi^*(\mathrm{Sub}_0(a, n)) \in \ulcorner \mathrm{Thm}(T) \urcorner.$$

Proposition 6.12. *Let T be a consistent σ-theory interpreting* Q *by π. Then for each recursive relation $R \subseteq \mathbb{N}$, there is $e \in \mathbb{N}$ such that $R = U_{\pi,T}(e)$. Furthermore, if T is recursive, then $U_{\pi,T}$ is Σ^0_1.*

Proof. The second statement follows from the definition of $U_{\pi,T}$ and 6.5. For the first statement, let $\varphi(v_0)$ be a formula representing R in Q (there is always one with the free variable being v_0), and thus for all $n \in \mathbb{N}$,

$$\begin{array}{ccccc} n \in R & \Longrightarrow & \mathrm{Q} \vdash \varphi(\Delta(n)) & \Longrightarrow & T \vdash \pi(\varphi(\Delta(n))) \\ n \notin R & \Longrightarrow & \mathrm{Q} \vdash \neg\varphi(\Delta(n)) & \Longrightarrow & T \vdash \neg\pi(\varphi(\Delta(n))). \end{array}$$

Since T is consistent, we get

$$n \in R \iff T \vdash \pi(\varphi(\Delta(n))),$$

and therefore, letting $e = \ulcorner \varphi(v_0) \urcorner$, we have

$$n \in R \iff U_{\pi,T}(e, n). \qquad \square$$

If we take $T = \mathrm{Q}$ and $\pi = \mathrm{id}$ in the above proposition, then, denoting $U_{\mathrm{id},\mathrm{Q}}$ by U_{Q}, we get an even stronger result:

Proposition 6.13. *The relation U_{Q} is Σ^0_1, and for every Σ^0_1 relation $P \subseteq \mathbb{N}$, there is $e \in \mathbb{N}$ with $P = U_{\mathrm{Q}}(e)$. Thus U_{Q} is a universal Σ^0_1 relation.*

Proof. This is left as a homework problem. $\square$

If T is recursive, we know that $U_{\pi,T}$ is Σ^0_1, but is it recursive? The answer is NO, and we show it by the diagonalization method.

Lemma 6.14 (Cantor)**.** *For a set A and a relation $R \subseteq A^2$, let $P \subseteq A$ be denote its antidiagonal, i.e. $P := \{a : \neg R(a, a)\}$. Then P is not equal to $R(a)$ for any $a \in A$.*

Proof. Assume for contradiction that $P = R(a)$, for some $a \in A$. Then we get a contradiction because

$$\neg R(a, a) \iff P(a) \iff R(a, a). \qquad \square$$

Corollary 6.15. *For every consistent σ-theory T interpreting Q by π, the relation $U_{\pi,T}$ is not recursive.*

Proof. If $U_{\pi,T}$ were recursive, so would be its antidiagonal P and thus, by 6.12, there is $a \in \mathbb{N}$ such that $P = U_{\pi,T}(a)$, contradicting 6.14. $\square$

As a corollary, we get the following important result:

Theorem 6.16 (Church, 1936)**.** *Any consistent σ-theory T interpreting Q is undecidable.*

Proof. Let π be an interpretation of Q in T. If T were decidable, i.e. $\ulcorner \mathrm{Thm}(T) \urcorner$ were recursive, then $U_{\pi,T}$ would be recursive as well, contradicting 6.15. $\square$

In particular, Q and PA are undecidable. Also, ZFC is undecidable unless it is inconsistent. Church's theorem also has the following rather surprising consequence based on the fact that Q is finite:

Corollary 6.17. *The empty σ_{arthm}-theory is undecidable, i.e. $\mathrm{Thm}_{\sigma_{\mathrm{arthm}}}(\emptyset)$ is not recursive.*

Proof. Let φ_{Q} be the conjunction of the axioms of Q (here is where we use that Q is finite!). Then, by the Deduction theorem, for any σ_{arthm}-sentence θ,

$$\mathrm{Q} \vdash \theta \iff \emptyset \vdash \varphi_{\mathrm{Q}} \to \theta.$$

Thus, letting $e = \ulcorner \varphi_{\mathrm{Q}} \urcorner$, we get that for all $a \in \mathbb{N}$,

$$a \in \ulcorner \mathrm{Thm}(\mathrm{Q}) \urcorner \iff \langle \mathrm{SN}(\to), e, a \rangle \in \ulcorner \mathrm{Thm}_{\sigma_{\mathrm{arthm}}}(\emptyset) \urcorner.$$

Hence, $\ulcorner \mathrm{Thm}_{\sigma_{\mathrm{arthm}}}(\emptyset) \urcorner$ cannot be recursive since otherwise $\ulcorner \mathrm{Thm}(\mathrm{Q}) \urcorner$ would also be recursive, contradicting 6.16. $\square$

7. Quantifier elimination

7.A. **Definitions and technicalities**

Fix a signature σ.

Definition 7.1. We say that a σ-theory T admits *quantifier elimination* (*q.e.*), if for every formula $\varphi(\vec{x})$, there is a quantifier-free (q.f.) formula $\psi(\vec{x})$ such that

$$T \vdash \forall \vec{x} (\varphi(\vec{x}) \leftrightarrow \psi(x)). \qquad (*)$$

Assuming that σ is finite, we say that T admits *effective quantifier elimination* if there is recursive function $h : \mathbb{N} \to \mathbb{N}$ such that for every formula $\varphi(\vec{x})$, $h(\ulcorner \varphi(\vec{x}) \urcorner)$ is a code of a q.f. formula $\psi(\vec{x})$ such that $(*)$ holds. We say that a σ-structure $\boldsymbol{A}$ admits (*effective*) *q.e.* if so does $\mathrm{Th}(\boldsymbol{A})$.

Note that for a σ-theory T to even have a chance to admit q.e., there would have to exist a quantifier-free sentence. To ensure that such always exists, we enrich $\mathbb{FOL}(\sigma)$ with propositional symbols for Truth and Falsity. More precisely, just like we always include the binary relation symbol $=$ in $\mathbb{FOL}(\sigma)$, we include 0-ary relation symbols $\top$ and $\bot$, together with the following axioms

(10)Truth: $\top \leftrightarrow \forall x(x = x)$

(11)Falsity: $\bot \leftrightarrow \neg T$

Below, we work with this enriched version of $\mathbb{FOL}(\sigma)$.

7.B. **Connection with decidability**

There is a strong connection between q.e. and decidability. To see this, consider the set $\mathrm{QFThm}(T) := \{\psi : \psi \text{ is a q.f. sentence and } T \vdash \psi\}$. In many interesting cases, this set (i.e. the set of the codes) is recursive. For example, for $T := \mathrm{Th}(\mathbb{R}, 0, 1, +, -, \cdot, <)$ or $T := \mathrm{ACF}$, a q.f. sentence is just a Boolean combination of (in)equalities about terms made out of $0, 1$ using $+, -, \cdot$, and hence it is (at least intuitively) clear that $\mathrm{QFThm}(T)$ is recursive (in fact primitive recursive); same is true for $T := \mathrm{Th}(\mathbb{N}, 0, S, +, \cdot)$.

Proposition 7.2. *Let σ be a finite signature and T a σ-theory such that $\mathrm{QFThm}(T)$ is recursive. If T admits effective q.e. then it is decidable.*

Proof. Let $h : \mathbb{N} \to \mathbb{N}$ be a recursive function as in Definition 7.1, then for every $n \in \mathbb{N}$,

$$n \in \ulcorner \mathrm{Thm}(T) \urcorner \iff h(n) \in \ulcorner \mathrm{QFThm}(T) \urcorner.$$

Thus, $\ulcorner \mathrm{Thm}(T) \urcorner$ is recursive since so is the right hand side. □

It is also important to note[11] that the effectiveness of q.e. comes for free if the theory is decidable; more precisely:

Proposition 7.3. *Let σ be a finite signature and T a σ-theory. If T admits q.e. and is decidable (e.g. when complete), then it actually admits effective q.e.*

Proof. Left as an exercise. □

Here are some famous q.e. results.

Theorem 7.4 (Tarski)**.** *The structure $(\mathbb{R}, 0, 1, +, -, \cdot, <)$ admits effective quantifier elimination and hence its theory is decidable.*

The above result is also known as *the decidability of Euclidean geometry*.

For p prime or 0, because ACF_p is decidable because it is complete. But here is a stronger result:

Theorem 7.5 (Robinson, Tarski, possibly others)**.** ACF *admits effective quantifier elimination.*

To appreciate this theorem, let $X = (x_{ij})_{i,j=1}^n$ be a matrix of variables and let $\varphi(X)$ be a σ_{ring}-formula expressing that X is invertible, i.e. $\varphi(X)$ says that there is a matrix of variables Y such that when multiplying by X one gets the identity matrix (this is a conjunction of n^2 equations). Clearly $\varphi(X)$ is an existential formula, but we know from linear algebra that there is a q.f. equivalent to it, namely, the formula expressing that the determinant of X is nonzero. This is not an entirely trivial fact, is it (think about coming up with the definition of determinant)? The above theorem implies this for every formula.

Recall the following reduct of $\boldsymbol{N}$: $\boldsymbol{N}_+ := (\mathbb{N}, 0, S, +)$. In one of the previous sections, we defined a (resp. primitive recursive) axiomatization T_+ for $\mathrm{Th}(\boldsymbol{N}_+)$ is stated that it is complete (and hence decidable). The completeness of T_+ is a consequence of the following.

Theorem 7.6 (Presburger)**.** T_+ *admits quantifier elimination.*

To conclude the completeness of T_+ from this note that any model $\boldsymbol{M}$ of T_+ has a standard part, i.e. $\boldsymbol{N} \subseteq \boldsymbol{M}$. Hence $\boldsymbol{M}$ and $\boldsymbol{N}$ believe the same q.f. sentences. But every sentence is equivalent (in T_+) to a q.f. sentence, and thus $\boldsymbol{N} \equiv \boldsymbol{M}$.

For the rest of the section, we will develop a model-theoretic criterion for q.e. using which we will show that ACF admits q.e. As an application, we will prove Hilbert's Nullstellensatz.

[11] Many thanks to William Balderrama for pointing this out.

7.C. **Syntactic approach**

Lemma 7.7 (Quantifier elimination test)**.** *A σ-theory T admits (effective) q.e. if and only if for every σ-formula of the form $\exists y \varphi(\vec{x}, y)$, where φ is q.f., there is a q.f. formula $\psi(\vec{x})$ such that $T \models \forall \vec{x} \Big(\big[\exists y \varphi(\vec{x}, y) \big] \leftrightarrow \psi(\vec{x}) \Big)$.*

Proof. Every formula is logically (i.e. in the empty theory) equivalent to one of the form:

$$\mathsf{Q}_1 y_1 \mathsf{Q}_2 y_2 \dots \mathsf{Q}_k y_k \varphi(\vec{x}, \vec{y}),$$

where each Q_i is either $\exists$ or $\forall$. Because $\forall$ is the same as $\neg\exists\neg$ and negation of a q.f. formula is still quantifier free, we can replace the quantifiers above with a sequence of $\exists$ and $\neg$, and eliminate the existential quantifiers one-by-one (more formally, by induction on k). □

Proposition 7.8. DLO *admits effective q.e.*

Proof. By the previous lemma, we have to describe a recursive procedure of getting rid of the existential quantifier from a formula of the form $\exists y \varphi(\vec{x}, y)$, where φ is q.f. Note that φ is a Boolean combination of equalities, inequalities and negations thereof. First note that we can get rid of negations: in DLO, $u \neq v$ is equivalent to $u < v \vee v < u$. Also, $u \not< v$ is equivalent in DLO to $u = v \vee v < u$. Thus, using the distributivity of $\wedge$ over $\vee$, we may assume that φ is a disjunction of conjunctions of equalities and inequalities. Finally, $\exists y$ distributes over disjunction and can be omitted from formulas with no other occurrences of y, so we may assume that φ is just a conjunction of equalities and inequalities, i.e. is of the form

$$\left(\bigwedge_{i \in I} y = x_i \right) \wedge \left(\bigwedge_{j \in J} y < x_j \right) \wedge \left(\bigwedge_{k \in K} x_k < y \right),$$

where $I, J, K \subseteq \{0, 1, \dots, |\vec{x}| - 1\}$.

Case: $I \neq \emptyset$. To obtain a q.f. equivalent, we fix $i \in I$ and simply replace every occurrences of y with x_i.

We now assume that $I = \emptyset$.

Case: $J = \emptyset$ or $K = \emptyset$. Say $J = \emptyset$. Then, φ is equivalent, in DLO, to $\top$ because DLO asserts that there is no maximum element, so a y satisfying $\bigwedge_{k \in K} x_k < y$ would exist in every model of DLO.

Case: $J \neq \emptyset$ and $K \neq \emptyset$. Because our linear ordering is required to be dense, such a y would exist in every model as long as $\max\{x_k : k \in K\} < \min\{x_j : j \in J\}$. Thus, in DLO, φ is equivalent to

$$\bigwedge_{j \in J, k \in K} x_k < x_j.$$

□

7.D. **Semantic approach**

Let σ be a signature and $\boldsymbol{A}$ be a σ-structure. For $B \subseteq A$, put $\sigma(B) := \sigma \cup B$, where elements of B are treated as new constant symbols. We define the natural expansion of $\boldsymbol{A}$ to a $\sigma(B)$-structure $\boldsymbol{A}(B)$ by interpreting symbols in B by themselves, i.e. $\forall b \in B$, $b^{\boldsymbol{A}(B)} = b$.

Definition 7.9. For a σ-structure $\boldsymbol{A}$ and $B \subseteq A$, define $\mathrm{Diag}(\boldsymbol{A}, B)$ as the set of all quantifier free $\sigma(B)$-sentences that are true in $\boldsymbol{A}(B)$, i.e.

$$\mathrm{Diag}(\boldsymbol{A}, B) := \{\psi : \psi \text{ is a q.f. } \sigma(B)\text{-sentence and } \boldsymbol{A}(B) \models \psi\}.$$

When $B = A$, we simply write $\mathrm{Diag}(\boldsymbol{A})$ instead of $\mathrm{Diag}(\boldsymbol{A}, A)$.

The following definition gives an equivalent (semantic) condition to quantifier elimination.

Definition 7.10. A σ-theory T is called diagram-complete if for any model $\boldsymbol{A}$ of T and any $\vec{a} \in A^n$ (for any n), the $\sigma(\vec{a})$-theory $T \cup \mathrm{Diag}(\boldsymbol{A}, \vec{a})$ is complete.

The term was chosen by me since I couldn't find an already existing name (although the notion is equivalent to substructure-completeness).

Proposition 7.11. *Suppose σ has at least one constant symbol c. Then a σ-theory T admits q.e. if and only if it is diagram-complete.*

Proof. $\Rightarrow$: Put $S := T \cup \mathrm{Diag}(\boldsymbol{A}, \vec{a})$ and let $\varphi(\vec{x})$ be a σ-formula with $\vec{x} = (x_1, \dots, x_n)$. We need to show that S proves either $\varphi(\vec{a})$ or $\neg\varphi(\vec{a})$. By q.e. there is a q.f. formula $\psi(\vec{x})$ such that $T \vdash \varphi(\vec{x}) \leftrightarrow \psi(x)$. By definition, $\psi(\vec{a}) \in \mathrm{Diag}(\boldsymbol{A}, \vec{a})$ or $\neg\psi(\vec{a}) \in \mathrm{Diag}(\boldsymbol{A}, \vec{a})$, and hence $S \vdash \varphi(\vec{a})$ or $S \vdash \neg\varphi(\vec{a})$.
$\Leftarrow$: Assume the right hand side and let $\varphi(\vec{x})$ be a σ-formula with $\vec{x} = (x_1, \dots, x_n)$. Take new constant symbols $\vec{d} = (d_1, \dots, d_n)$ and put

$$\Gamma(\vec{d}) := \left\{\psi(\vec{d}) : \psi \text{ is a q.f. } \sigma\text{-formula and } T \vdash \varphi(\vec{d}) \to \psi(\vec{d})\right\}.$$

Claim. $T \cup \Gamma(\vec{d}) \vdash \varphi(\vec{d})$.

Proof of Claim. Suppose for contradiction that $T \cup \Gamma(\vec{d}) \nvdash \varphi(\vec{d})$. Then $S(\vec{d}) := T \cup \Gamma(\vec{d}) \cup \left\{\neg\varphi(\vec{d})\right\}$ is consistent, so it has a model $\boldsymbol{A}(\vec{d})$, where $\boldsymbol{A}$ is its reduct to a σ-structure; in particular, $\mathrm{Diag}(\boldsymbol{A}, \vec{d}) \supseteq \Gamma(\vec{d})$. Since $\boldsymbol{A} \models T$ and T is diagram-complete, $S'(\vec{d}) := T \cup \mathrm{Diag}(\boldsymbol{A}, \vec{d})$ is a complete $\sigma(\vec{d})$-theory, so $S'(\vec{d})$ proves every sentence in $S(\vec{d})$ because $\boldsymbol{A}(\vec{d}) \models S'(\vec{d})$ and $\boldsymbol{A}(\vec{d}) \models S(\vec{d})$; in particular, $S'(\vec{d}) \vdash \neg\varphi(\vec{d})$. Because proofs are finite and $\mathrm{Diag}(\boldsymbol{A}, \vec{d})$ is closed under conjunctions, there is $\psi(\vec{d}) \in \mathrm{Diag}(\boldsymbol{A}, \vec{d})$ such that $T \vdash \psi(\vec{d}) \to \neg\varphi(\vec{d})$ (in case T alone proves $\neg\varphi(\vec{d})$, take $\psi(\vec{d}) \doteq \top$). Taking the contrapositive, it follows that $T \vdash \varphi(\vec{d}) \to \neg\psi(\vec{d})$, so $\neg\psi(\vec{d}) \in \Gamma(\vec{d}) \subseteq \mathrm{Diag}(\boldsymbol{A}, \vec{d})$, contradicting the consistency of $\mathrm{Diag}(\boldsymbol{A}, \vec{d})$. ⊠

Since proofs are finite and $\Gamma(\vec{d})$ is closed under conjunctions, there is $\psi \in \Gamma(\vec{d})$ such that $T \vdash \psi(\vec{d}) \to \varphi(\vec{d})$. $T \vdash \psi(\vec{d}) \to \varphi(\vec{d})$. On the other hand, by virtue of $\psi(\vec{d})$ being in $\Gamma(\vec{d})$, $T \vdash \varphi(\vec{d}) \to \psi(\vec{d})$. Therefore, $T \vdash \psi(\vec{d}) \leftrightarrow \varphi(\vec{d})$, and an application of the Constant Substitution Lemma 2.14 and Generalization Axiom (5) now finishes the proof. □

Note that in the definition of diagram-completeness, the model $\boldsymbol{A}$ is somewhat irrelevant, it is only there to make sure that $\mathrm{Diag}(\boldsymbol{A}, \vec{a})$ is consistent and contains $\psi(\vec{a})$ or $\neg\psi(\vec{a})$ for every q.f. formula $\psi(\vec{x})$. We make this precise in the lemma below.

Definition 7.12. Let $\vec{d}$ be a vector of distinct constant symbols that do not occur in σ. A set $\Gamma(\vec{d})$ of quantifier free $\sigma(\vec{d})$-sentences is called a T-diagram if $T \cup \Gamma(\vec{d})$ is consistent and for every q.f. $\sigma(\vec{d})$-sentence ψ, $\psi \in \Gamma(\vec{d})$ or $\neg\psi \in \Gamma(\vec{d})$.

Lemma 7.13. *A σ-theory T is diagram-complete if and only if for any $\vec{d}$ (of any length) and any T-diagram $\Gamma(\vec{d})$, $T \cup \Gamma(\vec{d})$ is a complete $\sigma(\vec{d})$-theory.*

Proof. $\Leftarrow$ follows from the Soundness of $\mathbb{FOL}$ and $\Rightarrow$ follows from the Completeness of $\mathbb{FOL}$. □

7.E. **Quantifier elimination for** ACF

In this subsection we prove that ACF is diagram-complete. The only method for showing completeness that we have learnt so far is the Łoś–Vaught test, and that is what we will use.

The proof of the following proposition is almost the same as of 4.7.

Proposition 7.14. *For every* ACF*-diagram $\Gamma(\vec{d})$, $\mathrm{ACF} \cup \Gamma(\vec{d})$ is a κ-categorical $\sigma_{ring}(\vec{d})$-theory, for every uncountable cardinal κ.*

Proof. Let $\boldsymbol{K}_1, \boldsymbol{K}_2 \models \mathrm{ACF} \cup \Gamma(\vec{d})$ with $|K_1| = |K_2| = \kappa$. Note that $\boldsymbol{K}_1, \boldsymbol{K}_2$ have the same characteristic since it is expressible by a q.f. σ_{ring}-sentence which must be contained in $\Gamma(\vec{d})$. Let p be the characteristic ($p = 0$ or p is prime).

For $i = 1, 2$, let F_i be the base field of $\boldsymbol{K}_i$, i.e. the substructures of $\boldsymbol{K}_i$ generated by $\emptyset$. (If $p = 0$, then F_i is a copy of $\mathbb{Q}$; otherwise it is a copy of $\mathbb{Z}/p\mathbb{Z}$.) Since F_1 and F_2 are clearly isomorphic (as rings), we can assume without loss of generality that $F_1 = F_2 =: F$. Let $\vec{a} = \vec{d}^{\boldsymbol{K}_1}$, $\vec{b} = \vec{d}^{\boldsymbol{K}_2}$, and denote by $F(\vec{a})$, $F(\vec{b})$ the fields inside K_1, K_2, generated by $\vec{a}$,$\vec{b}$ over F, respectively.

Claim. $F(\vec{a})$ and $F(\vec{b})$ are isomorphic.

Proof of Claim. Elements of $F(\vec{a})$ are of the form $\frac{p(\vec{a})}{q(\vec{a})}$, where p, q are polynomials over F and $q(\vec{a}) \neq 0$. Define $h : F(\vec{a}) \to F(\vec{b})$ by $\frac{p(\vec{a})}{q(\vec{a})} \mapsto \frac{p(\vec{b})}{q(\vec{b})}$. This is well-defined because if $q(\vec{a}) \neq 0$, then $q(\vec{b}) \neq 0$ since $\vec{a}$ and $\vec{b}$ have the same diagram $\Gamma(\vec{d})$ and $q(\vec{d}) \neq 0$ is a q.f. $\sigma_{\text{ring}}(\vec{d})$-sentence, which must be in $\Gamma(\vec{d})$ since $\vec{a}$ satisfies it. It is easy to verify that h is a field homomorphism and hence is injective, and it is surjective because elements of $F(\vec{b})$ are of the form $\frac{p(\vec{b})}{q(\vec{b})}$, for some polynomials p, q over F. ⊠

Without loss of generality, we can identify $F(\vec{a})$ and $F(\vec{b})$, i.e. assume that $L := F(\vec{a}) = F(\vec{b})$. Let B_i be transcendence base over L in K_i. (Transcendence base is a maximal collection of algebraically independent elements over L.) Now it is not hard to see that $K_i = \overline{L(B_i)}$, where $L(B_i)$ denotes the field generated by B_i over L and $\overline{L(B_i)}$ denotes its algebraic closure in K_i.

Because L is countable, $|K_i| = |B_i| \cdot \aleph_0 + |L|$. If B_i is countable then so is $|B_i| \cdot \aleph_0 + |L|$, but K_i is uncountable, and hence B_i is uncountable. Then, by basic cardinal arithmetic, $|B_i| \cdot \aleph_0 + |L| = |B_i|$ and so $\kappa = |K_i| = |B_i|$. Hence, there is a bijection $f : B_1 \to B_2$, which uniquely extends to an isomorphism of $L(B_1)$ onto $L(B_2)$ by a map similar to the one in the proof of the claim above. This isomorphism in its turn extends (not necessarily uniquely) to an isomorphism of $K_1 = \overline{L(B_1)}$ onto $K_2 = \overline{L(B_2)}$. □

Corollary 7.15. ACF *admits quantifier elimination.*

Proof. Follows from 7.13 and 7.11. □

Corollary 7.16. *The definable subsets of an algebraically closed field are finite or cofinite.*

Proof. Let K be an algebraically closed field. By q.e., every definable set $S \subseteq F$ is defined by a q.f. formula $\varphi(x)$. For the base case $\varphi(x) \doteq (t_1(x) = t_2(x))$, the statement is clear since $t_i(x)$ is a polynomial in x with coefficients in K and the polynomial $t_1(x) - t_2(x)$ has only finitely-many roots. The step case is also clear since the set of finite and cofinite subsets of K is closed under finite unions (corresponding to $\wedge$) and complements (corresponding to $\neg$). □

Remark 7.17. One can also show using a similar argument that the theory of vector spaces over a countable field admits q.e. and conclude that the definable subsets of a vector space are only the finite and cofinite ones. In general, structures with only definable subsets being finite or cofinite are called strongly minimal. It turns out that in those structures one can always define an abstract model-theoretic operation that generalizes *algebraic closure* (for fields) and *span* (for vector spaces), and this operation allows to define notions of basis and dimension such that the rest of the structure is "free" over a basis in the sense that any bijection between bases extends to a (not necessarily unique) isomorphism between the structures.

7.F. **Model-completeness**

The following is a very useful notion that is slightly weaker than quantifier elimination.

Definition 7.18. A σ-theory T is called *model-complete* if $\boldsymbol{A} \subseteq \boldsymbol{B}$ implies $\boldsymbol{A} \preceq \boldsymbol{B}$, for all $\boldsymbol{A}, \boldsymbol{B} \models T$.

Proposition 7.19. *Quantifier elimination implies model-completeness.*

Proof. Suppose T admits q.e. and $\boldsymbol{A} \subseteq \boldsymbol{B}$, where $\boldsymbol{A}, \boldsymbol{B} \models T$. Because $\boldsymbol{A}$ and $\boldsymbol{B}$ agree on the q.f. formulas about the elements of A, and every formula is equivalent to a q.f. formula (in T), $\boldsymbol{A}$ and $\boldsymbol{B}$ agree on all formulas about the elements of A. A σ-structure $\boldsymbol{M}$ is called *model-complete* if such is $\mathrm{Th}(\boldsymbol{M})$. □

It can be shown that $(\mathbb{R}, 0, 1, +, -, \cdot)$ is model-complete but it does not admit q.e.

Recalling that we simply write $\mathrm{Diag}(\boldsymbol{A})$ for $\mathrm{Diag}(\boldsymbol{A}, A)$, the following proposition justifies the terminology with regards to diagram-completeness and highlights the difference with quantifier elimination.

Proposition 7.20. *For a σ-theory T, the following are equivalent:*

(1) *T is model-complete.*

(2) *For every model $\boldsymbol{A} \models T$, $T \cup \mathrm{Diag}(\boldsymbol{A})$ is a complete $\sigma(A)$-theory.*

(3.a) *Every σ-formula $\varphi(\vec{x})$ is equivalent in T to a universal formula.*

(3.b) *Every σ-formula $\varphi(\vec{x})$ is equivalent in T to an existential formula.*

Proof. All implications are easy, except for (2)⇒(3.a). The proof of the latter follows the same idea as that of Proposition 7.11 and we leave it as a (good) exercise. □

7.G. **Hilbert's Nullstellensatz**

Recall that ACF admits q.e. and hence is model-complete. As a nice application of the latter fact, we deduce what would be the first theorem in algebraic geometry.

Hilbert's Nullstellensatz 7.21 (Weak Form)**.** *Let F be an algebraically closed field and I be a proper ideal in the polynomial ring $F[t_1,\dots,t_n]$. Then the polynomials in I have a common root in F, i.e. there is $\vec{a}\in F^n$ such that $f(\vec{a})=0$ for all $f(t_1,\dots,t_n)\in F[t_1,\dots,t_n]$.*

Proof. Take a maximal ideal M containing I (exists by Zorn's lemma) and put

$$K := F[t_1,\dots,t_n]/M.$$

Since M is maximal, K is a field. Note that now every polynomial in M has a root in K in the following sense: for $f(t_1,\dots,t_n)\in M$, let $f(x_1,\dots,x_n)$ be the polynomial obtained from $f(t_1,\dots,t_n)$ by replacing t_i with variables x_i of $\mathbb{FOL}(\sigma_{\text{ring}})$. Then, by the definition of K, for all such $f\in M$, $f(\vec{b})=0$, where $\vec{b}=(t_1+M,\dots,t_n+M)\in K$. (This is why we moved from F to K: to artificially create a common root).

Let L be an algebraic closure of K. Since $K\subseteq L$, there is still a common root in L for all polynomials in M. Now we want to use the model-completeness of ACF to transfer this statement down to F to obtain a common root in F. However, expressing (in a first-order way) the statement that all polynomials in M have a common root seems to be a problem because there are infinitely-many polynomials in M (while formulas are finite). Luckily, Hilbert's Basis theorem says that any ideal in $F[t_1,\dots,t_n]$ is finitely generated, so M is generated by some $f_1,\dots,f_m\in F[t_1,\dots,t_n]$. Thus all polynomials in M having a common root is equivalent to $f_1,\dots,f_m$ having a common root. Put

$$\varphi(\vec{a}) := \exists\vec{x}\bigwedge_{i=1}^{m}(f_i(\vec{x})\doteq 0),$$

where $\vec{a}\in F^k$ is a tuple containing all coefficients of $f_1,\dots,f_m$. By model-completeness of ACF, because $\boldsymbol{F}\subseteq\boldsymbol{L}$ and $\boldsymbol{F},\boldsymbol{L}\models$ ACF, we have $\boldsymbol{F}\preceq\boldsymbol{L}$. Hence $\boldsymbol{F}\models\varphi(\vec{a})$ because $\boldsymbol{L}\models\varphi(\vec{a})$, and thus $f_1,\dots,f_m$ have a common root in F. □

From this form of the Nullstellensatz, we can derive its strong form using the so-called Rabinowitsch trick; this does not use any model theory, but we do it here anyway for recreation. First we introduce some notation. For a ring R, let $\mathcal{I}(R)$ denote the set of its ideals. For a field F, $\vec{a}\in F^n$, and $J\in\mathcal{I}(F[\vec{x}])$, we say that $\vec{a}$ *annihilates* J, written $J(\vec{a})=0$, if for each $f\in J$, $f(\vec{a})=0$. Put $C(J) := \{\vec{a}\in F^n : J(\vec{a})=0\}$. Similarly, for $A\subseteq F^n$, put $I(A) := \{f\in F[\vec{x}] : (\forall\vec{a}\in A)\, f(\vec{a})=0\}$. Clearly, $I(C(J))\supseteq\sqrt{J}$, where $\sqrt{J}$ is the *radical* of J, i.e.

$$\sqrt{J} := \{f\in F[\vec{x}] : f^m\in J \text{ for some } m\in\mathbb{N}\}.$$

Hilbert's Nullstellensatz 7.22 (Strong Form)**.** *Let F be an algebraically closed field. For any $J\in\mathcal{I}(F[\vec{x}])$, $I(C(J))=\sqrt{J}$.*

Proof. Let $f\in I(C(J))$, so every $\vec{a}\in F^n$ that annihilates J, also annihilates f. Let t be a new indeterminant variable and note that there is no element of K^{n+1} that annihilates both J and $1-tf$. Thus, by the weak form of Hilbert's Nullstellensatz, the ideal generated by $J\cup\{1-tf\}$ in $F[\vec{x},t]$ must be equal to $F[\vec{x},t]$. Hence, there are some $f_1,\dots,f_k\in J$ and $g_1(t),\dots,g_{k+1}(t)\in F[\vec{x},t]$ such that

$$g_1(t)f_1+\dots+g_k(t)f_k+g_{k+1}(t)(1-tf)=1.$$

Assuming that $f\neq 0$ (otherwise, we are done), plug in $t=1/f$ and get

$$g_1(1/f)f_1+\dots+g_k(1/f)f_k=1.$$

Multiplying both sides with f^m for large enough $m\in\mathbb{N}$, we get

$$\tilde{g}_1f_1+\dots+\tilde{g}_kf_k=f^m,$$

where $\tilde{g}_1,\dots,\tilde{g}_k$ are some polynomials in $F[\vec{x}]$, which shows that $f\in\sqrt{J}$. □

References

[End01] H. B. Enderton, *A Mathematical Introduction to Logic*, 2nd ed., Academic Press, 2001.
[Mar02] D. Marker, *Model Theory: An Introduction*, Graduate Texts in Mathematics, Springer, 2002.
[Mos08] Y. N. Moschovakis, *Informal notes full of errors*, unpublished, 2008.
[vdD10] L. van den Dries, *Mathematical Logic: Lecture Notes*, unpublished, 2010.

Department of Mathematics, University of Illinois at Urbana-Champaign, IL, 61801, USA
E-mail address: anush@illinois.edu