FIRST ORDER LOGIC AND GÖDEL INCOMPLETENESS

ANUSH TSERUNYAN

Contents

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].

1. Introduction

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 three parts: basic model theory, basic recursion theory, and more.

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. Then, we'll apply the developed techniques to concrete examples such as the structure of natural numbers and algebraically closed fields; the latter will yield 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). We will also discuss applications of the Compactness theorem in 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.

Basic recursion theory

At the beginning of the 20th 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.

And more

Diving more into model theory, we will study quantifier elimination and model completeness, and, as a quick application, give a transparent proof of Hilbert's Nullstellensatz. Then, changing gears, we will learn two completely different (set-theoretic and combinatorial) constructions of structures from existing ones: ultraproducts and Fraïssé limits. The former will involve a rather measure-theoretic introduction to ultrafilters, while the latter will touch base with probabilistic objects like the random graph.

2. 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...

2.A. **Structures**

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

Examples 2.1.

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

(b) A *partial ordering* is a pair $\mathbf{P} = (P, \leq)$, 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 triple $\mathbf{G} = (G, e, \cdot)$, where G is a set, e is a fixed element of G (a constant) and $\cdot$ is a binary operation on G 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) $\forall x \in G$, $e \cdot x = x \cdot e = x$,
(iii) (Inverse) $\forall x \in G$ $\exists y \in G$, $xy = yx = e$.

(d) An *ordered field* is a 6-tuple $\mathbf{F} = (F, 0, 1, +, \cdot, <)$, 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 (d)) defined on it. Let's formalize this and give an abstract definition of a *mathematical structure*.

Definition 2.2. A *structure* is a quadruple $\mathbf{S} = (S, \mathcal{C}, \mathcal{F}, \mathcal{R})$, 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 "types" of constants, functions and relations. 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, $\mathcal{C}$ and $\mathcal{F}$ are singletons, $\mathcal{R} = \varnothing$, and the operation in $\mathcal{F}$ is binary, but we also have to require that conditions (i)-(iii) of Example 2.1(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 e refers to the unique element in $\mathcal{C}$ and $\cdot$ refers to the unique element in $\mathcal{F}$. So why don't we first fix a set of names (like $\{e, \cdot\}$) 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 2.3. A *signature* is a quadruple

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

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

$$a : \mathcal{F} \cup \mathcal{R} \to \mathbb{N}^{>0}.$$

(Here $\mathbb{N}^{>0}$ denotes the set of positive natural numbers because, in mathematical logic, $\mathbb{N}$ includes 0.)

A relation or function symbol P (i.e. an element of $\mathcal{F} \cup \mathcal{R}$) is said to be n-ary if $a(P) = n$. The sets $\mathcal{C}, \mathcal{F}, \mathcal{R}$ should be thought of as *names* for constant elements, relations and functions (operations), and not the actual constant elements, relations and functions themselves! It is also good to keep in mind that any of the sets $\mathcal{C}, \mathcal{F}, \mathcal{R}$ can be empty.

Examples 2.4.

(a) The signature for graphs is

$$\tau_{\text{graph}} = (\varnothing, \varnothing, \{E\}, (E \mapsto 2)),$$

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

$$\tau_{\text{graph}} = (E),$$

and then specify that E is a binary relation symbol.

(b) The signature for monoids is

$$\tau_{\text{mon}} = (e, \cdot),$$

where $\cdot$ is a binary function symbol and e is a constant symbol.

(c) The signature for rings is

$$\tau_{\text{ring}} = (0, 1, +, -, \cdot),$$

where $+, -, \cdot$ are binary function symbols and $0, 1$ are constant symbols.

(d) The signature for arithmetic is

$$\tau_{\text{arthm}} = (0, S, +, \cdot),$$

where 0 is a constant symbol, S is a unary function symbol (S stands for "successor"), and $+, \cdot$ are binary function symbols.

(e) The signature for sets is

$$\tau_{\text{set}} = (\in),$$

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 $\tau = (\mathcal{C}, \mathcal{R}, \mathcal{F})$.

Definition 2.5. A *τ-structure* is a pair $\mathbf{S} = (S, \iota)$, where S is a set and ι is a map (correspondence) that assigns

- to each constant symbol c in τ a member $\iota(c)$ of S;
- to each n-ary relation symbol R in τ an n-ary function $\iota(R) \subseteq S^n$;
- to each n-ary function symbol f in τ an n-ary function $\iota(f) : S^n \to S$.

We call S the universe of the structure $\mathbf{S}$. The choice of the letter ι is because we think of ι as the *interpretation* of the symbols of τ in the structure $\mathbf{S}$. To simplify the notation, we write $q^{\mathbf{S}}$ instead of $\iota(q)$, for all symbols q in τ, so instead of (S, ι), we write

$$\mathbf{S} = (S, \{c^{\mathbf{S}}\}_{c \in \mathcal{C}}, \{R^{\mathbf{S}}\}_{R \in \mathcal{R}}, \{f^{\mathbf{S}}\}_{f \in \mathcal{F}}).$$

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

Examples 2.6.

(a) A complete graph on n vertices is a τ_{graph}-structure

$$\mathbf{K_n} = (\Gamma, E^{\mathbf{K_n}}),$$

where Γ is a set of n elements and $E^{\mathbf{K_n}} = \Gamma^2$.

(b) $\mathbb{Z}$, as a group, is a τ_{mon}-structure

$$\mathbf{Z} = (\mathbb{Z}, e^{\mathbf{Z}}, \cdot^{\mathbf{Z}}),$$

where $e^{\mathbf{Z}}$ is $0 \in \mathbb{Z}$ and $\cdot^{\mathbf{Z}}$ is the usual addition operation.

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

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

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

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

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

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

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

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

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

Since it is annoying to keep writing $\mathbf{S}$ in the superscript to denote the interpretation of symbols of τ in a τ-structure $\mathbf{S}$, 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 $\mathbf{R} = (\mathbb{R}, 0, 1, +, -, \cdot)$ instead of $\mathbf{R} = (\mathbb{R}, 0^{\mathbf{R}}, 1^{\mathbf{R}}, +^{\mathbf{R}}, -^{\mathbf{R}}, \cdot^{\mathbf{R}})$ if it is the structure in Example 2.6(c), but we won't use this shorthand notation with anything like Example 2.6(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.

Definition 2.7. For τ-structures $\mathbf{A}, \mathbf{B}$, we say that $\mathbf{A}$ is a *substructure* of $\mathbf{B}$ and write $\mathbf{A} \subseteq \mathbf{B}$ if $A \subseteq B$ and the interpretations of τ by $\mathbf{A}$ and $\mathbf{B}$ coincide on A, more precisely:

- $c^{\mathbf{A}} = c^{\mathbf{B}}$, for any constant symbol c in τ,
- $f^{\mathbf{A}} = f^{\mathbf{B}} \restriction_{A^n}$ for any n-ary function symbol f in τ, i.e. $f^{\mathbf{A}}(\vec{a}) = f^{\mathbf{B}}(\vec{a})$ for all $\vec{a} \in A^n$,
- $R^{\mathbf{A}} = R^{\mathbf{B}} \cap A^n$ for any n-ary relation symbol R in τ, i.e. $R^{\mathbf{A}}(\vec{a}) \Leftrightarrow R^{\mathbf{B}}(\vec{a})$ for all $\vec{a} \in A^n$.

For example, $(\mathbb{N}, 0, +)$ is a substructure of $(\mathbb{Z}, 0, +)$, $\mathbf{Z} = (\mathbb{Z}, 0, 1, +, \cdot)$ is a substructure of $\mathbf{R} = (\mathbb{R}, 0, 1, +, \cdot)$. If τ only contains relation symbols, then any subset is (a universe of) a substructure. For example, if (Γ, E) is a graph and $\Delta \subseteq \Gamma$, then $(\Delta, E \cap \Delta^2)$, i.e. the induced subgraph on Δ, is a substructure of (Γ, E). However, note that being a subgraph is not the same as being a substructure of a graph: indeed, a subgraph of a graph (Γ, 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 (Γ, E).

Note that the intersection of substructures of the same structure is again a substructure. Let $\mathbf{B}$ be a τ-structure and $S \subseteq B$. The *substructure generated by* S is the smallest substructure containing S, i.e. it is the intersection of all substructures of $\mathbf{B}$ that contain S. We denote this fact by $\mathbf{A} = \langle S \rangle_{\mathbf{B}}$. Note that the universe of $\langle S \rangle_{\mathbf{B}}$ is obtained from S by throwing in the constants of $\mathbf{B}$ and closing it under the functions $\mathbf{B}$. For example, the substructure of $\mathbf{R} = (\mathbb{R}, 0, 1, +, \cdot)$ generated by $\varnothing$ is $(\mathbb{N}, 0, 1, +, \cdot)$ (why?).

For a τ-structure $\mathbf{B}$ and $A \subseteq B$, we say that A is *a universe of a substructure* of $\mathbf{B}$ if the universe of $\langle A \rangle_{\mathbf{B}}$ is A (in other words, A already contains all of the constants of $\mathbf{B}$ and is closed under the functions of $\mathbf{B}$). For example, if τ only has relation symbols, then any subset $A \subseteq B$ is a universe of a substructure.

Definition 2.8. Let $\mathbf{A}, \mathbf{B}$ be τ-structures. A function $h : A \to B$ is called a *τ-homomorphism* (or just homomorphism) if h respects the interpretation of τ, more precisely:

- $h(c^{\mathbf{A}}) = c^{\mathbf{B}}$, for any constant symbol c in τ,
- $h(f^{\mathbf{A}}(\vec{a})) = f^{\mathbf{B}}(h(\vec{a}))$, for any n-ary function symbol f in τ and for all $\vec{a} \in A^n$,
- $R^{\mathbf{A}}(\vec{a}) \Rightarrow R^{\mathbf{B}}(h(\vec{a}))$, for any n-ary relation symbol R in τ and for all $\vec{a} \in A^n$,

where for $\vec{a} = (a_1, ..., a_n)$, $h(\vec{a}) := (h(a_1), ..., h(a_n))$. Denote this by $h : \mathbf{A} \to \mathbf{B}$.

Note that in this definition, we only require $\Rightarrow$ for relations. This asymmetry is justified by the fact that if we look at the graphs of functions $f^{\mathbf{A}}$ and $f^{\mathbf{B}}$ as $(n+1)$-ary relations $R_f^{\mathbf{A}}$ and $R_f^{\mathbf{B}}$, then, putting $b = f^{\mathbf{A}}(\vec{a})$, the condition $h(f^{\mathbf{A}}(\vec{a})) = f^{\mathbf{B}}(h(\vec{a}))$ is equivalent to $R_f^{\mathbf{A}}(\vec{a}, b) \Rightarrow R_f^{\mathbf{B}}(h(\vec{a}), h(b))$.

It is straightforward to verify that for $h : \mathbf{A} \to \mathbf{B}$, $h(A)$ is a universe of a substructure of $\mathbf{B}$ (i.e. $h(A)$ contains all of the constants of $\mathbf{B}$ and is closed under applications of the functions of $\mathbf{B}$).

Definition 2.9. Let $\mathbf{A}, \mathbf{B}$ be τ-structures. A function $h : A \to B$ is called a *τ-isomorphism* (or just isomorphism) if h is bijective and both h and h^{-1} are τ-homomorphisms; in this case we write $h : \mathbf{A} \xrightarrow{\sim} \mathbf{B}$. The structures $\mathbf{A}, \mathbf{B}$ are called isomorphic if there is an isomorphism between them; denote this by $\mathbf{A} \simeq \mathbf{B}$.

Definition 2.10. Let $\mathbf{A}, \mathbf{B}$ be τ-structures and $h : \mathbf{A} \to \mathbf{B}$. Recalling that $h(A)$ is the universe of $\langle h(A) \rangle_{\mathbf{B}}$, call h a *τ-embedding* (or just embedding) if h is an isomorphism between $\mathbf{A}$ and $\langle h(A) \rangle_{\mathbf{B}}$. We denote this by $h : \mathbf{A} \hookrightarrow \mathbf{B}$.

Note that if $\mathbf{A} \subseteq \mathbf{B}$ 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 2.11. Let τ, τ' be signatures with $\tau \subseteq \tau'$, let $\mathbf{A}$ be a τ-structure and $\mathbf{B}$ be a τ'-structure. We say that $\mathbf{A}$ is a *reduct* of $\mathbf{B}$ (or $\mathbf{B}$ an expansion of $\mathbf{A}$) and write $\mathbf{A} = \mathbf{B}\downharpoonright_\tau$ if $\mathbf{A}$ and $\mathbf{B}$ have the same underlying set and the same interpretations of the symbols of τ.

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

2.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 2.1(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 τ denote a signature for the rest of the section.

Definition 2.12. The *alphabet* $\mathbb{FOL}(\tau)$ of the first order language in the signature τ comprises of the symbols in τ and the following additional symbols:

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

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

Definition 2.13. A *τ-term* (or a term in $\mathbb{FOL}(\tau)$) is a word formed by the following recursive rules:

(i) each constant symbol is a term;
(ii) each variable is a term;
(iii) if $t_1, \ldots, t_n$ are terms and $f \in \tau$ is an n-ary function symbol, then $f(t_1, \ldots, t_n)$ is a term.

Examples 2.14.

(a) $(v_0 \cdot 1) \cdot v_3$ is a term in $\mathbb{FOL}(\tau_{\text{group}})$. Note that the way this term is written is technically incorrect, we should have written $\cdot(\cdot(v_0, 1), v_3)$, 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 $\mathbb{FOL}(\tau_{\text{arthm}})$ (the language of arithmetic).

(c) Variables $v_0, v_1, \ldots$ are the only terms in $\mathbb{FOL}(\tau_{\text{graph}})$.

We also casually use letters different than $v_0, v_1, ...$ to denote variables, e.g. v, u, x, y, z. Below, a vector of variables $(v_1, ..., v_n)$ is denoted by $\vec{v}$ and we let $|\vec{v}|$ denote its length n.

Definition 2.15 (Interpretation of terms)**.** Let $\mathbf{M}$ be a τ-structure and t be a τ-term build using variables from $\vec{v}$. We define the *interpretation* of $t(\vec{v})$ in $\mathbf{M}$ as a function $t^{\mathbf{M}} : M^{|\vec{v}|} \to M$ by induction on the construction of t as follows: for $\vec{a} = (a_1, ..., a_{|\vec{v}|}) \in M^{|\vec{v}|}$,

(i) if $t = c$, where c is a constant symbol in τ, then $t^{\mathbf{M}}(\vec{a}) = c^{\mathbf{M}}$;
(ii) if $t = v_i$, then $t^{\mathbf{M}}(\vec{a}) = a_i$;
(iii) if $t = f(t_1, ..., t_k)$, where $t_1, ..., t_k$ are terms and f is an k-ary function symbol in τ, then $t^{\mathbf{M}}(\vec{a}) = f^{\mathbf{M}}(t_1^{\mathbf{M}}(\vec{a}), ..., t_k^{\mathbf{M}}(\vec{a}))$.

So one should think of the term $t(\vec{v})$ as a name of the function $t^{\mathbf{M}}$. 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). This is exactly what we do with polynomials for example: we write $p(x, y) = x^2 + 1$ to mean that this is a polynomial in two variables x and y although it doesn't depend on y.

Convention 2.16. For a term t and a vector of variable $\vec{v}$, whenever we write $t(\vec{v})$, we mean that the variables appearing in t are among those in $\vec{v}$.

Definition 2.17. A *τ-formula* (or a formula in $\mathbb{FOL}(\tau)$) is a word formed by the following recursive rules:

(i) if s, t are terms then $s = t$ is a formula;
(ii) if $t_1, ..., t_n$ are terms and $R \in \tau$ is an n-ary relation symbol, then $R(t_1, ..., t_n)$ is a formula;
(iii) if φ and ψ are formulas, then $\neg(\varphi)$, $(\varphi) \wedge (\psi)$, $(\varphi) \vee (\psi)$, $(\varphi) \to (\psi)$ are formulas;
(iv) if φ is a formula and v a variable symbol, then $\forall v \varphi$, $\exists v \varphi$ are formulas.

The formulas in (i) and (ii) of the above definition are called *atomic*. Also, if a formula is formed without using (iv), then it does not have any quantifiers, so we call it *quantifier free* (or q.f. for short).

According to this definition, $(\forall x(x = y)) \ \wedge \ (x \neq z)$ is a valid formula (in any signature), although 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 2.18. We say that the variable v is *quantified* in the formula φ if $\forall v \psi$ or $\exists v \psi$, for some formula ψ, occurs in some stage of the recursive construction of φ. We make the convention that each variable v can be used with a quantifier only once, i.e. a subword of the form $\mathtt{Q} v \psi$ occurs at most once, where $\mathtt{Q}$ is either $\forall$ or $\exists$, and if it does, then v is not allowed to be used elsewhere other than in ψ.

This convention makes things like $(\forall x(x = y)) \ \wedge \ (x \neq z)$ invalid, and one should write $(\forall t(t = y)) \ \wedge \ (x \neq z)$ 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 language of set theory).

We interpret formulas in a given structure $\mathbf{M}$ as n-ary relations on M, for some M, or, equivalently, as functions from M^n to $\{\mathtt{true,\ false}\}$. Just like we did with terms, we define interpretation for $\varphi(\vec{v})$ (as opposed to just φ), for a vector of variables $\vec{v} = (v_1, ..., v_n)$, as long as the free variables of φ are among $v_1, ..., v_n$ and none of $v_1, ..., v_n$ is quantified in φ.

Convention 2.19. For a formula φ and a vector of variable $\vec{v} = (v_1, ..., v_n)$, whenever we write $\varphi(\vec{v})$, we mean that all of the free variables of φ are among $v_1, ..., v_n$ and none of $v_1, ..., v_n$ is quantified in φ.

Definition 2.20 (Interpretation of formulas)**.** Let $\mathbf{M}$ be a τ-structure and $\varphi(\vec{v})$ a τ-formula. For $\vec{a} = (a_1, ...a_{|\vec{v}|}) \in M^{|\vec{v}|}$, we define the relation $\mathbf{M} \vDash \varphi(\vec{a})$ by induction on the construction of φ as follows:

(i) if φ is $t_1 = t_2$, then $\mathbf{M} \vDash \varphi(\vec{a})$ if $t_1^{\mathbf{M}}(\vec{a}) = t_2^{\mathbf{M}}(\vec{a})$;
(ii) if φ is $R(t_1, ..., t_k)$, then $\mathbf{M} \vDash \varphi(\vec{a})$ if $R^{\mathbf{M}}(t_1^{\mathbf{M}}(\vec{a}), ..., t_k^{\mathbf{M}}(\vec{a}))$, i.e. $(t_1^{\mathbf{M}}(\vec{a}), ..., t_k^{\mathbf{M}}(\vec{a})) \in R^{\mathbf{M}}$;
(iii) if φ is $\neg\psi$, then $\mathbf{M} \vDash \varphi(\vec{a})$ if $\mathbf{M} \nvDash \varphi(\vec{a})$;
(iv) if φ is $\psi \wedge \theta$, then $\mathbf{M} \vDash \varphi(\vec{a})$ if $\mathbf{M} \vDash \psi(\vec{a})$ and $\mathbf{M} \vDash \theta(\vec{a})$;
(v) if φ is $\psi \vee \theta$, then $\mathbf{M} \vDash \varphi(\vec{a})$ if $\mathbf{M} \vDash \psi(\vec{a})$ or $\mathbf{M} \vDash \theta(\vec{a})$;
(vi) if φ is $\forall u \psi(\vec{v}, u)$ (hence u is not in $\vec{v}$ by our assumption), then $\mathbf{M} \vDash \varphi(\vec{a})$ if for all $b \in M$, $\mathbf{M} \vDash \psi(\vec{a}, b)$;
(vii) if φ is $\exists u \psi(\vec{v}, u)$, then $\mathbf{M} \vDash \varphi(\vec{a})$ if there exists $b \in M$, $\mathbf{M} \vDash \psi(\vec{a}, b)$.

We read $\mathbf{M} \vDash \varphi(\vec{a})$ as φ is true (holds) about $\vec{a}$ in $\mathbf{M}$. Note that the above definition applies when φ is a sentence and $n = 0$. In this case, we read $\mathbf{M} \vDash \varphi$ as φ is true/valid (holds) in $\mathbf{M}$. For a vector of variables $\vec{v} = (v_1, ..., v_n)$, we say that a formula $\varphi(\vec{v})$ is valid in $\mathbf{M}$ and write $\mathbf{M} \vDash \varphi(\vec{v})$ if $\mathbf{M} \vDash \forall \vec{v} \varphi(\vec{v})$, where $\forall \vec{v}$ abbreviates $\forall v_1 \forall v_2 ... \forall v_n$.

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.

Examples 2.21.

(a) $\mathbf{N} \vDash S(S(0)) = 2$.

(b) Let $\mathbf{N}_{\exp} = (\mathbb{N}, 0, S, +, \cdot, \exp)$, 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 $\mathbf{N}_{\exp} \vDash \forall n \forall x \forall y \forall z[(n \geq 3 \wedge \exp(x, n) + \exp(y, n) = \exp(z, n)) \to (x = 0 \vee y = 0)]$, where $n \geq 3$ stands for $n \neq 0 \wedge n \neq S(0) \wedge n \neq S(S(0))$.

(c) $\mathbf{R} \vDash \exists y(a = y \cdot y)$ holds for all non-negative $a \in \mathbb{R}$.

Convention 2.22. Because the symbol = that is part of $\mathbb{FOL}(\tau)$, it may get confusing to use it also in our regular mathematical notation to express that two terms or two formulas are equal (literally the same). Thus, we adopt the convention to use the symbol $\doteq$ instead of =. For example, instead of writing "let $\phi = x = x$", we write "let $\phi \doteq x = x$".

Lemma 2.23. *Let* $\mathbf{A}, \mathbf{B}$ *be two* τ*-structures. If* $h : \mathbf{A} \to \mathbf{B}$ *is a homomorphism, then for any term* $t(\vec{v})$ *and* $\vec{a} \in A^{|\vec{v}|}$,

$$h(t^{\mathbf{A}}(\vec{a})) = t^{\mathbf{B}}(h(\vec{a})),$$

where $h(\vec{a}) = (h(a_1), ..., h(a_{|\vec{v}|}))$.

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

- If $t \doteq c$, for a constant symbol c in τ, 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 $t \doteq v_i$, for a variable v_i, then $t^{\mathbf{A}}(\vec{a}) = a_i$ and hence we have
$$h(t^{\mathbf{A}}(\vec{a})) = h(a_i) = t^{\mathbf{B}}(h(\vec{a})).$$
- If $t \doteq f(t_1, ..., t_k)$, for a function symbol f in τ, then
$$\begin{aligned} h(t^{\mathbf{A}}(\vec{a})) &= h(f^{\mathbf{A}}(t_1^{\mathbf{A}}(\vec{a}), ..., t_k^{\mathbf{A}}(\vec{a}))) \\ [h \text{ is a homomorphism}] \quad &= f^{\mathbf{B}}(h(t_1^{\mathbf{A}}(\vec{a})), ..., h(t_k^{\mathbf{A}}(\vec{a}))) \\ [\text{by the induction hypothesis}] \quad &= f^{\mathbf{B}}(t_1^{\mathbf{B}}(h(\vec{a})), ..., t_k^{\mathbf{A}}(h(\vec{a}))) \\ &= t^{\mathbf{B}}(h(\vec{a})). \end{aligned}$$

□

Proposition 2.24. *Let* $\mathbf{A}, \mathbf{B}$ *be two* τ*-structures. If* $h : \mathbf{A} \to \mathbf{B}$ *is an isomorphism, then for any formula* $\varphi(\vec{v})$ *and* $\vec{a} \in A^{|\vec{v}|}$,
$$\mathbf{A} \vDash \varphi(\vec{a}) \iff \mathbf{B} \vDash \varphi(h(\vec{a})).$$

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

- If $\varphi \doteq t_1 = t_2$, then
$$\begin{aligned} \mathbf{A} \vDash \varphi(\vec{a}) \quad &\iff t_1^{\mathbf{A}}(\vec{a}) = t_2^{\mathbf{A}}(\vec{a}) \\ [h \text{ is injective}] \quad &\iff h(t_1^{\mathbf{A}}(\vec{a})) = h(t_2^{\mathbf{A}}(\vec{a})) \\ [\text{by Lemma 2.23}] \quad &\iff t_1^{\mathbf{B}}(h(\vec{a})) = t_2^{\mathbf{B}}(h(\vec{a})) \\ &\iff \mathbf{B} \vDash \varphi(h(\vec{a})). \end{aligned}$$
- If $\varphi \doteq R(t_1, ..., t_k)$, then the calculation is similar to the previous case (also uses Lemma 2.23).
- If $\varphi \doteq \neg\psi$, then
$$\begin{aligned} \mathbf{A} \vDash \varphi(\vec{a}) \quad &\iff \mathbf{A} \nvDash \psi(\vec{a}) \\ [\text{by the induction hypothesis}] \quad &\iff \mathbf{B} \nvDash \psi(\vec{a}) \\ &\iff \mathbf{B} \vDash \varphi(h(\vec{a})). \end{aligned}$$
- If $\varphi \doteq \psi_1 \wedge \psi_2$, then the calculation is similar to the previous case.
- If $\varphi \doteq \exists v \psi$, then
$$\begin{aligned} \mathbf{A} \vDash \varphi(\vec{a}) \quad &\iff \exists a' \in A, \mathbf{A} \vDash \psi(\vec{a}, a') \\ [\text{by the induction hypothesis}] \quad &\iff \exists a' \in A, \mathbf{B} \vDash \psi(h(\vec{a}), h(a')) \\ [\text{use surjectivity of } h \text{ for } \Longleftarrow] \quad &\iff \exists b \in B, \mathbf{B} \vDash \psi(h(\vec{a}), b) \\ &\iff \mathbf{B} \vDash \varphi(h(\vec{a})). \end{aligned}$$

□

Proposition 2.25. *If a τ-structure* $\mathbf{A}$ *is a reduct of a τ'-structure* $\mathbf{B}$*, then for every τ-formula $\varphi(\vec{v})$ and $\vec{a} \in A^n$ $(= B^n)$,*

$$\mathbf{A} \vDash \varphi(\vec{a}) \iff \mathbf{B} \vDash \varphi(\vec{a}).$$

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

2.C. **Definability**

Definition 2.26 (Definability)**.** Let $\mathbf{M}$ be a τ-structure and $A \subseteq M$. $D \subseteq M^n$ is called A-definable (or definable from A) in $\mathbf{M}$ if there is a formula $\varphi(\vec{x}, \vec{y})$, where $\vec{x} = (x_1, ..., x_n)$ and $\vec{y} = (y_1, ..., y_m)$ (for some $m \geq 0$), and $\vec{a} \in M^m$ such that $\forall \vec{b} \in M^n$

$$\vec{b} \in D \Leftrightarrow \mathbf{M} \vDash \varphi(\vec{b}, \vec{a}).$$

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

Note that the set $\mathcal{D}_n(A)$ of A-definable subsets of M^n is an algebra, i.e. it is closed under finite unions and complements and contains $\varnothing$ and M^n. It is very useful to consider the topology $\mathcal{T}_n(A)$ on M^n generated by $\mathcal{D}_n(A)$. It is clear that $\mathcal{D}_n(A)$ is actually a base for that topology. Note that this topology might not be Hausdorff (see Example 2.27(c) below) and whether it is compact or not is tightly related to a property called saturation, which however is outside the scope of this course.

Examples 2.27.

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

(b) In $\mathbf{R}_<$ the set $\{r \in \mathbb{R} : r < \pi\}$ is definable by the formula $x < \pi$. It turns out that this set is not 0-definable. This follows from the fact that π is transcendental and a famous theorem of Tarski that $\mathbf{R}_<$ admits "quantifier elimination", which implies that all 0-definable sets are just finite unions of intervals with algebraic (or infinite) endpoints.

(c) In $\mathbf{C} := (\mathbb{C}, 0, 1, +, \cdot)$, the set $\{\sqrt{2}, -\sqrt{2}\}$ is 0-definable by $\varphi(z) \doteq z^2 - 2 = 0$, where z^2 and 2 are the abbreviation for $z \cdot z$ and $1 + 1$, respectively. However, $\sqrt{2}$ itself isn't 0-definable! This follows from the fact that $\mathbf{C}$ admits "quantifier elimination" (as we will see later), so the only definable sets are those defined by polynomials and Boolean combinations thereof. In particular, the topology $\mathcal{T}_1(\varnothing)$ of 0-definable sets isn't Hausdorff (not even T_0) as any 0-definable set containing $\sqrt{2}$ also contains $-\sqrt{2}$.

(d) In any graph $\mathbf{\Gamma} := (\Gamma, E)$, the set

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

is 0-definable by the formula

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

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

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

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

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

is not even definable in some (actually most) graphs. We will prove this later on in the course after proving the Compactness theorem.

The definable subsets of $\mathbf{N} := (\mathbb{N}, 0, S, +, \cdot)$ are called *arithmetical*. It is easy to see that a set is definable in $\mathbb{N}$ if and only if it is 0-definable.

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

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

Definition 2.28. We say that a **nonempty** τ-structure $\mathbf{M}$ *satisfies* (or *models*) a τ-theory T and write $\mathbf{M} \vDash T$ if $\mathbf{M} \vDash \varphi$, for every $\varphi \in T$. Equivalently, we also say that $\mathbf{M}$ is a *τ-model* (or just *model*) of T.

Notation 2.29. For a τ-theory T, let $\mathcal{M}_\tau(T)$ denote the class[1] of its τ-models, i.e. nonempty τ-structures that satisfy it.

Definition 2.30. For a class $\mathcal{C}$ of τ-structures that is invariant under τ-isomorphism[2], τ-theory T is called an *axiomatization* of $\mathcal{C}$ if $\mathcal{M}_\tau(T) = \mathcal{C}$. A class $\mathcal{C}$ is called (resp. *finitely*) *axiomatizable* if it admits a (resp. finite) axiomatization. A τ-theory S is called an *axiomatization* of T if it is an axiomatization of $\mathcal{M}_\tau(T)$, and we call T *finitely axiomatizable* if it admits a finitely axiomatization.

Here are examples of axiomatizations for various classes of structures.

Examples 2.31.

(a) Graphs (undirected with no loops): Letting $\tau_{\text{graph}} := (E)$, the class of undirected graphs (τ_{graph}-structures) 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.

(b) Infinite graphs: The class of undirected infinite graphs (τ_{graph}-structures) with no loops is axiomatized by the theory GRAPHS_∞ consisting of the following axioms:

$$\text{GRAPHS}_\infty := \text{GRAPHS} \cup \left\{ \exists v_1 \exists v_2 \ldots \exists v_n \bigwedge_{i<j} v_i \neq v_j : n \geq 2 \right\}.$$

We will show later on in the course that this class is not finitely axiomatizable.

[1] We will show later that if a theory has an infinite model, then it has models of arbitrary large cardinalities, so indeed, $\mathcal{M}_\tau(T)$ is too large to be a set, so it is a proper class.

[2] This means that if $\mathcal{C}$ contains a structure, then it also contains all of its isomorphic copies.

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

(d) Groups: Letting $\tau_{\text{mon}} := (e, \cdot)$, the class of groups ($\tau_{\text{mon}}$-structures) is axiomatized by the theory GROUPS consisting of the following axioms:
(G1) (Associativity) $\forall x \forall y \forall z[x \cdot (y \cdot z) = (x \cdot y) \cdot z]$,
(G2) (Identity) $\forall x[e \cdot x = x \cdot e = x]$,
(G3) (Inverse) $\forall x \exists y[xy = yx = e]$.

(e) Rings and fields: Similarly, one defines the theory RINGS of rings in the signature $\tau_{\text{ring}} := (0, 1, +, -, \cdot)$ (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) $\forall x \forall y[x \cdot y = y \cdot x]$,
(F3) (Multiplicative inverse) $\forall x \exists y[xy = yx = 1]$,

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

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

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

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

(h) Characteristic 0 fields:

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

(i) 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, such as connected graphs, disconnected graphs, and cyclic groups, are not axiomatizable!

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

Example 2.32. The theory PA of arithmetic, called Peano Arithmetic, in the signature $\tau_{\text{arthm}} := (0, S, +, \cdot)$, consists of the following (infinitely many) axioms:

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

$$[\varphi(0, \vec{y}) \wedge \forall x(\varphi(x, \vec{y}) \to \varphi(x + 1, \vec{y}))] \to \forall x \varphi(x, \vec{y}).$$

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

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

Example 2.33. The Zermelo-Fraenkel set theory, ZFC, is a theory in the signature $\tau_{\text{set}} := (\in)$, in which all of the mathematics is derived. Its list of axiom schemas 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.

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

Definition 2.34. We say that a τ-theory T satisfies a τ-sentence φ and write $T \models \varphi$, if every model of T satisfies φ, i.e. $\forall \mathbf{M} \models T(\mathbf{M} \models \varphi)$. Equivalently, we say that T semantically implies φ.

Examples 2.35.

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

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

$$\mathrm{FIELDS}_p \models \underbrace{1 + 1 + \ldots + 1}_{n} = 0 \iff p \text{ divides } n.$$

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

Definition 2.36. A τ-theory T is said to be

- *satisfiable* (or *semantically consistent*) if it has a model.
- *semantically τ-complete* (or just semantically complete if τ is understood) if for every τ-sentence φ, $T \models \varphi$ or $T \models \neg\varphi$.

Let $\top$ denote the sentence $\forall x(x = x)$ and set $\bot \doteq \neg\top$. Note that T is satisfiable if and only if $T \not\models \bot$.

Definition 2.37 (Elementary equivalence)**.** Let $\mathbf{A}$ and $\mathbf{B}$ be τ-structures. We say that $\mathbf{A}$ and $\mathbf{B}$ are called *elementarily equivalent*, and write $\mathbf{A} \equiv \mathbf{B}$, if $\mathrm{Th}(\mathbf{A}) = \mathrm{Th}(\mathbf{B})$.

By Proposition 2.24, 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 (in fact, $(\mathbb{Q}, <) \prec (\mathbb{R}, <)$), but they clearly cannot be isomorphic simply because of cardinality considerations.

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

Proposition 2.38 (Semantic completeness, rephrased)**.** *A τ-theory T is semantically complete if and only if for any $\mathbf{A}, \mathbf{B} \vDash T$, $\mathbf{A} \equiv \mathbf{B}$.*

Proof. Left as an exercise. □

In the light of this, it is easy to see that most of the examples of 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 (i.e. satisfies the sentence $\exists x(x \cdot x \cdot x = e)$) 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.

Definition 2.39. A τ-theory T is said to be *fully τ-complete* (or just fully complete if τ is understood) if for every τ-sentence φ, $\varphi \in T$ or $\neg\varphi \in T$.

Note that $\mathrm{Th}(\mathbf{M})$ is satisfiable and fully complete, for any τ-structure $\mathbf{M}$. Thus, every satisfiable theory admits a satisfiable full completion.

2.F. **Elementarity**

Let $\mathbf{B}$ be a τ-structure and $\mathbf{A}$ a substructure of $\mathbf{B}$. It is an interesting question as to which formulas $\mathbf{A}$ and $\mathbf{B}$ agree on. The following is all we can say for general $\mathbf{A} \subseteq \mathbf{B}$.

Proposition 2.40. *Substructures agree on quantifier free formulas; more precisely, for τ-structures $\mathbf{A} \subseteq \mathbf{B}$, any quantifier free τ-formula γ and $\vec{a} \in A^n$, we have*

$$\mathbf{A} \vDash \gamma(\vec{a}) \iff \mathbf{B} \vDash \gamma(\vec{a}).$$

Proof. Easy induction on the construction of γ which only involves the cases $\gamma \doteq \neg\varphi$, $\gamma \doteq \varphi \wedge \psi$, and $\gamma \doteq t_1 = t_2$, where for the latter case one has to use Lemma 2.23 and the fact that the inclusion map $A \hookrightarrow B$ is a homomorphism. □

However, the opinions of a structure and a substructure about 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 object for which φ holds (which we refer to as *a witness*) may not be in the universe of the substructure. For example, in the signature τ_{group}, 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 following definitions isolate those substructures which agree with the ambient structure on all of the statements about the elements of the substructure.

Definition 2.41 (Elementary embedding)**.** Let $\mathbf{A}, \mathbf{B}$ be τ-structures. An *embedding* $f : \mathbf{A} \to \mathbf{B}$ is called *elementary* if for all formulas $\varphi(\vec{x})$ and tuples $\vec{a} \in A^n$,

$$\mathbf{A} \vDash \varphi(\vec{a}) \iff \mathbf{B} \vDash \varphi(f(\vec{a})).$$

If such f exists, we say that $\mathbf{A}$ elementarily embeds into $\mathbf{B}$ and write $\mathbf{A} \hookrightarrow_e \mathbf{B}$.

Definition 2.42 (Elementary substructure)**.** A *substructure* $\mathbf{A}$ of a τ-structure $\mathbf{B}$ is called *elementary* if the inclusion map is elementary. We denote this by $\mathbf{A} \prec \mathbf{B}$.

Proposition 2.43 (Tarski–Vaught test)**.** *Let* $\mathbf{A}$ *be a substructure of a* τ*-structure* $\mathbf{B}$*.* $\mathbf{A}$ *is an elementary substructure of* $\mathbf{B}$ *if and only if for every formula* $\varphi(\vec{x}, y)$ *and* $\vec{a} \in A^n$*,*

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

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

$$\begin{array}{rrcl} [\text{elementarity}] & \mathbf{B} \vDash \exists y \varphi(\vec{a}, y) & \iff & \mathbf{A} \vDash \exists y \varphi(\vec{a}, y) \\ [\text{definition of } \vDash] & & \iff & \exists a' \in A \text{ such that } \mathbf{A} \vDash \varphi(\vec{a}, a') \\ [\text{elementarity}] & & \iff & \exists a' \in A \text{ such that } \mathbf{B} \vDash \varphi(\vec{a}, a'). \end{array}$$

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

$$\mathbf{A} \vDash \varphi(\vec{a}) \iff \mathbf{B} \vDash \varphi(\vec{a}).$$

Since Proposition 2.40 takes care of the atomic formulas and the cases $\varphi \doteq \neg\psi$ and $\varphi \doteq \psi_0 \wedge \psi_2$ are straightforward, so we only consider the case $\varphi(\vec{x}) \doteq \exists y \psi(\vec{x}, y)$. Fix $\vec{a} \in A^n$ and check:

$$\begin{array}{rrcl} [\text{Tarski–Vaught condition}] & \mathbf{B} \vDash \exists y \psi(\vec{a}, y) & \iff & \exists a' \in A \text{ such that } \mathbf{B} \vDash \psi(\vec{a}, a') \\ [\text{induction}] & & \iff & \exists a' \in A \text{ such that } \mathbf{A} \vDash \psi(\vec{a}, a') \\ [\text{definition of } \vDash] & & \iff & \mathbf{A} \vDash \exists y \psi(\vec{a}, y). \end{array}$$

□

Given a τ-structure $\mathbf{B}$ and $S \subseteq B$, we could define a substructure generated by S as the smallest substructure containing S mainly because intersection of substructures is still a substructure. However, intersection of elementary substructures may not be elementary, so we cannot define "the elementary substructure generated by S" like we did with substructures.

Let us re-examine how the substructure generated by S is produced: we have to throw in all the constants of $\mathbf{B}$ into S and close the resulting set S_1 under the functions of $\mathbf{B}$. The closing is an iterative process that one has to repeat $\aleph_0$-many[3] times, that is, assuming S_n is defined, let $S_{n+1} := S_n \cup \bigcup_{f \in \mathcal{F}_{\mathbf{B}}} f^{\mathbf{B}}[S_n]$, where $\mathcal{F}$ is the set of function symbols of τ and if $f \in \mathcal{F}_{\mathbf{B}}$ is a k-ary function, then $f^{\mathbf{B}}[S_n]$ stands for $f^{\mathbf{B}}(S_n^k)$. Finally put $A := \bigcup_{n=1}^{\infty} S_n$ and this A will be closed under all functions in $\mathcal{F}_{\mathbf{B}}$, hence will be a universe of a substructure. It is worth noting here that $|A| \leq \max(|S|, |\tau|, \aleph_0)$.

Now according to the Tarski–Vaught test, to be a universe of an *elementary* substructure, a set has to also contain witnesses to all formulas that claim existence of an object and are valid in $\mathbf{B}$. So in our procedure above, at every step, we have to additionally throw in these witnesses together with values of functions, and that's all. In fact, only throwing witnesses will also add the values of functions because for every k-ary function symbol f and $\vec{a} \in B^k$, the unique witness to the formula $\exists y f(\vec{a}) = y$ is exactly $f^{\mathbf{B}}(\vec{a})$. Same is true for constant symbols.

[3] $\aleph_0$ denotes the cardinality of $\mathbb{N}$.

The following theorem summarizes this discussion:

Theorem 2.44 (Löwenheim–Skolem). *Let* $\mathbf{B}$ *be a* τ*-structure and* $S \subseteq B$. *There exists* $\mathbf{A} \prec \mathbf{B}$ *with* $A \supseteq S$ *such that* $|A| \leq \max(|S|, |\tau|, \aleph_0)$.

Proof. We start by choosing witnesses for the formulas that claim existence of an object in B. For each τ-formula $\varphi(\vec{x}, y)$, where $\vec{x} = (x_1, ..., x_k)$, define a partial[4] function $f_{\varphi,k} : B^k \rightharpoonup B$ by $\vec{b} \mapsto$ one of the witnesses to $\mathbf{B} \vDash \exists y \varphi(\vec{b}, y)$ if such exist; more precisely, if $\mathbf{B} \vDash \exists y \varphi(\vec{b}, y)$ then $f_{\varphi,k}(\vec{b})$ is defined and is equal to one of the elements[5] $b' \in B$ for which $\mathbf{B} \vDash \varphi(\vec{b}, b')$. These $f_{\varphi,k}$ are often called *Skolem functions.*

Now we recursively construct an increasing sequence $(S_n)_{n \in \mathbb{N}}$ of subsets of B as follows: put $S_0 \coloneqq S$, and assuming S_n is defined, let

$$S_{n+1} \coloneqq S_n \cup \bigcup_{\varphi,k} f_{\varphi,k}(S_n^k).$$

Finally, let $A \coloneqq \bigcup_{n \in \mathbb{N}} S_n$ 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. In the definition of the Skolem functions above, it was possible that the same formula had multiple witnesses and we were free to choose any of them. Depending on this choice, the resulting substructure may be different (i.e. it is not canonical), and that is why there is no notion of "the elementary substructure generated by S".

3. 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 τ for this section and everything below is assumed to be in this signature.

3.A. **The axioms of $\mathbb{FOL}(\tau)$**

Unlike the definition of a τ-theory, the axioms of $\mathbb{FOL}(\tau)$ include formulas with free variables. Indeed, 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 3.1. 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 3.2. Below, whenever we write $\varphi(t/v)$, it is assumed that t is free for v in φ.

[4] A partial function $f : X \rightharpoonup Y$ is a function whose domain is a (possibly empty) subset of X.

[5] We are using the Axiom of Choice here.

Convention 3.3. From now on, we treat $\varphi \vee \psi; \varphi \wedge \psi; \exists v \varphi$ as abbreviations for $\neg\varphi \to \psi; \neg(\varphi \to \neg\psi); \neg\forall v \neg\varphi$.

The following are the *axioms* (or *axiom schemes*) of $\mathbb{FOL}(\tau)$.

Propositional axioms. For all τ-formulas φ, ψ, χ, we have:

Axioms for $\to$:

(1) $\varphi \to (\psi \to \varphi)$
(2) $(\varphi \to \psi) \to [(\varphi \to (\psi \to \chi)) \to (\varphi \to \chi)]$

Axioms for $\neg$:

(3) $(\varphi \to \psi) \to [(\varphi \to \neg\psi) \to \neg\varphi]$
(4) $\neg\neg\varphi \to \varphi$
(5a) $(\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$; (5b) $(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)$

Axioms for $\wedge$:

(6) $\varphi \to [\psi \to (\varphi \wedge \psi)]$
(7a) $(\varphi \wedge \psi) \to \varphi$; (7b) $(\varphi \wedge \psi) \to \psi$
(8) $(\varphi \wedge \neg\varphi) \to \psi$

Axioms for $\vee$:

(9) $(\varphi \to \chi) \to [(\psi \to \chi) \to ((\varphi \vee \psi) \to \chi)]$
(10a) $\varphi \to (\varphi \vee \psi)$; (10b) $\psi \to (\varphi \vee \psi)$

Quantifier axioms. For all τ-formulas φ, ψ, all τ-terms t, and all variables u, v, the following $\mathbb{FOL}(\tau)$ words are taken as axioms of $\mathbb{FOL}(\tau)$ whenever they form valid τ-formulas[6]:

(11) $\big(\forall v(\psi \to \varphi)\big) \to \big(\psi \to \forall u \varphi(u/v)\big)$ [whenever v does not occur in ψ]

Remark 3.4. What we really want to write here is

$$\big(\forall v(\psi \to \varphi)\big) \to \big(\psi \to \forall v \varphi(v)\big),$$

which is more readable and makes sense immediately, but this is not a (valid) formula according to Convention 2.18.

(12) Instantiation: $\forall v \varphi \to \varphi(t/v)$
(13) Generalization: $\varphi \to \forall u \varphi(u/v)$

Remark 3.5. Here again, 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 has content and will be used), $\varphi \to \forall v \varphi(v)$ is not be a (valid) formula according to Convention 2.18.

(14) $\exists$-elimination: $\big(\varphi \to \psi\big) \to \big((\exists v \varphi) \to \psi\big)$ [whenever v does not occur in ψ]

Remark 3.6. This axiom is used in proving a statement of the form $(\exists v \varphi) \to \psi$; for example, the statement "a convergent sequence (x_n) is bounded" is of this form as we can write

$$\big(\exists L \in \mathbb{R}, \lim_{n \to \infty} x_n = L\big) \to \big((x_n) \text{ is bounded}\big).$$

[6] This would be the case if we assume that u does not occur in either of φ and ψ, v does not occur in t, and t is free for v in φ.

In the course of the proof, we assume the hypothesis $\exists v\varphi$ in order to prove the conclusion ψ. To use the hypothesis, we need to get rid of (eliminate) the existential quantifier $\exists v$, which we do by saying "let v denote an object for which $\varphi(v)$ is true"; indeed, in our example, we would write "let L be the limit of (x_n)". Then we use the statement $\varphi(v)$ to prove ψ, so in the end, what we would actually proven is $\varphi(v) \to \psi$. But what we initially wanted was $(\exists v\varphi) \to \psi$, and the $\exists$-elimination axiom is what allows us to derive the latter statement from the former.

(15) Instance-to-existence: $\varphi(t/v) \to \exists v\varphi$

Equality axioms. For every $n \in \mathbb{N}$, each n-ary relation symbol R and n-ary function symbol f in τ, we have:

(16) $v = v$; $v = v' \to v' = v$; $(v = v' \wedge v' = v'') \to v = v''$
(17) $(\bigwedge_{i=1}^{n} v_i = w_i) \to (R(v_1, ..., v_n) \to R(w_1, ..., w_n))$
(18) $(\bigwedge_{i=1}^{n} v_i = w_i) \to (f(v_1, ..., v_n) = f(w_1, ..., w_n))$

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

Rule of inference. For every τ-formula φ, ψ, we have:

(19) Modus Ponens: $\varphi, \varphi \to \psi \implies \psi$

Definition 3.7. Let φ be a τ-formula and $\vec{v}$ be the vector of free variables of φ[7], so $\forall\vec{v}\varphi$ is a sentence. We say that φ is *satisfied/true* in a τ-structure $\mathbf{A}$, and write $\mathbf{A} \vDash \varphi$, if $\mathbf{A} \vDash \forall\vec{v}\varphi$.

The proof of the following lemma is an easy but tedious verification:

Lemma 3.8. *All of the axioms above are true in every τ-structure and Modus Ponens preserves the truth.*

3.B. **Formal proofs**

Definition 3.9. Let T be a theory and φ be a formula. A *proof* of φ from T is a finite sequence $\varphi_1, \varphi_2, ...\varphi_n$ of formulas such that $\varphi_n \doteq \varphi$ and for each i

either φ_i is an axiom of $\mathbb{FOL}(\tau)$,
or $\varphi_i \in T$,
or φ_i follows from the previous φ_j-s by Modus Ponens, i.e. for some $j, k < i$ (not necessarily $j < k$), $\varphi_k \doteq \varphi_j \to \varphi_i$; in this case, we say that φ_i is obtained by Modus Ponens from φ_j, φ_k.

Definition 3.10. We say that T *proves* φ, and write $T \vdash \varphi$, if there exists a proof of φ from T. When $T = \varnothing$, we just write $\vdash \varphi$.

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

Example 3.11. Here is a formal proof of $\theta \to \theta$ from the empty theory, for all formulas θ:

(i) $(\theta \to (\theta \to \theta)) \to [(\theta \to ((\theta \to \theta) \to \theta)) \to (\theta \to \theta)]$ [Axiom (2) for $\varphi \doteq \chi \doteq \theta$ and $\psi \doteq (\theta \to \theta)$],
(ii) $\theta \to (\theta \to \theta)$ [Axiom (1) for $\varphi \doteq \psi \doteq \theta$],
(iii) $(\theta \to ((\theta \to \theta) \to \theta)) \to (\theta \to \theta)$ [Modus Ponens (i), (ii)],

[7] We mean that $\vec{v}$ includes all of the free variables of φ and no other variables.

(iv) $\theta \to ((\theta \to \theta) \to \theta)$ [Axiom (1) for $\varphi \doteq \theta$ and $\psi \doteq (\theta \to \theta)$],
(v) $\theta \to \theta$ [Modus Ponens (iii), (iv)].

The following proposition justifies why we introduced a proof system and formal proofs:

Proposition 3.12 (Soundness)**.** *If $T \vdash \varphi$ then $T \vDash \varphi$.*

Proof. This follows by induction on the length of the formal proof of φ and Lemma 3.8. □

Lemma 3.13 (Deduction theorem)**.** *For a theory T, a sentence χ and a formula φ,*

$$T, \chi \vdash \varphi \iff T \vdash \chi \to \varphi.$$

Proof. $\Leftarrow$: Follows by an application of Modus Ponens.
$\Rightarrow$: Letting $\varphi_1, ..., \varphi_n$ with $\varphi_n \doteq \varphi$ be a proof of φ from $T \cup \{\chi\}$, we show that $T \vdash \chi \to \varphi$ by induction on n.

Case $n = 1$: $\varphi \doteq \varphi_1$ is an axiom of $\mathbb{FOL}(\tau)$ or is in T. Then $T \vdash \varphi$, and by Axiom (1), $T \vdash \varphi \to (\chi \to \varphi)$, so Modus Ponens gives $T \vdash \chi \to \varphi$.

Case $n = 1$: $\varphi \doteq \varphi_1 \doteq \chi$. Then $T \vdash \chi \to \varphi$ is the same as $T \vdash \varphi \to \varphi$, which is done in Example 3.11.

Case $n \Rightarrow n+1$: $\varphi \doteq \varphi_{n+1}$ is obtained by Modus Ponens. Then there are $i, j \leq n$ such that $\varphi_j \doteq \varphi_i \to \varphi$. By the inductive 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, we get $T \vdash \chi \to \varphi$. □

Let $\mathcal{S}$ be a set of symbols neither of which is in τ. Then we denote by $\tau(S)$ the extension of τ obtained by adding to it the symbols in S as constant symbols. If $S = \{s_1, ..., s_n\}$ is finite, we just write $\tau(s_1, ..., s_n)$.

Lemma 3.14 (Constant Substitution)**.** *Let c be a symbol that is not in τ and let u be free in a τ-formula λ. For a τ-theory T,*

$$T \vdash \lambda(c/u) \iff T \vdash \lambda,$$

where in the first statement T is viewed as a $\tau(c)$-theory.

Proof. The direction $\Leftarrow$ follows by applications of the generalization (13) and instantiation (12) 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 $\lambda(c/u)$ is an axiom, then so is λ, which is only worth checking for quantifier axioms. □

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

In this subsection, we define analogues of the notions defined in Subsection 2.E using $\vdash$ instead of $\vDash$.

Definition 3.15. A τ-theory T is said to be

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

Note that a satisfiable theory is consistent by the Soundness of the proof system. Also, any inconsistent theory is automatically complete because one can easily show that $\vdash \bot \to \varphi$ for any τ-formula φ. Clearly, $\mathrm{Th}(\mathbf{M})$ is complete for any τ-structure $\mathbf{M}$.

Lemma 3.16 (About consistency)**.** *Let T be a τ-theory.*

(a) T is consistent if and only if there is a sentence χ such that $T \nvdash \chi$.
(b) T is consistent if and only if every finite subset of T is consistent.
(c) For any sentence χ, $T \cup \{\chi\}$ is inconsistent if and only if $T \vdash \neg\chi$.
(d) If T is consistent, then for any sentence χ, at least one of $T \cup \{\chi\}$ and $T \cup \{\neg\chi\}$ is consistent.
(e) If $\exists v \varphi(v)$ is a sentence, $T \cup \{\exists v \varphi(v)\}$ is consistent, and c is a constant symbol that does not occur in $T \cup \{\exists v \varphi(v)\}$, then $T \cup \{\varphi(c)\}$ is consistent.

Proof. Part (a) just expresses the fact that once a theory proves a contradiction, then it proves every sentence. (b) follows from the fact that proofs are finite. We prove the rest in detail.

The right-to-left direction of (c) is immediate, and we show the other direction. Assume $T \cup \{\chi\}$ is inconsistent and hence $T, \chi \vdash \bot$. By the Deduction theorem (this is where we really need this theorem), $T \vdash \chi \to \bot$, and hence, by Axiom (4) and Modus Ponens, we get $T \vdash \top \to \neg\chi$. But $\top$ is an axiom (more precisely, it follows from Axioms (16) and (13), and Modus Ponens), so $T \vdash \top$, and hence by applying Modus Ponens again, we get $T \vdash \neg\chi$.

For (d), we prove the contrapositive. Assume both $T \cup \{\chi\}$ and $T \cup \{\neg\chi\}$ are inconsistent. Then by (c), $T \vdash \neg\chi$ and $T \vdash \neg\neg\chi$. Thus $T \vdash \chi \wedge \neg\chi$ and hence is inconsistent.

For (e), we also prove the contrapositive. Assume $T \cup \{\varphi(c)\}$ is inconsistent. Then by (c), $T \vdash \neg\chi(c)$. By the constant substitution lemma (Lemma 3.14), $T \vdash \neg\varphi(v)$, and by Axiom (13), $T \vdash \forall v \neg\varphi(v)$, so $T \vdash \neg\exists v \varphi(v)$. Thus, by (c), $T \cup \{\exists v \varphi(v)\}$ is inconsistent. □

Note the following "compactness" phenomenon: if $T \vdash \varphi$, then there is a finite $T_0 \subseteq T$ with $T_0 \vdash \varphi$. This is an immediate consequence of the fact that formal proofs are finite and hence they only use finitely many axioms from T. This "compactness" statement is actually equivalent to the fact that the following topological space is compact: let $\mathcal{T}$ be the set of all consistent fully 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 τ-sentences. The proof of the equivalence uses the following lemma and is left as an exercise.

Lemma 3.17. *Any (syntactically) consistent τ-theory T has a consistent full completion, i.e. there exists a consistent fully complete τ-theory $T' \supseteq T$.*

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

Proof for countable τ. In this case there are only countably many formulas, so we can enumerate all sentences $(\varphi_n)_{n\in\mathbb{N}}$. 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 $T_{n+1} := T_n \cup \{\varphi_n\}$ if $T_n \nvdash \neg\varphi_n$, and put $T_{n+1} := T_n \cup \{\neg\varphi_n\}$, otherwise. It follows from (c) of Lemma 3.16 that T_{n+1} is consistent. Finally, put $T' := \bigcup_n T_n$. Note that $T' \supseteq T$ and T' is consistent: indeed, if it was inconsistent, then, by (b) of Lemma 3.16 some finite subset $F \subseteq T'$ 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 T' is fully complete.

Proof for arbitrary τ. By (b) of Lemma 3.16, inconsistent theories have inconsistent finite subsets, so arbitrary increasing unions of consistent theories are consistent. Thus, by Zorn's lemma, there is a $\subseteq$-maximal consistent theory $T' \supseteq T$ and it remains to show that it is fully complete. Indeed, for any τ-sentence φ, one of $T' \cup \{\varphi\}$ or $T' \cup \{\neg\varphi\}$ is consistent by (d) of Lemma 3.16, so, by maximality, T' must already contain φ or $\neg\varphi$. □

4. Completeness of FOL and its consequences

Proposition 3.12 (the soundness of the proof system) says that if we have a "first order (finite) certificate" that something is true (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 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 this section is devoted.

4.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 and rules of inference that we threw in are enough to prove any statement that is semantically implied by T.

The following table compares the notions we have defined.

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

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

Remark 4.2 (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 FOL, the second means Incompleteness of PA (as a theory).

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

Corollary 4.3 (Syntactic-semantic duality)**.** *For a τ-theory T and a τ-sentence φ,*

$$T \vdash \varphi \iff T \vDash \varphi.$$

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

Proof. We only prove that $T \vDash \varphi$ 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 (c) of Lemma 3.16, so the Completeness theorem gives a model $\mathbf{M} \vDash T \cup \{\neg\varphi\}$, and hence, $T \nvDash \varphi$. □

Remark 4.4. If one somehow manages to prove a first-order statement φ about all models of T using methods from outside of FOL, the syntactic-semantic duality implies that there is a first-order proof of φ from T and using external methods was an overkill.

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

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.

4.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 τ 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 structure $\mathbf{A} := (A, \tau)$, 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 τ-structure $\mathbf{A}' := (C_A, \tau)$ isomorphic to $\mathbf{A}$, but the objects in the underlying set of $\mathbf{A}'$ 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 $\mathrm{Th}(\mathbf{A})$, we usually cannot recover the structure $\mathbf{A}$ even if we know the underlying set A. However, we can upgrade our signature τ so that we can by adding names for elements of A.

Definition 4.6. For a τ-structure $\mathbf{A}$, define a new signature

$$\tau_A := \tau \cup \{c_a : a \in A\},$$

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

Now the structure $\mathbf{A}' := (A, \tau_A)$ is such that every element in its underlying set A has a name in the signature τ_A, so $\mathrm{Th}(\mathbf{A}')$ will tell us exactly how the constant symbols, function symbols and relation symbols in τ are interpreted in $\mathbf{A}$; for example, if $a_1, a_2, a_3 \in A$ and $f^{\mathbf{A}'}(a_1, a_2) = a_3$, then $\mathrm{Th}(\mathbf{A}')$ would contain the τ_A-sentence $f(c_{a_1}, c_{a_2}) = c_{a_3}$; for groups this would correspond to the multiplication table. Moreover, $\mathrm{ElDiag}(\mathbf{A})$ also includes quantified statements about the elements of A. In particular, $\mathbf{A} \vDash \exists v \varphi$ if and only if $\exists v \varphi \in \mathrm{ElDiag}(\mathbf{A})$. Furthermore, the latter holds if and only if $\varphi(c/v) \in \mathrm{ElDiag}(\mathbf{A})$, for some constant symbol $c \in \tau_A$. 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 τ so they agree with T. Therefore, it would be really nice if T could tell us how exactly to define those interpretations because if we do exactly as T says, then we would naturally end up with a τ-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 τ-structure, so we can take that τ-structure as our model. The following definition makes all this precise.

Definition 4.7. 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

either $H \nvdash \exists v \varphi$,
or for some constant symbol $c \in \sigma$, $H \vdash \varphi(c/v)$.

A σ-theory H is called a *σ-Henkin theory* (or just a Henkin theory) if H is consistent, fully complete, and admits Henkin witnesses for every σ-formula of the form $\exists v \varphi$.

As expected, for a τ-structure $\mathbf{A}$, $\mathrm{ElDiag}(\mathbf{A})$ is an example of a Henkin theory (in the signature τ_A).

Note that the existence of a σ-Henkin theory implies that σ has at least one constant symbol. Our initial signature τ may not contain enough constants to be used as Henkin witnesses, so we artificially create them and throw them into τ; more precisely, letting $\kappa := \max\{|\tau|, \aleph_0\}$, we take a set $D := \{d_n\}_{n \le \kappa}$ of distinct constant symbols not in τ and put $\bar{\tau} := \tau \cup D$.

Lemma 4.8 (Constructing a Henkin theory)**.** *Any consistent τ-theory T admits a $\bar{\tau}$-Henkin extension $H \supseteq T$.*

Proof. We will prove this assuming τ is countable to make the exposition easier to understand for those readers who are not familiar with the ordinals and cardinals. However, the readers who are familiar are invited to prove this for general τ. Note that for countable τ, $\kappa = \aleph_0$, so $D := \{d_n\}_{n \in \mathbb{N}}$.

Since $\bar{\tau}$ is countable, there are exactly $\aleph_0$-many $\bar{\tau}$-sentences, so we enumerate them: $(\varphi_n)_{n \in \mathbb{N}}$. We emphasize, that these φ_n range over all $\bar{\tau}$-sentences, not just τ-sentences. Letting $\tau_n := \tau \cup \{d_i : i < n\}$, for $n \geq 0$, we recursively construct an increasing sequence $(H_n)_{n \in \mathbb{N}}$, where each H_n is a consistent fully τ_n-complete theory, as follows: let H_0 be a τ_0-completion of T and suppose that H_n is defined and satisfies the required conditions, i.e. it is a consistent fully τ_n-complete theory. If H_n is already a τ_n-Henkin theory, then let H_{n+1} be a consistent full τ_{n+1}-completion of H_n. Otherwise, let $m \in \mathbb{N}$ be the least index such that $\varphi_m \in H$ and $\varphi_m \doteq \exists v \varphi$ but H doesn't contain a Henkin witness for it. Because $\exists v \varphi \in H_n$ and H_n is consistent, so must be $H_n \cup \{\varphi(d_n)\}$ by (e) of Lemma 3.16 because d_n does not occur in H_n. Thus, we let H_{n+1} be a consistent full τ_{n+1}-completion of $H_n \cup \{\varphi(d_n)\}$, and this finishes the construction of the sequence $(H_n)_{n \in \mathbb{N}}$. Finally, taking $H := \bigcup_{n \in \mathbb{N}} H_n$, we leave it as an exercise to verify that H is a $\bar{\tau}$-Henkin theory. □

Having constructed a $\bar{\tau}$-Henkin theory H, we now construct a model of H, i.e. a $\bar{\tau}$-structure satisfying H and then take its reduct to the signature τ (i.e. forget the names of Henkin witnesses).

Lemma 4.9. *Let σ be a signature and H a σ-Henkin theory. For any σ-term t with no variables, there is a constant symbol $c \in \sigma$ with $t = c \in H$.*

Proof. We aim at getting such $c \in \sigma$ as a Henkin witness to $\exists v \varphi(v)$, where $\varphi(v) \doteq t = v$, so it is enough to show that $H \vdash \exists v \varphi$. But, by Axioms (16), (13), and (12), $H \vdash t = t$, and the latter sentence is precisely $\varphi(t/v)$. Thus, $H \vdash \varphi(t/v)$, so using Axiom (15), we get $H \vdash \exists v \varphi$. □

Lemma 4.10 (Constructing a model for a Henkin theory)**.** *If H is a Henkin theory in a signature σ, then it has a model. In fact, it has a model whose cardinality is at most the cardinality of the set of constants in σ.*

Proof. As our first attempt, we take the set of constant symbols C of σ as the universe of our future model $\mathbf{C}$ with the following interpretations: for all $e_1, ..., e_n, e \in C$,

$$\begin{array}{rcll} c^{\mathbf{C}} & = & c, & \text{for every constant symbol } c \text{ in } \sigma \\ R^{\mathbf{C}}(e_1, ..., e_n) & \iff & R(e_1, ..., e_n) \in H, & \text{for every } n\text{-ary relation symbol } R \text{ in } \sigma \\ f^{\mathbf{C}}(e_1, ..., e_n) = e & \iff & f(e_1, ..., e_n) = e \in H, & \text{for every } n\text{-ary function symbol } f \text{ in } \sigma. \end{array}$$

This construction almost works except that it may well be that $c = c' \in H$, for distinct constant symbols c and c' in C. Because of this, $\mathbf{C}$ is not even a σ-structure since the last clause defines a multi-valued function. Even if we managed to choose a single valued branch for $f^{\mathbf{C}}$, $\mathbf{C}$ would still not be a model of H because it would not satisfy $c = c'$. So what we do is we mod out C by the equivalence relation $c = c' \in H$. More precisely, for all $c, c' \in C$, define

$$c \sim c' \iff c = c' \in H.$$

It follows from Axioms (12) for equality that $\sim$ is an equivalence relation on C.

Put $M := C/\sim$, so $M = \{[c] : c \in C\}$, where $[c]$ denotes the equivalence class of c. We define a σ-structure $\mathbf{M}$ with universe M and the following interpretations: for all $e_1, ..., e_n, e \in C$,

$$\begin{array}{rcll} c^{\mathbf{M}} & = & [c], & \text{for every constant symbol } c \text{ in } \sigma \\ R^{\mathbf{M}}([e_1], ..., [e_n]) & \iff & R(e_1, ..., e_n) \in H, & \text{for every } n\text{-ary relation symbol } R \text{ in } \sigma \\ f^{\mathbf{M}}([e_1], ..., [e_n]) = e & \iff & f(e_1, ..., e_n) = e \in H, & \text{for every } n\text{-ary function symbol } f \text{ in } \sigma. \end{array}$$

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

Proof of Claim. One has to prove that the definitions of $R^{\mathbf{M}}$ and $f^{\mathbf{M}}$ do not depend on the choice of the representatives of the equivalence classes, but this immediately follows from Axioms (17) and (18) of $\mathbb{FOL}(\bar{\tau})$. Moreover, for f as above, one has to verify that for all $e_1, ..., e_n \in C$, there exists $e \in C$ such that $f(e_1, ..., e_n) = e \in H$, but this is an instance of Lemma 4.9. ⊣

Claim 2. For every σ-term t with no variables and $c \in C$, $t^{\mathbf{M}} = [c]$ if and only if $t = c \in H$.

Proof of Claim. We do induction on the construction (length) of t. The case of t being a variable is excluded, so the only base case is $t \doteq e$, for a constant symbol $e \in \sigma$ follows from the definitions of $\sim$ and M.

Now assume that $t \doteq f(t_1, ..., t_n)$. Let $c_1, ..., c_n \in \sigma$ such that $t_i^{\mathbf{M}} = [c_i]$. By induction, we have that $t_i = c_i \in H$, so by Axiom (18), we also have that $H \vdash f(t_1, ..., t_n) = f(c_1, ..., c_n)$. But then, the definition and well-definedness of $\mathbf{M}$ gives

$$f^{\mathbf{M}}([c_1], ..., [c_n]) = [c] \iff f(c_1, ..., c_n) = c \in H,$$

and hence, $H \vdash f(t_1, ..., t_n) = c$, by Axiom (16). ⊣

Claim 3. $\mathbf{M} \vDash H$.

Proof of Claim. We show that for every σ-formula φ and $c_1, ..., c_n \in C$,

$$\mathbf{M} \vDash \varphi([c_1], ..., [c_n]) \iff \varphi(c_1, ..., c_n) \in H,$$

by structural induction on the construction of φ. The case of equality is handled by the previous claim, and the case of a relation symbol follows from the same claim and the definition of $\mathbf{M}$ using Axioms (16) and (17). The case of $\neg$ follows easily from the induction hypothesis and the consistency and full completeness of H, while the case of $\wedge$ follows from the induction hypothesis and the $\mathbb{FOL}(\bar{\tau})$ axioms for $\wedge$. We now handle the remaining case of $\varphi([c_1], ..., [c_n]) \doteq \exists v \psi([c_1], ..., [c_n], v)$ as follows:

$$\begin{aligned} \mathbf{M} \vDash \varphi([c_1], ..., [c_n]) &\iff \text{there is } [b] \in M \text{ such that } \mathbf{M} \vDash \psi([c_1], ..., [c_n], [b]) \\ \text{[by induction]} &\iff \text{there is } b \in C \text{ such that } \psi(c_1, ..., c_n, b) \in H \\ &\iff \exists v \psi(c_1, ..., c_n, v) \in H, \end{aligned}$$

where in the last equivalence, $\Longrightarrow$ is by Axiom (15) and $\Longleftarrow$ is because H admits a Henkin witness for $\exists v \psi(c_1, ..., c_n, v)$. ⊣

The last claim finishes the proof of the lemma. □

Proof of the Completeness Theorem 4.1 (Henkin, 1949). By Lemma 4.8, there is a $\bar{\tau}$-Henkin theory $H \supseteq T$. Now applying Lemma 4.10 to $\sigma := \bar{\tau}$ and H, we get a model $\mathbf{M}$ of H of cardinality at most $|\sigma|$ and hence at most $\kappa := \max\{|\tau|, \aleph_0\}$. Finally, take the reduct of $\mathbf{M}$ to the signature τ. □

From now on, we will not differentiate between the syntactic and semantic notions in the table in Subsection 4.A above.

4.C. **The Skolem "paradox"**

The Completeness theorem has the following striking consequence: if ZFC is consistent (which we really hope it is), then it has a countable model. This is maybe strange because that countable model $\mathbf{M}$ believes that there is an uncountable set since Cantor's theorem that $\mathbb{R}$ is uncountable is true in $\mathbf{M}$. Does this imply that ZFC is inconsistent?

The answer is of course NO and here are the two reasons why (the main reason is (2)):

(1) It may well be that $M = \mathbb{N}$ with a binary relation $\in^{\mathbf{M}}$ defined on it. So what if somehow $\mathbf{M}$ satisfies the statement that reads "there is an uncountable set"? It is just some statement about this binary relation $\in^{\mathbf{M}}$ 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 $\in^{\mathbf{M}}$ was the true $\in$, then the countability of M would simply imply that $\mathbf{M}$'s version of the real numbers, $\mathbb{R}^{\mathbf{M}}$, is indeed countable (for us), i.e. there is a bijection of $\mathbb{R}^{\mathbf{M}}$ with $\mathbb{N}$. This bijection is a set (any function is a set of pairs), but it may not be an element of M. In fact, since $\mathbf{M}$ satisfies the statement "$\mathbb{R}^{\mathbf{M}}$ is uncountable", we conclude that NO bijection of $\mathbb{R}^{\mathbf{M}}$ with $\mathbb{N}$ is an element of M. In other words, M does not "see" the countability of $\mathbb{R}^{\mathbf{M}}$ and thus thinks that $\mathbb{R}^{\mathbf{M}}$ 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 $\mathbf{M}$ never will.

4.D. **Upward Löwenheim–Skolem theorem**

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

Theorem 4.11 (Upward Löwenheim–Skolem, weak version)**.** *If a τ-theory T has an infinite model, then it has a model of any cardinality $\kappa \geq \max\{|\tau|, \aleph_0\}$.*

Proof. Put $\bar{\tau} \coloneqq \tau \cup \{c_\alpha\}_{\alpha<\kappa}$, where c_α are constant symbols that are not in τ. Define

$$T' \coloneqq T \cup \{c_\alpha \neq c_\beta : \alpha \neq \beta, \alpha, \beta < \kappa\}.$$

T' is finitely satisfiable since it has an infinite model. Thus, by the Compactness theorem, T' has a model $\mathbf{M}$ of cardinality at most κ since $|\bar{\tau}| = \kappa \geq \aleph_0$. On the other hand, $|M| \geq \kappa$ since $c_\alpha^{\mathbf{M}} \neq c_\beta^{\mathbf{M}}$ for distinct $\alpha, \beta < \kappa$. Thus $|M| = \kappa$. □

This theorem implies for example that PA has uncountable models!

Recall that the Löwenheim–Skolem theorem gave us an elementary substructure $\mathbf{A}$ of a given τ-structure $\mathbf{B}$ of any cardinality $\kappa \leq |B|$ as long as $\kappa \geq \max\{|\tau|, \aleph_0\}$. We would like to also get an upward version of this, i.e. start with a τ-structure $\mathbf{A}$ and get an elementary extension $\mathbf{B} \succ \mathbf{A}$ of any cardinality $\geq \max\{|A|, |\tau|, \aleph_0\}$. To achieve this, we may consider applying the previous theorem to $\mathrm{Th}(\mathbf{A})$. However, this would only give us a structure $\mathbf{B}$ that is elementarily equivalent to $\mathbf{B}$, i.e. $\mathbf{A} \equiv \mathbf{B}$, whereas we want $\mathbf{A} \hookrightarrow_e \mathbf{B}$. So instead, we apply the previous theorem to the elementary diagram $\mathrm{ElDiag}(\mathbf{A})$ of $\mathbf{A}$ (see Definition 4.6), and the following lemma tells us why.

Lemma 4.12. *For τ-structures $\mathbf{A}, \mathbf{B}$, if an expansion $\mathbf{B}'$ of $\mathbf{B}$ is a model of* $\mathrm{ElDiag}(\mathbf{A})$*, then $\mathbf{A} \hookrightarrow_e \mathbf{B}$. In particular, there is an isomorphic copy of $\mathbf{B}$ containing $\mathbf{A}$ as an elementary substructure.*

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

Theorem 4.13 (Upward Löwenheim–Skolem)**.** *Any infinite τ-structure $\mathbf{A}$ has an elementary extension of any cardinality $\kappa \geq \max\{|A|, |\tau|, \aleph_0\}$; more precisely, there is a τ-structure $\mathbf{B}$ such that $|B| = \kappa$ and $\mathbf{A} \prec \mathbf{B}$.*

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

4.E. Nonstandard models of arithmetic

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

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

For each $n \in \mathbb{N}$, recursively define a τ_{arthm}-term $\Delta(n)$ as follows:

$$\begin{cases} \Delta(0) \doteq 0 \\ \Delta(n+1) \doteq S(\Delta(n)) \end{cases}.$$

Note that for every $n \in \mathbb{N}$, $\mathbf{N} \models \Delta(n) = n$ and hence $\mathbb{N} = \{\Delta(n)^{\mathbf{N}} : n \in \mathbb{N}\}$.

Proposition 4.14. *The theory $\mathrm{Th}(\mathbf{N})$, and hence also PA, admits a countable nonstandard model.*

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

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

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 $\mathbf{N}$ to a σ-structure with w being interpreted as $n+1$ satisfies T_0. Thus, by the Compactness theorem, T has a countable model $\mathbf{M}$.

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

4.F. From finite to infinite and back

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

4.F.1. *From finite to infinite.* 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 such is the entire graph. Similarly, if every finite subgraph admits a perfect matching, then so does the entire subgraph. In both of these statements, P is of the form $\exists$ a relation R with property Q. Proofs of such statements involve giving names to elements of the structure (i.e. adding constant symbols to the signature) and to the relation we are after (i.e. adding a relation symbol R to the signature). Then, as long as Q is a first-order expressible property, we can write a sentence for each finite subset F of our structure that states that R has the property Q on F. The resulting theory would be finitely satisfiable by our hypothesis, and thus satisfiable by the Compactness theorem. This yields a relation R with the property Q being satisfied on the elements of our original structure.

In other words, the Compactness theorem allows a switch of quantifiers: from $\forall\exists$ to $\exists\forall$. Indeed, we are given that *for all* finite subsets F *there is* a certain object R that "works" for F, and what we get is that *there is* a certain object R that "works" *for all* F at once.

4.F.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, no 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 S, let $[S]^2$ denote the set of two element subsets of S (think of it as the set of edges of the undirected complete graph on S). Given a 2-coloring of $[\mathbb{N}]^2$, i.e. a function $c : [\mathbb{N}]^2 \to \{0,1\}$, a set $E \subseteq [\mathbb{N}]^2$ is said to be monochromatic if all elements of E have the same color, i.e. $c\restriction_E$ is constant. A set $A \subseteq \mathbb{N}$ is called monochromatic if $[A]^2$ is monochromatic.

Theorem 4.15 (Infinite Ramsey)**.** *For any* 2*-coloring of* $[\mathbb{N}]^2$*, there exists an infinite monochromatic subset of* $\mathbb{N}$*.*

Proof. For $a \in \mathbb{N}$ and $A \subseteq \mathbb{N}$, put $(a, A) := \{\{a, a'\} : a' \in A \setminus \{a\}\}$. Set $A_0 := \mathbb{N}$ and take sequences $a_n \in \mathbb{N}$ and $A_n \subseteq \mathbb{N}$ satisfying:

(i) $a_n \in A_n$,
(ii) $A_{n+1} \subseteq A_n$ is infinite and (a_n, A_{n+1}) is monochromatic.

It is easy to see that such sequences $(a_n)_{n\in\mathbb{N}}$ and $(A_n)_{n\in\mathbb{N}}$ exist (define them recursively). Call a_n red if all elements of (a_n, A_{n+1}) have color 0, otherwise call it blue. Clearly, there is a subsequence $(a_{n_k})_{k\in\mathbb{N}}$ with all a_{n_k} having the same color (red or blue). Now it is straightforward to check that $A := \{a_{n_k}\}_{k\in\mathbb{N}}$ is monochromatic. □

Example 4.16. 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).

Let $\bar{n} := \{0, 1, ..., n-1\}$.

Theorem 4.17 (Finite Ramsey)**.** *For every* $m \in \mathbb{N}$*, there exists* $n \in \mathbb{N}$ *such that for any* 2*-coloring of* $[\bar{n}]^2$*, there exists a monochromatic subset* $A \subseteq \bar{n}$ *of cardinality* m*.*

Proof. Let τ 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 $R(x, y)$ and 0 otherwise). Fix $m \in \mathbb{N}$, and for each $n \in \mathbb{N}$, let φ_n be a τ-sentence expressing that $c_0, c_1, ..., c_{n-1}$ are pairwise distinct and the set $\{c_0, c_1, ..., c_{n-1}\}$ does not have a monochromatic subset of cardinality m (there are only finitely many such subsets, so we can express it).

Now assume for contradiction that for any n, 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 $\mathbf{M}$. Let $C := \{c_n^{\mathbf{M}} : n \in \mathbb{N}\}$. By Infinite Ramsey theorem, C has an infinite monochromatic subset A, i.e. either for all distinct $a, a' \in A$, $R^{\mathbf{M}}(a, a')$ or for all distinct $a, a' \in A$, $\neg R^{\mathbf{M}}(a, a')$. Let n be large enough so that $A \cap \{c_i : i < n\}$ has at least m elements. Then it is clear that $\mathbf{M} \nvDash \varphi_n$, a contradiction. □

4.G. **Nonaxiomatizable classes**

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

Proposition 4.18. *Let $\mathcal{C}$ be a class of τ-structures. If the cardinalities of the structures in $\mathcal{C}$ are bounded, 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 4.11. □

Example 4.19. *Cyclic groups.* By the last proposition, the class of cyclic groups is not axiomatizable.

Let τ be a signature and $\vec{x}$ a vector of d variables.

Notation 4.20. A set S of τ-formulas, we write $T(\vec{x})$ to mean that the free variables of each formula $\varphi \in T$ are among $\vec{x}$ (so it makes sense to write $\varphi(\vec{x})$). Also, put

- $\exists\vec{x}T(\vec{x}) := \{\exists\vec{x}\varphi(\vec{x}) : \varphi(\vec{x}) \in T(\vec{x})\}$,
- $\forall\vec{x}T(\vec{x}) := \{\forall\vec{x}\varphi(\vec{x}) : \varphi(\vec{x}) \in T(\vec{x})\}$,
- $\neg T(\vec{x}) := \{\neg\varphi(\vec{x}) : \varphi(\vec{x}) \in T(\vec{x})\}$.

Lastly, if $T(\vec{x})$ is finite, put

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

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

Proposition 4.21. *Let $\mathcal{C}$ be a class of τ-structures defined as follows: for some τ-theory T' and set $T(\vec{x})$ of τ-formulas, we have that for every τ-structure $\mathbf{A}$,*

$$\mathbf{A} \in \mathcal{C} \iff \mathbf{A} \vDash T' \text{ and } (\forall\vec{a} \in A^d)(\exists\varphi \in T)\ \mathbf{A} \vDash \varphi(\vec{a}). \tag{4.22}$$

Then $\mathcal{C}$ is not axiomatizable, unless for some finite $T_0 \subset T$, the theory

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

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 d-many new constant symbols and note that the theory

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

is finitely satisfiable; indeed, otherwise, by (c) of Lemma 3.16, for some finite $T_0 \subseteq T$, $S \vdash \bigvee T_0(\vec{c})$, and hence $S \vdash \forall\vec{x} \bigvee T_0(\vec{x})$ by the Constant Substitution Lemma 3.14 and Generalization Axiom (13). Since every model of $T' \cup \left\{\forall\left(\vec{x} \bigvee T_0(\vec{x})\right)\right\}$ is in $\mathcal{C}$, it follows that $T' \cup \big\{\forall \vec{x}\big(\bigvee T_0(\vec{x})\big)\big\}$ 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 (4.22), a contradiction. □

Examples 4.23.

(a) *Nonbipartite graphs.* Let $T := \{\varphi_n : n \in \mathbb{N} \text{ odd}\}$, where φ_n expresses that there is a cycle of length n. Clearly, for a graph $\mathbf{\Gamma} := (\Gamma, E)$,

$$\mathbf{\Gamma} \text{ is nonbipartite } \iff \mathbf{\Gamma} \models \text{GRAPHS and } (\exists n \in \mathbb{N})\ \mathbf{\Gamma} \models \varphi_n,$$

and the hypothesis of Proposition 4.21 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 $\mathbf{\Gamma} := (\Gamma, E)$,

$$\mathbf{\Gamma} \text{ is connected } \iff \mathbf{\Gamma} \models \text{GRAPHS and } (\forall u,v \in \Gamma)(\exists n \in \mathbb{N})\ \mathbf{\Gamma} \models \varphi_n(u,v),$$

and the hypothesis of Proposition 4.21 is met, so this class is not axiomatizable.

Proposition 4.24. *Let $\mathcal{C}$ be a class of τ-structures defined as follows: for some set $T(\vec{x})$ of τ-formulas, we have that for every τ-structure $\mathbf{A}$,*

$$\mathbf{A} \in \mathcal{C} \iff (\exists \vec{a} \in A^d)(\forall \varphi \in T)\ \mathbf{A} \models \varphi(\vec{a}). \tag{4.25}$$

Then, for any τ-sentence χ, if every structure $\mathbf{A} \in \mathcal{C}$ satisfies χ, then there is finite $T_0 \subseteq T$ with $\exists \vec{x}\big(\bigwedge T_0(\vec{x})\big) \models \chi$. In particular, $\mathcal{C}$ is not axiomatizable, unless the theory

$$\big\{\exists \vec{x}\big(\bigwedge T_0(\vec{x})\big) : \textit{finite } T_0 \subseteq T\big\}$$

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 d-many new constant symbols. Note that, by (4.25), the τ-reducts of models of $T(\vec{c})$ are in $\mathcal{C}$, so in particular, they all satisfy χ, and hence $T(\vec{c}) \models \chi$. By compactness, there is a finite $T_0(\vec{c}) \subseteq T(\vec{c})$ with $T_0(\vec{c}) \models \chi$, so $\models \big(\bigwedge T_0(\vec{c})\big) \to \chi$. By the Constant Substitution Lemma 3.14 and Exists Elimination Axiom (14), we get $\models \big(\exists \vec{x} \bigwedge T_0(\vec{x})\big) \to \chi$. □

Example 4.26. *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 at most n. Clearly, for a graph $\mathbf{\Gamma} := (\Gamma, E)$,

$$\mathbf{\Gamma} \text{ is disconnected } \iff (\exists u,v \in \Gamma)\ (\forall \varphi \in \text{GRAPHS} \cup T)\ \mathbf{\Gamma} \models \varphi(u,v),$$

and the hypothesis of Proposition 4.24 is met, so this class is not axiomatizable.

5. 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 $\text{Th}(\mathbf{N})$ 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 $\text{Th}(\mathbf{N})$ or not. For example, is the Twin Prime

Conjecture in $\mathrm{Th}(\mathbf{N})$? 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 "easy-to-verify" means 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 Examples 2.31 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.

5.A. **The Łoś–Vaught test**

Definition 5.1. Let κ be a cardinal. A τ-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 5.2.

(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 $\tau := (<)$:
(i) Antireflexivity: $\forall x (x \nless x)$
(ii) Antisymmetry[8]: $\forall x \forall y (x < y \to y \nless x)$
(iii) Transitivity: $\forall x \forall y \forall z \big[(x < y \wedge y < z) \to x < z \big]$
(iv) Linearity: $\forall x \forall y \big[(x \neq y \wedge x \nless y) \to y < x \big]$
(v) Density: $\forall x \forall y \big[x < y \to \exists z (x < z < y) \big]$
(vi) 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 τ-structure $\mathbf{A}$, $\mathrm{Th}(\mathbf{A})$ is *absolutely categorical*, i.e. any two models $\mathbf{B}, \mathbf{B}' \vDash \mathrm{Th}(\mathbf{A})$ are isomorphic.

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

Proof. Let $\mathbf{A}, \mathbf{B} \vDash T$ and we need to show that $\mathbf{A} \equiv \mathbf{B}$, by Proposition 2.38. Since $\mathbf{A}$ and $\mathbf{B}$ are infinite, we can apply the weak version of Löwenheim–Skolem (4.11) and get $\mathbf{A}' \vDash \mathrm{Th}(\mathbf{A})$

[8] This is redundant since it follows from antireflexivity and transitivity.

and $\mathbf{B}' \models \mathrm{Th}(\mathbf{B})$ such that $|A'| = \kappa = |B'|$. Because T is κ-categorical, $\mathbf{A}' \cong \mathbf{B}'$ and hence $\mathbf{A}' \equiv \mathbf{B}'$. Thus, $\mathbf{A} \equiv \mathbf{A}' \equiv \mathbf{B}' \equiv \mathbf{B}$. □

This immediately gives that the theories $\mathrm{VEC}_{\mathbb{Q}}$ and DLO are complete.

One cannot help mentioning the following very important theorem that started the modern model theory:

Theorem (Morley, 1965)**.** *Let T be a theory in a countable signature τ. 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.

Lastly, we would like to 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:

Vaught's conjecture. Let τ be a countable signature and T be a complete τ-theory having infinite models. If T has uncountably many nonisomorphic countable models, does it have continuum many nonisomorphic countable models?

5.B. Algebraically closed fields and the Lefschetz Principle

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

Lemma 5.4. *Every algebraically closed field is infinite.*

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

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 if (s)he does not feel like remembering field theory.

Proposition 5.5. *For p prime or 0, ACF_p is κ-categorical for any uncountable cardinal κ.*

Proof. Let $\mathbf{K}_1, \mathbf{K}_2 \models \mathrm{ACF}_p$ with $|K_1| = |K_2| = \kappa$. For $i = 1, 2$, let F_i be the base field of $\mathbf{K}_i$, i.e. the substructure of $\mathbf{K}_i$ generated by $\varnothing$. (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 transcendence basis over F[9] in $\mathbf{K}_i$. Now it is not hard to see that $K_i = \overline{F(B_i)}$, where $F(B_i)$ denotes the field generated by B_i over F and $\overline{F(B_i)}$ denotes its algebraic closure in K_i.

Thus, because F is countable, $|K_i| = |B_i| \cdot \aleph_0$. If B_i is countable then so is $|B_i| \cdot \aleph_0$, but K_i is uncountable, and hence B_i is uncountable. Then, by basic cardinal arithmetic, $|B_i| \cdot \aleph_0 = |B_i|$, so $\kappa = |K_i| = |B_i|$. Hence, there is a bijection $f : B_1 \xrightarrow{\sim} B_2$. This f uniquely extends to an

[9] *A transcendence basis over F* is a maximal collection of algebraically independent elements over F.

isomorphism of $F(B_1)$ onto $F(B_2)$, which in its turn extends (not necessarily uniquely) to an isomorphism of $K_1 = \overline{F(B_1)}$ onto $K_2 = \overline{F(B_2)}$. □

Corollary 5.6. ACF_p *is complete, for any prime p and for $p = 0$.*

Proof. Follows from the Łoś–Vaught test (5.3) and 5.5, 5.4. □

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

(1) $\mathbf{C} \vDash \varphi$.
(2) $\mathbf{K} \vDash \varphi$, *for some* $\mathbf{K} \vDash \mathrm{ACF}_0$.
(3) $\mathrm{ACF}_0 \vDash \varphi$.
(4) For sufficiently large primes p, $\mathrm{ACF}_p \vDash \varphi$.
(5) For infinitely many primes p, there is $\mathbf{K} \vDash \mathrm{ACF}_p$ *such that* $\mathbf{K} \vDash \varphi$.

Proof. (1) $\iff$ (2) $\iff$ (3): Follows from the completeness of ACF_0.
(3) $\implies$ (4): $\mathrm{ACF}_0 \vDash \varphi$ implies $\mathrm{ACF}_0 \vdash \varphi$ by the Completeness theorem. Hence, because proofs are finite, there is a finite $T \subseteq \mathrm{ACF}_0$ such that $T \vdash \varphi$. But then, by the definitions of ACF_0 and ACF_p, for sufficiently large prime p, $T \subseteq \mathrm{ACF}_p$. Thus $\mathrm{ACF}_p \vdash \varphi$ and hence $\mathrm{ACF}_p \vDash \varphi$.

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

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

Corollary 5.8 (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 τ_{ring}-sentence $\phi_{n,d}$, and hence, instead of proving it for the field $\mathbb{C}$, by the Lefschetz principle, it is enough to prove $\phi_{n,d}$ 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 a more or less straightforward exercise in algebra to check that $\overline{\mathbb{F}}_p$ is an increasing union over $k \geq 1$ of the finite fields $\mathbb{F}_{p^k}$ of p^k elements (unique up to isomorphism). Thus, letting $k_0 \geq 1$ be large enough so that all of the coefficients involved in the definition of f are in $\mathbb{F}_{p^{k_0}}$, we can write:

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

But then, because $\mathbb{F}_{p^k}$ is a field and the definition of f only uses field operations and elements of $\mathbb{F}_{p^k}$, $f(\mathbb{F}_{p^k}^n) \subseteq \mathbb{F}_{p^k}^n$, for all $k \geq k_0$. Because f is injective, the Pigeonhole Principle (yay!) gives $f(\mathbb{F}_{p^k}^n) = \mathbb{F}_{p^k}^n$, so

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

□

5.C. **Reducts of arithmetic**

Definition 5.9. Let T be a τ-theory. A τ-theory T' is called an axiomatization for T if for all τ-sentences,

$$T \vdash \tau \iff T' \vdash \varphi.$$

PA was constructed as an attempt to "conveniently" axiomatize $\mathrm{Th}(\mathbf{N})$, where "convenient" means that there is a computer program recognizing the axioms (we will make this more in the next section). However, as we will see, Gödel's Incompleteness theorem states that PA is incomplete. In fact, there is no convenient axiomatization for $\mathrm{Th}(\mathbf{N})$, i.e. any subtheory $T \subseteq \mathrm{Th}(\mathbf{N})$ is either incomplete or inconvenient.

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

We start with $\mathbf{N}_S \coloneqq (\mathbb{N}, 0, S)$. Let $\tau_S \coloneqq (0, S)$. Here is our first (and last) attempt of axiomatizing $\mathrm{Th}(\mathbf{N}_S)$. 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: $\forall x \forall y (S(x) = S(y) \to x = y)$.
(S3) Any nonzero number is a successor of something: $\forall x (x \neq 0 \to \exists y (x = S(y)))$.
(S4) For all $n \in \mathbb{N}$, there are no n-loops: $\forall x (S^n(x) \neq x)$, where S^n stands for the n-fold composition of S.

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 $\mathbf{M}$ of T_S has a standard part $\bar{\mathbb{N}} \coloneqq \{\Delta(n)^{\mathbf{M}} : n \in \mathbb{N}\}$, where $\Delta(n) \coloneqq S^n(0)$. Define a binary relation $\sim$ on M as follows: for all $a, b \in M$,

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

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

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

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

Proposition 5.10. *T_S is κ-categorical, for any uncountable cardinal κ.*

Proof. Let $\mathbf{A}, \mathbf{B} \models T_S$ with $|A| = |B| = \kappa$. By above, $\lambda_{\mathbf{A}} = |A| = \kappa = |B| = \lambda_{\mathbf{B}}$. Thus, there is a bijection $f : \Lambda_{\mathbf{A}} \to \Lambda_{\mathbf{B}}$. Now the standard parts of $\mathbf{A}$ and $\mathbf{B}$ are clearly isomorphic. Moreover, any $\mathbb{Z}$-chain $C \in \Lambda_{\mathbf{A}}$ 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 of $\mathbf{A}$ onto $\mathbf{B}$. □

From this and the Łoś–Vaught test, we get

Corollary 5.11. *T_S is complete.*

Now we turn to $\mathbf{N}_+ \coloneqq (\mathbb{N}, 0, S, +)$. Let $\tau_+ \coloneqq (0, S, +)$ and let T_+ be the theory consisting of all of the axioms of PA except for the ones involving multiplication (hence it is a convenient theory). The proof of the following theorem will be omitted since it uses the technique of quantifier elimination, which is not covered in these notes.

Theorem 5.12 (Presburger, 1929). *T_+ is complete.*

Thus, as we will see, the incompleteness phenomenon starts with $\mathbf{N} \coloneqq (\mathbb{N}, 0, S, +, \cdot)$.

6. Incomplete theories

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

Definition 6.1 (Informal). A τ-theory T is called *recursive* if there is a computer program such that given a τ-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 primitive recursive axiomatizations. However, the situation changes once we add multiplication because it enables prime numbers and makes it possible to code tuples of natural numbers into a single number, and we have the following ground-breaking theorem:

Theorem 6.2 (Incompleteness; Gödel, 1931). *Any recursive theory $T \subseteq \mathrm{Th}(\mathbf{N})$ 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.

6.A. Sketch of proof of the Incompleteness theorem

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 will give a more rigorous version of this proof later, when we develop the basics of recursion theory.

Definition 6.3 (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 so is its indicator function.

First thing one shows is that recursive functions are arithmetical. Thus, any function we can write a computer program for is expressible in the language of arithmetic.

For a finite signature τ, whose symbols are $s_0, ...s_n$ we enumerate the symbols of $\mathbb{FOL}(\tau)$ as follows:

$$s_0 \; s_1 \; ... \; s_n \;\; = \neg \;\wedge\; \vee \;\to\; \forall \;\exists \;,\; (\;)\; v_0,\; v_1,\; v_2,\; ...$$

and call the index of a symbol its *code*. For example, the code of s_0 is 0, the code of $=$ is $n+1$ and the code of v_i is $n+11+i$. Using prime numbers and the fact that prime number factorization is unique, we can code a tuple of natural numbers into a single natural number ($\langle n_1, ..., n_k \rangle \coloneqq p_1^{n_1+1} \cdot ... \cdot p_k^{n_k+1}$), and so we can code formulas since they are just tuples of symbols of $\mathbb{FOL}(\tau)$. In fact, we can make sure that the coding and decoding operations are recursive (think of computer programs that would do this).

Thus let $\ulcorner t \urcorner$ and $\ulcorner \varphi \urcorner$ denote the codes of a τ-term t and a τ-formula φ, respectively. It is now not hard to see that a τ-theory T is recursive if and only if the set of codes of its axioms is recursive (as a subset of $\mathbb{N}$).

Now let τ be the signature of arithmetic, i.e. $\tau := \tau_{\text{arthm}}$, and thus we have the above coding since τ_{arthm} is finite. For every $n \in \mathbb{N}$, set $\Delta(n) \doteq S^n(0)$. 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 τ_{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 words, 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 Δm.

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

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

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

Lemma 6.4 (Fixed point for $\mathbf{N}$). *For each τ_{arthm}-formula $\varphi(v)$ there is a τ_{arthm}-sentence θ such that*

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

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

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

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 $\mathbf{N}$. 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 6.5 (Tarski, 1939). *$Th(\mathbf{N})$ is not arithmetical, i.e. the set* $\ulcorner \text{Th}(\mathbf{N}) \urcorner := \{\ulcorner \varphi \urcorner : \varphi \in \text{Th}(\mathbf{N})\}$ *is not definable in* $\mathbf{N}$.

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* τ_{arthm}-theory T, it is straightforward to check that the following relation is recursive if such is T: for $a, e \in \mathbb{N}$,

$$\text{Proof}_T(a, e) \iff a \text{ is a code of a } \tau_{\text{arthm}}\text{-formula } \varphi \text{ and } e \text{ is a code of a proof of } \varphi \text{ from } T.$$

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 $\mathbb{FOL}(\tau_{\text{arthm}})$, 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 τ_{arthm}-formula $\mathbf{Proof}_T(x,y)$ such that for all $a,b \in \mathbb{N}$,

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

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

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

and hence, for any τ-formula φ,

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

Proof of the Incompleteness theorem 6.2. We let $T \subseteq \mathrm{Th}(\mathbf{N})$ be recursive and show that it is incomplete by finding a sentence that $\mathbf{N}$ satisfies but T does not prove.

Applying the Fixed Point lemma to

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

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

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

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 $T \subseteq \mathrm{Th}(\mathbf{N})$, we have

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

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

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

Another proof of the Incompleteness theorem 6.2. If T was recursive and complete, then the formula $\mathbf{Provable}_T(x)$ would define the set $\ulcorner \mathrm{Th}(\mathbf{N}) \urcorner$ in $\mathbf{N}$ because, by the completeness of T, for every sentence φ, φ is provable from T if and only if $\ulcorner \varphi \urcorner \in \ulcorner \mathrm{Th}(\mathbf{N}) \urcorner$. Thus $\ulcorner \mathrm{Th}(\mathbf{N}) \urcorner$ would be arithmetical, contradicting Tarski's theorem (6.5). □

6.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[10]. In this subsection, we will write such a program using informal pseudocode in the hope of obtaining a better (hands-on) understanding of how the self-reference is implemented via the substitution function.

[10] 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.

To write 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. Now we take the diagonal of this: 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 (which is a string of symbols) of x. It remains to feed the program PrintDiagSub(x) its own code: Quine() := PrintDiagSub(“the code of PrintDiagSub(x)”).

We now write this more explicitly using a pseudo-code, whose syntax resembles that of the programming language C; in our pseudo-code, $\doteq$ is the command that assigns a value to a variable. We start by writing a program without input that assigns the variable x some string (e.g. `mathx≐is≐xfunx≐`) using the command $x \doteq$ “`mathx≐is≐xfunx≐`”, and then, it iterates through the content of x and prints every symbol in it; however, whenever it encounters the pattern `x`$\doteq$, it, in addition, prints the opening quote symbol “, then the *content* of x, then the closing quote symbol ”.

```
NotYetQuine()
  {
     $x \doteq$ “mathx≐is≐xfunx≐”;
     for($i \doteq 0; i < \text{length}(x); i \doteq i + 1$)
     {
        Print($x[i]$);
        if($i \geq 1 \wedge x[i-1] =$ ‘x’ $\wedge\, x[i-1] =$ ‘≐’)
        {
           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 `mathx≐is≐xfunx≐` with the code above, excluding the substring “`mathx≐is≐xfunx≐`”.

```
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 − 1] = ‘≐’)
                     {
                       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 − 1] = ‘≐’)
       {
         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 $\mathrm{Th}(\mathbf{N})$, but also to theories (in an arbitrary finite signature τ), which have PA "encoded" in them; for example, $\mathrm{PA} \cup \{\neg\gamma_{\mathrm{PA}}\}$ and ZFC.

6.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 6.6. For a relation $R \subseteq \mathbb{N}^{k+1}$, define a partial function $f_R : \mathbb{N}^k \rightharpoonup \mathbb{N}$ by $\vec{a} \mapsto \mu x(R(\vec{a}, x))$, where $\mu x(R(\vec{a}, x))$ is the smallest $x \in \mathbb{N}$ for which $R(\vec{a}, x)$ holds, if such x exists, and it is undefined, otherwise, in which case we write $\mu x(R(\vec{a}, x)) =\perp$. The partial function f_R is said to be obtain by applying the *search operation* to R.

For example, $\mu x(x^2 > 7) = 3$. This operation is also called *minimalization*.

Definition 6.7 (Recursive functions)**.** A function $f : \mathbb{N}^k \to \mathbb{N}$ is called recursive (or computable) if it is obtained by inductively applying the following rules:

(R1)
- $+ : \mathbb{N}^2 \to \mathbb{N}$ and $\cdot : \mathbb{N}^2 \to \mathbb{N}$ are recursive;
- $\mathbb{1}_\leq : \mathbb{N}^2 \to \mathbb{N}$ is recursive, where $\mathbb{1}_\leq$ is the characteristic function of $\leq$, i.e. $\mathbb{1}_\leq(x,y) = 1$ if $x \leq y$, and 0, otherwise;
- The projection functions $P_i^n(x_1,...,x_n) = x_i$ are recursive, for all $i = 1,...,n$ and $n \in \mathbb{N}$;

(R2) Composition: if $g : \mathbb{N}^m \to \mathbb{N}$ and $h_1,...,h_m : \mathbb{N}^k \to \mathbb{N}$ are recursive, then so is the composition function $f = g(h_1,...,h_2) : \mathbb{N}^k \to \mathbb{N}$ defined by

$$f(\vec{a}) = g(h_1(\vec{a}),...,h_m(\vec{a}));$$

(R3) Well-defined search: if $g : \mathbb{N}^{n+1} \to \mathbb{N}$ is recursive and for all $\vec{a} \in \mathbb{N}^n$ there is $x \in \mathbb{N}$ with $g(\vec{a},x) = 0$, then the function $f : \mathbb{N}^n \to \mathbb{N}$ defined by

$$f(\vec{a}) = \mu x(g(\vec{a},x) = 0)$$

is recursive.

A relation $R \subseteq \mathbb{N}^n$ is called recursive if so is its characteristic function $\mathbb{1}_R : \mathbb{N}^n \to \mathbb{N}$.

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 those is the operation of *primitive recursion*, which 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).

The following proposition provides some closure properties of the class of recursive functions together with some examples.

Lemma 6.8.

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

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

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

(d) If n-ary relations P, Q on $\mathbb{N}^n$ are recursive, then so are the following

$$\neg P := \mathbb{N}^n \smallsetminus P, P \wedge Q := P \cap Q, P \vee Q := P \cup Q.$$

(e) (Definition by Cases) Let R_1, ..., $R_k \subseteq \mathbb{N}^n$ be recursive such that for each $\vec{a} \in \mathbb{N}^n$ exactly one of $R_1(\vec{a}),...,R_k(\vec{a})$ holds, and suppose that $g_1,...,g_k : \mathbb{N}^n \to \mathbb{N}$ are recursive. Then $g : \mathbb{N}^n \to \mathbb{N}$ given by

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

is recursive.

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

We prove (b) by induction on k. For $k = 0$, observe that $c_0^n(\vec{a}) = \mu x(P_{n+1}^{n+1}(\vec{a},x) = 0)$. Assume c_k^n is recursive and note that

$$c_{k+1}^n(\vec{a}) = \mu x(c_k^n(\vec{a}) < x) = \mu x(\mathbb{1}_\geq(c_k^{n+1}(\vec{a},x), P_{n+1}^{n+1}(\vec{a},x)) = 0).$$

For (c), just note that $S(a) = a + c_1^1(a)$.

For (d), observe that $\neg P(\vec{a}) \iff \mathbb{1}_P(\vec{a}) = c_0^n(\vec{a})$ and $\mathbb{1}_{P\wedge Q}(\vec{a}) = \mathbb{1}_P(\vec{a}) \cdot \mathbb{1}_Q(\vec{a})$. Thus $\neg P$ and $P \wedge Q$ are recursive if so are P and Q. Recursiveness of the rest of the Boolean combinations follows from this because they are expressible in terms of $\wedge$ and $\neg$.

Part (e) is left to the reader. □

Lemma 6.9. *Let $R \subseteq \mathbb{N}^{n+1}$ be recursive such that for all $\vec{a} \in \mathbb{N}^n$ there exists $x \in \mathbb{N}$ with $(\vec{a}, x) \in R$. Then the function $f : \mathbb{N}^n \to \mathbb{N}$ given by*

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

is recursive.

Proof. Note that $f(\vec{a}) = \mu x(\mathbb{1}_{\neg R}(\vec{a}, x) = 0)$. □

Using this we get the following convenient property for verifying recursiveness of functions:

Proposition 6.10 (Graph property)**.** *Let $f : \mathbb{N}^n \to \mathbb{N}$. Then f is recursive if and only if so is its graph (as a subset of $\mathbb{N}^{n+1}$).*

Proof. let $R \subseteq \mathbb{N}^{n+1}$ be the graph of f. Then for all $\vec{a} \in N^n$ and $b \in \mathbb{N}$,

$$R(\vec{a}, b) \iff f(\vec{a}) = b,$$

and hence

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

from which the proposition follows immediately. □

Definition 6.11. Let $g : \mathbb{N}^k \to \mathbb{N}$ and $h : \mathbb{N}^{k+2} \to \mathbb{N}$. We say that $f : \mathbb{N}^{k+1} \to \mathbb{N}$ is defined by *primitive recursion* from g, h if for all $\vec{a} \in \mathbb{N}^k$ and $n \in \mathbb{N}$,

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

We aim at showing that the class of recursive functions is closed under this operation. For that, we first convert the recursive definition into an explicit (iterative) one as follows.

Proposition 6.12 (Dedekind's analysis of recursion)**.** *If $f : \mathbb{N}^{k+1} \to \mathbb{N}$ is defined by primitive recursion from g, h as in 6.11, then for all $\vec{a} \in \mathbb{N}^k$, $n \in \mathbb{N}$ and $w \in \mathbb{N}$,*

$$\begin{aligned} f(\vec{a}, n) = w \iff & \text{ there exists a sequence } (w_0, ..., w_n) \text{ such that} \\ & w_0 = g(\vec{a}) \wedge (\forall i < n)[w_{i+1} = h(\vec{a}, i, w_i)] \wedge w_n = w. \end{aligned}$$

Proof. Obvious. □

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

Chinese Remainder Theorem 6.13. *Let $d_0, ..., d_n$ be pairwise coprime and put $d = d_0 d_1 ... d_n$. Then the natural projection map*

$$h : \mathbb{Z}/d\mathbb{Z} \to \mathbb{Z}/d_0\mathbb{Z} \times ... \times \mathbb{Z}/d_n\mathbb{Z}$$

defined by

$$[a]_d \mapsto ([a]_{d_0}, ..., [a]_{d_n})$$

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 Pigeon Hole 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 6.14.

(a) *If relation* $R \subseteq \mathbb{N}^{k+1}$ *is recursive, then so are the relations*

$$P(\vec{a}, y) \iff \exists x_{<y} R(\vec{a}, x), Q(\vec{a}, y) \iff \forall x_{<y} R(\vec{a}, x),$$

for all $\vec{a} \in \mathbb{N}^k$, $y \in \mathbb{N}$.

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

(c) *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.*

(d) *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.

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

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

are recursive.

Proof. We leave parts (a),(b) and (c) to the reader. For (d), $\mathrm{Pair}(x, y) = \mu z(2z \dot{-} (x+y)(x+y+1) = 0) + x$ and hence is recursive. It is a bijection because it enumerates pairs (x, y) as follows:

$$\underbrace{(0,0)}_{x+y=0}\underbrace{(0,1)(1,0)}_{x+y=1}\underbrace{(0,2)(1,1)(2,0)}_{x+y=2}\dots$$

For (e), observe that $\mathrm{Left}(z) = \mu x(\exists y_{<z+1} \mathrm{Pair}(x, y) = z)$ and similarly for Right. □

Lemma 6.15 (Gödel's β-function)**.** *The function* $\beta : \mathbb{N}^2 \to \mathbb{N}$ *defined by*

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

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

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

Proof. The fact that β is recursive follows from 6.14, so we prove the second statement. Let $s = \max\{n, w_0, w_1, \dots, w_n\}$, set $b = s!$ and verify that

$$d_0 = 1 + (0+1)b, d_1 = 1 + (1+1)b, \dots, d_n = 1 + (n+1)b$$

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

By the Chinese Remainder Theorem, there is $a < d_0 \cdot ... \cdot d_n$ such that $\mathrm{Rem}(a, d_i) = w_i$. Thus setting $w = \mathrm{Pair}(a, b)$, we get

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

□

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

- $\langle a_0, ..., a_{n-1}\rangle := \mu x(\beta(x,0) = n \wedge \bigwedge_{i=1}^{n} \beta(x,i) = a_{i-1})$. Note that $<>= 0$ (as a nullary function).
- $\mathrm{lh} : \mathbb{N} \to \mathbb{N}$ by $\mathrm{lh}(a) = \beta(a, 0)$.
- $(a)_i := \beta(a, i+1)$. Note that $(\langle a_0, ..., a_{n-1}\rangle)_i = a_i$.
- $\mathrm{InitSeg}(a, i) = \mu x(\mathrm{lh}(x) = i \wedge \forall j_{<i} (x)_j = (a)_j)$. Thus $\mathrm{InitSeg}(\langle a_0, ..., a_n\rangle, i) = \langle a_0, ... a_{i-1}\rangle$.
- $a * b = \mu x(\mathrm{lh}(x) = \mathrm{lh}(a) + \mathrm{lh}(b) \wedge \forall i_{<\mathrm{lh}(a)})(x)_i = (a)_i \wedge \forall i_{<\mathrm{lh}(b)})(x)_{\mathrm{lh}(a)+i} = (b)_i$. Thus $\langle a_0, ... a_{n-1}\rangle * \langle b_0, ... b_{m-1}\rangle = \langle a_0, ... a_{n-1}, b_0, ..., b_{m-1}\rangle$.

Proposition 6.16. *Recursive functions are closed under the operation of primitive recursion, i.e. if g, h, f are as in Definition 6.11 and g, h are recursive, then so is f.*

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, so is f. □

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 6.17. The class of *primitive recursive functions* is the smallest class containing the successor function $S : \mathbb{N} \to \mathbb{N}$, the constant functions $C_k^n : \mathbb{N}^n \to \mathbb{N}$, $k, n \in \mathbb{N}$ and the projection functions $P_i^n(x_1, ..., x_n) = x_i$, $i \leq n, n \in \mathbb{N}$, and is closed under composition and primitive recursion. A relation $R \subseteq \mathbb{N}^n$ is called primitive recursive if so is its characteristic function $\mathbb{1}_R : \mathbb{N}^n \to \mathbb{N}$.

The reader can verify that the functions in (R1) of the definition of recursive functions are primitive recursive. It is also easy to check that Lemma 6.8 holds with *recursive* replaced by *primitive recursive*.

The following makes it easy to verify that Lemmas 6.14 and 6.15 also hold with *recursive* replaced by *primitive recursive*.

Lemma 6.18 (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 6.15 yields a primitive recursive function $B : \mathbb{N} \to \mathbb{N}$, defined by $B(N) = \prod_{i<n}(1 + (1+i)N!)$, such that for every $n \in \mathbb{N}$ and $\vec{a} \in \mathbb{N}^n$,

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

Using this together with 6.18 one can easily show that the coding/decoding functions $\langle a_0, ..., a_{n-1}\rangle$, $\mathrm{lh}(a)$, $(a)_i$, $\mathrm{InitSeg}(a, i)$, $a * b$ 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 6.19 (Complete primitive recursion)**.** *For $f : \mathbb{N}^{n+1} \to \mathbb{N}$, let*

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

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) and leave (b) to the reader.
$\Leftarrow$: Put $f(\vec{a}, n) = (\bar{f}(\vec{a}, n+1))_n$.
$\Rightarrow$: 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 here is why:

Proposition 6.20. *There exists a recursive function $\varphi : \mathbb{N}^2 \to \mathbb{N}$ such that $\varphi_n := \varphi(n, \cdot)$ enumerates all the primitive recursive functions (possibly with repetitions), i.e. for every n, φ_n is primitive recursive and for every primitive recursive function f, there is n such that $f = \varphi_n$. Moreover, any such function φ is not primitive recursive.*

Proof. A proof of the existence of such φ is outlined in one of the homework problems and here we show that such φ is not primitive recursive by applying Cantor's diagonalization[11]

[11] As van den Dries suggests, perhaps *antidiagonalization* would be a better name.

method. Assume for contradiction that φ is primitive recursive. Then so is the function $\psi(n) = \varphi(n,n) + 1$, for all n, and thus there is $n_0 \in \mathbb{N}$ such that $\varphi_{n_0} = \psi$. But then we have

$$\psi(n_0) = \varphi_{n_0}(n_0) = \varphi(n_0, n_0)$$

on one hand, and

$$\psi(n_0) = \varphi(n_0, n_0) + 1$$

on the other, which is a contradiction. □

Note that the same proof shows that there is no recursive enumeration of recursive functions. Similarly, the set of codes of recursive functions 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 6.21. *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 6.10) does not hold for primitive recursive functions.

6.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 $\mathbf{N}$. 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 6.22 (Representability)**.** Let T be a τ_{arthm}-theory in the signature τ_{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 $\vec{a} \in \mathbb{N}^n$,

  $$R(\vec{a}) \implies T \vDash \varphi(\Delta(\vec{a})) \text{ and } \neg R(\vec{a}) \implies T \vDash \neg\varphi(\Delta(\vec{a})),$$

  where $\Delta(\vec{a}) = (\Delta(a_1), ..., \Delta(a_n))$. 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 $\vec{a} \in \mathbb{N}^n$,

  $$T \vDash \forall y \Big[\varphi(\Delta(\vec{a}), y) \leftrightarrow y = \Delta(f(\vec{a})) \Big].$$

  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 τ_{arthm}-term $t(\vec{x})$ such that for all $\vec{a} \in \mathbb{N}^n$,
$$T \vDash t(\Delta(\vec{a})) = \Delta(f(\vec{a})).$$
Such t is said to represent f in T.

Proposition 6.23. *Let T be a τ_{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 $m, k \in \mathbb{N}$, $T \vDash \Delta(m) \neq \Delta(k)$. Then, if f is representable in T, then so is its graph.*

Proof. For part (a), letting $t(\vec{x})$ be a term representing f, it is straightforward to check that the formula $\varphi(\vec{x}, y) \doteq t(\vec{x}) = y$ represents f. As for (b), we check that any formula representing f also represents its graph. Indeed, 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
$$T \vDash \varphi(\Delta(\vec{a}, \Delta(b))) \leftrightarrow \Delta(b) = \Delta(f(\vec{a})).$$
Thus, it is clear that if $f(\vec{a}) = b$ then $T \vDash \varphi(\Delta(\vec{a}), \Delta(b))$, and if $f(\vec{a}) \neq b$ then the additional hypothesis on T guarantees that $T \vDash \neg\varphi(\Delta(\vec{a}), \Delta(b))$. □

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

Proposition 6.24. *If T is a τ_{arthm}-theory such that $T \vDash \Delta(1) \neq 0$ and $R \subseteq \mathbb{N}^n$, then*
$$R \text{ is representable in } T \text{ if and only if } \mathbb{1}_R \text{ is representable in } T.$$

Proof. $\Rightarrow$: Let $\varphi(\vec{x})$ represent R in T and put
$$\psi(\vec{x}, y) \doteq \big[\varphi(\vec{x}) \wedge y = \Delta(1)\big] \vee \big[\neg\varphi(\vec{x}) \wedge y = 0\big].$$
We show that $\psi(\vec{x}, y)$ represents $\mathbb{1}_R$ 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 $T \vDash \varphi(\Delta(\vec{a}))$, $\mathbb{1}_R(\vec{a}) = 1$, and we have to show
$$T \vDash \forall y\Big[\psi(\Delta(\vec{a}), y) \leftrightarrow y = \Delta(1)\Big].$$
Fixing a model $\mathbf{M} \vDash T$ and an arbitrary $y \in M$, we see that, since $\mathbf{M} \vDash \varphi(\Delta(\vec{a}))$,
$$\mathbf{M} \vDash \psi(\Delta(\vec{a}), y) \iff \mathbf{M} \vDash \big[\varphi(\Delta(\vec{a})) \wedge y = \Delta(1)\big] \iff \mathbf{M} \vDash y = \Delta(1).$$
A similar argument handles the case $\neg R(\vec{a})$.
$\Leftarrow$: Let $\varphi(\vec{x}, y)$ represent $\mathbb{1}_R$ 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
$$T \vDash \varphi(\Delta(\vec{a}), \Delta(1)) \leftrightarrow \Delta(1) = \Delta(\mathbb{1}_R(\vec{a})).$$
Thus, it is clear that if $R(\vec{a})$ holds then $T \vDash \varphi(\Delta(\vec{a}), \Delta(1))$, and if $R(\vec{a})$ fails then $T \vDash \Delta(1) \neq 0$ guarantees that $T \vDash \neg\varphi(\Delta(\vec{a}), \Delta(1))$. □

Proposition 6.25. *All recursive functions and relations are representable in* PA*.*

Proof. By Lemma 6.24, it is enough to show for functions.

Because the standard part of any model of PA is isomorphic to $\mathbf{N}$, the terms $t_+(x,y) \doteq x+y$, $t_\cdot(x,y) \doteq x \cdot y$ and $t_i^{(n)}(x_1,\dots,x_n) = x_i$ represent, respectively, the addition, multiplication and the projection functions. For the same reason, the formula $x \leq y \doteq \exists z(z+x=y)$ represents the relation $\leq$, and hence $\mathbb{1}_\leq$ is representable as well by 6.24. It remains to show that representability is closed under composition (R2) and safe search $R(3)$ operations.

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

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

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

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

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 $b \in \mathbb{N}$ such that $g(\vec{a},b) = 0$. 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 $f(\vec{a}) = \mu x(g(\vec{a},x) = 0)$. Fix $\vec{a} \in \mathbb{N}^n$ and let $b = f(\vec{a})$. We have to show that

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

By definition, $g(\vec{a},i) = d_i \neq 0$ for all $i < b$, and $g(\vec{a},b) = 0$. Thus,

$$\mathrm{PA} \models \varphi(\vec{a},\Delta(b),0) \text{ and } \mathrm{PA} \models \varphi(\vec{a},\Delta(i),\Delta(d_i)), \text{ for all } i < b.$$

Because $\mathrm{PA} \models \forall u\Big[u < \Delta(b) \to \bigvee_{i<b} u = \Delta(i)\Big]$, it follows that $\mathrm{PA} \models \psi(\Delta(\vec{a}),\Delta(b))$. On the other hand, for any $\mathbf{M} \models \mathrm{PA}$ and $z \in \mathbf{M}$ with $z \neq \Delta(b)$, we have that either $z < \Delta(b)$, so $z = \Delta(i)$ for some $i < b$ and hence $\mathbf{M} \models \neg\varphi(\Delta(\vec{a}),z,0)$, or $z > \Delta(b)$ and taking $u := \Delta(b)$ witnesses the failure of $\forall u\big[u < z \to \neg\varphi(\vec{x},u,0)\big]$ in $\mathbf{M}$. Hence, for all $z \in \mathbf{M}$,

$$\mathbf{M} \models \psi(\Delta(\vec{a}),z) \leftrightarrow z = \Delta(b).$$

□

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

6.E. **Robinson's system** Q

Now we describe a finite subtheory of $\mathrm{Th}(\mathbf{N})$, namely Robinson's[12] system Q, which is much weaker than PA, but still rich enough to represent recursive functions. The advantage of it over PA is that it is finite, and we will use this later in proving that the empty τ_{arthm}-theory

[12] This is due to Raphael Robinson and not Abraham or Julia Robinsons as I falsely thought.

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 $\mathrm{Th}(\mathbf{N})$.

Definition 6.26 (Robinson's system Q)**.** The following are the axioms of Q:
(Q1) $\forall x[\neg S(x) = 0]$,
(Q2) $\forall x \forall y[S(x) = S(y) \to x = y]$,
(Q3) $\forall x[x + 0 = x]$,
(Q4) $\forall x \forall y[S(x + y) = x + S(y)]$,
(Q5) $\forall x[x \cdot 0 = 0]$,
(Q6) $\forall x \forall y[x \cdot S(y) = x \cdot y + x]$,
(Q7) $\forall x(x \neq 0 \to \exists y(x = S(y)))$.

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 6.27.
(a) For any model $\mathbf{M}$ *of* Q*, there is a unique homomorphism* $f : \mathbf{N} \to \mathbf{M}$*. In fact, this* f *is a* τ_{arthm}*-embedding and hence we can view* $\mathbf{N}$ *as a substructure of* $\mathbf{M}$*.*
(b) For any quantifier free formula $\varphi(\vec{x})$ *and* $\vec{a} \in \mathbb{N}^k$*,*

$$\mathbf{N} \vDash \varphi(\vec{a}) \iff \mathrm{Q} \vdash \varphi(\Delta(\vec{a})),$$

where $\Delta(\vec{a}) = (\Delta(a_1), ..., \Delta(a_k))$*.*

Proof. Part (b) follows from (a) since for $\mathbf{M} \vDash \mathrm{Q}$, $\mathbf{N} \subseteq \mathbf{M}$ and hence

$$\mathbf{N} \vDash \varphi(\vec{a}) \iff \mathbf{M} \vDash \varphi(\Delta(\vec{a})),$$

because φ is quantifier free. Because $\mathbf{M}$ 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 $f(\Delta(n)^{\mathbf{N}}) = \Delta(n)^{\mathbf{M}}$. This function is injective because $S^{\mathbf{M}}$ is injective and $0^{\mathbf{M}}$ 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 $\mathbf{M}$ satisfies axiom (Q4). □

Let $x \leq y$ and $x < y$ abbreviate the formulas $\exists z(z + x = y)$ and $x \neq y \wedge \exists z(z + x = y)$, respectively. Keep in mind that $z + x$ may not be equal to $x + z$ in a model of Q. Since the statement $x \leq y$ is not quantifier free, it does not follow from the previous lemma that a model of Q and $\mathbf{N}$ have to agree on the ordering of natural numbers (the standard part of $\mathbf{M}$). However, it turns out to still be true:

Lemma 6.28 (Q preserves the ordering on $\mathbb{N}$)**.** *For all* $n, m \in \mathbb{N}$*,*

(a) $\mathrm{Q} \vdash x \leq \Delta(n) \to \bigvee_{i=0}^{n} x = \Delta(i)$*;*
(b) $n \leq m \iff \mathrm{Q} \vdash \Delta(n) \leq \Delta(m)$*;*

(c) $\neg n \leq m \iff \mathrm{Q} \vdash \neg\Delta(n) \leq \Delta(m)$*;*
(d) $\mathrm{Q} \vdash x \leq \Delta(n) \vee \Delta(n+1) \leq x$*;*
(e) $\mathrm{Q} \vdash x \leq \Delta(n) \vee \Delta(n) < x$*.*

Proof. For part (b), the right-to-left direction follows immediately from (a). As for the other direction, if $n \leq m$, then let $k = m - n$ and thus $\mathbf{N} \vDash \Delta(k) + \Delta(n) = \Delta(m)$. By (b) of 6.27, $\mathrm{Q} \vdash \Delta(k) + \Delta(n) = \Delta(m)$ and thus $\mathrm{Q} \vdash \Delta(n) \leq \Delta(m)$.

For (e), first consider $n = 0$. Then by (Q3), $\mathrm{Q} \vdash 0 \leq x$, 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), $\mathrm{Q} \vdash x \leq \Delta(m) \vee \Delta(n) \leq x$. Thus, arguing in Q and using (a), either $x = \Delta(k)$ for some $k < n$, or $x = \Delta(n)$, or $\Delta(n) \geq x$. Hence, again using (a) and the definition of the formula $y < z$, we get that either $x \leq \Delta(n)$ or $\Delta(n) < x$.

We leave the proofs of (c) and (d) to the reader, and we prove (a) by induction on n. Let $\mathbf{M} \vDash \mathrm{Q}$. For $n = 0$, assume $a \in M$ and $\mathbf{M} \vDash a \leq 0$. Thus, there is $b \in M$ such that $\mathbf{M} \vDash b + a = 0$. Now if $a \neq 0^{\mathbf{M}}$, then a has a predecessor, i.e. for some $c \in \mathbf{M}$, $\mathbf{M} \vDash a = S(c)$ and thus $\mathbf{M} \vDash b + S(c) = 0$. Arguing inside $\mathbf{M}$, $0 = b + S(c) = S(b + c)$, which contradicts the fact that 0 is not a successor. Thus $a = 0$.

Now assume the statement is true for n and assume $\mathbf{M} \vDash a \leq \Delta(n+1)$. Hence there is $b \in M$ such that $b + a = \Delta(n+1)$ (arguing inside $\mathbf{M}$). 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 $c \leq \Delta(n)$. By the induction hypothesis, c is equal to one of $\Delta(i)$ for $i = 0, ..., n$ and thus a is equal to one of $\Delta(j)$ for $j = 1, ..., n+1$. □

Proposition 6.29. *All recursive functions and relations are representable in* Q*.*

Proof. The proof is word-by-word the same as for Proposition 6.25 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 6.27 and (a,c,d) of 6.28. □

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

6.F. **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 τ be a finite signature.

- We code the symbols of $\mathbb{FOL}(\tau)$ as follows: for $s \in \tau \cup \{\text{logical symbols}\} \cup \{v_0, v_1, ...\}$, assign a number $\mathrm{SN}(s)$ as follows: put $\mathrm{SN}(s) = 2i$ if $s = v_i$ and assign an odd number to each of the remaining symbols (finitely many) such that different symbols get different numbers.
- For a τ-term t, define its Gödel code $\ulcorner t \urcorner$ as follows

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

Note that for a variable or a constant symbol s, $\ulcorner s \urcorner$ may not be equal to $\mathrm{SN}(s)$.

- For a τ-formula φ, define its Gödel code $\ulcorner\varphi\urcorner$ as follows

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

Lemma 6.30. *The following subsets of* $\mathbb{N}$ *are primitive recursive:*

(i) Variable $:= \{\ulcorner x\urcorner : x$ *is a variable*$\}$
(ii) Term $:= \{\ulcorner t\urcorner : t$ *is a* τ*-term*$\}$
(iii) Formula $:= \{\ulcorner\varphi\urcorner : \varphi$ *is a* τ*-formula*$\}$

Proof. In all proofs we use complete primitive recursion (Lemma 6.19).
(i) $a \in$ Variable if and only if $\mathrm{lh}(a) = 1$ and $(a)_0$ is even.
(ii) Term(a) if and only if 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 $\forall i < n, \mathrm{Term}((a)_{i+1})$.
(iii) is left to the reader. It gets messy if one wants to also check our convention about quantified variables. □

Lemma 6.31. *There is a primitive recursive function* Sub : $\mathbb{N}^3 \to \mathbb{N}$ *such that for any* τ*-formula* φ*, variable* v *and* τ*-term* t *that is free for* v *in* φ*,*

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

Proof. Define Sub$(a, m, k) =$

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

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

Lemma 6.32. *The following relations are primitive recursive:*

(1) FreeVar $:= \{(\ulcorner\varphi\urcorner, \mathrm{SN}(v)) : v$ *occurs free in* $\varphi\} \subseteq \mathbb{N}^2$
(2) FreeSub $:= \{(\ulcorner\varphi\urcorner, \ulcorner t\urcorner) : t$ *is free for* $\varphi\} \subseteq \mathbb{N}^3$
(3) Sentence $:= \{\ulcorner\varphi\urcorner : \varphi$ *is a sentence*$\} \subseteq \mathbb{N}$
(4) Axiom $:= \{\ulcorner\varphi\urcorner : \varphi$ *is an axiom of* $\mathbb{FOL}(\tau)\} \subseteq \mathbb{N}$
(5) MP $:= \{(\ulcorner\varphi\urcorner, \ulcorner\varphi \to \psi\urcorner, \psi) : \varphi, \psi$*are* τ*-formulas*$\} \subseteq \mathbb{N}^3$

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

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 6.33. For a τ-theory T, define binary relations $\mathrm{Proof}_T, \mathrm{Refute}_T \subseteq \mathbb{N}^2$ by

$$\mathrm{Proof}_T := \{(\langle \ulcorner\varphi_1\urcorner, ..., \ulcorner\varphi_n\urcorner\rangle, \ulcorner\varphi\urcorner) : (\varphi_1, ..., \varphi_n) \text{ is a proof of } \varphi \text{ from } T\},$$
$$\mathrm{Refute}_T := \{(\langle \ulcorner\varphi_1\urcorner, ..., \ulcorner\varphi_n\urcorner\rangle, \ulcorner\varphi\urcorner) : (\varphi_1, ..., \varphi_n) \text{ is a proof of } \neg\varphi \text{ from } T\},$$

where φ_i and φ vary over τ-formulas.

For a τ-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 6.34. *If a τ-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 $\exists i_{<k} \exists j_{<k} \mathrm{MP}((a)_i, (a)_j, (a)_k)$. □

6.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 τ_{arthm}-formula θ, put $[\theta] := \Delta(\ulcorner\theta\urcorner)$.

Lemma 6.35 (Fixed point for Q)**.** *For every τ_{arthm}-formula $\varphi(v)$, there is a τ_{arthm}-sentence θ such that*

$$\mathrm{Q} \vDash \theta \leftrightarrow \varphi([\theta]).$$

Proof. Let $\mathbf{Sub}_0(x, y, z)$ be a τ_{arthm}-formula representing Sub_0 in Q. We can assume without loss of generality that the variable v_0 does not appear in $\mathbf{Sub}_0$ and φ. Put

$$\psi(v_0) \doteq \exists z(\mathbf{Sub}_0(v_0, v_0, z) \wedge \varphi(z)),$$

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,

$$\mathrm{Q} \vDash \mathbf{Sub}_0(\Delta(m), \Delta(m), z) \leftrightarrow z = [\theta]. \tag{i}$$

In particular,

$$\mathrm{Q} \vDash \mathbf{Sub}_0(\Delta(m), \Delta(m), [\theta]). \tag{ii}$$

Therefore, we have

$$\begin{array}{rcl} \mathrm{Q} \vDash \theta & \iff & \mathrm{Q} \vDash \psi(\Delta(m)) \\ & \iff & \mathrm{Q} \vDash \exists z(\mathbf{Sub}_0(\Delta(m), \Delta(m), z) \wedge \varphi(z)) \\ [\implies \text{ is because of (i)}] & \iff & \mathrm{Q} \vDash \mathbf{Sub}_0(\Delta(m), \Delta(m), [\theta]) \wedge \varphi([\theta]) \\ [\impliedby \text{ is because of (ii)}] & \iff & \mathrm{Q} \vDash \varphi([\theta]). \end{array}$$

□

Now we are ready to prove the Incompleteness theorem for all τ_{arthm}-theories $T \supseteq \mathrm{Q}$. However, we would like to prove a slightly stronger version that applies to theories in signatures other than τ_{arthm} that are rich enough to encode Q in them. We make this precise in the following definition.

Definition 6.36. Let T_1, T_2 be theories in finite signatures τ_1, τ_2, respectively. An *interpretation* of T_1 in T_2 is a map π from the set of τ_1-sentences to the set of τ_2-sentences such that

(i) $T_1 \vDash \theta \implies T_2 \vDash \pi(\theta)$,
(ii) $T_2 \vDash \pi(\neg\theta) \leftrightarrow \neg\pi(\theta)$,
(iii) $T_2 \vDash \pi(\varphi \wedge \psi) \leftrightarrow \pi(\varphi) \wedge \pi(\psi)$,
(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 τ_1-sentences, and in the last equality, $\ulcorner\ \urcorner$ denotes the coding function of $\mathbb{FOL}(\tau_1)$ on the left and of $\mathbb{FOL}(\tau_2)$ 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 τ be a finite signature.

Lemma 6.37. *Let T be a (resp. primitive) recursive τ-theory that interprets* Q *and let π be an interpretation of* Q *in T. Then the following relations are (resp. primitive) recursive:*

$$\begin{aligned}\mathrm{Proof}_{\pi,T}(a,b) &\iff \textit{b is an } \mathbb{FOL}(\tau_{\mathrm{arthm}})\textit{-code of a } \tau_{\mathrm{arthm}}\textit{-sentence } \varphi \textit{ and} \\ &\qquad \textit{a is an } \mathbb{FOL}(\tau)\textit{-code of a proof of } \pi(\varphi) \textit{ from } T, \\ \mathrm{Refute}_{\pi,T}(a,b) &\iff \textit{b is an } \mathbb{FOL}(\tau_{\mathrm{arthm}})\textit{-code of a } \tau_{\mathrm{arthm}}\textit{-sentence } \varphi \textit{ and} \\ &\qquad \textit{a is an } \mathbb{FOL}(\tau)\textit{-code of a proof of } \pi(\neg\varphi) \textit{ from } T.\end{aligned}$$

Proof. Observe that

$$\begin{aligned}\mathrm{Proof}_{\pi,T}(a,b) &\iff \mathrm{Sentence}_{\tau_{\mathrm{arthm}}}(b) \text{ and } \mathrm{Proof}_T(a, \pi^*(b)), \\ \mathrm{Refute}_{\pi,T}(a,b) &\iff \mathrm{Sentence}_{\tau_{\mathrm{arthm}}}(b) \text{ and } \mathrm{Proof}_T(a, \pi^*(\langle \mathrm{SN}(\neg), b\rangle)).\end{aligned}$$

□

First Incompleteness Theorem 6.38 (Rosser's form)**.** *Any consistent recursive τ-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 $T \subseteq \mathrm{Th}(\mathbf{N})$, 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 $\mathbf{Proof}_{\pi,T}(x,y)$ be a τ_{arthm}-formula representing $\mathrm{Proof}_{\pi,T}$ in Q. Then by the fixed point lemma for Q, we get a τ_{arthm}-sentence γ such that

$$\mathrm{Q} \vDash \gamma \leftrightarrow \forall x \neg \mathbf{Proof}_{\pi,T}(x, [\gamma]). \tag{$*$}$$

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 $\mathrm{Q} \vDash \mathbf{Proof}_{\pi,T}(\Delta(a), [\gamma])$. But then by $(*)$, $\mathrm{Q} \vDash \neg\gamma$ and thus $T \vDash \pi(\neg\gamma)$, so $T \vDash \neg\pi(\gamma)$, contradicting the consistency of T.

However, we don't get any contradiction if we assume $T \vDash \neg\pi(\gamma)$. 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 $a \in \mathbb{N}$, $\mathrm{Q} \vDash \neg\mathbf{Proof}_{\pi,T}(\Delta(a), [\gamma])$. Unfortunately, this does NOT imply that $\mathrm{Q} \vDash \forall x \neg\mathbf{Proof}_{\pi,T}(x, [\gamma])$ because there may well be a model $\mathbf{M}$ of Q with a nonstandard element $w \in M \smallsetminus \mathbb{N}$ such that $\mathbf{M} \vDash \mathbf{Proof}_{\pi,T}(w, [\gamma])$ 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[13]:

[13] The author has heard this joke from Itay Neeman in the context of searching for an apartment to rent in LA.

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 Theorem 6.38. Let π be an interpretation of Q in T, and let $\mathbf{Proof}_{\pi,T}(x,y)$ and $\mathbf{Refute}_{\pi,T}(x,y)$ be τ_{arthm}-formulas representing $\text{Proof}_{\pi,T}$ and $\text{Refute}_{\pi,T}$ in Q. Then by the fixed point lemma for Q, we get a τ_{arthm}-sentence ρ such that

$$\mathrm{Q} \vDash \rho \leftrightarrow \forall x(\mathbf{Proof}_{\pi,T}(x,[\rho]) \to (\exists u < x)\mathbf{Refute}_{\pi,T}(u,x)). \tag{1}$$

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 $T \vdash \pi(\rho)$ nor $T \vdash \neg\pi(\rho)$.
Case 1: suppose $T \vdash \pi(\rho)$. Then there is a code $m \in \mathbb{N}$ of a proof of $\pi(\rho)$ from T and hence

$$\mathrm{Q} \vdash \mathbf{Proof}_{\pi,T}(\Delta(m),[\rho]). \tag{2}$$

Because T is consistent, $T \nvdash \neg\pi(\rho)$, and hence, by the definition of interpretation, $T \nvdash \pi(\neg\rho)$. Thus $\forall k \in \mathbb{N}$, $\neg\text{Refute}_{\pi,T}(k,\ulcorner\rho\urcorner)$ and hence $\mathrm{Q} \vdash \neg\mathbf{Refute}_{\pi,T}(\Delta(k),[\rho])$; in particular, this is true for all $k < m$. Therefore, by (a) of Lemma 6.28,

$$\mathrm{Q} \vdash (\forall u < \Delta m)\neg\mathbf{Refute}_{\pi,T}(u,[\rho]). \tag{3}$$

From (2) and (3), we get

$$\mathrm{Q} \vdash \exists x(\mathbf{Proof}_{\pi,T}(x,[\rho]) \wedge (\forall u < x)\neg\mathbf{Refute}_{\pi,T}(u,x)),$$

which implies $\mathrm{Q} \vdash \neg\rho$ by (1). Therefore, $T \vdash \pi(\neg\rho)$ and hence $T \vdash \neg\pi\rho$, contradicting the consistency of T.

Case 2: suppose $T \vdash \neg\pi(\rho)$. Thus $T \vdash \pi(\neg\rho)$, so there is a code $k \in \mathbb{N}$ of a proof of $\pi(\neg\rho)$ from T. Hence $\text{Refute}_{\pi,T}(k,\ulcorner\rho\urcorner)$ holds and by representability in Q,

$$\mathrm{Q} \vdash \mathbf{Refute}_{\pi,T}(\Delta(k),[\rho]). \tag{4}$$

Also, for any $n \in \mathbb{N}$, $\neg\text{Proof}_{\pi,T}(n,\ulcorner\rho\urcorner)$ holds by the consistency of T, and thus

$$\mathrm{Q} \vdash \neg\mathbf{Proof}_{\pi,T}(\Delta(n),[\rho]). \tag{5}$$

We argue in models, so fix $\mathbf{M} \vDash \mathrm{Q}$. By (e) of Lemma 6.28, for every $a \in M$, $a \leq \Delta(k)$ or $\Delta(k) < a$. In the first case, by (a) of Lemma 6.28, we get that $a = \Delta(n)$ for some $n \leq k$, and thus $\mathbf{M} \vDash \neg\mathbf{Proof}_{\pi,T}(a,[\rho])$, by (5). In the second case, i.e. if $\Delta(k) < a$,

$$\mathbf{M} \vDash (\exists u < a)\mathbf{Refute}_{\pi,T}(u,[\rho]),$$

by (4). Therefore, for all $a \in M$,

$$\mathbf{M} \vDash \mathbf{Proof}_{\pi,T}(a,[\rho]) \to (\exists u < a)\mathbf{Refute}_{\pi,T}(u,[\rho]).$$

Thus

$$\mathrm{Q} \vdash \forall x(\mathbf{Proof}_{\pi,T}(x,[\rho]) \to (\exists u < x)\mathbf{Refute}_{\pi,T}(u,x)),$$

and hence $\mathrm{Q} \vdash \rho$, by (1). But then $T \vdash \pi(\rho)$, contradicting the consistency of T. □

6.H. **The Second Incompleteness Theorem and Löb's theorem**

Let τ be a finite signature and let T be a recursive τ-theory. Recall (see Definition 6.33) that the relations $\mathrm{Proof}_T, \mathrm{Refute}_T \subseteq \mathbb{N}^2$ are recursive. Let $\mathbf{Proof}_T(x,y)$ and $\mathbf{Refute}_T(x,y)$ be τ_{arthm}-formulas representing them in Q, and put $\mathbf{Provable}_T(y) \doteq \exists x \mathbf{Proof}_T(x,y)$. Also recall that by $\perp$ we denote the sentence $\exists x(x \neq x)$.

Definition 6.39. For T as above, we define a τ_{arthm}-sentence that expresses the consistency of T as follows:

$$\mathbf{Con}_T \doteq \neg \mathbf{Provable}_T([\perp]).$$

Lemma 6.40. *Let T be a recursive τ-theory interpreting* PA *and let π be an interpretation. Also, let ρ_T be the Rosser sentence for T as in the proof of 6.38 above. Then* PA $\vdash \mathbf{Con}_T \to \rho_T$.

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 $T \nvdash \pi(\rho_T)$:

$$\mathrm{PA} \vdash \mathbf{Con}_T \to \forall x \neg \mathbf{Proof}_{\pi,T}(x, [\rho_T]).$$

On the other hand, it follows from the definition of ρ_T that

$$\mathrm{PA} \vdash \forall x \neg \mathbf{Proof}_{\pi,T}(x, [\rho_T]) \to \rho_T.$$

Therefore, PA $\vdash \mathbf{Con}_T \to \rho_T$. □

From this we immediately get yet another foundational theorem by Gödel:

Second Incompleteness Theorem 6.41. *Let T be a recursive τ-theory interpreting* PA *and let π be an interpretation. Then $T \nvdash \pi(\mathbf{Con}_T)$, 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

$$T \vdash \pi(\mathbf{Con}_T) \to \pi(\rho_T).$$

Thus, if $T \vdash \pi(\mathbf{Con}_T)$ then $T \vdash \pi(\rho_T)$, which is a contradiction. □

Lemma 6.42. *Let τ be a finite signature and T a recursive τ-theory. For any τ-sentences φ, θ, the following statements are provable in* PA:

(a) The Deduction theorem: $\mathbf{Provable}_{T \cup \{\theta\}}([\varphi]) \leftrightarrow \mathbf{Provable}_T([\theta \to \varphi])$.
(b) Proof by contradiction: $\mathbf{Provable}_T([\neg\theta \to \perp]) \leftrightarrow \mathbf{Provable}_T([\theta])$.
(c) Lemma about consistency: $\mathbf{Con}_{T \cup \{\neg\theta\}} \leftrightarrow \neg \mathbf{Provable}_T(\theta)$.

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 $\mathbf{N}$ is a model of PA, we know that whatever PA proves is true about the natural numbers, in other words, for every τ_{arthm}-sentence θ,

$$\mathbf{N} \vDash \mathbf{Provable}_{\mathrm{PA}}([\theta]) \to \theta.$$

Does PA know this? That is: does it prove $\mathbf{Provable}_{\mathrm{PA}}([\theta]) \to \theta$ for all θ? Here is the answer:

Theorem 6.43 (Löb, 1955)**.** *For every τ_{arthm}-sentence θ,* PA *does not prove* $\mathbf{Provable}_{\mathrm{PA}}([\theta]) \to \theta$ *unless it proves θ itself, i.e.*

$$\mathrm{PA} \vdash \mathbf{Provable}_{\mathrm{PA}}([\theta]) \to \theta \iff \mathrm{PA} \vdash \theta.$$

Proof. We prove the left-to-right direction since the other one is trivial. Assume for contradiction that $\mathrm{PA} \vdash \mathbf{Provable}_{\mathrm{PA}}([\theta]) \to \theta$ yet $\mathrm{PA} \nvdash \theta$. Thus the theory $S := \mathrm{PA} \cup \{\neg\theta\}$ is consistent. By contrapositive, $\mathrm{PA} \vdash \neg\theta \to \neg\mathbf{Provable}_{\mathrm{PA}}([\theta])$ and hence,

$$S \vdash \neg\mathbf{Provable}_{\mathrm{PA}}([\theta]). \qquad (*)$$

By (c) of Lemma 6.42, we have

$$\mathrm{PA} \vdash \mathbf{Con}_S \leftrightarrow \neg\mathbf{Provable}_{\mathrm{PA}}(\theta),$$

thus also

$$S \vdash \mathbf{Con}_S \leftrightarrow \neg\mathbf{Provable}_{\mathrm{PA}}(\theta),$$

so, by Modus Ponens with $(*)$, we get $S \vdash \mathbf{Con}_S$, contradicting the Second Incompleteness theorem. □

7. Undecidable theories

Fix a finite signature τ.

Definition 7.1. For a τ-theory T, let $\mathrm{Thm}(T)$ denote the set of its theorems, i.e. $\mathrm{Thm}(T) := \{\varphi : T \vdash \varphi\} \subseteq \mathbb{N}$, where φ ranges over all τ-sentences. If $\ulcorner \mathrm{Thm}(T) \urcorner$ 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.

7.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 7.2. For each $k \in \mathbb{N}$, put $\Gamma(\mathbb{N}^k) := \{A \subseteq \mathbb{N}^k : A \in \Gamma\}$, so $\Gamma = \bigcup_{k \in \mathbb{N}} \Gamma(\mathbb{N}^k)$. Also, put

$$\neg\Gamma := \left\{\mathbb{N}^k \setminus A : A \in \Gamma(\mathbb{N}^k), k \in \mathbb{N}\right\}$$
$$\exists^{\mathbb{N}}\Gamma := \left\{\mathrm{proj}_{k+1}(R) : R \in \Gamma(\mathbb{N}^{k+1}), k \in \mathbb{N}\right\}$$
$$\forall^{\mathbb{N}}\Gamma := \neg\exists^{\mathbb{N}}\neg\Gamma.$$

Definition 7.3. 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 7.4. (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 7.5. *For a* τ*-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. $\Pi^0_1 = \{\neg R : R \in \Sigma^0_1\}$, and let $\Delta^0_1 := \Sigma^0_1 \cap \Pi^0_1$. Also, let Recursive denote the set of recursive relations.

Lemma 7.6 (Kleene's theorem)**.** $\Delta^0_1 = \text{Recursive}$.

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 7.7. *Every complete recursive* τ*-theory* T *is decidable.*

Proof. Using the fact that for every τ-sentence φ, $\varphi \notin \mathrm{Thm}(T) \iff \neg\varphi \in \mathrm{Thm}(T)$, we get that for every $a \in \mathbb{N}$,

$$a \notin \ulcorner \mathrm{Thm}(T) \urcorner \iff a \notin \mathrm{Sentence}_\tau \text{ or } \langle \mathrm{SN}(\neg), a \rangle \in \ulcorner \mathrm{Thm}(T) \urcorner.$$

By Lemma 7.5, $\ulcorner \mathrm{Thm}(T) \urcorner$ is Σ^0_1. Because $\neg\mathrm{Sentence}_\tau$ is recursive (hence Σ^0_1) and Σ^0_1 is closed under recursive preimages and finite unions (7.4), 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[14] are decidable.

Another corollary of Kleene's theorem is a strengthening of Proposition 6.10.

Corollary 7.8. *For a function $f : \mathbb{N}^k \to \mathbb{N}$, the following are equivalent:*

(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 6.10, 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. □

7.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 $R(a) \coloneqq \{b \in B : (a, b) \in R\}$. 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 6.25.

Proposition 7.9. *Let T be a recursive consistent τ_{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 6.23), therefore, by the first statement, the graph is recursive, and hence such is f (Proposition 6.10).

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

$$\vec{a} \in R \iff T \vdash \varphi(\Delta(\vec{a})) \iff \ulcorner \varphi(\Delta(\vec{a})) \urcorner \in \ulcorner \mathrm{Thm}(T) \urcorner.$$

By 7.5, $\mathrm{Thm}(T)$ 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 7.4.

[14] As it is written, 7.7 applies only to finite signatures and if a countable field F is not finite, the signature τ_F of the theory of vector spaces over F is infinite. However, we can still assign codes to symbols in τ_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 τ_F as well.

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 6.25 and 6.29, gives the following characterization of recursive functions.

Corollary 7.10. *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 7.11. Recall the primitive recursive function $\mathrm{Sub}_0(a, n)$ that has the property that for every τ_{arthm}-formula φ,

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

For a τ-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 7.12. *Let T be a consistent τ-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 7.5. 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 $n \in \mathbb{N}$,

$$\begin{array}{rcccc} n \in R & \Longrightarrow & \mathrm{Q} \vdash \varphi(\Delta(n)) & \Longrightarrow & T \vdash \pi(\varphi(\Delta(n))) \\ n \notin R & \Longrightarrow & \mathrm{Q} \vdash \neg\varphi(\Delta(n)) & \Longrightarrow & T \vdash \neg\pi(\varphi(\Delta(n))). \end{array}$$

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

$$n \in R \iff U_{\pi,T}(e, n).$$

□

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 7.13. *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. □

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 7.14 (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

$$\neg R(a,a) \iff P(a) \iff R(a,a).$$

□

Corollary 7.15. *For every consistent τ-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 7.12, there is $a \in \mathbb{N}$ such that $P = U_{\pi,T}(a)$, contradicting 7.14. □

As a corollary, we get the following important result:

Theorem 7.16 (Church, 1936)**.** *Any consistent τ-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 7.15. □

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 7.17. *The empty τ_{arthm}-theory is undecidable, i.e.* $\mathrm{Thm}_{\tau_{\mathrm{arthm}}}(\varnothing)$ *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 τ_{arthm}-sentence θ,

$$\mathrm{Q} \vdash \theta \iff \varnothing \vdash \varphi_{\mathrm{Q}} \to \theta.$$

Thus, letting $e = \ulcorner \varphi_{\mathrm{Q}} \urcorner$, we get that for all $a \in \mathbb{N}$,

$$a \in \ulcorner \mathrm{Thm}(\mathrm{Q}) \urcorner \iff \langle \mathrm{SN}(\to), e, a \rangle \in \ulcorner \mathrm{Thm}_{\tau_{\mathrm{arthm}}}(\varnothing) \urcorner.$$

Hence, $\ulcorner \mathrm{Thm}_{\tau_{\mathrm{arthm}}}(\varnothing) \urcorner$ cannot be recursive since otherwise $\ulcorner \mathrm{Thm}(\mathrm{Q}) \urcorner$ would also be recursive, contradicting 7.16. □

8. Quantifier elimination

8.A. Definitions and technicalities

Fix a signature τ.

Definition 8.1. We say that a τ-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

$$T \vdash \forall \vec{x} (\varphi(\vec{x}) \leftrightarrow \psi(x)). \tag{$*$}$$

Assuming that τ 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 τ-structure $\mathbf{A}$ admits (*effective*) *q.e.* if so does $\mathrm{Th}(\mathbf{A})$.

Note that for a τ-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 $\mathbb{FOL}(\tau)$ with propositional symbols for Truth and Falsity. More precisely, just like we always include the binary relation symbol $=$ in $\mathbb{FOL}(\tau)$, we include 0-ary relation symbols T and F, together with the following axioms

(20) Truth: $\mathsf{T} \leftrightarrow \forall x(x = x)$
(21) Falsity: $\mathsf{F} \leftrightarrow \neg T$

Below, we work with this enriched version of $\mathbb{FOL}(\tau)$.

8.B. **Connection with decidability**

There is a strong connection between q.e. and decidability. To see this, consider the set $\mathrm{QFThm}(T) := \{\psi : \psi$ 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 8.2. *Let τ be a finite signature and T a τ-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 8.1, 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[15] that the effectiveness of q.e. comes for free if the theory is decidable; more precisely:

Proposition 8.3. *Let τ be a finite signature and T a τ-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 8.4 (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 8.5 (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 τ_{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 $\mathbf{N}$: $\mathbf{N}_+ := (\mathbb{N}, 0, S, +)$. In one of the previous sections, we defined a (resp. primitive recursive) axiomatization T_+ for $\mathrm{Th}(\mathbf{N}_+)$ is stated that it is complete (and hence decidable). The completeness of T_+ is a consequence of the following.

[15] Many thanks to William Balderrama for pointing this out.

Theorem 8.6 (Presburger)**.** *T_+ admits quantifier elimination.*

To conclude the completeness of T_+ from this note that any model $\mathbf{M}$ of T_+ has a standard part, i.e. $\mathbf{N} \subseteq \mathbf{M}$. Hence $\mathbf{M}$ and $\mathbf{N}$ believe the same q.f. sentences. But every sentence is equivalent (in T_+) to a q.f. sentence, and thus $\mathbf{N} \equiv \mathbf{M}$.

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.

8.C. **Syntactic approach**

Lemma 8.7 (Quantifier elimination test)**.** *A τ-theory T admits (effective) q.e. if and only if for every τ-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:

$$\mathsf{Q}_1 y_1 \mathsf{Q}_2 y_2 ... \mathsf{Q}_k y_k \varphi(\vec{x}, \vec{y}),$$

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 8.8. 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 $I, J, K \subseteq \{0, 1, ..., |\vec{x}| - 1\}$.

Case $I \neq \varnothing$: 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 $I = \varnothing$.

Case $J = \varnothing$ or $K = \varnothing$: Say $J = \varnothing$. 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 $J \neq \varnothing$ and $K \neq \varnothing$: 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

$$\bigwedge_{j \in J, k \in K} x_k < x_j.$$

□

8.D. **Semantic approach**

Let τ be a signature and $\mathbf{A}$ be a τ-structure. For $B \subseteq A$, put $\tau(B) \coloneqq \tau \cup B$, where elements of B are treated as new constant symbols. We define the natural expansion of $\mathbf{A}$ to a $\tau(B)$-structure $\mathbf{A}(B)$ by interpreting symbols in B by themselves, i.e. $\forall b \in B$, $b^{\mathbf{A}(B)} = b$.

Definition 8.9. For a τ-structure $\mathbf{A}$ and $B \subseteq A$, define $\mathrm{Diag}(\mathbf{A}, B)$ as the set of all quantifier free $\tau(B)$-sentences that are true in $\mathbf{A}(B)$, i.e.

$$\mathrm{Diag}(\mathbf{A}, B) \coloneqq \{\psi : \psi \text{ is a q.f. } \tau(B)\text{-sentence and } \mathbf{A}(B) \models \psi\}.$$

When $B = A$, we simply write $\mathrm{Diag}(\mathbf{A})$ instead of $\mathrm{Diag}(\mathbf{A}, A)$.

The following definition gives an equivalent (semantic) condition to quantifier elimination.

Definition 8.10. A τ-theory T is called diagram-complete if for any model $\mathbf{A}$ of T and any $\vec{a} \in A^n$ (for any n), the $\tau(\vec{a})$-theory $T \cup \mathrm{Diag}(\mathbf{A}, \vec{a})$ 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 8.11. *Suppose τ has at least one constant symbol c. Then a τ-theory T admits q.e. if and only if it is diagram-complete.*

Proof. $\Rightarrow$: Put $S \coloneqq T \cup \mathrm{Diag}(\mathbf{A}, \vec{a})$ and let $\varphi(\vec{x})$ be a τ-formula with $\vec{x} = (x_1, ..., x_n)$. 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, $\psi(\vec{a}) \in \mathrm{Diag}(\mathbf{A}, \vec{a})$ or $\neg\psi(\vec{a}) \in \mathrm{Diag}(\mathbf{A}, \vec{a})$, 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 τ-formula with $\vec{x} = (x_1, ..., x_n)$. Take new constant symbols $\vec{d} = (d_1, ..., d_n)$ and put

$$\Gamma(\vec{d}) \coloneqq \left\{\psi(\vec{d}) : \psi \text{ is a q.f. } \tau\text{-formula and } T \vdash \varphi(\vec{d}) \to \psi(\vec{d})\right\}.$$

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 $S(\vec{d}) \coloneqq T \cup \Gamma(\vec{d}) \cup \left\{\neg\varphi(\vec{d})\right\}$ is consistent, so it has a model $\mathbf{A}(\vec{d})$, where $\mathbf{A}$ is its reduct to a τ-structure; in particular, $\mathrm{Diag}(\mathbf{A}, \vec{d}) \supseteq \Gamma(\vec{d})$. Since $\mathbf{A} \models T$ and T is diagram-complete, $S'(\vec{d}) \coloneqq T \cup \mathrm{Diag}(\mathbf{A}, \vec{d})$ is a complete $\tau(\vec{d})$-theory, so $S'(\vec{d})$ proves every sentence in $S(\vec{d})$ because $\mathbf{A}(\vec{d}) \models S'(\vec{d})$ and $\mathbf{A}(\vec{d}) \models S(\vec{d})$; in particular, $S'(\vec{d}) \vdash \neg\varphi(\vec{d})$. Because proofs are finite and $\mathrm{Diag}(\mathbf{A}, \vec{d})$ is closed under conjunctions, there is $\psi(\vec{d}) \in \mathrm{Diag}(\mathbf{A}, \vec{d})$ such that $T \vdash \psi(\vec{d}) \to \neg\varphi(\vec{d})$ (in case T alone proves $\neg\varphi(\vec{d})$, take $\psi(\vec{d}) \doteq \mathsf{T}$). Taking the contrapositive, it follows that $T \vdash \varphi(\vec{d}) \to \neg\psi(\vec{d})$, so $\neg\psi(\vec{d}) \in \Gamma(\vec{d}) \subseteq \mathrm{Diag}(\mathbf{A}, \vec{d})$, contradicting the consistency of $\mathrm{Diag}(\mathbf{A}, \vec{d})$. $\dashv$

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 3.14 and Generalization Axiom (13) now finishes the proof. □

Note that in the definition of diagram-completeness, the model $\mathbf{A}$ is somewhat irrelevant, it is only there to make sure that $\mathrm{Diag}(\mathbf{A}, \vec{a})$ 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 8.12. Let $\vec{d}$ be a vector of distinct constant symbols that do not occur in τ. A set $\Gamma(\vec{d})$ of quantifier free $\tau(\vec{d})$-sentences is called a T-diagram if $T \cup \Gamma(\vec{d})$ is consistent and for every q.f. $\tau(\vec{d})$-sentence ψ, $\psi \in \Gamma(\vec{d})$ or $\neg\psi \in \Gamma(\vec{d})$.

Lemma 8.13. *A τ-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 $\tau(\vec{d})$-theory.*

Proof. $\Leftarrow$ follows from the Soundness of $\mathbb{FOL}$ and $\Rightarrow$ follows from the Completeness of $\mathbb{FOL}$. □

8.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 5.5.

Proposition 8.14. *For every* ACF*-diagram $\Gamma(\vec{d})$,* ACF $\cup\, \Gamma(\vec{d})$ *is a κ-categorical $\tau_{ring}(\vec{d})$-theory, for every uncountable cardinal κ.*

Proof. Let $\mathbf{K}_1, \mathbf{K}_2 \vDash \mathrm{ACF} \cup \Gamma(\vec{d})$ with $|K_1| = |K_2| = \kappa$. Note that $\mathbf{K}_1, \mathbf{K}_2$ have the same characteristic since it is expressible by a q.f. τ_{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 $\mathbf{K}_i$, i.e. the substructures of $\mathbf{K}_i$ generated by $\varnothing$. (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 $\vec{a} = \vec{d}^{\mathbf{K}_1}$, $\vec{b} = \vec{d}^{\mathbf{K}_2}$, 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. $\tau_{\mathrm{ring}}(\vec{d})$-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 8.15. ACF *admits quantifier elimination.*

Proof. Follows from 8.13 and 8.11. □

Corollary 8.16. *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 8.17. 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.

8.F. **Model-completeness**

The following is a very useful notion that is slightly weaker than quantifier elimination.

Definition 8.18. A τ-theory T is called *model-complete* if $\mathbf{A} \subseteq \mathbf{B}$ implies $\mathbf{A} \preceq \mathbf{B}$, for all $\mathbf{A}, \mathbf{B} \vDash T$.

Proposition 8.19. *Quantifier elimination implies model-completeness.*

Proof. Suppose T admits q.e. and $\mathbf{A} \subseteq \mathbf{B}$, where $\mathbf{A}, \mathbf{B} \vDash T$. Because $\mathbf{A}$ and $\mathbf{B}$ agree on the q.f. formulas about the elements of A, and every formula is equivalent to a q.f. formula (in T), $\mathbf{A}$ and $\mathbf{B}$ agree on all formulas about the elements of A. A τ-structure $\mathbf{M}$ is called *model-complete* if such is $\mathrm{Th}(\mathbf{M})$. □

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 $\mathrm{Diag}(\mathbf{A})$ for $\mathrm{Diag}(\mathbf{A}, A)$, the following proposition justifies the terminology with regards to diagram-completeness and highlights the difference with quantifier elimination.

Proposition 8.20. *For a τ-theory T, the following are equivalent:*

(1) *T is model-complete.*
(2) *For every model $\mathbf{A} \vDash T$, $T \cup \mathrm{Diag}(\mathbf{A})$ is a complete $\tau(A)$-theory.*
(3.a) *Every τ-formula $\varphi(\vec{x})$ is equivalent in T to a universal formula.*
(3.b) *Every τ-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 8.11 and we leave it as a (good) exercise. □

8.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 8.21 (Weak Form). *Let F be an algebraically closed field and I be a proper ideal in the polynomial ring $F[t_1,...,t_n]$. 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 $f(t_1,...,t_n) \in F[t_1,...,t_n]$.*

Proof. Take a maximal ideal M containing I (exists by Zorn's lemma) and put

$$K := F[t_1,...,t_n]/M.$$

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 $f(t_1,...,t_n) \in M$, let $f(x_1,...,x_n)$ be the polynomial obtained from $f(t_1,...,t_n)$ by replacing t_i with variables x_i of $\mathbb{FOL}(\tau_{\text{ring}})$. Then, by the definition of K, for all such $f \in M$, $f(\vec{b}) = 0$, where $\vec{b} = (t_1 + M,...,t_n + M) \in K$. (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 $F[t_1,...t_n]$ is finitely generated, so M is generated by some $f_1,...,f_m \in F[t_1,...t_n]$. Thus all polynomials in M having a common root is equivalent to $f_1,...,f_m$ having a common root. Put

$$\varphi(\vec{a}) \doteq \exists\vec{x}\bigwedge_{i=1}^{m}(f_i(\vec{x}) = 0),$$

where $\vec{a} \in F^k$ is a tuple containing all coefficients of $f_1,...,f_m$. By model-completeness of ACF, because $\mathbf{F} \subseteq \mathbf{L}$ and $\mathbf{F}, \mathbf{L} \models$ ACF, we have $\mathbf{F} \preceq \mathbf{L}$. Hence $\mathbf{F} \models \varphi(\vec{a})$ because $\mathbf{L} \models \varphi(\vec{a})$, and thus $f_1,...,f_m$ 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, and write $J(\vec{a}) = 0$, 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 8.22 (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 $f_1,...,f_k \in J$ and $g_1(t),...,g_{k+1}(t) \in F[\vec{x}, t]$ such that

$$g_1(t)f_1 + ... + g_k(t)f_k + g_{k+1}(t)(1 - tf) = 1.$$

Assuming that $f \neq 0$ (otherwise, we are done), plug in $t = 1/f$ and get

$$g_1(1/f)f_1 + ... + g_k(1/f)f_k = 1.$$

Multiplying both sides with f^m for large enough $m \in \mathbb{N}$, we get

$$\tilde{g}_1 f_1 + ... + \tilde{g}_k f_k = f^m,$$

where $\tilde{g}_1, ..., \tilde{g}_k$ 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