Foundations of applied arithmetic

Eric Johannesson

February 4, 2026

Abstract

In order to apply theories of pure mathematics for reasoning about non-mathematical objects (e.g. tables and chairs), various bridge principles need to be assumed, connecting claims about the mathematical realm with the non-mathematical realm. A theory of applied mathematics is comprised of both pure mathematics and such bridge principles. Following Hartry Field, I will assume that what distinguishes applied mathematics from, say, physics, is that applied mathematics is conservative over any set of non-mathematical assumptions, in the following sense: if a non-mathematical conclusion logically follows from a set of non-mathematical assumptions together with a theory of applied mathematics, it already follows from the non-mathematical assumptions alone. This is the sense in which mathematics can be regarded as a priori - or, rather, as arbitrary. The usefulness of applied mathematics can be understood in terms of how difficult it is to reach those conclusions without it - more specifically, in terms of the length of the shortest proofs. I will investigate bridge principles for applied arithmetic in particular, to see how they may facilitate reasoning about non-mathematical objects in the above sense. Ultimately, the goal is to understand how arithmetic can be useful despite being arbitrary, and if perhaps there is something special about standard arithmetic that makes it particularly useful.

1 Preliminary observations

Consider the following problem of elementary school arithmetic:

(1) Suppose that Alice buys exactly one apple, and that Ben buys exactly two apples. Provided that no apple is bought by both Alice and Ben, exactly how many apples are bought by either Alice or Ben?

Ironically, no arithmetic is required for solving this problem. It can be construed as a purely logical one. Let 'A' and 'B' be unary predicates, where 'A' applies to all and only apples bought by Alice, and 'B' applies to all and only apples bought by Ben. The assumptions of the problem can then be expressed in first-order logic:

(2) a. $\exists x(Ax \wedge \forall y(Ay \rightarrow x = y))$
b. $\exists x \exists y(x \neq y \wedge Bx \wedge By \wedge \forall z(Bz \rightarrow x = z \vee y = z))$
c. $\neg \exists x(Ax \wedge Bx)$

Moreover, the answer (namely that exactly three apples are bought by either Alice or Ben) can be derived as logical consequence of these assumptions:

(3) $\exists x \exists y \exists z(x \neq y \wedge x \neq z \wedge y \neq z \wedge (Ax \vee Bx) \wedge (Ay \vee By) \wedge (Az \vee Bz) \wedge \forall u(Au \vee Bu \rightarrow x = u \vee y = u \vee z = u))$

Nevertheless, elementary school children are expected to solve the problem by applying knowledge of the arithmetical fact that $1 + 2 = 3$. Some philosophers have taken this to suggest that the sentence '$1 + 2 = 3$' really is nothing but a logically valid principle in disguise.[1] Remaining within the confines of first-order logic, the sentence could perhaps be regarded as an abbreviation of the theory consisting of – for each formula φ and ψ and variable α – the universal closure of

(4) $\ulcorner \exists_{=1} \alpha \varphi \wedge \exists_{=2} \alpha \psi \wedge \neg \exists \alpha(\varphi \wedge \psi) \rightarrow \exists_{=3} \alpha(\varphi \vee \psi) \urcorner$

Using this logically valid theory, the answer to the problem can be derived from the assumptions in one step by modus ponens. Hence, this interpretation of '$1 + 2 = 3$' explains how knowledge of the fact that $1 + 2 = 3$ would be applicable to the problem of Alice and Ben.

More generally, the idea would be to interpret each arithmetical sentence $\ulcorner a + b = c \urcorner$ (where a, b, and c are numerals) as the theory consisting of (for each formula φ and ψ and variable α) the universal closure of

(5) $\ulcorner \exists_{=a} \alpha \varphi \wedge \exists_{=b} \alpha \psi \wedge \neg \exists \alpha(\varphi \wedge \psi) \rightarrow \exists_{=c} \alpha(\varphi \vee \psi) \urcorner$

Alternatively, leaving the confines of first-order logic, the idea would be to interpret each sentence $\ulcorner a + b = c \urcorner$ as the following second-order sentence instead:

(6) $\ulcorner \forall X \forall Y[\exists_{=a} x Xx \wedge \exists_{=b} x Yx \wedge \neg \exists x(Xx \wedge Yx) \rightarrow \exists_{=c} x(Xx \vee Yx)] \urcorner$

Either way, there is a problem of how to extend the interpretation to arbitrary arithmetical sentences. The problem pertains not only to quantified sentences like 'Every natural number is either odd or even', but arises already for the quantifier-free fragment of arithmetic. Assuming that all arithmetical truths are to be interpreted as logical truths (or as logically true theories), the problem is that the extension cannot be compositional. In particular, the interpretation of the negation of an arithmetical sentence cannot always be the negation of the interpretation of the negated sentence. In the first-order case, there is of course already the issue of negating an infinite set of sentences, but the problem remains even in the second-order case. According to the recipe given, as Wang (1990, p. 35) once pointed out, a false arithmetical sentence like '$1 + 2 = 2$' will be interpreted as the second-order sentence

(7) $\forall X \forall Y[\exists_{=1} x Xx \wedge \exists_{=2} x Yx \wedge \neg \exists x(Xx \wedge Yx) \rightarrow \exists_{=2} x(Xx \vee Yx)]$

---

[1] I think Husserl (2003, pp. 191–192) can be taken to say something to that effect when he writes:

> $5 + 5 = 10$ means the same as: a group - anyone, whichever it may be - falling under the concept *five*, and any *other* group falling under that same concept, when united yield a group falling under the concept *ten*.

Obviously, in order for this claim to be even remotely plausible, 'other' must be taken to mean 'disjoint'.

This sentence, however, is not logically false. In fact, it is true in every model with at most two elements in the domain of the first-order quantifiers. Assuming compositionality, the arithmetical truth '$1 + 2 \neq 2$' (the negation of '$1 + 2 = 2$') will then be interpreted as the negation of (7), which is not a logical truth.

Given these difficulties, we shall proceed by taking arithmetical sentences at face value, namely as statements about natural numbers. But we still need an account of how arithmetical facts pertain to facts about Alice and Ben. To make the arithmetical connection more conspicuous, we may reformulate the problem:

(8) Suppose that the number of apples bought by Alice is 1, and that the number of apples bought by Ben is 2. Provided that no apple is bought by both Alice and Ben, what is the number of apples bought by either Alice or Ben?

For a natural formalization of this problem, we may want to extend the syntax of first-order logic with a term-forming variable-binding operator '#': for any variable α and formula φ, we declare that $\ulcorner \# \alpha \varphi \urcorner$ is a term with the same free variable occurrences as φ, except for those of α, which are bound by term's leftmost occurrence of '#'. Although $\ulcorner \# \alpha \varphi \urcorner$ shall be intended to refer to the number of objects that, when assigned to α, satisfy φ, we assume only that the logical rules of inference apply to it just as they apply to any other term. In addition, we need a language for speaking about natural numbers. Let 'N' be a unary predicate intended to apply to all and only natural numbers, let '0' be a constant intended to refer to the number 0, let 's' be a unary function symbol intended to refer to the successor function, and let '+' and '×' be binary function symbols intended to refer to the addition and multiplication function, respectively. The numerals in this language (corresponding to the decimals '0', '1', '2', etc.) are the expressions '0', '$s(0)$', '$s(s(0))$', etc. The assumptions in (8) may then be expressed as follows:

(9) a. $\# x Ax = s(0)$
b. $\# x Bx = s(s(0))$
c. $\neg \exists x (Ax \wedge Bx)$

But these assumptions do not logically entail the answer, namely that the number of apples bought by either Alice or Ben is 3:

(10) $\# x (Ax \vee Bx) = s(s(s(0)))$

As we shall see, given that the classical rules of inference apply to '#'-terms just as they apply to any other terms, one can find an interpretation of our formal language (including the '#'-terms) consistent with the rules of inference under which the assumptions are true but the conclusion is false. Hence, in order to logically derive the answer, we need to make some further assumptions. But a pure theory of arithmetic will not help, at least not by itself (for a proof of this claim, see Remark 3.2 below). Assuming, for instance, that $1 + 2 = 3$, i.e.

(11) $s(0) + s(s(0)) = s(s(s(0)))$

we still only get

(12) $\# x Ax + \# x Bx = s(s(s(0)))$

What we need, in addition, are assumptions allowing us to derive things like

(13) $\#x(Ax \vee Bx) = \#xAx + \#xBx$

What might these assumptions be?

Before we start making suggestions, we first need to acknowledge the possibility that not all #-terms refer to natural numbers. Consider, for instance, the term $\#x(x = x)$. If there are infinitely many things, the numbers of things that are self-identical will not be a natural number.

Secondly, we need some idea of what a pure theory of arithmetic might look like. Let N be a unary predicate whose intended interpretation is *being a natural number*, and consider the following theory of pure arithmetic in the vocabulary $\{N, 0, s, +, \times\}$:

Pure Peano Arithmetic:

1. $N0$
2. $\forall x(Nx \to Ns(x))$
3. $\forall x(Nx \to s(x) \neq 0)$
4. $\forall x \forall y(Nx \wedge Ny \wedge s(x) = s(y) \to x = y)$
5. $\forall x \forall y(Nx \wedge Ny \to N(x + y))$
6. $\forall x(Nx \to x + 0 = x)$
7. $\forall x \forall y(Nx \wedge Ny \to x + s(y) = s(x + y))$
8. $\forall x \forall x(Nx \wedge Ny \to N(x \times y))$
9. $\forall x \forall y(Nx \wedge Ny \to x \times 0 = 0)$
10. $\forall x \forall y(Nx \wedge Ny \to x \times s(y) = s(x \times y) + x)$
11. $\varphi(0) \wedge \forall x(Nx \to (\varphi(x) \to \varphi(s(x)))) \to \forall x(Nx \to \varphi(x))$

Let us call it PA_N, for short, and let us refer to its i:th axiom as $\mathsf{PA}_N(i)$. Moreover, for each natural number n, let us define a unique term $\underline{n}$ in our formal language intended to refer to that number (its *numeral*, as it were). We do this recursively:

$$\underline{0} = 0$$
$$\underline{n+1} = s(\underline{n})$$

Returning to the problem of Alice and Ben, a general principle of additivity naturally comes to mind. The principle (which is really a set of first-order sentences) includes, for each formula φ, ψ and variable α, the universal closure of the following formula:

(14) $\ulcorner N\#\alpha\varphi \wedge N\#\alpha\psi \wedge \neg\exists\alpha(\varphi \wedge \psi) \to \#\alpha(\varphi \vee \psi) = \#\alpha\varphi + \#\alpha\psi \urcorner$

Together with PA_N, the answer can now be derived. As an instance of (14), we have

(15) $N\#xAx \wedge N\#xBx \wedge \neg\exists x(Ax \wedge Bx) \to \#x(Ax \vee Bx) = \#xAx + \#xBx$

The antecedent follows from (9), $\mathsf{PA}_N(1)$ and $\mathsf{PA}_N(2)$, and the consequent is (13). Moreover, since (11) follows from PA_N, and (12) follows from (11), (9-a) and (9-b), we obtain (10).

Principles such as (14), whose instances may contain a mix of both arithmetical and non-arithmetical vocabulary, are often called *bridge principles*. As the name suggest, they are required for establishing logical connections between, on the one hand, assumption of pure arithmetic, and, on the other hand, assumptions involving non-arithmetical matters (e.g. apples). As we have seen, it is only by assuming such principles that we may apply pure arithmetic to non-mathematical objects, such as apples. Simply put, applied arithmetic consists of pure arithmetic plus bridge principles. Although no arithmetic is *required* for deriving (3) from (2), using a theory of applied arithmetic could (at least in principle) make it easier.

Here are some obvious candidates (including the one already mentioned), presented schematically:

Extensionality: $\forall x(\varphi \leftrightarrow \psi) \rightarrow \#x\varphi = \#x\psi$

Conjunctive comprehension: $N\#x\varphi \rightarrow N\#x(\varphi \wedge \psi)$

Disjunctive comprehension: $N\#x\varphi \wedge N\#x\psi \rightarrow N\#x(\varphi \vee \psi)$

Zero: $\#x(x \neq x) = 0$

Successor: $N\#x\varphi \rightarrow \forall y(\neg\varphi(y/x) \rightarrow \#x(\varphi \vee x = y) = s(\#x\varphi))$, with y not in φ.

Additivity: $N\#x\varphi \wedge N\#x\psi \wedge \neg\exists x(\varphi \wedge \psi) \rightarrow \#x(\varphi \vee \psi) = \#x\varphi + \#x\psi$

What the bridge principles listed above all have in common is that they are true when we assume the intended interpretation of the '#'-operator and the vocabulary of arithmetic, without assuming anything about the interpretation of the non-arithmetical vocabulary. In fact, having bridge principles with this property will ensure that applied arithmetic is *conservative* over purely non-arithmetical claims, in the following sense: If N is a set of pure non-arithmetical claims, and C is a pure non-arithmetical claim, A is a pure theory of arithmetic, and B is a set of bridge principles, then C follows from $A + B + N$ only if C already follows from N.[2]

Here are two more candidates with this property:

Equinumerosity: $\forall x(\varphi \rightarrow \exists! y(\psi \wedge \chi)) \wedge \forall y(\psi \rightarrow \exists! x(\varphi \wedge \chi)) \rightarrow \#x\varphi = \#y\psi$, with x not free in ψ, and y not free in φ.

Correspondence: $\exists_{=n} x\varphi \leftrightarrow \#x\varphi = \underline{n}$.

If we allow the '#'-operator to bind finite sequences of variables (with the intended interpretation of $\ulcorner\#\alpha_1 \ldots \alpha_n\varphi\urcorner$ being $\ulcorner$*the number of sequences of objects* $\alpha_1, \ldots, \alpha_n$ *such that* $\varphi\urcorner$), the following principle for multiplication naturally comes to mind:

Multiplicativity: $N\#x\varphi \wedge N\#y\psi \rightarrow \#xy(\varphi \wedge \psi) = \#x\varphi \times \#y\psi$, with x not free in ψ, and y not free in φ.

[2] According to Field (1980), conservativity is what distinguishes applied mathematics from, say, physics, and it is precisely in this sense that mathematics may be regarded as a priori.

So do the following comprehension principles:

Universal comprehension: $N\#xy\varphi \to N\#x\forall y\varphi$.

Existential comprehension: $N\#xy\varphi \to N\#x\exists y\varphi$.

General comprehension: $N\#xy\varphi \to N\#x\varphi$.

At this point, we should pause and ask ourselves: how many bridge principles do we need? And, perhaps more importantly, what do we need them for? One could ask these questions about about arithmetic as well. For instance, do the axioms of Peano arithmetic suffice to settle every question of pure arithmetic? Gödel famously showed that they do not. More generally, he showed that *true arithmetic* (the set of sentences of first-order arithmetic that are true in the standard model) is not *axiomatizable*: there is no decidable set of sentences from which all and only arithmetical truths can be derived.

Let us stipulate that a sentence is **standardly valid** just in case it is true under every interpretation that is standard with respect to the arithmetical vocabulary and the #-operator (for a precise definition, see section 5). The principles listed so far should all strike us a standardly valid. Moreover, the set of standard validities contains all the truths of pure arithmetic. Since being a sentence of pure arithmetic is a decidable matter, it follows from Gödel's theorem that the standard validities are not axiomatizable either. Still, there is an interesting question to be asked concerning their *relative* axiomatizability:

(16) Is there a decidable set of bridge principles from which, together with the truths of pure arithmetic, all and only standard validities can be derived?

The answer to the question, however, is negative. To see why, let $L_A = \{N, 0, s, +, \times\}$, let L be a vocabulary such that $L \cap L_A = \{N\}$, and let $L\#$ be the extension of L containing all the new function symbols simulating the '#'-operator. Thus, $L\#$ is the vocabulary of a language in the standard syntax interpreting a language with vocabulary L in the extended syntax. Let T be the set of $L\#$-sentences that are true under every interpretation of $L\#$ that is standard with respect to # and N. Clearly, T is just the set of $L\#$-sentences that are true under every interpretation $\mathcal{M}$ of $L\#$ such that, for each $L\#$-formular $\varphi(x)$, we have that $N\#x\varphi(x)$ is true in $\mathcal{M}$ just in case only finitely many elements of $\mathcal{M}$ satisfy $\varphi(x)$. Hence, T is essentially the theory of the 'finitely many'-quantifier for L. In the analytical hierarchy, this theory (i.e. the corresponding set of Gödel numbers) is known to be Π^1_1-hard (i.e. every Π^1_1-problem can be reduced to it).[3] Clearly, a sentence belongs to T just in case it is a standardly valid $L\#$-sentence. It follows that standard validity is also Π^1_1-hard. The truths of pure arithmetic, on the other hand, is already in Δ^1_1, i.e. both in Π^1_1 and Σ^1_1. Since some Π^1_1-problems cannot be reduced to Δ^1_1-problems, it follows that standard validity is not a Δ^1_1-problem. Since relations in Δ^1_1 are closed under definitions with boolean operators and first-order quantifiers, this contradicts the claim that there is a decidable set of bridge principles from which, together with the truths of pure arithmetic, all and only standard validities can be derived.

[3] van Benthem and Icard (2023)

2 Pure arithmetic

The subject matter of arithmetic is the natural numbers $0, 1, 2, 3$, etc. Intuitively, a *pure* theory of arithmetic is one that only talks natural numbers. We can make this notion of pureness precise by introducing a unary predicate N, whose intended interpretation is *being a natural number*. Let L be a first-order vocabulary containing N.

Definition 2.1 (Reduct)**.** Let $L \subseteq L'$, and let $\mathcal{M}$ be an L'-model. The *L-reduct* of $\mathcal{M}$ (written $\mathcal{M}|L$) is the L-model with the same domain as $\mathcal{M}$ such that, for any symbol $u \in L$, we have $u^{\mathcal{M}|L} = u^{\mathcal{M}}$.

Definition 2.2 (Part)**.** Let $\mathcal{M}$ be an L-model. The N-part of $\mathcal{M}$ (written $\mathcal{M} \upharpoonright N$), is *defined* just in case

(i) $\mathcal{M} \vDash \exists x Nx$,

(ii) for any constant $c \in L$, we have $\mathcal{M} \vDash Nc$, and

(iii) for any n-place function symbol $f \in L$, we have $\mathcal{M} \vDash \forall x_1 \ldots \forall x_n (Nx_1 \wedge \ldots \wedge Nx_n \rightarrow Nf(x_1, \ldots, x_n))$,

Furthermore, provided the above three conditions are satisfied, we define $\mathcal{M} \upharpoonright N$ as follows:

(i) Let $|\mathcal{M} \upharpoonright N| = N^{\mathcal{M}}$.

(ii) For any constant $c \in L$, let $c^{\mathcal{M}\upharpoonright N} = c^{\mathcal{M}}$.

(iii) For any n-place function symbol $f \in L$, let $f^{\mathcal{M}\upharpoonright N} = f^{\mathcal{M}} \cap (|\mathcal{M} \upharpoonright N|^n \times |\mathcal{M}_N|)$.

(iv) For any n-place predicate symbol $P \in L$, let $P^{\mathcal{M}\upharpoonright N} = P^{\mathcal{M}} \cap |\mathcal{M} \upharpoonright N|^n$.

Definition 2.3 (Pure theory of arithmetic)**.** We say that a first-order L-theory T is a *pure theory of arithmetic* just in case, for any L-model $\mathcal{M}$,

(i) if $\mathcal{M} \vDash T$, then $\mathcal{M} \upharpoonright N$ is defined, and

(ii) if $\mathcal{M} \upharpoonright N$ is defined, we have $\mathcal{M} \vDash T$ if and only if $\mathcal{M} \upharpoonright N \vDash T$.

Somewhat less precise, one might say that the truth of a pure theory of arithmetic in a model only depends on the natural number part of that model.

By soundness and completeness, we may characterize this notion of pureness syntactically.

Definition 2.4 (Relativization)**.** For any L-formula φ, we define its *relativization* $[\varphi]_N$ to N recursively:

(i) $[s = t]_N = s = t$

(ii) $[P\bar{t}]_N = P\bar{t}$

(iii) $[\neg\varphi]_N = \neg[\varphi]_N$

(iv) $[\varphi \to \psi]_N = [\varphi]_N \to [\psi]_N$

(v) $[\forall x\varphi]_N = \forall x(Nx \to [\varphi]_N)$

(vi) $[\exists x\varphi]_N = \exists x(Nx \wedge [\varphi]_N)$

For instance, we have

$$[\forall x(Px \to \exists yQxy)]_N = \forall x(Nx \to (Px \to \exists y(Ny \wedge Qxy)))$$

If there is no risk of ambiguity, we may write φ_N instead of $[\varphi]_N$. For any L-theory T, we define

$$\begin{aligned} T_N =& \{\exists x Nx\} \cup \{Nc : c \in L\} \\ & \cup \{\forall x_1 \ldots \forall x_n(Nx_1 \wedge \ldots \wedge Nx_n \to Nf(x_1, \ldots, x_n)) : f \in L\} \\ & \cup \{\varphi_N : \varphi \in T\} \end{aligned}$$

Lemma 2.1. *Let $\mathcal{M}$ be an L-model for which $\mathcal{M} \restriction N$ is defined, and let φ be an L-sentence. Then we have $\mathcal{M} \vDash \varphi_N$ just in case $\mathcal{M} \restriction N \vDash \varphi$.*

Proof. Let $\mathcal{M}$ be an L-model for which $\mathcal{M} \restriction N$ is defined. Hence,

(17) a. $\mathcal{M} \vDash \exists x Nx$,
b. for any constant $c \in L$, we have $\mathcal{M} \vDash Nc$, and
c. for any n-place function symbol $f \in L$, we have $\mathcal{M} \vDash \forall x_1 \ldots \forall x_n(Nx_1 \wedge \ldots \wedge Nx_n \to Nf(x_1, \ldots, x_n))$,

and $\mathcal{M} \restriction N$ is defined by

(18) a. $|\mathcal{M} \restriction N| = N^{\mathcal{M}}$,
b. for any constant $c \in L$, $c^{\mathcal{M}\restriction N} = c^{\mathcal{M}}$,
c. for any n-place function symbol $f \in L$, $f^{\mathcal{M}\restriction N} = f^{\mathcal{M}} \cap (|\mathcal{M} \restriction N|^n \times |\mathcal{M} \restriction N|)$, and
d. for any n-place predicate symbol $P \in L$, $P^{\mathcal{M}\restriction N} = P^{\mathcal{M}} \cap |\mathcal{M} \restriction N|^n$.

Let X be the set of variables. First we show

(19) For any assignment $g : X \to |\mathcal{M} \restriction N|$ and L-term t, we have $t^{\mathcal{M},g} = t^{\mathcal{M}\restriction N,g}$.

by induction on the complexity of t. If t is a variable or constant, the claim obviously holds, in the latter case by (17-b) and (18-b). Assume, as induction hypothesis, that the claim holds for $t_1, \ldots, t_n$. Let $f \in L$ be an n-place function symbol. We get

$$\begin{aligned} f(t_1, \ldots, t_n)^{\mathcal{M},g} =& f^{\mathcal{M}}(t_1^{\mathcal{M},g}, \ldots, t_n^{\mathcal{M},g}) \\ =& f^{\mathcal{M}}(t_1^{\mathcal{M}\restriction N,g}, \ldots, t_n^{\mathcal{M}\restriction N,g}) && \text{by ind. hyp.} \\ =& f^{\mathcal{M}\restriction N}(t_1^{\mathcal{M}\restriction N,g}, \ldots, t_n^{\mathcal{M}\restriction N,g}) && \text{by (17-c) and (18-c)} \\ =& f(t_1, \ldots, t_n)^{\mathcal{M}\restriction N,g} \end{aligned}$$

Next, we show that

(20) For any assignment $g : X \to |\mathcal{M} \upharpoonright N|$ and L-formula φ, we have $\mathcal{M}, g \vDash \varphi_N$ iff $\mathcal{M} \upharpoonright N, g \vDash \varphi$.

by induction on the complexity of φ. For the base cases, we have

$$\begin{aligned} \mathcal{M}, g \vDash [s = t]_N &\text{ iff } \mathcal{M}, g \vDash s = t \\ &\text{ iff } s^{\mathcal{M},g} = t^{\mathcal{M},g} \\ &\text{ iff } s^{\mathcal{M}\upharpoonright N,g} = t^{\mathcal{M}\upharpoonright N,g} && \text{by (19)} \\ &\text{ iff } \mathcal{M} \upharpoonright N, g \vDash s = t \end{aligned}$$

and

$$\begin{aligned} \mathcal{M}, g \vDash [Pt_1 \ldots t_n]_N &\text{ iff } \mathcal{M}, g \vDash Pt_1 \ldots t_n \\ &\text{ iff } P^{\mathcal{M}}(t_1^{\mathcal{M},g}, \ldots, t_n^{\mathcal{M},g}) \\ &\text{ iff } P^{\mathcal{M}}(t_1^{\mathcal{M}\upharpoonright N,g}, \ldots, t_n^{\mathcal{M}\upharpoonright N,g}) && \text{by (19)} \\ &\text{ iff } P^{\mathcal{M}\upharpoonright N}(t_1^{\mathcal{M}\upharpoonright N,g}, \ldots, t_n^{\mathcal{M}\upharpoonright N,g}) && \text{by (18-d) and (19)} \\ &\text{ iff } \mathcal{M} \upharpoonright N, g \vDash Pt_1 \ldots t_n \end{aligned}$$

Assume, as induction hypothesis, that the claim holds for formulas φ and ψ. We get

$$\begin{aligned} \mathcal{M}, g \vDash [\neg\varphi]_N &\text{ iff } \mathcal{M}, g \vDash \neg[\varphi]_N \\ &\text{ iff } \mathcal{M}, g \nvDash [\varphi]_N \\ &\text{ iff } \mathcal{M} \upharpoonright N, g \nvDash \varphi && \text{by ind. hyp.} \\ &\text{ iff } \mathcal{M} \upharpoonright N, g \vDash \neg\varphi \end{aligned}$$

and

$$\begin{aligned} \mathcal{M}, g \vDash [\varphi \wedge \psi]_N &\text{ iff } \mathcal{M}, g \vDash [\varphi]_N \wedge [\psi]_N \\ &\text{ iff } \mathcal{M}, g \vDash [\varphi]_N \text{ and } \mathcal{M}, g \vDash [\psi]_N \\ &\Leftrightarrow \mathcal{M} \upharpoonright N, g \vDash \varphi \text{ and } \mathcal{M} \upharpoonright N, g \vDash \psi && \text{by ind. hyp.} \\ &\Leftrightarrow \mathcal{M} \upharpoonright N, g \vDash \varphi \wedge \psi \end{aligned}$$

and

$$\begin{aligned} \mathcal{M}, g \vDash [\forall x\varphi]_N &\text{ iff } \mathcal{M}, g \vDash \forall x(Nx \to \varphi_N) \\ &\text{ iff } \mathcal{M}, g_{a\to x} \vDash Nx \to \varphi_N \text{ for all } a \in |\mathcal{M}| \\ &\text{ iff } \mathcal{M}, g_{a\to x} \vDash \varphi_N \text{ for all } a \in |\mathcal{M} \upharpoonright N| && \text{by (18-a)} \\ &\text{ iff } \mathcal{M} \upharpoonright N, g_{a\to x} \vDash \varphi \text{ for all } a \in |\mathcal{M} \upharpoonright N| && \text{by ind. hyp.} \\ &\text{ iff } \mathcal{M} \upharpoonright N, g \vDash \forall x\varphi \end{aligned}$$

It now follows from (20) that, for any L-sentence φ, we have $\mathcal{M} \vDash \varphi_N$ iff $\mathcal{M} \upharpoonright N \vDash \varphi$. □

N-relativized sentences only talk about natural numbers, in the following precise sense:

Corollary 2.1. *Let $\mathcal{M}$ and $\mathcal{M}'$ be L-models for which $\mathcal{M} \upharpoonright N$ and $\mathcal{M}'_N$ are defined, and let φ be an L-sentence. If $\mathcal{M} \upharpoonright N = \mathcal{M}'_N$, we have $\mathcal{M} \models \varphi_N$ just in case $\mathcal{M}' \models \varphi_N$.*

Moreover, we can characterize pureness syntactically:

Theorem 2.1. *An L-theory T is a pure theory of arithmetic just in case T and T_N are logically equivalent.*

Proof. Assume that T is a pure theory of arithmetic, and let $\mathcal{M}$ be an L-model. If $\mathcal{M} \upharpoonright N$ is defined, we have $\mathcal{M} \models T$ just in case $\mathcal{M} \upharpoonright N \models T$, which by Lemma holds just in case $\mathcal{M} \models T_N$. If $\mathcal{M} \upharpoonright N$ is not defined, we have $\mathcal{M} \not\models T$ and $\mathcal{M} \not\models T_N$. Hence, T and T_N are logically equivalent. For the other direction, assume that T and T_N are logically equivalent, and let $\mathcal{M}$ be an L-model. If $\mathcal{M} \upharpoonright N$ is defined, we have $\mathcal{M} \models T$ just in case $\mathcal{M} \models T_N$, which by Lemma holds just in case $\mathcal{M} \upharpoonright N \models T$. If $\mathcal{M} \upharpoonright N$ is not defined, we have $\mathcal{M} \not\models T_N$ and thus $\mathcal{M} \not\models T$. Hence, T is a pure theory of arithmetic. □

We also observe that

Theorem 2.2. *T_N is logically equivalent to $\{\varphi_N : T \vdash \varphi\}$.*

Proof. For left to right, assume that $\mathcal{M} \models T_N$. Then $\mathcal{M} \upharpoonright N$ is defined and, by Lemma, $\mathcal{M} \upharpoonright N \models T$. If follows that $\mathcal{M} \upharpoonright N \models \{\varphi : T \vdash \varphi\}$. Hence, by Lemma, $\mathcal{M} \models \{\varphi_N : T \vdash \varphi\}$.

For right to left, assume that $\mathcal{M} \models \{\varphi_N : T \vdash \varphi\}$. Since

- $T \vdash \exists x(x = x)$,
- $T \vdash \{\exists x(x = c) : c \in L\}$, and
- $T \vdash \{\forall x_1 \ldots x_n \exists y f(x_1, \ldots, x_n) = y : f \in L\}$,

we get

- $\exists x(Nx \wedge x = x) \in \{\varphi_N : T \vdash \varphi\}$,
- $\{\exists x(Nx \wedge x = c) : c \in L\} \subseteq \{\varphi_N : T \vdash \varphi\}$, and
- $\{\forall x_1 \ldots x_n(Nx_1 \wedge \ldots \wedge Nx_n \to \exists y(Ny \wedge f(x_1, \ldots, x_n) = y)) : f \in L\} \subseteq \{\varphi_N : T \vdash \varphi\}$,

from which it follows that

- $\{\varphi_N : T \vdash \varphi\} \vdash \exists x Nx$,
- $\{\varphi_N : T \vdash \varphi\} \vdash \{Nc : c \in L\}$, and
- $\{\varphi_N : T \vdash \varphi\} \vdash \{\forall x_1 \ldots \forall x_n(Nx_1 \wedge \ldots \wedge Nx_n \to Nf(x_1, \ldots, x_n)) : f \in L\}$.

Hence, $\mathcal{M} \models T_N$. □

3 Interpreting the extended syntax

In this section, we shall find an interpretation of the extended syntax with respect to which the rules of inference are sound and complete. The idea is to translate the extended syntax into the standard syntax. We achieve this by extending any given vocabulary L to a vocabulary $L\#$ containing infinitely (but countably) many new function symbols. We shall then define a translation τ from L-formulas in the extended syntax to $L\#$-formulas in the standard syntax, and show that, for any set of L-sentences Γ and sentence φ in the extended syntax, we have

(21) $\Gamma \vdash \varphi$ just in case $\tau[\Gamma] \vdash \tau[\varphi]$

where $\tau[\Gamma] = \{\tau[\varphi] : \varphi \in \Gamma\}$. By soundness and completeness of the standard syntax and semantics, this will allow us to conclude that

(22) $\Gamma \vdash \varphi$ if and only if $\tau[\Gamma] \vDash \tau[\varphi]$

3.1 Extending the vocabulary

Say that an occurrence o of a term is **free** in an expression e just in case no subterm of o is a variable bound in e by a quantifier outside o. Furthermore, say that o is **salient** in e just in case o is (i) free in e, and (ii) no proper superterm of o is free in e. Suppose that there are exactly n salient occurrences of terms in e. If $n = 0$, let $\overline{e}$ be the empty sequence. Otherwise, for each $1 \leq i \leq n$, let e_i be the term with the i:th salient occurrence in e, counting from left to right, and let $\overline{e}$ be the (possibly repetitive) sequence $e_1, \ldots, e_n$ of terms. Let $\underline{e}$ be the result of replacing each salient occurrence of a term in e with the low dash symbol _. For instance, if

$$e = \forall x(Pxy \to f(x, y) = g(y, z))$$

we get $e_1 = y$, $e_2 = y$, $e_3 = g(y, z)$, and thus

$$\underline{e} = \forall x(Px\_ \to f(x, \_) = \_)$$

We stipulate that, if $\#x\varphi$ is an L-term in the extended syntax with n salient occurrences of terms, then $f_{\underline{\#x\varphi}}$ is an n-place function symbol. Finally, we define the extension $L\#$ of L by

$$L\# = L \cup \{f_{\underline{\#x\varphi}} : \varphi \text{ an } L\text{-formula in the extended syntax}\}$$

3.2 Translation

We define a translation τ, from L-expressions in the extended syntax to $L\#$-expressions in the standard syntax, recursively:

- If t is a variable or a constant, then $\tau[t] = t$.
- If $f \in L$ is a function symbol, then $\tau[f(\bar{t})] = f(\tau[\bar{t}])$, where $\bar{t} = \langle t_1, \ldots, t_n \rangle$ and $\tau[\bar{t}] = \langle \tau[t_1], \ldots, \tau[t_n] \rangle$.
- $\tau[\#x\varphi] = f_{\underline{\#x\varphi}}(\tau[\overline{\#x\varphi}])$.

- $\tau[s = t] = \tau[s] = \tau[t]$.
- $\tau[P\bar{t}] = P\tau[\bar{t}]$.
- $\tau[\neg\varphi] = \neg\tau[\varphi]$.
- $\tau[\varphi \to \psi] = \tau[\varphi] \to \tau[\psi]$.
- $\tau[\forall x\varphi] = \forall x\tau[\varphi]$.

For instance, with $P, f, g \in L$, we have

$$\tau[\#x(Pxy \to f(x, y) = g(y, z))] = f_{\#x(Px\_\to f(x,\_)=\_)}(y, y, g(y, z))$$

First we observe that, for any expression e,

(23) e and $\tau[e]$ have the same constants and variables occurring freely.

If x a variable and t is a closed term, let $e(t/x)$ be the result of replacing all free occurrences of x in e with t. Since substitution of a free occurrence of a variable in an expression always takes place inside a salient occurrence of a term in that expression, we also have

(24) $\underline{e(t/x)} = \underline{e}$

Using this fact, we show that

(25) $\tau[e(t/x)] = \tau[e](\tau[t]/x)$

Proof. By induction on the complexity of e. For the base cases, we have

- $\tau[x(t/x)] = \tau[t] = x(\tau[t]/x) = \tau[x](\tau[t]/x)$
- $\tau[c(t/x)] = \tau[c] = c(\tau[t]/x) = \tau[c](\tau[t]/x)$

Assume, as induction hypothesis, that the claim holds for the immediate sub-expressions. We consider the following cases:

- If $f \in L$ is a function symbol, we get

$$\begin{aligned} \tau[f(\bar{s})(t/x)] &= \tau[f(\bar{s}(t/x))] \\ &= f(\tau[\bar{s}(t/x)]) && \text{by definition of } \tau \\ &= f(\tau[\bar{s}](\tau[t]/x)) && \text{by induction hypothesis} \\ &= f(\tau[\bar{s}])(\tau[t]/x) \\ &= \tau[f(\bar{s})](\tau[t]/x) && \text{by definition of } \tau \end{aligned}$$

-

$$\begin{aligned} \tau[\#y\varphi(t/x)] &= f_{\underline{\#y\varphi(t/x)}}(\tau[\overline{\#y\varphi(t/x)}]) && \text{by definition of } \tau \\ &= f_{\underline{\#y\varphi(t/x)}}(\tau[\overline{\#y\varphi}](\tau[t]/x)) && \text{by induction hypothesis} \\ &= f_{\underline{\#y\varphi(t/x)}}(\tau[\overline{\#y\varphi}])(\tau[t]/x) \\ &= f_{\underline{\#y\varphi}}(\tau[\overline{\#y\varphi}])(\tau[t]/x) && \text{by (24)} \\ &= \tau[\#y\varphi](\tau[t]/x) && \text{by definition of } \tau \end{aligned}$$

- $$\begin{aligned}\tau[\neg\varphi(t/x)] &= \neg\tau[\varphi(t/x)] && \text{by definition of } \tau\\ &= \neg\tau[\varphi](\tau[t]/x) && \text{by induction hypothesis}\\ &= \tau[\neg\varphi](\tau[t]/x) && \text{by definition of } \tau\end{aligned}$$

- $$\begin{aligned}\tau[(\varphi\wedge\psi)(t/x)] &= \tau[\varphi(t/x)\wedge\psi(t/x)]\\ &= \tau[\varphi(t/x)]\wedge\tau[\psi(t/x)] && \text{by definition of } \tau\\ &= \tau[\varphi](\tau[t]/x)\wedge\tau[\psi](\tau[t]/x) && \text{by induction hypothesis}\\ &= (\tau[\varphi]\wedge\tau[\psi])(\tau[t]/x)\\ &= \tau[\varphi\wedge\psi](\tau[t]/x) && \text{by definition of } \tau\end{aligned}$$

- If $x = y$, we have trivially that

  $$\tau[\forall y\varphi(t/x)] = \tau[\forall y\varphi] = \tau[\forall y\varphi](\tau[t]/x)$$

  If $x \neq y$, we get

  $$\begin{aligned}\tau[\forall y\varphi(t/x)] &= \tau[\forall y(\varphi(t/x))]\\ &= \forall y\tau[\varphi(t/x)] && \text{by definition of } \tau\\ &= \forall y(\tau[\varphi](\tau[t]/x)) && \text{by induction hypothesis}\\ &= \forall y\tau[\varphi](\tau[t]/x)\\ &= \tau[\forall y\varphi](\tau[t]/x) && \text{by definition of } \tau\end{aligned}$$

□

Lastly, we show that τ is injective:

(26) If $\tau[e] = \tau[e']$ then $e = e'$.

Proof. By induction on the complexity of e. The base cases are obvious, since $\tau[t] = t$ if t is a variable or a constant. Assume, as induction hypothesis, that the claim holds for any immediate sub-expressions. We consider the following cases:

- If $f \in L$ is a function symbol, then $\tau[f(\bar{t})] = \tau[e']$ implies

  $$f(\tau[\bar{t}]) = f(\tau[t_1],\ldots,\tau[t_n]) = \tau[e'] = f(\tau[t'_1],\ldots,\tau[t'_n])$$

  where $e' = f(t'_1,\ldots,t'_n)$ and $\tau[t_1] = \tau[t'_1],\ldots,\tau[t_n] = \tau[t'_n]$. By induction hypothesis, we get $t_1 = t'_1,\ldots,t_n = t'_n$. Hence, $f(\bar{t}) = f(t'_1,\ldots,t'_n) = e'$.

- $\tau[\#x\varphi] = \tau[e']$ implies

  $$f_{\underline{\#x\varphi}}(\tau[\overline{\#x\varphi}]) = \tau[e'] = f_{\underline{\#x'\varphi'}}(\tau[\overline{\#x'\varphi'}])$$

  where $e' = \#x'\varphi'$, $\underline{\#x\varphi} = \underline{\#x'\varphi'}$ and $\tau[\overline{\#x\varphi}] = \tau[\overline{\#x'\varphi'}]$. By induction hypothesis, we get $\overline{\#x\varphi} = \overline{\#x'\varphi'}$. Since $\#x\varphi = \underline{\#x\varphi}(\overline{\#x\varphi}/\_)$ and $\#x'\varphi' = \underline{\#x'\varphi'}(\overline{\#x'\varphi'}/\_)$, it follows that $\#x\varphi = \#x'\varphi' = e'$.

□

3.3 Rules of inference

We define the classical provability relation $\vdash$ inductively, letting it apply to the extended syntax as well. For any sentences (closed formulas) φ, ψ, χ, and for any sets Γ, Δ, Σ of sentences:

- $\dfrac{\varphi \in \Gamma}{\Gamma \vdash \varphi}$ A
- $\dfrac{\Gamma \vdash \varphi \qquad \Delta \vdash \psi}{\Gamma, \Delta \vdash \varphi \wedge \psi}$ $\wedge$I
- $\dfrac{\Gamma \vdash \varphi \wedge \psi}{\Gamma \vdash \varphi \text{ and } \Gamma \vdash \psi}$ $\wedge$E
- $\dfrac{\Gamma \vdash \varphi}{\Gamma \vdash \varphi \vee \psi \text{ and } \Gamma \vdash \psi \vee \varphi}$ $\vee$I
- $\dfrac{\Gamma \vdash \varphi \vee \psi \qquad \Delta, \varphi \vdash \chi \qquad \Sigma, \psi \vdash \chi}{\Gamma, \Delta, \Sigma \vdash \chi}$ $\vee$E
- $\dfrac{\Gamma, \varphi \vdash \psi}{\Gamma \vdash \varphi \to \psi}$ $\to$I
- $\dfrac{\Gamma \vdash \varphi \to \psi \qquad \Delta \vdash \varphi}{\Gamma, \Delta \vdash \psi}$ $\to$E
- $\dfrac{\Gamma, \varphi \vdash \psi \qquad \Delta, \varphi \vdash \neg\psi}{\Gamma, \Delta \vdash \neg\varphi}$ $\neg$I
- $\dfrac{\Gamma, \neg\varphi \vdash \psi \qquad \Delta, \neg\varphi \vdash \neg\psi}{\Gamma, \Delta \vdash \varphi}$ $\neg$E

For any formula φ with free occurrences of at most one variable x:

- $\dfrac{\Gamma \vdash \varphi(c/x) \qquad c \text{ a constant not in } \Gamma \text{ or } \varphi}{\Gamma \vdash \forall x \varphi}$ $\forall$I
- $\dfrac{\Gamma \vdash \forall x \varphi \qquad t \text{ a closed term}}{\Gamma \vdash \varphi(t/x)}$ $\forall$E
- $\dfrac{\Gamma \vdash \varphi(t/x)}{\Gamma \vdash \exists x \varphi}$ $\exists$I
- $\dfrac{\Gamma \vdash \exists x \varphi \qquad \Delta, \varphi(c/x) \vdash \psi \qquad c \text{ a constant not in } \Delta, \varphi \text{ or } \psi}{\Gamma, \Delta \vdash \psi}$ $\exists$E
- $\dfrac{t \text{ a closed term}}{\Gamma \vdash t = t}$ =I
- $\dfrac{\Gamma \vdash \varphi(t/x) \qquad \Delta \vdash t = t' \text{ or } \Delta \vdash t' = t \qquad t \text{ and } t' \text{ closed terms}}{\Gamma, \Delta \vdash \varphi(t'/x)}$ =E

Remark 3.1. Perhaps this is not what the elimination rule for identity usually looks like. However, if $\varphi(t, t'/x)$ is the result of replacing some free occurrences of x in φ with t, and the rest with t', one can derive the perhaps more standard rule

$$\frac{\Gamma \vdash \varphi(t/x) \qquad \Delta \vdash t = t' \qquad t \text{ and } t' \text{ closed terms}}{\Gamma, \Delta \vdash \varphi(t, t'/x)}$$

as follows. Let ψ be the result of replacing only some free occurrences of x in φ with t, so that $\psi(t'/x) = \varphi(t, t'/x)$. Then we also have $\psi(t/x) = \varphi(t/x)$. Hence, we get

$$\dfrac{\Gamma \vdash \varphi(t/x) \qquad \dfrac{\Delta \vdash t = t' \qquad \dfrac{\neg\psi(t'/x) \in \{\neg\psi(t'/x)\}}{\neg\psi(t'/x) \vdash \neg\psi(t'/x)}\text{AI}}{\Delta, \neg\psi(t'/x) \vdash \neg\psi(t/x)\ (= \neg\varphi(t/x))}=\text{E}}{\Gamma, \Delta \vdash \psi(t'/x)\ (= \varphi(t, t'/x))}\neg\text{E}$$

First we show that

Lemma 3.1. *For any set of L-sentences Γ and L-sentence φ in the extended syntax such that $\Gamma \vdash \varphi$, we have $\tau[\Gamma] \vdash \tau[\varphi]$.*

Proof. By induction on the complexity of proofs. For the base case, assume that we have $\Gamma \vdash \varphi$ by A, with $\varphi \in \Gamma$. Clearly, we then have $\tau[\varphi] \in \tau[\Gamma]$. By A, we get $\tau[\Gamma] \vdash \tau[\varphi]$.

Assume, as induction hypothesis, that the claim holds for any immediate sub-proofs. We consider the following cases:

- We have $\Gamma, \Delta \vdash \varphi \wedge \psi$ by $\wedge$I, with $\Gamma \vdash \varphi$ and $\Delta \vdash \psi$. By induction hypothesis, we have $\tau[\Gamma] \vdash \tau[\varphi]$ and $\tau[\Delta] \vdash \tau[\psi]$. Hence, by $\wedge$I, we get $\tau[\Gamma], \tau[\Delta] \vdash \tau[\varphi] \wedge \tau[\psi]$, which is the same as $\tau[\Gamma, \Delta] \vdash \tau[\varphi \wedge \psi]$.

- We have $\Gamma \vdash \forall x\varphi$ by $\forall$I, with $\Gamma \vdash \varphi(c/x)$ and c a constant not in Γ or φ. By induction hypothesis, we have $\tau[\Gamma] \vdash \tau[\varphi(c/x)]$. By (25), this is the same as $\tau[\Gamma] \vdash \tau[\varphi](\tau[c]/x)$, which is the same as $\tau[\Gamma] \vdash \tau[\varphi](c/x)$. By (23), c does not occur $\tau[\Gamma]$ or $\tau[\varphi]$. By $\forall$I, we get $\tau[\Gamma] \vdash \forall x\tau[\varphi]$, which is the same as $\tau[\Gamma] \vdash \tau[\forall x\varphi]$.

- We have $\Gamma \vdash \varphi(t/x)$ by $\forall$E, with $\Gamma \vdash \forall x\varphi$ and t a closed term. By induction hypothesis, we have $\tau[\Gamma] \vdash \tau[\forall x\varphi]$, which is the same as $\tau[\Gamma] \vdash \forall x\tau[\varphi]$. By (23), $\tau[\varphi]$ has free occurrences of at most one variable x. Hence, by $\forall$E, we get $\tau[\Gamma] \vdash \tau[\varphi](\tau[t]/x)$. By (25), this is the same as $\tau[\Gamma] \vdash \tau[\varphi(t/x)]$.

- We have $\Gamma \vdash t = t$ by =I. Since $\tau[t] = \tau[t]$, we get $\tau[\Gamma] \vdash \tau[t] = \tau[t]$ by =I, which is the same as $\tau[\Gamma] \vdash \tau[t = t]$.

- We have $\Gamma, \Delta \vdash \varphi(t'/x)$ by =E, with $\Gamma \vdash \varphi(t/x)$, $\Delta \vdash t = t'$ or $\Delta \vdash t' = t$, where t and t' are closed terms. By induction hypothesis, we have $\tau[\Gamma] \vdash \tau[\varphi(t/x)]$ and $\tau[\Delta] \vdash \tau[t = t']$, which is the same as $\tau[\Delta] \vdash \tau[t] = \tau[t']$. By (23), $\tau[t]$ and $\tau[t']$ are closed terms. By (25), we have $\tau[\Gamma] \vdash \tau[\varphi](\tau[t]/x)$. By =E, we get $\tau[\Gamma], \tau[\Delta] \vdash \tau[\varphi](\tau[t']/x)$, which is the same as $\tau[\Gamma, \Delta] \vdash \tau[\varphi](\tau[t']/x)$. By (25), this is the same as $\tau[\Gamma, \Delta] \vdash \tau[\varphi(t'/x)]$.

□

Remark 3.2. Using the result above, we can already show that no amount of pure arithmetic will, by itself, help solve Example **??**. Let A be a pure theory of arithmetic in the vocabulary $L_A = \{N, 0, s, +, \times\}$. We assume that A is consistent, as it otherwise will imply the answer trivially. Let $L_C = \{M, T\}$, let $L = L_A \cup L_C$, and let Γ be the following set of L-sentences in the extended syntax:

(a) $\#xMx = \underline{1}$

(b) $\#x(\neg Mx \wedge Tx) = \underline{2}$

(c) $\#x(Mx \wedge \neg Tx) = \underline{0}$

(d) $\#xTx \neq \underline{3}$

Their translations are given by the set of $L\#$-sentences $\tau[\Gamma]$:

(a) $f_{\#xMx} = \underline{1}$

(b) $f_{\#x(\neg Mx \wedge Tx)} = \underline{2}$

(c) $f_{\#x(Mx \wedge \neg Tx)} = \underline{0}$

(d) $f_{\#xTx} \neq \underline{3}$

Let $\mathcal{M}_A$ be an L_A-model of A. We extend it to an $L\#$-model $\mathcal{M}$ of $A \cup \tau[\Gamma]$ as follows. Let D be a non-empty set disjoint from $|\mathcal{M}_A|$, let $|\mathcal{M}| = |\mathcal{M}_A| \cup D$, and let $a \in D$. Furthermore, let

(a) $[f_{\#xMx}]^{\mathcal{M}} = [\underline{1}]^{\mathcal{M}_A}$

(b) $[f_{\#x(\neg Mx \wedge Tx)}]^{\mathcal{M}} = [\underline{2}]^{\mathcal{M}_A}$

(c) $[f_{\#x(Mx \wedge \neg Tx)}]^{\mathcal{M}} = [\underline{0}]^{\mathcal{M}_A}$

(d) $[f_{\#xTx}]^{\mathcal{M}} = a$

The interpretation of the rest of $L\#$ in $\mathcal{M}$ can be chosen arbitrarily. In any case, we have $\mathcal{M} \vDash \tau[\Gamma]$. Moreover, since A is a pure theory of arithmetic, and $\mathcal{M} \restriction N = (\mathcal{M}_A) \restriction N$, we have $\mathcal{M} \vDash A$. By soundness, it follows that

$$A \cup \{f_{\#xMx} = \underline{1}, f_{\#x(\neg Mx \wedge Tx)} = \underline{2}, f_{\#x(Mx \wedge \neg Tx)} = \underline{0}\} \nvdash f_{\#xTx} = \underline{3}$$

and, by Lemma 3.1, that

$$A \cup \{\#xMx = \underline{1}, \#x(\neg Mx \wedge Tx) = \underline{2}, \#x(Mx \wedge \neg Tx) = \underline{0}\} \nvdash \#xTx = \underline{3}$$

Next, we show that

Lemma 3.2. *For any set of L-sentences Γ and L-sentence φ in the extended syntax such that $\tau[\Gamma] \vdash \tau[\varphi]$, we have $\Gamma \vdash \varphi$.*

Proof. By induction on the complexity of proofs. For the base case, assume that we have $\tau[\Gamma] \vdash \tau[\varphi]$ by A, with $\tau[\varphi] \in \tau[\Gamma]$. By (26), we then have $\varphi \in \Gamma$. By A, we get $\Gamma \vdash \varphi$.

Assume, as induction hypothesis, that the claim holds for any immediate sub-proofs. We consider the following cases:

- We have $\tau[\Gamma], \tau[\Delta] \vdash \tau[\varphi] \wedge \tau[\psi]$ by $\wedge$I, with $\tau[\Gamma] \vdash \tau[\varphi]$ and $\tau[\Delta] \vdash \tau[\psi]$. By induction hypothesis, we have $\Gamma \vdash \varphi$ and $\Delta \vdash \psi$. Hence, by $\wedge$I, we get $\Gamma, \Delta \vdash \varphi \wedge \psi$.

- We have $\tau[\Gamma] \vdash \forall x \tau[\varphi]$ by ∀I, with $\tau[\Gamma] \vdash \tau[\varphi](c/x)$ and c a constant not in $\tau[\Gamma]$ or $\tau[\varphi]$. By (25), we have $\tau[\varphi(c/x)] = \tau[\varphi](\tau[c]/x) = \tau[\varphi](c/x)$. Hence, by induction hypothesis, we have $\Gamma \vdash \varphi(c/x)$. By (23), c does not occur Γ or φ. By ∀I, we get $\Gamma \vdash \forall x \varphi$.

- We have $\tau[\Gamma] \vdash \tau[\varphi](\tau[t]/x)$ by ∀E, with $\tau[\Gamma] \vdash \forall x \tau[\varphi]$ and $\tau[t]$ a closed term. By induction hypothesis, we have $\Gamma \vdash \forall x \varphi$. By (23), $\tau[\varphi]$ has free occurrences of at most one variable x, and t is a closed term. Hence, by ∀E, we get $\Gamma \vdash \varphi(t/x)$.

- We have $\tau[\Gamma] \vdash \tau[t] = \tau[t']$ by =I, with $\tau[t] = \tau[t']$. By (26), we have $t = t'$. By =I, we get $\Gamma \vdash t = t'$.

- We have $\tau[\Gamma], \tau[\Delta] \vdash \tau[\varphi](\tau[t']/x)$ by =E, with $\tau[\Gamma] \vdash \tau[\varphi](\tau[t]/x)$, $\tau[\Delta] \vdash \tau[t] = \tau[t']$ or $\tau[\Delta] \vdash \tau[t'] = \tau[t]$, where $\tau[t]$ and $\tau[t']$ are closed terms. By (25) and induction hypothesis, we have $\Gamma \vdash \varphi(t/x)$ and $\Delta \vdash t = t'$. By =E, we get $\Gamma, \Delta \vdash \varphi(t'/x)$.

□

The desired results now follows by Lemma 3.1 and 3.2:

Theorem 3.1. *For any set of L-sentences Γ and L-sentence φ in the extended syntax, we have $\Gamma \vdash \varphi$ just in case $\tau[\Gamma] \vdash \tau[\varphi]$.*

4 Numerical validity

Let $L_A = \{N, 0, s, +, \times\}$ be our arithmetical vocabulary, let L_E be a vocabulary disjoint from L_A containing a unary predicate O, and let $L = L_A \cup L_E$. Let $\mathcal{N}$ be the standard L_A-model, with $N^{\mathcal{N}} = |\mathcal{N}| = \mathbb{N}$, and let $Th(\mathcal{N})_N = \{\varphi_N : \mathcal{N} \vDash \varphi\}$ be the pure theory of true arithmetic.

Definition 4.1 (Numerical extensions)**.** Let $\mathcal{M}_E$ be a L_E-model such that $\mathcal{M}_E \vDash \forall x O x$. An $L\#$-model $\mathcal{M}$ is a *numerical extension* of $\mathcal{M}_E$ just in case the following obtains:

(i) $(\mathcal{M}|L_A) \upharpoonright N$ and $(\mathcal{M}|L_E) \upharpoonright O$ are defined.

(ii) $(\mathcal{M}|L_E) \upharpoonright O = \mathcal{M}_E$.

(iii) There is $c \in |\mathcal{M}| - N^{\mathcal{M}}$ such that the following obtains. Let φ be an L-formula in the extended syntax, and suppose that $\#x\varphi$ has n salient occurrences of terms. Let $\varphi(\bar{v}/\#x\varphi)$ be the result of replacing the occurrences of these terms in φ with n distinct variables $\bar{v} = \langle v_1, \ldots, v_n \rangle$ not occurring in $\#x\varphi$. If $a_1, \ldots, a_n \in |\mathcal{M}|$, let g be an assignment such that $g(v_1) = a_1, \ldots, g(v_n) = a_n$, and let

$$\kappa = |\{a \in |\mathcal{M}| : \mathcal{M}, g_{x \to a} \vDash \tau[\varphi(\bar{v}/\overline{\#x\varphi})]\}|$$

If $\kappa \in \mathbb{N}$, we then have

$$f^{\mathcal{M}}_{\underline{\#x\varphi}}(a_1, \ldots, a_n) = \underline{\kappa}^{\mathcal{M}}$$

and otherwise

$$f^{\mathcal{M}}_{\underline{\#x\varphi}}(a_1, \ldots, a_n) = c$$

Definition 4.2 (Numerical validity)**.** An $L\#$-sentence φ is *numerically valid* just in case, for any L_E-model $\mathcal{M}_E$ such that $\mathcal{M}_E \vDash \forall x O x$, and for any numerical $L\#$-extension $\mathcal{M}$ of $\mathcal{M}_E$, we have $\mathcal{M} \vDash \varphi$.

The set of numerical validities is arithmetically neutral, since every consistent theory of pure arithmetic together with any consistent theory of pure non-arithmetic has a numerical extension:

Lemma 4.1. *For every L_E-model $\mathcal{M}_E$ such that $\mathcal{M}_E \vDash \forall x O x$, and for every L_A-model $\mathcal{M}_A$ such that $\mathcal{M}_A \vDash \forall x N x$, if $|\mathcal{M}_E| \cap |\mathcal{M}_A| = \emptyset$, there is a numerical $L\#$-extension $\mathcal{M}$ of $\mathcal{M}_E$ with domain $|\mathcal{M}| = |\mathcal{M}_E| \cup |\mathcal{M}_A|$ such that $(\mathcal{M}|L_E) \restriction O = \mathcal{M}_E$ and $(\mathcal{M}|L_A) \restriction N = \mathcal{M}_A$.*

Proof. Once each L-symbol has received an interpretation in $\mathcal{M}$, the interpretation of each function symbol $f_{\underline{\#x\varphi}} \in L\#$ can be defined inductively on the complexity of φ, which is an L-formula in the extended syntax. In the base case, φ is just an L-formula. If $\#x\varphi$ is an L-formula in the extended syntax, we can assume as part of the induction hypothesis that all $L\#$-symbols in $\tau[\varphi]$ already have received an interpretation in $\mathcal{M}$. □

However, due to the following property and Trakhtenbrot's theorem, numerical validity is not axiomatizable whenever L_E contains at least one binary predicate:

Lemma 4.2. *Let $\mathcal{M}_E$ be an L_E-model such that $\mathcal{M}_E \vDash \forall x O x$, and let $\mathcal{M}$ be a numerical $L\#$-extension of $\mathcal{M}_E$. For any L-formula φ in the extended syntax, and for any assignment g, we then have $\mathcal{M}, g \vDash \tau[N\#x\varphi]$ just in case $\{a \in |\mathcal{M}| : \mathcal{M}, g_{x\to a} \vDash \tau[\varphi]\}$ is finite.*

In particular:

Lemma 4.3. *Let $\mathcal{M}_E$ be an L_E-model such that $\mathcal{M}_E \vDash \forall x O x$, and let $\mathcal{M}$ be a numerical $L\#$-extension of $\mathcal{M}_E$. Then we have $\mathcal{M} \vDash \tau[N\#xOx]$ just in case $\mathcal{M}_E$ is finite.*

Hence:

Lemma 4.4. *Let $L'_E = L_E - \{O\}$. For each L'_E-sentence φ, we have that $\tau[N\#xOx \to \varphi_O]$ is numerically valid just in case φ is true in all finite L'_E-models.*

Proof. For left to right, assume that $\tau[N\#xOx \to \varphi_O]$ is numerically valid. Let $\mathcal{M}$ be a finite L'_E-model, and expand it to an L_E-model $\mathcal{M}_E$ with $O^{\mathcal{M}_E} = |\mathcal{M}_E|$. Let $\mathcal{M}_\#$ be a numerical extension of $\mathcal{M}_E$. By Lemma 4.3, since $\mathcal{M}_E$ is finite, we have $\mathcal{M}_\# \vDash \tau[N\#xOx]$. By assumption, we get $\mathcal{M}_\# \vDash \varphi_O$, and thus $\mathcal{M}_\#|L_E \vDash \varphi_O$. By Lemma 2.1, since $(\mathcal{M}_\#|L_E) \restriction O = \mathcal{M}_E$, we have $\mathcal{M}_E \vDash \varphi$, and thus $\mathcal{M} \vDash \varphi$.

For right to left, assume that φ is true in all finite L'_E-models. Let $\mathcal{M}_E$ be an L_E-model such that $\mathcal{M}_E \vDash \forall x O x$, and let $\mathcal{M}$ be a numerical extension of $\mathcal{M}_E$. We get two cases, in either of which $\mathcal{M} \vDash \tau[N\#xOx \to \varphi_O]$:

1. $\mathcal{M}_E$ is finite. By assumption, we then have $\mathcal{M}_E \vDash \varphi$. By Lemma 2.1, since $(\mathcal{M}|L_E) \restriction O = \mathcal{M}_E$, we get $\mathcal{M} \vDash \varphi_O$. Hence, $\mathcal{M} \vDash \tau[N\#xOx \to \varphi_O]$.

2. $\mathcal{M}_E$ is infinite. By Lemma 4.3, we have $\mathcal{M} \vDash \tau[\neg N\#xOx]$. Hence, $\mathcal{M} \vDash \tau[N\#xOx \to \varphi_O]$.

Hence, $\tau[N\#xOx \to \varphi_O]$ is numerically valid. $\square$

Thus, due to Trakhtenbrot's theorem, if L_E contains a binary predicate, numerical validity in $L\#$ is not axiomatizable.

5 Standard validity

Definition 5.1 (Standard extensions)**.** Let $\mathcal{M}_E$ be an L_E-model such that $\mathcal{M}_E \vDash \forall x Ox$. An $L\#$-model $\mathcal{M}$ is a *standard extension* of $\mathcal{M}_E$ just in case

(i) $\mathcal{M}$ is a numerical extension of $\mathcal{M}_E$, and

(ii) $(\mathcal{M}|L_A) \restriction N$ is isomorphic to $\mathcal{N}$.

Remark 5.1. It should be reasonably clear that every L_E-model satisfying $\forall x Ox$ of has a standard extension.

Remark 5.2. Whenever two formulas φ and ψ are both satisfied by infinitely many elements, any standard (indeed, numerical) extension will satisfy $\#x\varphi = \#\psi$. Why, you may ask? The chief reason is that, with this arbitrary stipulation, we ensure that Equinumerosity (and Extensionality, which follows from the former) is true in all standard extensions, without having to concern ourselves with infinite cardinalities. Also, since we are dealing with first-order logic, we can assume without loss for all relevant purposes that all models are countable.

Definition 5.2 (Standard validity)**.** An $L\#$-sentence φ is *standardly valid* just in case, for any L_E-model $\mathcal{M}_E$ such that $\mathcal{M}_E \vDash \forall x Ox$, and for any standard $L\#$-extension $\mathcal{M}$ of $\mathcal{M}_E$, we have $\mathcal{M} \vDash \varphi$.

Suppose that φ is standardly valid. Assume that $\mathcal{M}$ is an $L\#$-model such that

(i) $(\mathcal{M}|L_A) \restriction N$ and $(\mathcal{M}|L_E) \restriction O$ are defined,

(ii) $\mathcal{M}$ satisfies Correspondence, and

(iii) $(\mathcal{M}|L_A) \restriction N$ is elementary equivalent to $\mathcal{N}$.

Does it follow that $\mathcal{M} \vDash \varphi$? Possible counterexample: Disjunctive comprehension. Perhaps even Conjunctive comprehension? Indeed. We can define an extension $\mathcal{M}$ of an L_E-model whose N-part is a non-standard model of $Th(\mathcal{N})_N$, as follows. Let c be a non-standard number in that model, and let $d \in |\mathcal{M}| - N^{\mathcal{M}}$. For any $a_1, \ldots, a_n \in |\mathcal{M}|$, let g be an assignment such that $g(v_1) = a_1, \ldots, g(v_n) = a_n$, let

$$K = |\{a \in |\mathcal{M}| : \mathcal{M}, g_{x \to a} \vDash \tau[\varphi(\bar{v}/\overline{\#x\varphi})]\}|$$

and let

$$f^{\mathcal{M}}_{\underline{\#x\varphi}}(a_1, \ldots, a_n) = \begin{cases} \underline{|K|}^{\mathcal{M}} & \text{if } K \text{ is finite} \\ c & \text{if } K \text{ is infinite and } |\mathcal{M}| - K \text{ is finite} \\ d & \text{otherwise} \end{cases}$$

Provided that $\mathcal{M} - N^{\mathcal{M}}$ is infinite, we have $\mathcal{M} \models \tau[N\#x(x = x)]$ but $\mathcal{M} \not\models \tau[N\#(x = x \wedge Nx)]$. To construct a counterexample to Disjunctive comprehension, just switch c and d. Then $\mathcal{M} \models \tau[N\#xNx \wedge N\#x\neg Nx]$ but $\mathcal{M} \not\models \tau[N\#x(Nx \vee \neg Nx)]$. This will also serve as a counterexample to Additivity.

As we saw earlier, numerical validity is not axiomatiable due to Trakhtenbrot's theorem concerning the unaxiomatizability of finite validity. We see that standard validity is not axiomatizable either, but only for the rather trivial reason that it contains true arithmetic. But finite validity, for instance, is axiomatizable relative to true arithmetic, since truth in all finite models can be decided by true arithmetic. There is an L_A-formula $\mathsf{FinVal}(x)$ such that, for any L'_E-sentence φ, we have that $\mathsf{FinVal}(\ulcorner\varphi\urcorner)$ is a theorem of true arithmetic just in case φ is true in all finite L'_E-models. Let

$$B = \{\mathsf{FinVal}(\ulcorner\varphi\urcorner) \to (\tau[N\#xOx] \to \varphi_O) : \varphi \text{ an } L'_E\text{-sentence}\}$$

First we observe that every element of B is standardly valid. To see why, let $\mathcal{M}$ be a standard extension of an L_E-model $\mathcal{M}_E$ such that $\mathcal{M}_E \models \forall x Ox$, and assume that $\mathcal{M} \models \mathsf{FinVal}(\ulcorner\varphi\urcorner) \wedge \tau[N\#xOx]$. Hence, φ is true in all finite L'_E-models. Moreover, by Lemma 4.3 and the second conjunct, $\mathcal{M}_E$ is finite. Hence, $\mathcal{M}_E \models \varphi_O$. By Lemma 2.1, we get $\mathcal{M} \models \varphi_O$.

Now, if φ is true in all finite L'_E-models, we clearly have $Th(\mathcal{N})_N \cup B \vdash \tau[N\#xOx] \to \varphi_O$. Does the converse hold? Yes. Assume that φ is false in some finite L'_E-model $\mathcal{M}$. Extend it to an L_E-model $\mathcal{M}_E$ with $O^{\mathcal{M}_E} = |\mathcal{M}_E|$. Let $\mathcal{M}_\#$ be a standard extension of $\mathcal{M}_E$. Thus, $\mathcal{M}_\# \models Th(\mathcal{N})_N$. By Lemma 4.3, we have $\mathcal{M}_\# \models \tau[N\#xOx]$. By Lemma 2.1, we have $\mathcal{M}_\# \models \neg\varphi_O$. Since every element of B is standardly valid, we also have $\mathcal{M}_\# \models B$. Hence, $Th(\mathcal{N})_N \cup B \not\vdash \tau[N\#xOx] \to \varphi_O$.

Let NV be the set of all numerically valid $L\#$-sentences, and let SV be the set of all standardly valid $L\#$-sentences. We have

$$Th(\mathcal{N})_N + NV \not\vdash SV$$

since, on any common L_A-definition of $<$,

$$\neg\exists y(Ny \wedge \forall z(Nz \to \#x(Nx \wedge x < z) \neq y)) \in SV$$

but

$$Th(\mathcal{N})_N + NV \not\vdash \neg\exists y(Ny \wedge \forall z(Nz \to \#x(Nx \wedge x < z) \neq y))$$

as witnessed by any non-standard numerical extension satisfying $Th(\mathcal{N})_N$.

6 Arithmetically neutral bridge principles

Lemma 6.1. *The following bridge principles are all numerically valid:*

1. *Equinumerosity*
2. *Conjunctive comprehension*
3. *Disjunctive comprehension*

4. *Zero*

5. *Successor*

Taken together, they are therefore arithmetically neutral.

Proof. Let $\mathcal{M}$ be a numerical extension.

- We show that $\mathcal{M}$ is a model of Zero. Let φ be the formula $x \neq x$. Then $\overline{\#x\varphi}$ is empty, $f_{\#x(x\neq x)}$ is a 0-place function symbol, and $\tau[\#x(x = x)] = f_{\#x(x\neq x)}$. Moreover, we have

$$|\{a \in |\mathcal{M}| : \mathcal{M}, g_{x\to a} \vDash \tau[\varphi(\bar{v}/\overline{\#x\varphi})]\}| = |\{a \in |\mathcal{M}| : \mathcal{M}, g_{x\to a} \vDash x \neq x\}| = 0$$

  in which case $f^{\mathcal{M}}_{\#x(x\neq x)} = \underline{0}^{\mathcal{M}}$. Since, by definition of numerals, $\underline{0} = 0$, we get $\mathcal{M} \vDash f_{\#x(x\neq x)} = 0$.

- We show that $\mathcal{M}$ is a model of Successor. Let $t_1, \ldots, t_n$ be the salient terms of $\#x\varphi$, i.e. $\overline{\#x\varphi} = \langle t_1, \ldots, t_n\rangle$. Let g be an assignment. By definition, we have

$$[\tau[\#x\varphi]]^{\mathcal{M},g} = [f_{\#x\varphi}(t_1, \ldots, t_n)]^{\mathcal{M},g} = [f_{\#x\varphi}]^{\mathcal{M}}([t_1]^{\mathcal{M},g}, \ldots, [t_n]^{\mathcal{M},g})$$

  Let h be an assignment such that $h(v_1) = [t_1]^{\mathcal{M},g}, \ldots, h(v_n) = [t_n]^{\mathcal{M},g}$, and let

$$\kappa = |\{a \in |\mathcal{M}| : \mathcal{M}, h_{x\to a} \vDash \tau[\varphi(\bar{v}/\overline{\#x\varphi})]\}|$$

  Observe that

$$\{a \in |\mathcal{M}| : \mathcal{M}, h_{x\to a} \vDash \tau[\varphi(\bar{v}/\overline{\#x\varphi})]\} = \{a \in |\mathcal{M}| : \mathcal{M}, g_{x\to a} \vDash \tau[\varphi]\}$$

  We get two cases:

  - $\kappa \in \mathbb{N}$. Then

$$[f_{\#x\varphi}]^{\mathcal{M}}([t_1]^{\mathcal{M},g}, \ldots, [t_n]^{\mathcal{M},g}) = [\underline{\kappa}]^{\mathcal{M}}$$

    Let $b \in |\mathcal{M}|$, and assume that $\mathcal{M}, g_{y\to b} \vDash \neg\varphi(y/x)$. Observe that

$$\overline{\#x(\varphi \vee x = y)} = \langle t_1, \ldots, t_n, y\rangle$$

    Moreover, since y can be assumed not to occur in $t_1, \ldots, t_n$, the assignment $h_{v_{n+1}\to b}$ satisfies

$$\begin{aligned} h_{v_{n+1}\to b}(v_1) &= t_1^{\mathcal{M},g_{y\to b}} \\ &\vdots \\ h_{v_{n+1}\to b}(v_n) &= t_n^{\mathcal{M},g_{y\to b}} \\ h_{v_{n+1}\to b}(v_{n+1}) &= y^{\mathcal{M},g_{y\to b}} \end{aligned}$$

    Let

$$\lambda = |\{a \in |\mathcal{M}| : \mathcal{M}, (h_{v_{n+1}\to b})_{x\to a} \vDash \tau[\varphi(\bar{v}/\overline{\#x\varphi})] \vee x = v_{n+1}\}|$$

Since, as noted earlier,

$$\{a \in |\mathcal{M}| : \mathcal{M}, h_{x\to a} \models \tau[\varphi(\bar{v}/\overline{\#x\varphi})]\} = \{a \in |\mathcal{M}| : \mathcal{M}, g_{x\to a} \models \tau[\varphi]\}$$

we get

$$\begin{aligned}
&|\{a \in |\mathcal{M}| : \mathcal{M}, (h_{v_{n+1}\to b})_{x\to a} \models \tau[\varphi(\bar{v}/\overline{\#x\varphi})] \vee x = v_{n+1}\}| = \\
&|\{a \in |\mathcal{M}| : \mathcal{M}, (h_{v_{n+1}\to b})_{x\to a} \models \tau[\varphi(\bar{v}/\overline{\#x\varphi})]\} \cup \\
&\ \{a \in |\mathcal{M}| : \mathcal{M}, (h_{v_{n+1}\to b})_{x\to a} \models x = v_{n+1}\}| = \\
&|\{a \in |\mathcal{M}| : \mathcal{M}, h_{x\to a} \models \tau[\varphi(\bar{v}/\overline{\#x\varphi})]\} \cup \{b\}| = \\
&|\{a \in |\mathcal{M}| : \mathcal{M}, g_{x\to a} \models \tau[\varphi]\} \cup \{b\}| = \kappa + 1
\end{aligned}$$

and thus $\lambda = \kappa + 1$. Hence,

$$\begin{aligned}
&\tau[\#x(\varphi \vee x = y)]^{\mathcal{M}, g_{y\to b}} = f_{\#x(\varphi\vee x=y)}(t_1, \ldots, t_n, y)^{\mathcal{M}, g_{y\to b}} = \\
&f^{\mathcal{M}}_{\#x(\varphi\vee x=y)}(t_1^{\mathcal{M}, g_{y\to b}}, \ldots, t_n^{\mathcal{M}, g_{y\to b}}, y^{\mathcal{M}, g_{y\to b}}) = \underline{\kappa + 1}^{\mathcal{M}} = \\
&s(\underline{\kappa})^{\mathcal{M}} = s^{\mathcal{M}}(\underline{\kappa}^{\mathcal{M}})
\end{aligned}$$

as required.

– $\kappa \notin \mathbb{N}$. Then

$$f^{\mathcal{M}}_{\#x\varphi}(t_1^{\mathcal{M},g}, \ldots, t_n^{\mathcal{M},g}) = c$$

Since $c \notin N^{\mathcal{M}}$, we get $\mathcal{M}, g \not\models N f_{\underline{\#x\varphi}}(t_1, \ldots, t_n)$.

- We show that $\mathcal{M}$ is a model of Conjunctive comprehension. Let $t_1, \ldots, t_n$ be the salient terms of $\#x\varphi$, and let $s_1, \ldots, s_m$ be the salient terms of $\#x\psi$. Observe that

$$\overline{\#x(\varphi \wedge \psi)} = \langle t_1, \ldots, t_n, s_1, \ldots, s_m\rangle$$

Let $v_1, \ldots, v_n, u_1, \ldots, u_m$ be distinct variables not occurring in $\varphi \wedge \psi$, let $\bar{v} = \langle v_1, \ldots, v_n\rangle$, let $\bar{u} = \langle u_1, \ldots, u_m\rangle$ and $\bar{w} = \langle v_1, \ldots, v_n, u_1, \ldots, u_m\rangle$. Let g be an assignment, let h be an assignment such that

$$h(v_1) = [t_1]^{\mathcal{M},g}, \ldots, h(v_n) = t_n^{\mathcal{M},g}$$

and

$$h(u_1) = [s_1]^{\mathcal{M},g}, \ldots, h(u_m) = s_m^{\mathcal{M},g}$$

Let

$$\kappa = |\{a \in |\mathcal{M}| : \mathcal{M}, h_{x\to a} \models \tau[\varphi(\bar{v}/\overline{\#x\varphi})]\}|$$

and let

$$\begin{aligned}
\lambda =& |\{a \in |\mathcal{M}| : \mathcal{M}, h_{x\to a} \models \tau[(\varphi \wedge \psi)(\bar{w}/\overline{\#x(\varphi \wedge \psi)})]\}| \\
=& |\{a \in |\mathcal{M}| : \mathcal{M}, h_{x\to a} \models \tau[\varphi(\bar{v}/\overline{\#x\varphi}) \wedge \psi(\bar{u}/\overline{\#x\psi})]\}|
\end{aligned}$$

We get two cases:

– $\kappa \in \mathbb{N}$. Clearly, we then have $\lambda \in \mathbb{N}$. By definition of $\mathcal{M}$, we get

$$f^{\mathcal{M}}_{\#x(\varphi\wedge\psi)}(t_1^{\mathcal{M},g}, \ldots, t_n^{\mathcal{M},g}, s_1^{\mathcal{M},g}, \ldots, s_m^{\mathcal{M},g}) = \underline{\lambda}^{\mathcal{M}}$$

and thus

$$\mathcal{M}, g \vDash Nf_{\#x(\varphi\wedge\psi)}(t_1, \ldots, t_n, s_1, \ldots, s_n)$$

– $\kappa \notin \mathbb{N}$. By definition of $\mathcal{M}$, we get

$$f^{\mathcal{M}}_{\#x\varphi}(t_1^{\mathcal{M},g}, \ldots, t_n^{\mathcal{M},g}) = c$$

Since $c \notin N^{\mathcal{M}}$, we get $\mathcal{M}, g \nvDash Nf_{\underline{\#x\varphi}}(t_1, \ldots, t_n)$.

- The case of Disjunctive comprehension is similar.
- We show that $\mathcal{M}$ is a model of Equinumerosity. Let g be an assignment, and assume that

$$\mathcal{M}, g \vDash \tau[\forall x(\varphi \to \exists! y(\psi \wedge \chi)) \wedge \forall y(\psi \to \exists! x(\varphi \wedge \chi))]$$

with x not free in ψ, and y not free in φ. Hence,

$$|\{a \in |\mathcal{M}| : \mathcal{M}, g_{x\to a} \vDash \tau[\varphi]\}| = |\{a \in |\mathcal{M}| : \mathcal{M}, g_{y\to a} \vDash \tau[\psi]\}|$$

Since $\mathcal{M}$ is a numerical extension, we get two cases, in either of which we have

$$\mathcal{M}, g \vDash \tau[\#x\varphi = \#y\psi]$$

□

Extensionality follows from Equinumerosity by taking $x = y$ as χ. Moreover, provided that we allow φ, ψ and χ to contain free variables other than x and y, and take Equinumerosity to be the universal closure of each such instance, we can establish the following:

Lemma 6.2. *Let $\mathcal{M}$ be an $L\#$-model satisfying Equinumerosity, let φ and ψ be L-formulas in the extended syntax with x not free in ψ and y not free in φ, and let g be an assignment. If $\{a \in |\mathcal{M}| : \mathcal{M}, g_{x\to a} \vDash \tau[\varphi]\}$ and $\{a \in |\mathcal{M}| : \mathcal{M}, g_{y\to a} \vDash \tau[\psi]\}$ are both finite and contain equally many elements, then $\mathcal{M}, g \vDash \tau[\#x\varphi = \#y\psi]$.*

Proof. Let $\mathcal{M}$ be an $L\#$-model satisfying Equinumerosity, let φ and ψ be L-formulas in the extended syntax with x not free in ψ and y not free in φ, and let g be an assignment. Suppose that

$$\{a \in |\mathcal{M}| : \mathcal{M}, g_{x\to a} \vDash \tau[\varphi]\} = \{a_1, \ldots, a_n\}$$

and

$$\{a \in |\mathcal{M}| : \mathcal{M}, g_{y\to a} \vDash \tau[\psi]\} = \{b_1, \ldots, b_n\}$$

Let $x_1, \ldots, x_n \neq x$ and $y_1, \ldots, y_n \neq y$ be distinct variables not occurring in φ or ψ, and let χ be the formula

$$(x = x_1 \wedge y = y_1) \vee \ldots \vee (x = x_n \wedge y = y_n)$$

Let h be an assignment just like g, except that $h(x_i) = a_i$ and $h(y_i) = b_i$ for each $i = 1, \dots, n$. Clearly,

$$\{a \in |\mathcal{M}| : \mathcal{M}, h_{x \to a} \vDash \tau[\varphi]\} = \{a_1, \dots, a_n\}$$

and

$$\{a \in |\mathcal{M}| : \mathcal{M}, h_{y \to a} \vDash \tau[\psi]\} = \{b_1, \dots, b_n\}$$

Since $\mathcal{M}$ satisfies Equinumerosity, we have

$$\mathcal{M}, h \vDash \tau[\forall x(\varphi \to \exists! y(\psi \wedge \chi)) \wedge \forall y(\psi \to \exists! x(\varphi \wedge \chi)) \to \# x\varphi = \# y\psi]$$

By assumption, we also have

$$\mathcal{M}, h \vDash \tau[\forall x(\varphi \to \exists! y(\psi \wedge \chi)) \wedge \forall y(\psi \to \exists! x(\varphi \wedge \chi))]$$

Hence,

$$\mathcal{M}, h \vDash \tau[\# x\varphi = \# y\psi]$$

and thus

$$\mathcal{M}, g \vDash \tau[\# x\varphi = \# y\psi]$$

□

Let B be the L-theory in the extended syntax consisting of the universal closure of every instance of Equinumerosity, Conjunctive comprehension, Zero, and Successor. We show that

Lemma 6.3. *For any n and L-formula φ in the extended syntax, we have*

$$B \vdash \exists_{=n} x\varphi \to \# x\varphi = \underline{n}$$

Proof. By induction on n. The base case is given by Zero and Extensionality. Assume, as induction hypothesis, that the claim holds for n. We thus assume that, for any formula φ, we have

$$B \vdash \exists_{=n} x\varphi \to \# x\varphi = \underline{n}$$

We will show that the same holds for $n + 1$:

$$B \vdash \exists_{=n+1} x\varphi \to \# x\varphi = \underline{n+1}$$

We observe that, as a matter of pure logic,

$$\vdash \exists_{=n+1} x\varphi \leftrightarrow \exists y(\varphi(y/x) \wedge \exists_{=n} x(\varphi \wedge x \neq y))$$

By induction hypothesis, it follows that

$$B \vdash \exists_{=n+1} x\varphi \to \exists y(\varphi(y/x) \wedge \# x(\varphi \wedge x \neq y) = \underline{n})$$

Hence, it suffices to establish that

$$B \vdash \exists y(\varphi(y/x) \wedge \# x(\varphi \wedge x \neq y) = \underline{n}) \to \# x\varphi = \underline{n+1}$$

We reason inside B. Assume c to be such that

$$\varphi(c/x) \wedge \#x(\varphi \wedge x \neq c) = \underline{n}$$

Let ψ be the formula $\varphi \wedge x \neq c$. By Conjunctive comprehension and the second conjunct of our assumption, we have $N\#x\psi$. As an instance of Successor, we have

$$N\#x\psi \rightarrow (\neg\psi(c/x) \rightarrow \#x(\psi \vee x = c) = s(\#x\psi))$$

Since $\neg\psi(c/x)$, we get

$$\#x(\psi \vee x = c) = s(\#x\psi)$$

Since $\varphi(c/x)$ by assumption, we also have

$$\forall x(\psi \vee x = c \leftrightarrow \varphi)$$

By Extensionality and our assumption that $\#x(\varphi \wedge x \neq c) = \underline{n}$, we finally get $\#x\varphi = s(\underline{n})$, which by definition is the same as $\#x\varphi = \underline{n+1}$. □

6.1 Adding some pure arithmetic

Let our pure theory of arithmetic A consist of $\mathsf{PA}_N(1)$–$\mathsf{PA}_N(4)$. Observe that

(27) For any natural number n, we have $A \vdash N\underline{n}$.

Let $T = A \cup B$. We show that

Lemma 6.4. *For any n and L-formula φ in the extended syntax, we have*

$$T \vdash \#x\varphi = \underline{n} \rightarrow \exists_{=n} x\varphi$$

Proof. By induction on n. For the base case, we need to establish that

$$T \vdash \#x\varphi = 0 \rightarrow \neg\exists x\varphi$$

We reason inside T. Assume that $\#x\varphi = 0$. Assume, towards contradiction, that there is c such that $\varphi(c/x)$. Let ψ be the formula $\varphi \wedge x \neq c$. By Conjunctive comprehension, $\mathsf{PA}_N(1)$ and our first assumption, we have $N\#x\psi$. As an instance of Successor, we have

$$N\#x\psi \rightarrow (\neg\psi(c/x) \rightarrow \#x(\psi \vee x = c) = s(\#x\psi))$$

Since $\neg\psi(c/x)$, we get

$$\#x(\psi \vee x = c) = s(\#x\psi)$$

Since $\varphi(c/x)$ by assumption, we also have

$$\forall x(\psi \vee x = c \leftrightarrow \varphi)$$

By Extensionality, we get $\#x\varphi = s(\#x\psi)$. By $\mathsf{PA}_N(3)$, it follows that $\#x\varphi \neq 0$, a contradiction.

Assume, as induction hypothesis, that the claim holds for n. We thus assume that, for any formula φ, we have

$$T \vdash \#x\varphi = \underline{n} \to \exists_{=n} x\varphi$$

We will show that the same holds for $n+1$:

$$T \vdash \#x\varphi = \underline{n+1} \to \exists_{=n+1} x\varphi$$

We observe that, as a matter of pure logic,

$$\vdash \exists_{=n+1} x\varphi \leftrightarrow \exists y(\varphi(y/x) \wedge \exists_{=n} x(\varphi \wedge x \neq y))$$

By induction hypothesis and Lemma 6.3, it follows that

$$T \vdash \exists_{=n+1} x\varphi \leftrightarrow \exists y(\varphi(y/x) \wedge \#x(\varphi \wedge x \neq y) = \underline{n})$$

Hence, it suffices to establish that

$$T \vdash \#x\varphi = \underline{n+1} \to \exists y(\varphi(y/x) \wedge \#x(\varphi \wedge x \neq y) = \underline{n})$$

We reason inside T. Assume that $\#x\varphi = \underline{n+1}$. If $\neg\exists x\varphi$, we have $\#x\varphi = 0$ by Lemma 6.3, contradicting $\mathsf{PA}_N(3)$. Hence, we can assume that there is c such that $\varphi(c/x)$. It remains to be shown that

$$\#x(\varphi \wedge x \neq c) = \underline{n}$$

Let ψ be the formula $\varphi \wedge x \neq c$. By Conjunctive comprehension, (27) and the second conjunct of our assumption, we have $N\#x\psi$. As an instance of Successor, we have

$$N\#x\psi \to (\neg\psi(c/x) \to \#x(\psi \vee x = c) = s(\#x\psi))$$

Since $\neg\psi(c/x)$, we get

$$\#x(\psi \vee x = c) = s(\#x\psi)$$

Since $\varphi(c/x)$ by assumption, we also have

$$\forall x(\psi \vee x = c \leftrightarrow \varphi)$$

By Extensionality, we get $\#x\varphi = s(\#x\psi)$. By our assumption that $\#x\varphi = s(\underline{n})$, we get $s(\#x\psi) = s(\underline{n})$. By (27), we get $Ns(\#x\psi)$. Since $N\#x\psi$, we get $\#x\psi = \underline{n}$ by $\mathsf{PA}_N(4)$. □

Theorem 6.1. *For any n and L-formula φ in the extended syntax, we have*

$$T \vdash \exists_{=n} x\varphi \leftrightarrow \#x\varphi = \underline{n}$$

Proof. By Lemma 6.3 and 6.4. □

References

Field, H. (1980). *Science Without Numbers.* Princeton University Press.

Husserl, E. (2003). *Philosophy of Arithmetic: Psychological and Logical Investigations - with Supplementary Texts From 1887-1901.* Springer Verlag.

van Benthem, J. and Icard, T. (2023). Interleaving logic and counting. *The Bulletin of Symbolic Logic*, 29(4):pp. 503–587.

Wang, H. (1990). *Process and Existence in Mathematics*, pages 30–46. Springer Netherlands, Dordrecht.