Finitist Axiomatic Truth

Walker, Jan

How to cite

WALKER, Jan. Finitist Axiomatic Truth. Doctoral Thesis, 2019. doi: 10.13097/archive-ouverte/unige:120650

This publication URL: https://archive-ouverte.unige.ch/unige:120650
Publication DOI: 10.13097/archive-ouverte/unige:120650

UNIVERSITÉ
DE GENÈVE

FACULTÉ DES LETTRES
Département de philosophie

Finitist Axiomatic Truth

Jan Walker

Doctoral Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

April 2019

Examining Committee: Professor Fabrice Correia (Thesis Advisor), Professor Thomas Strahm (Thesis Co-Advisor), Professor Laurent Cesalli (Committee President), Professor Andrea Cantini, Professor Volker Halbach, Professor Reinhard Kahle, Doctor Kentaro Sato

Acknowledgements

I am grateful to my supervisors, Fabrice Correia and Thomas Strahm, for making my research possible and for their invaluable guidance and support.

My most sincere gratitude also and especially goes to Kentaro Sato who was my informal mentor in the more advanced stages of my doctoral research. He helped me sharpen my proof-theoretic intuitions in countless discussions that made the writing of this thesis a valuable intellectual experience. The quality of my work has increased considerably due to Kentaro's help.

Moreover, I am indebted to Andrea Cantini, Volker Halbach, and Reinhard Kahle for carefully inspecting an earlier version of this manuscript and for sending me helpful feedback.

Thanks also to the Logic and Theory Group in Bern and the Eidos Group in Geneva for providing a very stimulating research environment.

Finally, I would like to thank my family and my girlfriend for their patience, support, and love.

My research was generously supported by the Swiss National Science Foundation (SNF).

Abstract

In this thesis we adapt several prominent methods and state consistent axiomatic theories of (type-free) arithmetical truth for the particular levels and the entire Grzegorczyk (primitive recursive) hierarchy and arithmetical hierarchy. More specifically, the considered theories include:

- a theory of *naive* truth (in the spirit of Heck [24]) for some basic level of the Grzegorczyk hierarchy;
- theories of *Friedman-Sheard* truth (due to [14] and Halbach [19]) for higher levels of the Grzegorczyk hierarchy;
- theories of *Kripke-Feferman* truth (going back to [32, 12]), *grounded* truth (also due to [32]), and *disquotational* truth (reminiscent of the Tarski schema from the classic [49]) for the entire Grzegorczyk hierarchy;
- *iterations* and *progressions* (via so-called reflection schemas) of theories of Kripke-Feferman truth for the particular levels and the entire arithmetical hierarchy.

The bulk of our work consists in proving – upon devising the required technical machinery and concepts – that the considered theories of truth are (in the specific proof-theoretic sense to be devised) interpretable in corresponding theories of arithmetic by employing recursion-theoretic tools.

A great deal of our proof-theoretic analysis relies on Parsons' theorem, according to which primitive recursive arithmetic is Π_2 equivalent to the first-order theory $\mathbf{I}(\Sigma_1)$.

Contents

1 Introduction

The concept of *truth* is (in many forms or facets) ubiquitous in mathematical logic and philosophy. In proof theory, most notably in contexts related to Gödel's famous Incompleteness Theorems, the focus is on *arithmetical* truth and the general approach is to investigate that concept very systematically, often in the form of various axiomatizations. (See Halbach and Leigh [23] for a brief summary of the axiomatic approach to truth.)

At the same time, fundamental questions that led to Gödel's theorems are closely tied to *finitist reasoning* in arithmetic and Hilbert's (failed) program. (For a brief summary of Hilbert's program, finitary methods, and their impact, see Zach [52].)

Roughly, Hilbert's program aims at a formalization of mathematics in axiomatic form and a proof that this formalization of mathematics is consistent. The consistency proof itself has to be given by finitary methods, the idea being, roughly, that finitary methods are philosophically secure with respect to their foundational status. (It is hard to make clear what "Hilbert's finitary methods" exactly are, but the rough idea is that these methods must not presuppose any infinite totalities, such as the set of all natural numbers.)

Even though Gödel's theorems made it ultimately clear that

Hilbert's program must fail, the program may today still be regarded as a huge success, because it lead to an enormous body of mathematical results which sharpened our understanding of the functioning, power, and limitations of formal systems. Similarly, we do not have to go into the utterly difficult question of what is a philosophically and historically appropriate reading of "Hilbert's finitary methods" in order to assure ourselves that, even though large parts of ordinary mathematics presuppose infinite totalities, the concepts of a finite formula, a finite derivation, a finite computation, and other cognates associated to "finitist reasoning" are ubiquitous and indispensable in all proof theory, recursion theory, and logic in general. In the words of Tait [48], p. 525:

> The special role of finitism consists in the circumstance that it is a minimal kind of reasoning presupposed by all nontrivial mathematical reasoning about numbers. And for this reason it is *indubitable* in a Cartesian sense that there is no preferred or even equally preferable ground on which to stand and criticize it. Thus finitism is fundamental to mathematics even if it is not a foundation in the sense Hilbert wished.

However, although the glaring significance of reasoning about truth and of finitist reasoning are well perceived in foundational contexts, the literature on their interplay is quite sparse (but see Section 1.2 for some related research). The general aim of the present thesis is to study the interplay between truth and finitist (arithmetical) reasoning more systematically. Informally put, the guiding question of the thesis is the following.

> Which principles of truth are in accordance with the principles of (sub-theories of) finitist arithmetic?

(Of course, we will have to fix a precise sense of "in accordance with".)

To contextualize our guiding question and to relate it to consistency proofs for formal systems and incompleteness results, it is illustrative to consider it (partly) as reversal of the *implicitness* program in the spirit of Feferman [12].

1.1 Implicitness

According to the standards of mathematical logic, any theory R about the (standard) natural numbers (which is formulated in a language of arithmetic or in some extension thereof) is *acceptable* only if (as necessary condition) it is *sound* in the following sense: For any statement A,

(1) A is provable in R

implies that

(2) A is true

in a clear-cut sense of "true" in the context of natural numbers. Furthermore, it is a standard of mathematical logic that, in the same sense of "true", a universal statement that

(3) every natural number n has property P

is true if and only if, for every natural number n, the particular instance that

(4) n has property P

is true. (Similarly for other logical operations.)

Since the soundness condition involves a notion of truth, the acceptance of R is, by being conditional on the soundness of R, conditional on some capacity to reason about the truth of certain statements. For example, by (1)–(4), this capacity might enable one to infer from the particular instances (for particular n) that

$(0 \leq n)$ is provable in R

the conclusion that the universal statement that

every natural number n satisfies $(0 \leq n)$

is true.

However, as is well known, by using one's capacity to reason about truth, one might impose a burden of *implicit commitments* on oneself. Namely, the acceptance of R might – in addition to the acceptance of statements that are provable in R – involve not only some implicit capacity to reason about truth, but also the acceptance of many further statements in the language of R that are not provable in R itself.

Take the statement that R is consistent. Suppose that R *is* sound (so consistent) and that R provides enough arithmetic to encode "A is a sentence in the language of R", "A is derivable in R", and other such syntactical concepts (in accordance with some usual requirements on the coding) in the natural numbers.

Under this (weak) supposition, it holds by Gödel's Second Incompleteness Theorem that R may not prove its own consistency and, thus, may not prove

(5) $(0 \neq 0)$ is not provable in R

in particular. (Due to the coding, (5) is a usual statement about natural numbers which may not be proved in R.) On the other hand, by applying one's implicit capacity to reason about truth, (5) may readily be proved: Namely, (5) is obtained from

(6) $(0 \neq 0)$ is not true

and soundness, by which (6) implies (5) by contraposition.

In a nutshell, not only does one's acceptance of R involve a capacity to reason about truth, but – by this very capacity – may also involve a commitment to statements from the language of R

which are not provable in R itself: for example, the (consistency) statement (5). We might say that, by accepting R, one commits oneself to such statements only *implicitly*.

1.2 Guiding Question

By making explicit some principles of truth (implicitly at play), we may also make explicit the commitments that are implicit in the acceptance of a theory R. Suppose that R is some theory of arithmetic. Let S be an extension of R in an extended language of truth, where S contains the principles of truth, now explicitly stated in the extended language.

Against this given background, the statements about natural numbers to which we (explicitly/implicitly) commit ourselves by accepting R may be regarded to be the statements about natural numbers (stated in the basic language of R) that are provable in extension S.

However, to determine the statements about natural numbers which are provable in (the yet unstudied) S, we may first bring into play a better understood extension T of R, a theory whose theorems about the natural numbers are already known, and we may try to establish (proof-theoretically) that S and T prove the same statements about natural numbers (from language of R).

That is, in the given sketch of implicitness, three theories are involved: some theory R that we initially accept, an extension S of R that (explicitly) provides some (implicit) principles of truth, and an extension T of R as reference to determine the implicit commitments that are derivable in S. Importantly, the aim is to identify the target (super-)theory T. As presupposition, S has to be equipped with specific principles of truth.

However, there is no indisputable, universally accepted set of such principles of truth. Rather, several alternatively motivated and equally elaborate accounts of (axiomatic arithmetical) truth have emerged in the literature. (Consult Halbach and Leigh [23] for a brief overview.)

In this thesis, we consider various principles of (arithmetical) truth which are well known in the literature and we investigate systematically how they may be modified and proved (by finitist means) to conform to principles of (sub-theories of) finitist arithmetic. Our method will be to assume certain principles of arithmetic that are already known to be (in some sense) reducible to principles of (the sub-theories of) finitist arithmetic and to proof-theoretically relate certain principles of truth to the presupposed principles of arithmetic. (Note that conformation with the principles of (sub-theories of) finitist arithmetic will not be a matter of transitivity: The proof-theoretic relationship between (sub-theories of) finitist arithmetic and presupposed (intermediate) theories of arithmetic will not be the same as the proof-theoretic relationship between the presupposed (intermediate) arithmetical theories and the investigated theories of truth.)

Let us specify the proof-theoretic relationship which we shall prove to hold between the theories of truth and the presupposed (intermediate) theories of arithmetic in terms of two requirements. The first requirement relies on the idea that, to conform to the principles of (sub-theories of) finitist arithmetic, the principles of truth must not enable us to prove certain statements about natural numbers that are not provable from principles of arithmetic (intermediately) presupposed in the reduction.

So, as first (but only a first) requirement, we consider conservativeness, which we may state in the form of reversing the implicitness program. That is, in the above wording, we start from presupposing some theory $\mathsf{T} = \mathsf{R}$ about natural numbers – thus, the theory that we initially accept already comprises all the implicit commitments – and we aim at stating a theory S of truth which meets the first requirement that:

> The principles of truth that are comprised in S must

> conform with the presumption that only the theorems of T ought to be (explicitly or implicitly) endorsed.

More precisely, the first requirement reads as follows.

1. *Conservativeness:* For every statement A in the language of T, A is provable in S if and only if A is provable in T.

In our particular case, as we have just stressed, the presumed theories T must comprise only principles of arithmetic which are known to be (in some sense) reducible to the principles of (sub-theories of) *finitist arithmetic.*

However, we shall opt for a more fine-grained proof-theoretic analysis and demand more than conservativeness: To conform to the principles of (the sub-theories of) finitist arithmetic, we also demand that truth principles must not have contents we cannot *make sense of* in terms of (intermediate) principles of arithmetic which are known to be (in some sense) reducible to principles of (sub-theories of) finitist arithmetic. Thus, as second requirement, we add that:

> The principles of truth (reasoning about truth) which are comprised (which is performed) in S ought to be interpretable in T.

More precisely, the second requirement might be put as:

2. *Interpretability:* There is some logic-preserving mapping $\circ$ from the language of S to the language of T such that, for every statement A in the language of S, the corresponding statement A° is a translation of A and fulfills that, if A is provable in S, then A° is provable in T. (If A belongs to language of T, then A° is provably equivalent to A in T.)

(Remark: The way in which we will establish conservativeness and interpretability will suggest that it could be formalized and proved in finitist arithmetic that both conditions are satisfied for theories under consideration. However, a precise treatment of the issue goes beyond the scope of the present thesis. Although it is not our intention to champion finitist views in this study, let us add that the formalization issue would be of more interest to a proponent of such finitist views.)

Thus, in sum, only if conservativeness and interpretability are met, we shall consider the principles of truth to be *in accordance with* the principles of (sub-theories of) finitist arithmetic.

(Indeed, if our focus were exclusively on conservativeness, we may borrow several insights from Leigh [33]. Namely, it follows by results in Leigh [33] that, if induction (no matter whether in schematic or truth-internal form) is *restricted to the arithmetical formulas that do not include a truth predicate*, the addition of the usual Tarskian compositional axioms of truth for an overall first-order language of arithmetic to one of our arithmetical reference theories would virtually always result in a theory of truth that is conservative over that reference theory.)

The proof-theoretic relation (called retractability) that we will choose for our investigation resembles the relation of truth definability by Fujimoto [15] and is tailored to satisfy specifically the two requirements of conservativeness and interpretability.

Let us close this section by drawing the reader's attention to related studies that were conducted in alternative frameworks by Cantini [7, 8], Eberhard [9], and Feferman and Strahm [13]. But when we say "related", we mean this in a rather loose sense. The only similarity is, first, that certain theories in [7, 8, 9] also come equipped with (the so-called Kripke-Feferman) axioms of partial truth that we will study in Chapter 8, and these theories also have only sub-finitist strength. Second, [13] may be seen as an instance of Feferman's implicitness program for finitist arithmetic, as it presents an unfolding of finitist arithmetic on the basis of a so-called schematic system (without explicit truth axioms).

Otherwise, the cited studies are plainly different from our approach. First of all, they are conducted in a completely different setting: While we will work in the *arithmetical* setting, the cited studies are based on the combinatory logic for untyped operations (abstraction). Relatedly, while we aim at *interpreting* theories of truth in theories of arithmetic, in the sense above, other proof-theoretic relations need to be established in the cited studies. In particular, the so-called realization techniques and asymmetric interpretations are typically used in these cited studies to compare theories according to which functions are provably total in them.

1.3 Axiomatic Approach

In our investigations, we follow the *axiomatic* approach, according to which a (target) theory S that provides the principles of truth takes the form of a so-called axiomatic theory of truth: The language of S has the primitive (unary) predicate symbol(s) $T(x)$ to express "x is true" and the principles of truth are formulated as (non-logical) axioms for the symbol(s) $T(x)$. Since a capacity to reason about truth will be conceived over theories about natural numbers in the present thesis, truth will also be conceived as a property of certain numbers, where the respective numbers figure as (Gödel) codes of expressions of a given language.

A comprehensive overview on the proof-theory and philosophy of axiomatic truth is given in the monograph by Halbach [21] and (in more condensed form) in the monograph by Horsten [27].

On the other hand, it is a common view in the philosophy of mathematics that the *finitist part of arithmetic* is best captured by the axioms/rules of *Primitive Recursive Arithmetic*, stated in a language without quantifiers. This view is rooted in a number of classical references, such as Skolem [46] and Hilbert and Bernays [26], and has received philosophical support by Tait [48].

However, it is also well known (due to Parsons [40]) that (the quantifier-free) Primitive Recursive Arithmetic is Π_2 equivalent to the theory $\mathbf{I}(\Sigma_1)$ of Σ_1 induction over first-order logic, that is, a theory in a language *with* quantifiers. Indeed, as Hilbert seems to have believed, the statements that are *finitistically meaningful* – in a sense of "checkable by finite computation" – can be seen to be comprised in the Π_2 fragment of arithmetical languages with quantifiers. (See, for example, Pohlers [41], Section 3.3, where an analogy to "experimentally checkable in physics" is drawn. For an elaboration of the Π_2 equivalence of theory $\mathbf{I}(\Sigma_1)$ and Primitive Recursive Arithmetic, consider Section 2.3 below, where Parsons' result is recalled.)

While we adhere to the common view that finitist arithmetic (reality) consists in the theorems of (the quantifier-free) Primitive Recursive Arithmetic, we shall nevertheless avail ourselves of the proof-theoretic results à la Parsons [40] that there are – modulo inter-translation – several ways to describe finitist arithmetic. In particular, we shall not attempt to interpret principles of truth in (sub-theories of) Primitive Recursive Arithmetic directly, but in (sub-theories of) the theory $\mathbf{I}(\Sigma_1)$ of Σ_1 induction because it is known that the theorems of $\mathbf{I}(\Sigma_1)$ are – if finitistically meaningful at all – exactly the same as the theorems of Primitive Recursive Arithmetic modulo inter-translation.

Thus, we might say that we use ideal means (quantifiers) and certain first-order principles (more complex than Π_2) as auxiliary tools in interpretations of the principles of truth only insofar as they are safe with regard to (sub-theories of) finitist arithmetic. (This is the sense of "reducible to principles of (sub-theories of) finitist arithmetic" we had in mind before.) Conversely, we may also use the same ideal means in our formulations of theories of truth because, if some truth theory is interpretable in a certain (first-order) arithmetical theory, then so is the same truth theory restated (without ideal means) in a restricted finitist fragment of its (first-order) language.

We will work in (the sub-theories of) $\mathbf{I}(\Sigma_1)$ for three reasons. First, conditions imposed by (sub-theories of) $\mathbf{I}(\Sigma_1)$ can be neatly turned into respective conditions on the truth theories from the literature (on the arithmetical setting). Second, by taking the opportunity (below in Section 2.3) to recall some of the main ideas behind the Π_2 equivalence of $\mathbf{I}(\Sigma_1)$ and Primitive Recursive Arithmetic, we may already get familiar with some of the main ideas behind finitist truth. Third, by working in (the sub-theories of) $\mathbf{I}(\Sigma_1)$, we facilitate the comparison of our results with the other results from proof theory, because it is common to use $\mathbf{I}(\Sigma_1)$ as reference theory for finitism in technical studies.

1.4 Axioms of Truth

As Tarski's profound study (in [49]) on the concept of truth in formal languages has shown, axioms of truth must be stated very carefully in order to prevent inconsistencies (or paradox). If some given truth axioms lead us into contradictions, the contradictions can usually be revealed by considering self-referential statements, obtained by the method of diagonalization, which are reminiscent of the old Liar Paradox (or variants thereof).

Let us use $\ulcorner t \urcorner$ and $\ulcorner A \urcorner$ to formally denote, respectively, the (Gödel) codes of term t and formula A. Moreover, let $T(x)$ be a primitive unary truth predicate. A so-called self-referential Liar sentence is a sentence D (for diagonalization) such that we have, provably,

$$D \leftrightarrow \neg T(\ulcorner D \urcorner) \tag{1.1}$$

where $\neg$ is negation and $\leftrightarrow$ is equivalence/biconditional. On the other hand, by Tarski's Adequacy Condition of Truth, or the so-called Tarski Schema for short, we (should) have, provably,

$$A \leftrightarrow T(\ulcorner A \urcorner) \tag{1.2}$$

where A is arbitrary. (As intuitive example, consider: The statement that $0 < 1$ is true if and only if $0 < 1$.)

To get an idea of the initial situation of our research, observe that, by the laws of classical logic, which are at work in (sub-theories of) $\mathbf{I}(\Sigma_1)$, (1.1) and (1.2) jointly entail a classical contradiction $D \leftrightarrow \neg D$ if A is *unrestrictedly* arbitrary in (1.2) and may be D in particular. So, we may ask, how shall we prevent such contradiction in our anticipated finitist setting?

Indeed, there are plenty of options, of which we will consider only a few in the present thesis. Specifically, we will restrict our research to theories of *type-free truth* because of their naturalness (as argued, for example, in Halbach [21], Section 11). That is, we will not prevent contradiction by imposing on (1.2) the type restriction that, very roughly, A must not involve the same truth predicate $T(x)$ that also applies to the code of A in (1.2). (The idea is that the problematic D involves this same truth predicate $T(x)$.)

We rather address the issue of consistency by considering the *complexity* of self-referential statements, such as the problematic D, and by restricting (1.2) only after that complexity has been determined.

From the first-order perspective of (the sub-theories of) $\mathbf{I}(\Sigma_1)$, there are two important (intertwined) measures of complexity: the quantifier complexity and the term complexity. Corresponding to the Σ_1 (or Π_2) fragment of the finitistically meaningful first-order formulas (sentences), we shall be interested in the Tarski Schema (1.2) only with regard to (the substitution instances of) formulas that involve only (sub-)finitist first-order operations. Indeed, the respective restriction on the quantifier complexity of A in (1.2) is an important precondition for the consistency of some of the axiomatizations of finitist type-free truth. However, restricting the quantifier complexity is not enough. Actually, to distinguish the alternative axiomatizations of type-free truth to be studied in the present thesis, it is more illustrative to focus on term complexity.

(Note also that a decrease in quantifier complexity might be compensated by a respective increase in term complexity.)

Suppose that we have at our disposal some function symbol $sub(x, y)$ to denote (via coding) the term or formula which is obtained, respectively, by uniformly substituting in x the only free variable by the term y. Moreover, let $\overline{n}$ be the numeral (some canonical term) to denote the number n and let $num(x)$ be some function symbol to denote the code of the x-th numeral, meaning that, provably, $num(\overline{n}) = \ulcorner \overline{n} \urcorner$.

Given this term machinery, we may define a Liar sentence D without any quantifiers as $\neg T(d(\overline{m}))$, where $d(y)$ is the term

$$sub(\ulcorner \neg T(x) \urcorner, sub(y, num(y)))$$

and we have, provably, $\overline{m} = \ulcorner d(y) \urcorner$ and therefore:

$$d(\overline{m}) = \ulcorner \neg T(d(\overline{m})) \urcorner \tag{1.3}$$

(We may also put (1.3) a bit more laboriously as

$$val(\ulcorner d(\overline{m}) \urcorner) = sub(\ulcorner \neg T(x) \urcorner, \ulcorner d(\overline{m}) \urcorner)$$

where $val(x)$ is a function symbol to denote the value/denotation of a term encoded as x.) By (1.3) and the laws of identity, the following instance of (1.1) (let D be $\neg T(d(\overline{m}))$) is immediate.

$$\neg T(\ulcorner \neg T(d(\overline{m})) \urcorner) \leftrightarrow \neg T(d(\overline{m}))$$

Against the background of (1.3), restricting the above Tarski Schema (1.2) to sentences A with low quantifier complexity is no remedy for inconsistency. For already the *atomic* instances

$$T(\ulcorner T(t) \urcorner) \leftrightarrow T(t) \tag{1.4}$$

of the Tarski Schema, where t is an arbitrary closed term, suffice to derive a contradiction from (1.3) – provided that we adhere to the classical negation principle that, for arbitrary sentences A:

$$T(\ulcorner \neg A \urcorner) \leftrightarrow \neg T(\ulcorner A \urcorner) \tag{1.5}$$

To see how we may arrive at a contradiction, note that by (1.3) and the laws of identity, like before:

$$T(\ulcorner \neg T(d(\overline{m}))\urcorner) \leftrightarrow T(d(\overline{m}))$$

So, by (1.5):

$$\neg T(\ulcorner T(d(\overline{m}))\urcorner) \leftrightarrow T(d(\overline{m}))$$

However, this contradicts (1.4), where $t = d(\overline{m})$.

Indeed, against the background of a finitist restriction of the Tarski Schema (1.2) to only *bounded* quantification, the different axiomatizations of type-free truth to be investigated in this thesis may be regarded as alternative reactions to the trilemma induced by *self-reference* (1.3) via terms, the *iteration principle* (1.4), and the *negation principle* (1.5). The reactions may be distinguished according to which condition (or horn) of the trilemma is denied.

1. Denial of (1.3): No self-reference via terms!
2. Denial of (1.4): No iteration principle!
3. Denial of (1.5): No negation principle!

Clearly, we may sustain at most two of the conditions (or horns) (1.3),(1.4),(1.5) in a consistent way. So, we are going to consider only the three maximally consistent reactions to the trilemma in the present thesis. Furthermore, (1.4) and (1.5) will be denied in accordance to the most popular respective theories of truth from the literature.

The first reaction – as Volker Halbach kindly brought to our attention after we had worked it out – is an elaboration of some ideas outlined by Heck [24]. It is based on the observation that, with regard to a particular (sub-)finitist fragment of a first-order language, we can have a consistent axiomatization of *naive type-free truth* with an associated Tarski Schema (1.2) for arbitrary A from the whole fragment, where (1.2) includes all instances of the iteration principle (1.4) and obeys the negation principle (1.5) at the same time. That is, if A belongs to the (sub-)finitist fragment at stake, then A may not be a Liar statement D satisfying (1.1) because A does not only have too low quantifier complexity (only bounded quantifiers), but also a term complexity too low to allow for (1.3).

Since Heck [24] already observed that (1.4) and (1.5) may be consistently combined in the absence of certain terms, let us call the first reaction just sketched the *Heck* axiomatization of finitist type-free truth. However, while Heck [24] focuses on the pair of principles (1.4) and (1.5), we take additional principles of truth into consideration. (See Chapter 6.)

Furthermore, in contrast to Heck [24], we consider terms from several (sub-)finitist fragments more systematically and we draw a clear-cut boundary between the (sub-)finitist fragments for which we may have a consistent axiomatization of naive type-free truth and the (sub-)finitist fragments for which we may not and must react differently to the above trilemma, as sketched next.

The second reaction is to allow – against the background of restricting the Tarski Schema (1.2) to bounded quantification as before – for self-reference (1.3) via terms, while requiring merely that the Tarski Schema (1.2) has to obey the negation principle (1.5), but may exclude certain instances of the iteration principle (1.4). However, in place of biconditionals (1.4), respective *truth rules* may be postulated to the effect that, most notably,

$$\frac{T(t)}{T(\ulcorner T(t) \urcorner)} \qquad \frac{T(\ulcorner T(t) \urcorner)}{T(t)}$$

or, for statements A (from given fragments) more generally, A is *provable* if and only if $T(\ulcorner A \urcorner)$ is *provable*. The idea of using such truth rules gives rise to what Halbach [19] called the *Friedman-Sheard* axiomatization of type-free truth in dependence on some study by Friedman and Sheard [14].

Halbach [19] studied an unrestricted Friedman-Sheard axiomatization, comprising compositional axioms and truth rules for an overall first-order language, on the basis of Peano Arithmetic and determined its proof-theoretic strength (in terms of the provable arithmetical statements) to be the same as the one of Ramified Analysis $\mathbf{RA}_{<\omega}$ up to level ω. ($\mathbf{RA}_{<\omega}$ is, roughly put, a system of ω times iterated arithmetical comprehension. For the details, the reader is referred to Feferman [12], Paragraph 4, and also [10], Paragraph 6.) We note that the Friedman-Sheard axiomatization studied by Halbach [19] is – in sharp contrast to its *finitist* variants (see below) – only consistent (in the regular sense), but *not* ω-consistent.

We are going to present Friedman-Sheard axiomatizations of *finitist* type-free truth in Chapter 7.

The third reaction is to deny that Tarski Schema (1.2) has to be closed under the negation principle (1.5), which allows us to sustain the requirement that it must include all instances of the iteration principle (1.4). Furthermore, by dropping (1.5), we may allow (via Δ_1 definitions) arbitrary finitist terms to achieve self-reference (1.3) and we may extend Tarski Schema (1.2) even to *unbounded* existential quantification. Note that, in the absence of the negation principle (1.5), negations may not be moved outside the scope of the truth predicate. So, if (1.5) is missing, we may axiomatize type-free truth like some *positive* inductive definition. Building on seminal work by Kripke [32], Feferman [12] came up with an axiomatization of positive type-free truth. It is usually called the *Kripke-Feferman* axiomatization.

Unrestricted Kripke-Feferman axiomatizations which comprise *positive* compositional truth (falsity) axioms for an overall first-order language have been investigated over Peano Arithmetic by Cantini [6]. The respective proof-theoretic strengths (in terms of the provable arithmetical statements) of these systems reach the one of Ramified Analysis $\mathbf{RA}_{<\omega^\omega}$ up to level ω^ω and even proof-theoretically stronger systems.

We shall present a Kripke-Feferman axiomatization of *finitist* type-free truth in Chapter 8.

1.5 Main Results

As indicated in the foregoing section, the Friedman-Sheard and Kripke-Feferman theories have been investigated extensively over Peano Arithmetic (and much stronger systems). Our contribution and the novelty of the current thesis is that we shall develop and systematically investigate respective Heck (naive), Friedman-Sheard, and Kripke-Feferman theories which are tailored for the (weaker) finitist setting.

The peculiarity which is shared by axiomatizations of *finitist* type-free truth to be studied in the present thesis is that $T(x)$, with the associated Tarski Schema

$$T(\ulcorner A \urcorner) \leftrightarrow A,$$

can be understood as normal form for the *computable* (recursive) valuation of *statements* A in the same way as formula $Val(x, y)$, with the associated schema

$$Val(\ulcorner t \urcorner, y) \leftrightarrow t = y,$$

serves as a (Kleene) normal form for the computable (recursive) valuation of (primitive recursive) *terms* t: Just as $Val(x, y)$ is a uniform description of how to compute the numerical value y of some term x in finitely many steps, $T(x)$ (in finitist terms) is a uniform description of how to compute (verify) in finitely many steps that a statement x has the Boolean value “true”.

Accordingly, if $T(x)$ is axiomatized as a finitist form of Heck truth or a finitist form of Friedman-Sheard truth, we get by the negation principle (1.4) that, for every statement A,

$$T(\ulcorner A \urcorner) \vee T(\ulcorner \neg A \urcorner),$$

expressing that both are forms of *decidable truth*. As will become visible, our finitist modification of Kripke-Feferman truth, on the other hand, is a form of *semi-decidable truth*.

In any case, a uniform description of how to compute a value "true" will be given by the respective interpretation $T(x)^{\circ}$ of the truth predicate $T(x)$ in what follows, where $\circ$ is a mapping into some language of arithmetic as explained above. Thus, we may summarize the main results of the present study by highlighting which formal theories S of finitist truth are interpretable in and conservative over which formal theories T of arithmetic.

Let us write $\mathsf{S} \rhd \mathsf{T}$ to denote that S is interpretable in and conservative over T. Moreover, let us write $\mathcal{L}(m,k)$ to denote a language which comprises the truth predicates

$$T_1, T_2, \ldots, T_m$$

and a symbol for every function of the k-th level of (as will be defined in Section 3.3 below) a smashed Grzegorczyk hierarchy. In particular, $m = 0$ indicates that we have some basic language of arithmetic without any truth predicates, where $\mathcal{L}(0,1)$ comprises some function symbol for addition and some function symbol for multiplication, $\mathcal{L}(0,2)$ also includes some symbol for the smash function, $\mathcal{L}(0,3)$ also comprises some symbol for exponentiation, $\mathcal{L}(0,4)$ also includes some symbol for tetration (hyper-4), and so on.

As a general recipe, we shall axiomatize arithmetical truth in accordance with the (sub-)finitist principles by formulating truth axioms only for the logical operations and terms which conform with these (sub-)finitist principles.

As was indicated before, the main result of Chapter 6 will be that we can have a consistent theory $\mathbf{FH}(1,1)$ of (as we will call it) Finitist Heck $\mathcal{L}(1,1)$ Truth, which is a theory of naive type-free truth. In particular, $\mathbf{FH}(1,1)$ will be endowed with axioms of Δ_0 (decidable, bounded) $\mathcal{L}(1,1)$ truth and an axiom of Truth Induction, where the latter states that truth conforms with the induction principle for natural numbers. (Informally put, Truth Induction states that, if $A(\overline{0})$ is true and, for any number x, the truth of $A(x)$ implies the truth of $A(x+\overline{1})$, then, for arbitrary x, statement $A(x)$ is true.) These truth axioms will be based on a theory $\mathbf{A}(0,3)$ which essentially is elementary arithmetic as the base (sub-)theory of $\mathbf{FH}(1,1)$ (in an auxiliary form). The main result of Chapter 6 will be the following.

$$\mathbf{FH}(1,1) \rhd \mathbf{A}(0,3)$$

By the conservativeness, this yields the consistency of $\mathbf{FH}(1,1)$: Since $(0 \neq 0)$ is not provable in $\mathbf{A}(0,3)$, neither is it in $\mathbf{FH}(1,1)$.

In Chapter 7, we will first establish that, in cases $k > 1$, we have for theories $\mathbf{FD}(1,k)$ that comprise axioms of type-free Δ_0 (decidable, bounded) $\mathcal{L}(0,k)$ truth – that is, axioms which range over $\mathcal{L}(1,k)$ expressions but do not have the biconditionals (1.4) and are not accompanied by any truth rules – together with the axiom of Truth Induction that

$$\mathbf{FD}(1,k) \rhd \mathbf{A}(0,k+1)$$

where $\mathbf{A}(0,k+1)$ is a theory about $(k+1)$-th level of (smashed) Grzegorczyk hierarchy. Second, we expand $\mathbf{FD}(1,k)$ by adding truth rules for the Δ_0 fragment of $\mathcal{L}(1,k)$ and establish that, for respective theories $\mathbf{FFS}(1,k)$ of Finitist Friedman-Sheard $\mathcal{L}(1,k)$ Truth, we have

$$\mathbf{FFS}(1,k) \underset{n}{\rhd} \mathbf{A}(0,k+1)$$

where n indicates that interpretations are *local* in the sense that they depend on the number n of applications of truth rules. In contrast to the unrestricted Friedman-Sheard axiomatization over Peano Arithmetic, as studied by Halbach [19], the finitist variants $\mathbf{FFS}(1,k)$ are not only consistent (in the regular sense), but also ω-consistent.

In Chapters 8–10, we will concentrate on a restricted language $\mathcal{L}(1,\mathrm{Q})$ of Robinson Arithmetic (for addition and multiplication) with a truth predicate. The key result of Chapter 8 will be that the theory $\mathbf{FKF}(1, \mathrm{Q})$ of Finitist Kripke-Feferman $\mathcal{L}(1, \mathrm{Q})$ Truth, which contains axioms of essentially Σ_1 (semi-decidable) positive $\mathcal{L}(1, \mathrm{Q})$ truth and the axiom of Truth Induction, is such that

$$\mathbf{FKF}(1, \mathrm{Q}) \rhd \mathbf{I}(\Sigma_1)$$

where $\mathbf{I}(\Sigma_1)$, remember, is Π_2 equivalent to the formalization of finitist arithmetic.

Chapter 9 is about the philosophically interesting subject of *grounded* truth, where the groundedness in non-semantic facts is conceived mathematically as membership in a *least* fixed point of a truth operator. While in non-finitist settings, such minimality requirement on fixed points of a truth operator may increase the proof-theoretic strength (in terms of provable arithmetical statements) of a system considerably, it does not in a finitist setting, where the same arithmetical tools that allow us to define truth operators also (and at the same time) allow us to interpret the (built-in) minimality constraint. Relatedly, we will also see that it may be proved by finitist means that (as truth is grounded in non-semantic states of affairs) the semantic extension of a truth predicate is consistent and does not comprise a Liar sentence.

In Chapter 10, we will show that $\mathbf{FKF}(1, \mathrm{Q})$ is interpretable also by applying only a Uniform Tarski Schema for essentially Σ_1 formulas in $\mathcal{L}(1, \mathrm{Q})$, where the truth predicate is only allowed to occur positively.

While Chapters 6 and 7 are associated with the Grzegorczyk hierarchy of finitist arithmetic, the final Chapters 11 and 12 are about the relationship between the *arithmetical* hierarchy and the Kripke-Feferman axiomatization of finitist truth. The subject of Chapter 11 is *iteration* of Finitist Kripke-Feferman Truth as an analogue of the Turing jump. For $n > 1$ and theories $\mathbf{FKF}(n, \mathrm{Q})$ extending $\mathbf{FKF}(1, \mathrm{Q})$ basically by axioms of Truth Induction for $T_i(x)$ $(0 < i \leq n)$, the Finitist Kripke-Feferman axioms for $T_i(x)$

$(0 < i \leq n)$, and – most importantly – *iteration* axioms

$$T_j([T_i(t)]) \leftrightarrow T_i(t) \qquad (i < j \leq n)$$

and

$$T_j([\neg T_i(t)]) \leftrightarrow \neg T_i(t) \qquad (i < j \leq n)$$

we are going to show that

$$\mathbf{FKF}(n, \mathrm{Q}) \rhd \mathbf{I}(\Sigma_n)$$

where $\mathbf{I}(\Sigma_n)$ is a theory with induction for the n-th level of the arithmetical hierarchy. While $T_1(x)$ (in finitist terms) is a uniform description of a (possibly non-terminating) verification algorithm without oracle, $T_2(x)$ also uniformly describes a (possibly non-terminating) computation, but one with an oracle which tells whether verifications at lower level 1 will eventually terminate or not, and so on.

The subject of the last Chapter 12 is the relationship between iteration of truth (that is the subject of Chapter 11) and *formal soundness* of a formal system in the form of a so-called *Uniform Reflection Schema*

$$Prov_{\mathsf{S}}(\ulcorner A(t_0, \ldots, t_k) \urcorner) \rightarrow A(t_0, \ldots, t_k)$$

expressing that, if $A(t_0, \ldots, t_k)$ is provable in the system S, then $A(t_0, \ldots, t_k)$ holds (or is true). Starting basically from a system $\mathsf{S}_1 = \mathbf{FKF}(1, \mathrm{Q})$ as investigated in Chapter 8, we will consider a progression

$$\mathsf{S}_1, \mathsf{S}_2, \mathsf{S}_3, \ldots$$

of theories, where $\mathsf{S}_n = \mathbf{FR}(n, \mathrm{Q})$ $(n > 1)$ has Finitist Kripke-Feferman axioms, but for only a single level n, and the Uniform Reflection Schema for every formula that may be proved in the predecessor theory S_{n-1} and also is provably equivalent to a Δ_0 formula in S_n. For $n > 1$, we will show that

$$\mathbf{FR}(n, \mathrm{Q}) \rhd \mathbf{FKF}(n, \mathrm{Q})$$

and, as direct consequence, that $\mathbf{FR}(n, \mathrm{Q})$ and $\mathbf{FKF}(n, \mathrm{Q})$ are in fact *identical theories.*

1.6 Outline

We conclude this introduction with a brief outline of the initial Chapters 2–5.

The general aim of the next four chapters is to provide the prerequisites for the study of finitist truth.

In Chapter 2, we will define some key concepts and establish several preliminary results for the study of finitist truth. More specifically, we shall

- define basic *logical* concepts,
- clarify the role of $\mathbf{I}(\Sigma_1)$ by relating it proof-theoretically to *Primitive Recursive Arithmetic*,
- elaborate on *retractability*, that is, the main proof-theoretic relation of our research.

A theory about the natural numbers qualifies as a theory of arithmetical truth only if it incorporates some (base sub-)theory that provides enough arithmetic to encode the target language of the truth predicate(s) in the natural numbers. In Chapter 3, we shall successively formulate such bases for the theories of finitist truth to be investigated in later chapters. Although we follow a common approach and take (for the most part) elementary arithmetic as convenient (albeit not very efficient) base for the coding of syntax, Chapter 3 will take quite some space for two reasons. The first is that Chapter 3 includes several basic results that are required in later chapters.

The second reason is heuristic and will be briefly discussed in Section 3.3: We want to simplify our work in our base theories considerably by making use of auxiliary function symbols (as, for example, in Chapters 4 and 5), but we also want to be flexible enough to get rid of auxiliary symbols whenever it is convenient. To achieve both aims at the same time, we shall apply common techniques from proof theory to show how the auxiliary function symbols may be inter-substituted for certain definitions of graphs in the basic language of $\mathbf{I}(\Sigma_1)$. In sum, we shall digress a bit in order to clarify the following.

- How to define base theories in the basic language of $\mathbf{I}(\Sigma_1)$ augmented by a (or several) truth predicate(s) only.
- How to simplify our work in the base theories by additional function symbols of just auxiliary status.

The use of auxiliary function symbols will be the subject of Sections 3.7 and 3.8.

While Sections 3.4–3.8 are of heuristic interest, as mentioned before, they also fit a core idea behind axiomatizations of type-free – conceived as natural – truth: Namely, once we have fixed (the Π_2 fragment of) a language for some description ($\mathbf{I}(\Sigma_1)$) of finitist arithmetic and once we have endowed that language with some truth predicate(s), we should be in a position to state the statements to which the truth predicate(s) apply and the axioms which involve the truth predicate(s) in that very same language (fragment) – and we should be in that position irrespectively of which layer of the finitist (Grzegorczyk) hierarchy we consider.

In Chapters 4 and 5, we will work in reasonable base theories (with smash function or exponentiation, respectively) and derive several theorems that will be used for the investigation of finitist truth. Chapter 4 focuses on the sequence coding and forms of bounded recursion. The focus of Chapter 5 is on syntax coding and valuation of terms as miniaturized truth theory.

Part I

Prerequisites

2 Key Concepts

The purpose of this chapter is to define some key concepts and to establish some preliminary results for later investigations. For starters, we fix the logical terminology.

2.1 Logic

The languages that are employed in this thesis are first-order and one-sorted. *Symbols* are restricted to the following.

- Individual Variables: $a_0, a_1, \ldots$
- (k-ary) Function Symbols: $f_0^k, f_1^k, \ldots$
- (k-ary) Relation Symbols: $R_0^k, R_1^k, \ldots$
- Propositional Connectives: $\sim, \vee, \wedge$
- Quantifiers: $\exists, \forall$
- Auxiliary Symbols: brackets, comma

In particular, symbols may include the following.

- (0-ary) Constants: $f_0^0, f_1^0, \ldots$
- (0-ary) Propositional Variables: $R_0^0, R_1^0, \ldots$

Remark 2.1. As a peculiarity, negation $\neg$ will not be used as a primitive logical constant, but will be defined from the primitive logical constants (in Definition 2.8) only later.

(Since we will frequently be concerned with computations as positive operations in this study, this treatment of negation will prove convenient.)

Let $\mathcal{L}$ be some class of function and relation symbols. Then $\mathcal{L}$ uniquely determines a *language*, conceived as the collection of all well-formed formulas, as follows.

Definition 2.2. $\mathcal{L}$ *terms* are defined recursively.

- Individual variables are $\mathcal{L}$ terms.
- Constants from $\mathcal{L}$ are $\mathcal{L}$ terms.
- If $t_0, \ldots, t_k$ are $\mathcal{L}$ terms and f is some $(k+1)$-ary function symbol from $\mathcal{L}$, then $f(t_0, \ldots, t_k)$ is an $\mathcal{L}$ term.

Definition 2.3. The *rank* $rk(t)$ of term t is defined recursively.

- $rk(t) := 0$ for variables t
- $rk(t) := 0$ for constants t from $\mathcal{L}$
- $rk(f(t_0, \ldots, t_k)) := \max\{rk(t_0), \ldots, rk(t_k)\} + 1$

Definition 2.4. $\mathcal{L}$ *atoms* are defined as follows.

- Propositional variables from $\mathcal{L}$ are $\mathcal{L}$ atoms.
- If $t_0, \ldots, t_k$ are $\mathcal{L}$ terms and R is some $(k+1)$-ary relation symbol from $\mathcal{L}$, then $R(t_0, \ldots, t_k)$ is an $\mathcal{L}$ atom.

Definition 2.5. $\mathcal{L}$ *literals* are defined as follows.

- $\mathcal{L}$ atoms are (positive) $\mathcal{L}$ literals.
- If A is an $\mathcal{L}$ atom, then ${\sim}A$ is a (negative) $\mathcal{L}$ literal.

Definition 2.6. $\mathcal{L}$ *formulas* are defined recursively. (Notice: In what follows, x is a variable).

- $\mathcal{L}$ literals are $\mathcal{L}$ formulas.
- If B, C are $\mathcal{L}$ formulas, then $(B \vee C)$ is an $\mathcal{L}$ formula.
- If B, C are $\mathcal{L}$ formulas, then $(B \wedge C)$ is an $\mathcal{L}$ formula.
- If B is an $\mathcal{L}$ formula, then $\exists x B$ is an $\mathcal{L}$ formula.
- If B is an $\mathcal{L}$ formula, then $\forall x B$ is an $\mathcal{L}$ formula.

Definition 2.7. The *rank* $rk(A)$ of formulas A is defined recursively.

- $rk(A) := 0 \quad \text{for literals } A$
- $rk(B \vee C) := \max\{rk(B), rk(C)\} + 1$
- $rk(B \wedge C) := \max\{rk(B), rk(C)\} + 1$
- $rk(\exists x B) := rk(B) + 1$
- $rk(\forall x B) := rk(B) + 1$

Definition 2.8. The *negation* $\neg A$ of formulas A is defined recursively.

- $\neg A := {\sim}A \quad \text{for positive literals } A$
- $\neg({\sim}B) := B$
- $\neg(B \vee C) := (\neg B \wedge \neg C)$
- $\neg(B \wedge C) := (\neg B \vee \neg C)$
- $\neg \exists x B := \forall x \neg B$
- $\neg \forall x B := \exists x \neg B$

Proposition 2.9. *$rk(\neg A)$ and $rk(A)$ are identical.*

Proposition 2.10. *$\neg\neg A$ and A are identical.*

To simplify terminology, it will often be convenient to not distinguish between a given class $\mathcal{L}$ of function and relation symbols and the language, conceived as class of all well-formed formulas, that is determined by $\mathcal{L}$.

The theories that will be studied in what follows – except the theory of Primitive Recursive Arithmetic that we will consider in Section 2.3 – are based on *classical first-order logic with identity.* Let $\mathcal{L}$ be a language that includes the binary relation symbol $=$. Let S be a class of pairs (Γ, A), where Γ is a finite class of $\mathcal{L}$ formulas and A is an $\mathcal{L}$ formula. Pairs (Γ, A) will also be called *inferences.* Whenever Γ is empty, they will be called *axioms.* To simplify notation, we will usually not distinguish between $(\varnothing, A)$ and A. In addition, we will often conceive $\{A\}$ simply as A.

Now, S uniquely determines a *theory*, conceived as a class of all theorems, as follows: A is a theorem if and only if there is a finite sequence $B_0, \dots, B_n$ so that B_n is A and, for every $i \leq n$, B_i belongs to a pair (Γ, B_i), where $\Gamma \subseteq \{B_j : j < i\}$ and the pair (Γ, B_i) is one of the

- axioms/inferences for the logical constants,
- axioms for the identity symbol,
- axioms/inferences from S.

The whole sequence $B_0, \dots, B_n$ will also be called a *derivation* or *proof* of length $n + 1$.

The axioms or inferences of the first two items constitute the *logical part* of a theory; they are the *logical inferences* (Γ, A) and *logical axioms* A, respectively. If the logical inferences and logical axioms of a theory contain only collections Γ of $\mathcal{L}$ formulas and $\mathcal{L}$ formulas A, the theory is called an $\mathcal{L}$ *theory.*

In what follows, we take it for granted that the axioms and inferences from the first two items are strongly sound and strongly complete: A is derivable from Γ by means of these axioms and inferences if and only if A semantically follows from Γ due to the semantics of classical first-order logic with identity.

The axioms and inferences from the third item constitute the *non-logical part* of a theory; they are the *non-logical axioms* and *non-logical inferences*, respectively.

To simplify our terminology, it will often be convenient to not distinguish between a class S of inferences (and axioms) and the theory, the class of all theorems, which is determined by S. We write $\mathsf{S} \vdash A$ to mean: A belongs to this class of theorems.

The following meta variables will be used.

- Individual Variables: x, y, z, u, v, w
- Terms: s, t
- Formulas: A, B, C, D, E, F, G, H
- Theories: $\mathsf{R}, \mathsf{S}, \mathsf{T}$

If necessary, subscripts may be added to meta variables.

To improve readability, we will usually omit outermost brackets and quantifier brackets of formulas. If helpful, we may even add redundant brackets. For function and relation symbols, we will frequently use the infix notation and write, for example, $t_0 + t_1$ rather than $+(t_0, t_1)$, where $+$ is the binary function symbol for addition and t_0, t_1 are terms.

We will indicate a sequence $t_0, \ldots, t_k$ of terms by notation $\vec{t}$. Given term t and variables $\vec{x}$, we will write $t(\vec{x})$ to indicate that all variables that occur (freely) in t are among $\vec{x}$. Similarly, for formulas A, we will write $A(\vec{x})$ to indicate that all the variables that occur *freely* in A are among $\vec{x}$.

For terms t, we will write

$$t(s_0, \ldots, s_k / x_0, \ldots, x_k)$$

or, if there is no danger of confusion, $t(s_0, \ldots, s_k)$ to denote the term that results by simultaneously substituting, for every $i \leq k$, every (free) occurrence of the variable x_i by the term s_i.

Simultaneous substitutions

$$A(s_0, \ldots, s_k/x_0, \ldots, x_k)$$

or simply $A(s_0, \ldots, s_k)$ are defined for the formulas in a similar fashion: They yield the formula which results by simultaneously substituting, for every $i \leq k$, every *free* occurrence of the variable x_i by the term s_i. Thus, for example,

$$\begin{array}{l} \forall x_0 B(x_0, x_1, x_2)(s_0, s_1/x_0, x_1) \\ \forall x_0 B(x_0, x_1, x_2)(s_1/x_1) \end{array}$$

both denote the formula $\forall x_0 B(x_0, s_1, x_2)$. Notice that we do not assume that, *by the definition of the substitution function*, bound variables are replaced if a variable that occurs (freely) in term s_i becomes bound in the formula $A(s_0, \ldots, s_k)$. This assumption is not needed because variable clashes may always be prevented by renaming bound variables *before* a substitution is performed. So, for example,

$$\forall x_0 B(x_0, x_1)(s(x_0)/x_1)$$

denotes $\forall x_0 B(x_0, s(x_0))$, but the variable clash may be prevented by first replacing x_0 by (say) x_2 in $\forall x_0 B(x_0, x_1)$ and performing

$$\forall x_2 B(x_2, x_1)(s(x_0)/x_1)$$

to yield $\forall x_2 B(x_2, s(x_0))$ only afterwards. We shall reconsider this issue very briefly in connection with the coding of syntax later in Remark 5.9.

Notation: To save horizontal space, we sometimes write

$$A\left(\frac{s_0, \ldots, s_k}{x_0, \ldots, x_k}\right)$$

in place of $A(s_0, \ldots, s_k/x_0, \ldots, x_k)$. Similarly for terms.

Finally, the following abbreviations will be adopted for every language that includes the binary relation symbols $=, \leq$.

$$s \neq t := \neg(s = t)$$
$$s < t := s \leq t \wedge s \neq t$$

Propositional connectives $\to$ and $\leftrightarrow$ are defined as follows.

$$A \to B := (\neg A \vee B)$$
$$A \leftrightarrow B := (A \to B) \wedge (B \to A)$$

Bounded quantifiers are defined as

$$(\exists x \leq t)A := \exists x(x \leq t \wedge A)$$
$$(\forall x \leq t)A := \forall x(x \leq t \to A)$$

where it is required that x does not occur in term t. Similarly for $(\exists x < t)A$ and $(\forall x < t)A$. Blocks of quantifiers are defined as follows.

$$(\exists x_0, \ldots, x_k)A := \exists x_0 \cdots \exists x_k A$$
$$(\forall x_0, \ldots, x_k)A := \forall x_0 \cdots \forall x_k A$$

2.2 Complexity

The following complexity notions will be used.

Definition 2.11. *Quantifier-free* (q.f.) formulas are defined recursively.

- Literals are q.f.
- If B, C are q.f., then $(B \vee C)$ is q.f.
- If B, C are q.f., then $(B \wedge C)$ is q.f.

Definition 2.12. $\Delta_0^{\mathcal{L}}$ formulas are defined recursively. (Note: In what follows, t is an $\mathcal{L}$ term.)

- $\mathcal{L}$ literals are $\Delta_0^{\mathcal{L}}$.
- If B, C are $\Delta_0^{\mathcal{L}}$, then $(B \vee C)$ is $\Delta_0^{\mathcal{L}}$.
- If B, C are $\Delta_0^{\mathcal{L}}$, then $(B \wedge C)$ is $\Delta_0^{\mathcal{L}}$.
- If B is $\Delta_0^{\mathcal{L}}$, then $(\exists x \leq t)B$ is $\Delta_0^{\mathcal{L}}$.
- If B is $\Delta_0^{\mathcal{L}}$, then $(\forall x \leq t)B$ is $\Delta_0^{\mathcal{L}}$.

Moreover, A is $\Delta_0^{\mathcal{L}}$ if and only if it is $\Sigma_0^{\mathcal{L}}$ if and only if it is $\Pi_0^{\mathcal{L}}$.

Definition 2.13. Formula A is $\Sigma_{n+1}^{\mathcal{L}}$ if and only if

- A is $\exists x B$,

where B is $\Pi_n^{\mathcal{L}}$; formula A is $\Pi_{n+1}^{\mathcal{L}}$ if and only if

- A is $\forall x B$,

where B is $\Sigma_n^{\mathcal{L}}$.

Definition 2.14. A formula A is *essentially* $\Sigma_0^{\mathcal{L}}$ (written $\mathrm{E}\Sigma_0^{\mathcal{L}}$) if and only if it is *essentially* $\Pi_0^{\mathcal{L}}$ (written $\mathrm{E}\Pi_0^{\mathcal{L}}$) if and only if it is simply $\Delta_0^{\mathcal{L}}$.

Definition 2.15. $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$ formulas are defined recursively. (Note: In what follows, t is an $\mathcal{L}$ term.)

- $\mathrm{E}\Pi_n^{\mathcal{L}}$ formulas are $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$.
- If B, C are $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$, then $(B \vee C)$ is $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$.
- If B, C are $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$, then $(B \wedge C)$ is $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$.
- If B is $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$, then $(\forall x \leq t)B$ is $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$.
- If B is $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$, then $\exists x B$ is $\mathrm{E}\Sigma_{n+1}^{\mathcal{L}}$.

$\mathrm{E}\Pi_{n+1}^{\mathcal{L}}$ formulas are defined dually.

Proposition 2.16. *For arbitrary $\mathcal{L}$ terms t, the following hold.*

- *If B is $\mathrm{E}\Sigma_n^{\mathcal{L}}$, then $(\exists x \leq t)B$ is $\mathrm{E}\Sigma_n^{\mathcal{L}}$.*

- *If B is* $\mathrm{E}\Pi_n^{\mathcal{L}}$*, then* $(\forall x \leq t)B$ *is* $\mathrm{E}\Pi_n^{\mathcal{L}}$*.*

Proof. Immediate for $n = 0$. As for the first item, given $n > 0$, notice that $(x \leq t)$ is $\mathrm{E}\Sigma_0^{\mathcal{L}}$ and apply the clauses for conjunction and existential quantification. Dually for the second item. □

Proposition 2.17. *For arbitrary n, the following hold.*

- *If A is* $\Sigma_n^{\mathcal{L}}$*, then A is* $\mathrm{E}\Sigma_n^{\mathcal{L}}$*.*
- *If A is* $\Pi_n^{\mathcal{L}}$*, then A is* $\mathrm{E}\Pi_n^{\mathcal{L}}$*.*

For the structure of certain proofs, we will sometimes tacitly apply:

Proposition 2.18. *A formula A is* $\mathrm{E}\Sigma_1^{\mathcal{L}}$ *if and only if one of the following holds, where t is an $\mathcal{L}$ term.*

- *A is an $\mathcal{L}$ literal.*
- *A is $(B \vee C)$, where B, C are* $\mathrm{E}\Sigma_1^{\mathcal{L}}$*.*
- *A is $(B \wedge C)$, where B, C are* $\mathrm{E}\Sigma_1^{\mathcal{L}}$*.*
- *A is $(\forall x \leq t)B$, where B is* $\mathrm{E}\Sigma_1^{\mathcal{L}}$*.*
- *A is $\exists x B$, where B is* $\mathrm{E}\Sigma_1^{\mathcal{L}}$*.*

2.3 Primitive Recursive Arithmetic

We next define the theory of *Primitive Recursive Arithmetic.* In accordance with some classic references, such as Skolem [46] and Hilbert and Bernays [26], we shall define the theory of Primitive Recursive Arithmetic in a language without quantifiers: no quantification over the *infinite* domain of the natural numbers. Note that Primitive Recursive Arithmetic will be the only theory in the current study which is not a theory in the sense of Section 2.1.

To highlight the difference, we write $\mathfrak{L}$ for classes of function and relation symbols if they determine a class of formulas not in the sense of Definition 2.6, but in the sense of 2.6 without the closure conditions for the quantifiers.

Similarly, we write S for classes of inferences (and axioms) if they determine a theory (class of theorems) not in the sense of a derivation as defined in Section 2.1, but in the sense of such a derivation in which no inferences (axioms) for the quantifiers are used.

We first define numerals.

Definition 2.19. For any extension of $\{\overline{0}, +\overline{1}\}$, where $\overline{0}$ (zero) is a constant and $+\overline{1}$ (successor) some unary function symbol, the *numeral* $\overline{n}$ is defined recursively as follows.

- $\overline{0} := \overline{0}$
- $\overline{n+1} := \overline{n} + \overline{1}$

The language $\mathfrak{L}_{\text{PRA}}$ and the respective $\mathfrak{L}_{\text{PRA}}$ theory PRA of *Primitive Recursive Arithmetic* are now defined simultaneously as follows.

Definition 2.20. $\mathfrak{L}_{\text{PRA}}$ is an extension of $\{\overline{0}, +\overline{1}, =\}$, where $\overline{0}$ (zero) is a constant, $+\overline{1}$ (successor) an unary function symbol, and $=$ (identity) a binary relation symbol. PRA extends the pair of axioms:

(1) $x + \overline{1} \neq \overline{0}$

(2) $x + \overline{1} = y + \overline{1} \rightarrow x = y$

The extensions are given as follows.

(3) For an unary function symbol f in $\mathcal{L}_{\text{PRA}}$, PRA contains the axiom:

$$f(x) = \overline{0}$$

(4) For every k and every $i \leq k$, there is some $(k+1)$-ary function symbol f in $\mathcal{L}_{\text{PRA}}$ so that PRA comprises the axiom:

$$f(x_0, \ldots, x_k) = x_i$$

(5) For every k-ary function symbol $f_0, \ldots, f_n$ in $\mathcal{L}_{\text{PRA}}$ and for every $(n+1)$-ary function symbol f_{n+1} in $\mathcal{L}_{\text{PRA}}$, there is some k-ary function symbol f in $\mathcal{L}_{\text{PRA}}$ so that PRA comprises the axiom:

$$f(x_1, \ldots, x_k) = f_{n+1}(f_0(x_1, \ldots, x_k), \ldots, f_n(x_1, \ldots, x_k))$$

(6) For any k-ary function symbol f_0 in $\mathcal{L}_{\text{PRA}}$ and $(k+2)$-ary function symbol f_1 in $\mathcal{L}_{\text{PRA}}$, there also is a $(k+1)$-ary function symbol f in $\mathcal{L}_{\text{PRA}}$ such that PRA comprises the axioms:

$$\begin{aligned} f(x_1, \ldots, x_k, \overline{0}) &= f_0(x_1, \ldots, x_k) \\ f(x_1, \ldots, x_k, y + \overline{1}) &= f_1(f(x_1, \ldots, x_k, y), x_1, \ldots, x_k, y) \end{aligned}$$

(7) For any (quantifier-free) $\mathfrak{L}_{\text{PRA}}$ formula $A(\vec{x}, y)$, where y is not among $\vec{x}$, PRA comprises the induction rule:

$$\frac{A(\vec{x}, \overline{0}) \wedge \big(A(\vec{x}, y) \rightarrow A(\vec{x}, y + \overline{1})\big)}{A(\vec{x}, y)}$$

The subject of the present section is to make clear that, with regard to finitistically meaningful Σ_1 formulas (Π_2 sentences) with quantifiers, the theory $\mathbf{I}(\Sigma_1)$ of Σ_1 induction that we shall define in every detail below in 3.4 is *reducible* to PRA. Technically, quantifier-free formulas from $\mathfrak{L}_{\text{PRA}}$ correspond to Σ_1 formulas in the language of $\mathbf{I}(\Sigma_1)$ if the additional (primitive recursive) function symbols from $\mathfrak{L}_{\mathrm{PRA}}$ are Σ_1 defined in language of $\mathbf{I}(\Sigma_1)$, as shown below. Intuitively, quantifiers can be conceived to be auxiliary means to express partial (or incomplete) statements. In the words of Hilbert and Bernays [26], pp. 32f.:

> Ein *Existenzsatz* über Ziffern, also ein Satz von der Form "es gibt eine Ziffer $\mathfrak{n}$ von der Eigenschaft $\mathfrak{A}(\mathfrak{n})$", ist finit aufzufassen als ein "Partialurteil", d.h. als eine unvollständige Mitteilung einer genauer bestimmten Aussage, welche entweder in der direkten Angabe einer Ziffer von der Eigenschaft $\mathfrak{A}(\mathfrak{n})$ oder der Angabe eines Verfahrens zur Gewinnung einer solchen Ziffer besteht, – wobei zur Angabe eines Verfahrens gehört, dass für die Reihe der auszuführenden Handlungen eine bestimmte Grenze aufgewiesen wird.
>
> In entsprechender Weise sind diejenigen Urteile finit zu interpretieren, in denen eine allgemeine Aussage mit einer Existenzbehauptung verknüpft ist. So hat man z.B. einen Satz von der Form "zu jeder Ziffer $\mathfrak{k}$ von der Eigenschaft $\mathfrak{A}(\mathfrak{k})$ gibt es eine Ziffer $\mathfrak{l}$, für welche $\mathfrak{B}(\mathfrak{k}, \mathfrak{l})$ gilt" finit aufzufassen als unvollständige Mitteilung von einem Verfahren, welches gestattet, zu jeder vorgelegten Ziffer $\mathfrak{k}$ von der Eigenschaft $\mathfrak{A}(\mathfrak{k})$ eine Ziffer $\mathfrak{l}$ zu finden, welche zu $\mathfrak{k}$ in der Beziehung $\mathfrak{B}(\mathfrak{k}, \mathfrak{l})$ steht.

We may translate this as follows:

> An *existential claim* about numbers, i.e., a statement of the form "there exists a number $\mathfrak{n}$ with the property $\mathfrak{A}(\mathfrak{n})$", is to be conceived finitistically as a "partial judgment", i.e., as an incomplete rendering of a more concrete claim[, where the concrete claim] consists in the direct specification of a number with property $\mathfrak{A}(\mathfrak{n})$ or the specification of a method to obtain such a number,

> – where the specification of a method involves that the series of actions to be taken be limited by a specific bound.
>
> Judgments in which a universal claim is combined with an existential claim are to be understood finitistically in an analogous way. For example, a sentence of the form "for every number $\mathfrak{t}$ with property $\mathfrak{A}(\mathfrak{t})$, there exists a number $\mathfrak{l}$ such that $\mathfrak{B}(\mathfrak{t}, \mathfrak{l})$" is to be conceived finitistically as incomplete rendering of a method which allows one to find, for every given number $\mathfrak{t}$ with property $\mathfrak{A}(\mathfrak{t})$, a number $\mathfrak{l}$ which stands in the relationship $\mathfrak{B}(\mathfrak{t}, \mathfrak{l})$ to $\mathfrak{t}$.

Specifically, we want to briefly recall a well-known theorem by Parsons [40] which shows the Π_2 equivalence of $\mathbf{I}(\Sigma_1)$ and PRA – in a precise sense to be explained. As we will see, the upshot of Parsons' result is that, with regard to the finitistically meaningful Π_2 sentences (Σ_1 formulas) in the language of $\mathbf{I}(\Sigma_1)$, $\mathbf{I}(\Sigma_1)$ has – modulo inter-translation – precisely the same theorems as PRA.

Note that, according to its Definition 2.20, the language $\mathfrak{L}_{\mathrm{PRA}}$ includes a binary symbol $+$ for addition, a binary symbol $\cdot$ for multiplication, and a binary symbol

$$f_{\leq}(x, y) := \overline{1} - (x - y)$$

that uses (modified) subtraction – and is such that

$$\begin{cases} \mathrm{PRA} \vdash f_{\leq}(\overline{m}, \overline{n}) = \overline{1} & : m \leq n \\ \mathrm{PRA} \vdash f_{\leq}(\overline{m}, \overline{n}) = \overline{0} & : m > n \end{cases}$$

As will be defined in all detail in 3.4, we take $\mathbf{I}(\Sigma_1)$ to be a theory with an induction schema for Σ_1 formulas in the language

$$\{\overline{0}, +\overline{1}, +, \cdot, =, \leq\}, \tag{2.1}$$

where $\leq$ is the binary symbol for the less-than relation.

With regard to the theories of truth to be studied later, it is illustrative to see that Σ_1 formulas in (2.1) may be used to characterize primitive recursive functions in the following sense: For every function symbol $f(\vec{x})$ in $\mathfrak{L}_{\mathrm{PRA}}$, there is such a Σ_1 formula $F(\vec{x}, y)$ such that, in an encompassing theory $\mathbf{U}$ that unifies PRA and $\mathbf{I}(\Sigma_1)$, we have

$$\mathbf{U} \vdash f(\vec{x}) = y \leftrightarrow F(\vec{x}, y) \tag{2.2}$$

and, in addition, we have

$$\mathbf{I}(\Sigma_1) \vdash \exists y F(\vec{x}, y) \tag{2.3}$$

$$\mathbf{I}(\Sigma_1) \vdash F(\vec{x}, y) \wedge F(\vec{x}, z) \rightarrow y = z \tag{2.4}$$

where (2.3) is the totality condition and (2.4) is the uniqueness condition.

To see how $F(\vec{x}, y)$ may be defined, note that Σ_1 formulas in (2.1) can be employed to define sequence numbers $\langle x_0, \ldots x_k \rangle$, projections $(x)_y$ on the y-th member of a sequence x, concatenations $x * y$ of sequences x and y, length $lh(x)$ of a sequence x, and the relation $x \in y$ that x is a member of the sequence y. Focussing on 2.20 (3)–(6), and using suggestive notation rather than official formal language, we may code function symbols f from $\mathfrak{L}_{\mathrm{PRA}}$, respectively, as

- Constant Function: $\langle \ulcorner 0 \urcorner \rangle$
- Projection: $\langle \ulcorner \mathrm{P} \urcorner, k, i \rangle$
- Composition: $\langle \ulcorner \mathrm{C} \urcorner, \ulcorner f_0 \urcorner, \ldots, \ulcorner f_n \urcorner, \ulcorner f_{n+1} \urcorner \rangle$
- Primitive Recursion: $\langle \ulcorner \mathrm{R} \urcorner, \ulcorner f_0 \urcorner, \ulcorner f_1 \urcorner \rangle$

where $\ulcorner\ \urcorner$ are specific code numbers. So, $\langle \langle \langle \ulcorner \mathrm{P} \urcorner, 1, 0 \rangle, \langle x_0, x_1 \rangle \rangle, x_0 \rangle$, for example, may be taken to express that projection $\langle \ulcorner \mathrm{P} \urcorner, 1, 0 \rangle$ applied to x_0, x_1 yields the value x_0.

To define an operator $\mathcal{V} : \mathbb{N} \rightarrow \mathbb{N}$ whose inputs/outputs are sequences of tuples $\langle \langle z, x \rangle, y \rangle$, intended to represent that y is the value of z applied to sequence x of arguments, we consider some disjunction $\mathbb{V}$ that expresses that, given a sequence $(s)_i$ of such tuples $\langle\langle z, x\rangle, y\rangle$, under which conditions r is such a tuple again. Hence, for example, the cases 2.20 (3)–(6) may be described as follows.

- Constant Function:

$$\begin{aligned}(\exists x, y, z)\big(r = \langle\langle z, x\rangle, y\rangle \wedge (z)_0 = \ulcorner 0 \urcorner \wedge \\ y = 0\big)\end{aligned}$$

- Projection:

$$\begin{aligned}(\exists x, y, z)\big(r = \langle\langle z, x\rangle, y\rangle \wedge (z)_0 = \ulcorner \mathrm{P} \urcorner \wedge \\ (\exists u \leq (z)_1)\big(u = (z)_2 \wedge y = (x)_u\big)\big)\end{aligned}$$

- Composition:

$$\begin{aligned}(\exists x, y, z)\big(r = \langle\langle z, x\rangle, y\rangle \wedge (z)_0 = \ulcorner \mathrm{C} \urcorner \wedge \\ \exists u\big((\forall w < lh(u))\big(\langle\langle (z)_{w+1}, x\rangle, (u)_w\rangle \in (s)_i\big) \wedge \\ \langle\langle (z)_{lh(z)-1}, u\rangle, y\rangle \in (s)_i\big)\end{aligned}$$

- Primitive Recursion:

$$\begin{aligned}(\exists x, y, z, u_1)\big(r = \langle\langle z, x * \langle u_1\rangle\rangle, y\rangle \wedge (z)_0 = \ulcorner \mathrm{R} \urcorner \wedge \\ \exists u_2\big(\langle\langle (z)_1, x\rangle, (u_2)_0\rangle \in (s)_i \wedge \\ (\forall w < u_1)\big(\langle\langle (z)_2, \langle (u_2)_w\rangle * x * \langle w\rangle\rangle, (u_2)_{w+1}\rangle \in (s)_i\big) \wedge \\ (u_2)_{u_1} = y\big)\big)\end{aligned}$$

For example, the last condition reads: If z is given by primitive recursion from $(z)_1, (z)_2$ and applied to $x * \langle u_1\rangle$, then its value y is the last member of a sequence of length $u_1 + 1$, with the 0-th member being the value of applying $(z)_1$ to x, and the $(w + 1)$-th member being the value of applying $(z)_2$ to the w-th member concatenated with $x * \langle w\rangle$.

Having defined this disjunction $\mathbb{V}$, we may next define finite valuation sequences ($\mathbb{V}Seq$): sequences s of $l+1$ steps of computing values on the basis of values that have been obtained by sub-computations at preceding steps. The definition of $\mathbb{V}Seq(s,l)$ is:

$$\begin{aligned}&Seq(s) \wedge lh(s) = l+1 \wedge\\ &(\forall i \leq l) Seq((s)_i) \wedge (s)_0 = \langle\ \rangle \wedge\\ &(\forall i < l)(\forall r < s)\big(r \in (s)_i \rightarrow r \in (s)_{i+1} \wedge r \in (s)_{i+1} \rightarrow \mathbb{V}\big)\end{aligned}$$

Now, $Val(\langle z, x\rangle, y)$ – expressing that y is the value of applying function z to sequence x of arguments – may be defined as: $\langle\langle z, x\rangle, y\rangle$ obtains in the final step of a valuation sequence. So, $Val(\langle z, x\rangle, y)$ is defined as follows.

$$(\exists s, l)\big(\mathbb{V}Seq(s,l) \wedge \langle\langle z, x\rangle, y\rangle \in (s)_l\big)$$

Eventually, for every function symbol f from $\mathfrak{L}_{\mathrm{PRA}}$, we may guarantee (2.2)–(2.4) above by taking $F(\vec{x}, y)$ to be the formula $Val(\langle\ulcorner f\urcorner, \langle\vec{x}\rangle\rangle, y)$, where $\ulcorner f\urcorner$ is a code of f as indicated above. Indeed, (2.2) restated as

$$f(\vec{x}) = y \leftrightarrow Val(\langle\ulcorner f\urcorner, \langle\vec{x}\rangle\rangle, y) \tag{2.5}$$

is an instance of a (Kleene) *normal form* for (finite/terminating) primitive recursive computations.

Together with (2.3) and (2.4), we may use (2.5) in order to mimic in $\mathbf{I}(\Sigma_1)$ any PRA induction – as stated in 2.20 (7) – by some Σ_1 induction, using the Σ_1 formulas by which the function symbols from $\mathfrak{L}_{\mathrm{PRA}}$ have been defined.

In a nutshell, the upshot is this: If an $\mathfrak{L}_{\mathrm{PRA}}$ formula has a derivation in PRA, that formula can be interpreted by some Σ_1 formula in (2.1) such that the interpretation is adequate in the sense of (2.2) and derivable in $\mathbf{I}(\Sigma_1)$.

Conversely, even though less illustrative for our investigations following below: If some Σ_1 formula in (2.1) is derivable inside

$\mathbf{I}(\Sigma_1)$, we may systematically assign to it a quantifier-free $\mathfrak{L}_{\mathrm{PRA}}$ formula which is adequate in the sense of (2.2) and provable in PRA. The assigned $\mathfrak{L}_{\mathrm{PRA}}$ formula expresses, very roughly, how introductions of existential quantifiers are witnessed by particular numbers in the course of a (partial cut-free) $\mathbf{I}(\Sigma_1)$ derivation. We do not go into details here, but refer the reader to, for example, Buss [4], Section 3.1.1.

Summing up: Even though $\mathbf{I}(\Sigma_1)$ is formulated in a language with quantifiers, it comprises – modulo an inter-translation – the same finitistically meaningful theorems as PRA.

2.4 Retractability

So far, we found that the arithmetical principles of $\mathbf{I}(\Sigma_1)$, even though they are partly ideal/meaningless, imply – modulo inter-translation – precisely the real/meaningful theorems of the finitist arithmetic. It is in this regard that we consider the arithmetical principles of (sub-theories of) $\mathbf{I}(\Sigma_1)$ to be *finitistically safe* and insouciantly apply them in arithmetical interpretations of principles of truth.

Relatedly, let us put our principle task of Section 1.2 in the following form, where T is (sub-)theory (of) $\mathbf{I}(\Sigma_1)$: Define a truth theory S and a logic-preserving mapping $\circ : \mathcal{L}_{\mathsf{S}} \to \mathcal{L}_{\mathsf{T}}$ from the language $\mathcal{L}_{\mathsf{S}}$ of S to the language $\mathcal{L}_{\mathsf{T}}$ of T, the former being an extension of the latter, so that the following hold.

- For all $\mathcal{L}_{\mathsf{S}}$ formulas A, $\mathsf{S} \vdash A$ implies $\mathsf{T} \vdash A^{\circ}$.
- For all $\mathcal{L}_{\mathsf{T}}$ formulas A, $\mathsf{S} \vdash A$ if and only if $\mathsf{T} \vdash A$.

(Also, for all $\mathcal{L}_{\mathsf{T}}$ formulas A, $\mathsf{T} \vdash A \leftrightarrow A^{\circ}$.)

Observe that the second item indeed yields the consistency of S: Since $\mathsf{T} \nvdash \overline{0} \neq \overline{0}$, we have by second item that $\mathsf{S} \nvdash \overline{0} \neq \overline{0}$.

To conclude this preliminary chapter, we introduce a proof-theoretic relation, called *the relation of retractability*, tailored to yield precisely (1) and (2). We present the following material in a heuristic way: That is, the relation of retractability is introduced successively by several propositions that show how a theory may actually be retracted to another theory.

We first consider the commutation with logical operators.

Definition 2.21. Let T be an $\mathcal{L}_\mathsf{T}$ theory and let $\mathcal{L}$ be a language. A mapping $\circ : \mathcal{L} \to \mathcal{L}_\mathsf{T}$ *commutes with the logical operators in* T if and only if, for all arbitrary $\mathcal{L}$ formulas B, C, the following are derivable in T.

- $(\sim B)^\circ \leftrightarrow \neg(B^\circ)$
- $(B \vee C)^\circ \leftrightarrow (B^\circ \vee C^\circ)$
- $(B \wedge C)^\circ \leftrightarrow (B^\circ \wedge C^\circ)$
- $(\exists x B)^\circ \leftrightarrow \exists x (B)^\circ$
- $(\forall x B)^\circ \leftrightarrow \forall x (B)^\circ$

If a mathematical induction is performed outside the theories under consideration, we usually indicate this by calling it a *meta* induction. (But see Remark 2.33.)

Proposition 2.22. *Let* T *be an* $\mathcal{L}_\mathsf{T}$ *theory and* $\mathcal{L}$ *a language. Suppose that* $\circ : \mathcal{L} \to \mathcal{L}_\mathsf{T}$ *commutes with logical operators in* T*. Then, for all* $\mathcal{L}$ *formulas* B, C*, the following are derivable in* T*.*

- $(\neg B)^\circ \leftrightarrow \neg(B^\circ)$
- $(B \to C)^\circ \leftrightarrow (B^\circ \to C^\circ)$
- $(B \leftrightarrow C)^\circ \leftrightarrow (B^\circ \leftrightarrow C^\circ)$

Proof. By (meta) induction on the rank of B. We consider only a few examples of the first item.

- B is a positive literal. Then $\neg B$ is $\sim B$. Hence, the claim is immediate by Definition 2.21.

- B is $\sim C$. Then apply Definition 2.8 to obtain that $\neg B$ is C. By Definition 2.21, moreover, T proves $(\sim C)^\circ \leftrightarrow \neg(C^\circ)$. So, by logic, T proves $\neg(\sim C)^\circ \leftrightarrow \neg\neg(C^\circ)$, that is, T proves $\neg(B^\circ) \leftrightarrow (\neg B)^\circ$.
- B is $(C \vee D)$. Then $\neg B$ is $\neg C \wedge \neg D$. Moreover, by Definition 2.21 and the induction hypothesis, T proves

$$\begin{aligned}(\neg C \wedge \neg D)^\circ &\leftrightarrow ((\neg C)^\circ \wedge (\neg D)^\circ)\\ &\leftrightarrow (\neg(C^\circ) \wedge \neg(D^\circ))\\ &\leftrightarrow \neg(C^\circ \vee D^\circ)\\ &\leftrightarrow \neg(C \vee D)^\circ\end{aligned}$$

which is $(\neg B)^\circ \leftrightarrow \neg(B^\circ)$, as required. ☐

Proposition 2.23. *Let* T *be an* $\mathcal{L}_\mathsf{T}$ *theory and* $\mathcal{L}$ *a language such that* $\mathcal{L}_\mathsf{T} \subseteq \mathcal{L}$. *Suppose that* $\circ : \mathcal{L} \to \mathcal{L}_\mathsf{T}$ *commutes with the logical operators in* T *and that, for any* $\mathcal{L}_\mathsf{T}$ *atom* A, $\mathsf{T} \vdash A \leftrightarrow A^\circ$. *Then, for any* $\mathcal{L}_\mathsf{T}$ *formula* B, *we have* $\mathsf{T} \vdash B \leftrightarrow B^\circ$.

Proof. By (meta) induction on the rank of B. ☐

Definition 2.24. Let T be an $\mathcal{L}_\mathsf{T}$ theory and $\mathcal{L}$ a language. A mapping $\circ : \mathcal{L} \to \mathcal{L}_\mathsf{T}$ *preserves logical structure in* T if and only if $\circ$ commutes with the logical operators in T and, for arbitrary $\mathcal{L}$ formulas B, C and $\mathcal{L}$ terms t, the following hold.

(1) $\mathsf{T} \vdash B(t/x)^\circ \to \exists x(B)^\circ$

(2) $\mathsf{T} \vdash \forall x(B)^\circ \to B(t/x)^\circ$

(3) $\mathsf{T} \vdash B^\circ \to C^\circ$ implies $\mathsf{T} \vdash \exists x(B)^\circ \to C^\circ$

(4) $\mathsf{T} \vdash C^\circ \to B^\circ$ implies $\mathsf{T} \vdash C^\circ \to \forall x(B)^\circ$

(5) $\mathsf{T} \vdash (x = x)^\circ$

(6) $\mathsf{T} \vdash (x = y)^\circ \to (B^\circ \to B(y/x)^\circ)$

For conditions (3) and (4), x is the eigenvariable and must not occur freely in C.

Notation: While $B^\circ(t/x)$ would mean the formula that is obtained from substituting term t for the variable x in the formula B°, $B(t/x)^\circ$ denotes the result of applying the function $\circ$ to the formula $B(t/x)$ that is obtained by substituting t for x in B.

Remark 2.25. For every $\mathcal{L}$ formula B and $\mathcal{L}$ term t such that $\mathsf{T} \vdash B(t/x)^\circ \leftrightarrow B^\circ(t/x)$, meaning that t also is an $\mathcal{L}_\mathsf{T}$ term, we have 2.24 (1) and (2) by logic.

Similarly, if $\circ : \mathcal{L} \to \mathcal{L}_\mathsf{T}$ is such that any variable that occurs freely in C° must be free in C itself – which is the case if $\circ$ does not introduce any free variables – we may simply rename bound variables, if needed, and obtain by logic from $\mathsf{T} \vdash B^\circ \to C^\circ$ that $\mathsf{T} \vdash \forall x(B^\circ \to C^\circ)$ and, so, $\mathsf{T} \vdash \exists x(B)^\circ \to C^\circ$. This ensures 2.24 (3) and, in similar fashion, (4).

Let Γ be a finite class of formulas. We write $\bigwedge \Gamma$ and $\bigvee \Gamma$ for the conjunction and disjunction, respectively, of all the formulas comprised in Γ. (For definiteness, we may conceive the members $A_0, \ldots, A_n$ of Γ in lexicographical order and stipulate that $\bigwedge \Gamma$ is

$$\big(\cdots \big((A_0 \wedge A_1) \wedge A_2\big) \wedge \cdots \wedge A_n\big).$$

Similarly for $\bigvee \Gamma$.)

In addition, if Γ is a (possibly infinite) class of formulas, we write Γ° for the class of all formulas A° so that A belongs to Γ. Similarly for other superscripts.

We now define the proof-theoretic relation of *retractability by function* $\circ$. Note that, since all theories considered in this thesis are about the natural numbers, we may proof-theoretically relate these theories without putting any restrictions on the domain of quantified formulas $\exists x A$ and $\forall x A$, such as by employing $D(x)$ in

$\exists x(D(x) \land A)$ and $\forall x(D(x) \to A)$, respectively. This is one peculiarity of the relation of retractability that is not shared by usual proof-theoretic interpretations (see Shoenfield [44], 4.7).

Definition 2.26. Let S be an $\mathcal{L}_\mathsf{S}$ theory and T an $\mathcal{L}_\mathsf{T}$ theory, where $\mathcal{L}_\mathsf{T} \subseteq \mathcal{L}_\mathsf{S}$. Furthermore, let the function $\circ : \mathcal{L}_\mathsf{S} \to \mathcal{L}_\mathsf{T}$ preserve logical structure in T. We define: S is *retractable to* T *by* $\circ$ if and only if the following hold.

(1) For all $\mathcal{L}_\mathsf{T}$ atoms A, $\mathsf{T} \vdash A \leftrightarrow A^\circ$.

(2) For all non-logical A in S, $\mathsf{T} \vdash A^\circ$.

(3) For all non-logical A in T, $\mathsf{S} \vdash A$.

(4) For all non-logical (Γ, A) in S, $\mathsf{T} \vdash \bigwedge \Gamma^\circ$ implies $\mathsf{T} \vdash A^\circ$.

(5) For all non-logical (Γ, A) in T, $\mathsf{S} \vdash \bigwedge \Gamma$ implies $\mathsf{S} \vdash A$.

The following proposition shows the core implications of our proof-theoretic relation. Its first condition (1) is about *interpretation.* Its second condition (2) is about *conservation.*

Proposition 2.27. *If* S *is retractable to* T *by* $\circ$*, the following hold.*

(1) *For all* $\mathcal{L}_\mathsf{S}$ *formulas* A*,* $\mathsf{S} \vdash A$ *implies* $\mathsf{T} \vdash A^\circ$.

(2) *For all* $\mathcal{L}_\mathsf{T}$ *formulas* A*,* $\mathsf{S} \vdash A$ *if and only if* $\mathsf{T} \vdash A$.

Proof. We only sketch the proof.

- For (1), use a (meta) induction on the length of derivations in S. If A is a logical axiom, then $\mathsf{T} \vdash A^\circ$, since $\circ$ preserves logical structure in T. Thus, by 2.26 (2), for all axioms A of S, the logical as well as the non-logical ones, we obtain that $\mathsf{T} \vdash A^{\circ}$. For the induction step, we use the induction hypothesis that, given some inference (Γ, A) in S, we have $\mathsf{T} \vdash \bigwedge \Gamma^{\circ}$, and apply 2.26 (4).

- For the right-to-left direction of (2), notice first that, since $\mathcal{L}_{\mathsf{T}} \subseteq \mathcal{L}_{\mathsf{S}}$, any logical axiom of T and any logical inference of T is, respectively, a logical axiom of S and a logical inference of S. Therefore, by a (meta) induction on the length of derivations in T, like before, 2.26 (3) and 2.26 (5) yield the claim.

  For the left-to-right direction of (2), let A be an $\mathcal{L}_{\mathsf{T}}$ formula and assume that $\mathsf{S} \vdash A$. Then, since $\mathcal{L}_{\mathsf{T}} \subseteq \mathcal{L}_{\mathsf{S}}$, we apply (1) to yield $\mathsf{T} \vdash A^{\circ}$. So, the claim follows by 2.26 (1) and Proposition 2.23. □

Remark 2.28. It may be worth to highlight that Fujimoto [15] defines a relation of *relative truth definability* between *theories of truth* which resembles our relation of retractability in certain respects. While our relation serves the purpose of comparing some given theory S (of truth or arithmetic) to some theory T (of truth or arithmetic) in some restricted language ($\mathcal{L}_{\mathsf{T}} \subseteq \mathcal{L}_{\mathsf{S}}$), where the restriction is not further qualified, Fujimoto [15] leaves the arithmetical expressions untouched and considers the truth predicates only. So, given two truth theories R_1 and R_2 that are based on arithmetical sub-theories B_1 and B_2, respectively, where $\mathcal{L}_1$ is the language of B_1 and $\mathcal{L}_2$ the language of B_2, Fujimoto [15] requires $\mathcal{L}_1 \subseteq \mathcal{L}_2$ and defines, roughly put: R_1 is (relatively) truth definable in R_2 if there is a translation (say $\circ$) from the language of R_1 onto the language of R_2 – which basically translates the truth predicates of R_1 and – that commutes with logical operators and fulfills the condition 2.27 (1) for $\mathsf{S} = \mathsf{R}_1$ and $\mathsf{T} = \mathsf{R}_2$ above. (Notice that, since Fujimoto [15] leaves arithmetical terms unchanged, commutation with logical operators is enough in his case.)

We state two consequences of 2.27.

Corollary 2.29. *Let* S *be retractable to* T *by* $\circ$ *and, in addition, suppose that* S, T *both are* $\mathcal{L}$ *theories. Then* S, T *are identical.*

Proof. By 2.27 (2). $\square$

Corollary 2.30. *Let* S *be retractable to* T *by* $\circ$ *and, in addition, suppose the following.*

For all $\mathcal{L}_{\mathsf{S}}$ *formulas* A*,* $\mathsf{S} \vdash A^\circ \rightarrow A$*.*

Then the following holds.

For all $\mathcal{L}_{\mathsf{S}}$ *formulas* A*,* $\mathsf{S} \vdash A$ *if and only if* $\mathsf{T} \vdash A^\circ$*.*

Proof. The left-to-right direction holds by 2.27 (1). For the other direction, let A be some $\mathcal{L}_{\mathsf{S}}$ formula and suppose $\mathsf{T} \vdash A^\circ$. Then, since A° is an $\mathcal{L}_{\mathsf{T}}$ formula, 2.27 (2) and $\mathsf{S} \vdash A^\circ \rightarrow A$ yield the claim. $\square$

In the technical proposition below, which we mention only for future reference, $(\circ \cdot \bullet) : \mathcal{L} \rightarrow \mathcal{L}_{\mathsf{T}}$ denotes the composition of $\circ : \mathcal{L}_{\mathsf{S}} \rightarrow \mathcal{L}_{\mathsf{T}}$ and $\bullet : \mathcal{L} \rightarrow \mathcal{L}_{\mathsf{S}}$, mapping $\mathcal{L}$ formulas A to $(A^\bullet)^\circ$.

Proposition 2.31. *Let* $\mathcal{L}_{\mathsf{T}} \subseteq \mathcal{L}_{\mathsf{S}} \subseteq \mathcal{L}$*. Suppose that* $\bullet : \mathcal{L} \rightarrow \mathcal{L}_{\mathsf{S}}$ *preserves logical structure in* S*, that* $\circ : \mathcal{L}_{\mathsf{S}} \rightarrow \mathcal{L}_{\mathsf{T}}$ *preserves logical structure in* T*, and that the following holds.*

For all $\mathcal{L}_{\mathsf{S}}$ *formulas* A*,* $\mathsf{S} \vdash A$ *if and only if* $\mathsf{T} \vdash A^\circ$*.*

Then $(\circ \cdot \bullet)$ *preserves logical structure in* T*.*

Proof. We consider only two examples.

- $\mathsf{T} \vdash ((\sim B)^\bullet)^\circ \leftrightarrow \neg((B^\bullet)^\circ)$, where B is a positive $\mathcal{L}$ literal. To show this, we first notice that, since $\bullet$ preserves logical structure in S, we have

  $$\mathsf{S} \vdash (\sim B)^\bullet \leftrightarrow \neg(B^\bullet).$$

  Thus, by assumption,

  $$\mathsf{T} \vdash ((\sim B)^\bullet \leftrightarrow \neg(B^\bullet))^\circ.$$

  Since $\circ$ preserves logical structure in T, we get by Proposition 2.22 the following in T.

  $$\begin{gathered} ((\sim B)^\bullet)^\circ \leftrightarrow (\neg(B^\bullet))^\circ \\ ((\sim B)^\bullet)^\circ \leftrightarrow \neg((B^\bullet)^\circ) \end{gathered}$$

- $\mathsf{T} \vdash (C^\bullet)^\circ \rightarrow (B^\bullet)^\circ$ implies that $\mathsf{T} \vdash (C^\bullet)^\circ \rightarrow \forall x(B^\bullet)^\circ$, where B, C are $\mathcal{L}$ formulas and x does not occur freely in C. To show this, suppose

  $$\mathsf{T} \vdash (C^\bullet)^\circ \rightarrow (B^\bullet)^\circ.$$

  By Proposition 2.22, we obtain

  $$\mathsf{T} \vdash (C^\bullet \rightarrow B^\bullet)^\circ.$$

  By assumption and preservation of logical structure by $\bullet$ in theory S:

  $$\begin{gathered} \mathsf{S} \vdash C^\bullet \rightarrow B^\bullet \\ \mathsf{S} \vdash C^\bullet \rightarrow \forall x(B^\bullet) \end{gathered}$$

  Therefore, by assumption and Proposition 2.22, we successively obtain the following in T.

  $$\begin{gathered} (C^\bullet \rightarrow \forall x(B^\bullet))^\circ \\ (C^\bullet)^\circ \rightarrow \forall x(B^\bullet)^\circ \end{gathered}$$

Other cases are treated accordingly. ☐

Finally, we define retractability (simpliciter) as follows.

Definition 2.32. S is *retractable to* T (simpliciter) if and only if, for some mapping $\circ : \mathcal{L}_{\mathsf{S}} \to \mathcal{L}_{\mathsf{T}}$, S is retractable to T by $\circ$.

Remark 2.33. Actually, even though we shall abandon this issue in this study, the theories S of truth to be studied later will not only satisfy 2.27 (1) and (2), where T is a (sub-)theory (of) $\mathbf{I}(\Sigma_1)$, but even more: *That* they satisfy these two conditions could be proved in $\mathbf{I}(\Sigma_1)$ by formalizing the proof that S is retractable to theory T.

To a large extend, the $\mathbf{I}(\Sigma_1)$ proofs of 2.27 (1) and (2) would consist in performing certain (meta) inductions formally in $\mathbf{I}(\Sigma_1)$, by coding involved (meta) concepts by Σ_1 formulas.

(In fact, this remark is the reason why we often indicate the meta status of an induction only in brackets.)

3 Base Theories

Any axiomatic theory of truth comprises two major components. First, it has certain axioms for a (or several) distinguished truth predicate(s). Second, it includes the base (sub-)theory about the syntax of the truth predicate(s)' target language. In the present setting, the syntax will be coded in the natural numbers.

Structurally, a theory qualifies as a base theory of syntax if it gives an account of how to code finite sequences of numbers:

$$(n_0, n_1, \ldots, n_k)$$

The aim of Sections 3.1–3.6 is to develop such a base theory in a rudimentary arithmetical language. The aim is not to present the most efficient coding of sequences, but just one that is efficient enough for the theories of truth to be investigated later. Section 3.2 and Sections 3.4–3.5 are, to a large extent, in line with Hájek and Pudlák [18], Chapter V, Section 3 and Chapter I, Section 2.

The procedure is the following. In Section 3.1, we present a theory $\mathbf{I}(\Sigma_0)$ of arithmetic for Cantor's pairing function. To avail ourselves of a stock of function symbols as expressive means of auxiliary status, as discussed in Section 3.3, we will use Cantor's pairing function in Section 3.2 to derive just the properties of our coding of sequences to define, in the respective base theories, the graphs of functions from the (smashed) Grzegorczyk hierarchy.

In particular, the graph of the exponentiation function will be defined in Section 3.4. The graphs of functions from the higher levels of the hierarchy will be defined in Section 3.10. Finally, in Sections 3.5–3.6, it will be shown how to put the definitions of graphs to work by means of the collection schema.

In Sections 3.7–3.8, we will establish how axioms/derivations, conveniently stated in languages with auxiliary function symbols, may be restated in the respective languages without the auxiliary function symbols by using the definitions of graphs from Sections 3.4 and 3.10. Thus, only after we have convinced ourselves that additional function symbols are merely auxiliary, we make use of them to simplify the presentation. We may then also get rid of these auxiliary function symbols whenever it is convenient.

Our choice of auxiliary function symbols is motivated by the convenience of the bounded μ operator and corresponds to a definition in Schwichtenberg and Wainer [43], Section 2.2.3

3.1 Pairing

According to usual practice, we will base our coding of finite sequences of numbers on a coding of pairs of numbers. So, we first pave the way for a theory of the latter.

Definition 3.1. $\mathcal{L}(0,\mathrm{Q})$ is the *language of Robinson Arithmetic*

$$\{\overline{0}, +\overline{1}, +, \cdot, =, \leq\},$$

where $\overline{0}$ is a constant, $+\overline{1}$ is an unary function symbol, $+, \cdot$ are binary function symbols, and $=, \leq$ are binary relation symbols.

The first index 0 of $\mathcal{L}(0,\mathrm{Q})$, as we will see, marks the basic status of $\mathcal{L}(0,\mathrm{Q})$ as a language of arithmetic without any truth predicates.

We now define Robinson Arithmetic (**Q**) in $\mathcal{L}(0,\mathrm{Q})$.

Definition 3.2. **Q** comprises the following axioms.

(Q1) $x + \overline{1} \neq \overline{0}$

(Q2) $x + \overline{1} = y + \overline{1} \rightarrow x = y$

(Q3) $x \neq \overline{0} \rightarrow \exists y(x = y + \overline{1})$

(Q4) $x + \overline{0} = x$

(Q5) $x + (y + \overline{1}) = (x + y) + \overline{1}$

(Q6) $x \cdot \overline{0} = \overline{0}$

(Q7) $x \cdot (y + \overline{1}) = (x \cdot y) + x$

(Q8) $x \leq y \leftrightarrow \exists z(x + z = y)$

Note that (Q1)–(Q3) axiomatize the symbol $x + \overline{1}$ for the successor function.

To improve on **Q**, we introduce the schema of mathematical induction.

Definition 3.3. Let Γ be some class of formulas. $\mathsf{IS}(\Gamma)$ is the *induction schema*

$$A(\vec{x}, \overline{0}) \wedge \forall y\big(A(\vec{x}, y) \rightarrow A(\vec{x}, y + \overline{1})\big) \rightarrow A(\vec{x}, z), \tag{3.1}$$

where y, z are not among $\vec{x}$ and $A(\vec{x}, y)$ belongs to Γ. Moreover, $A(\vec{x}, y)$ is the *induction formula* and y is the *induction variable.*

We will usually conceive $\mathsf{IS}(\Gamma)$ as a class of formulas: It comprises all formulas that can be obtained by replacing a formula from Γ for $A(\vec{x}, y)$ in (3.1). The same convention will be adopted for other schemas. So, the following definition makes sense.

Definition 3.4. $\mathbf{I}(\Gamma)$ is the theory $\mathbf{Q} \cup \mathsf{IS}(\Gamma)$.

Remark 3.5. Let $\mathsf{IS}^{<}(\Gamma)$ be the *bounded induction schema*

$$A(\vec{x}, \overline{0}) \wedge (\forall y < z)\big(A(\vec{x}, y) \rightarrow A(\vec{x}, y + \overline{1})\big) \rightarrow A(\vec{x}, z),$$

where y, z are not among $\vec{x}$ and Γ contains $A(\vec{x}, y)$. Accordingly, let $\mathbf{I}^{<}(\Gamma)$ be the theory $\mathbf{Q} \cup \mathsf{IS}^{<}(\Gamma)$. Assume our goal is to prove, for any language $\mathcal{L}$ that extends $\mathcal{L}(0, \mathrm{Q})$ and for any $n \geq 0$, that $\mathbf{I}(\Sigma_n^{\mathcal{L}})$ and $\mathbf{I}^{<}(\Sigma_n^{\mathcal{L}})$ are identical theories.

In the wording of Section 2: Our aim, by Corollary 2.29, is to show that $\mathbf{I}(\Sigma_n^{\mathcal{L}})$ is retractable to $\mathbf{I}^{<}(\Sigma_n^{\mathcal{L}})$ by the identity function $\circ$. So, we have to show two things:

- Any formula in $\mathsf{IS}(\Sigma_n^{\mathcal{L}})$ is derivable in $\mathbf{I}^{<}(\Sigma_n^{\mathcal{L}})$.
- Any formula in $\mathsf{IS}^{<}(\Sigma_n^{\mathcal{L}})$ is derivable in $\mathbf{I}(\Sigma_n^{\mathcal{L}})$.

For the second claim, assume in $\mathbf{I}(\Sigma_n^{\mathcal{L}})$ that

$$A(\vec{x}, \overline{0}) \wedge (\forall y < z)\big(A(\vec{x}, y) \rightarrow A(\vec{x}, y + \overline{1})\big), \tag{3.2}$$

where $A(\vec{x}, y)$ is some $\Sigma_n^{\mathcal{L}}$ formula. By logic, there is some $\Sigma_n^{\mathcal{L}}$ formula $B(\vec{x}, y, z)$ such that $\mathbf{I}(\Sigma_n^{\mathcal{L}})$ proves

$$\big(y \leq z \rightarrow A(\vec{x}, y)\big) \leftrightarrow B(\vec{x}, y, z). \tag{3.3}$$

So, we may apply $\mathsf{IS}(\Sigma_n^{\mathcal{L}})$ to $B(\vec{x}, y, z)$, where y is the induction variable, and rely on (3.3) to conclude $A(\vec{x}, z)$. (To establish

$$\big(y \leq z \rightarrow A(\vec{x}, y)\big) \rightarrow \big(y + \overline{1} \leq z \rightarrow A(\vec{x}, y + \overline{1})\big),$$

use $y + \overline{1} \leq z \rightarrow y < z$, which is derivable in $\mathbf{I}(\Sigma_n^{\mathcal{L}})$.) As for the first claim, that any formula in $\mathsf{IS}(\Sigma_n^{\mathcal{L}})$ is provable in $\mathbf{I}^{<}(\Sigma_n^{\mathcal{L}})$, use the logical part of $\mathbf{I}^{<}(\Sigma_n^{\mathcal{L}})$ to derive (3.2) from

$$A(\vec{x}, \overline{0}) \wedge \forall y\big(A(\vec{x}, y) \rightarrow A(\vec{x}, y + \overline{1})\big).$$

Then, apply $\mathsf{IS}^{<}(\Sigma_n^{\mathcal{L}})$.

Let us only add here that, since

$$\mathbf{Q} - \{(\mathrm{Q}3)\} \cup \mathsf{IS}(\Sigma_n^{\mathcal{L}}) \vdash (\mathrm{Q}3),$$

$\mathbf{I}(\Sigma_n^{\mathcal{L}})$ and $\mathbf{Q} - \{(\mathrm{Q}3)\} \cup \mathsf{IS}(\Sigma_n^{\mathcal{L}})$ are identical theories too.

To simplify notation for our basic $\mathcal{L}(0, \mathrm{Q})$, we let

$$\begin{aligned}\Sigma_n &:= \Sigma_n^{\mathcal{L}(0,\mathrm{Q})}\\ \Pi_n &:= \Pi_n^{\mathcal{L}(0,\mathrm{Q})}\end{aligned}$$

Similarly for $\mathrm{E}\Sigma_n$ and $\mathrm{E}\Pi_n$. For $k > 1$, we also write t^k as a shorthand for

$$\underbrace{t \cdot t \cdot \ \cdots \ \cdot t}_{k \text{ times}}$$

where t is a term. We will tacitly use similar notations, such as $\sum_{i=0}^{k} t_i$ to mean $t_0 + t_1 + \cdots + t_k$, if the reader can see from the context how they are to be understood.

We now exploit the idea behind Cantor's pairing bijection to define, by Σ_0 formula $\mathit{Pair}(x, y, z)$, that z is the pair (x, y) of x and y.

Proposition 3.6. *There is some Σ_0 formula $\mathit{Pair}(x, y, z)$ so that $\mathbf{I}(\Sigma_0)$ proves the following.*

(1) $(\exists z)\mathit{Pair}(x, y, z)$

(2) $\mathit{Pair}(x, y, u) \wedge \mathit{Pair}(x, y, v) \to u = v$

(3) $(\exists x, y)\mathit{Pair}(x, y, z)$

(4) $\mathit{Pair}(u_1, v_1, z) \wedge \mathit{Pair}(u_2, v_2, z) \to u_1 = u_2 \wedge v_1 = v_2$

(5) $\mathit{Pair}(x, y, z) \to x \leq z \wedge y \leq z$

(6) $\mathit{Pair}(x, y, z) \wedge x \leq v \wedge y \leq v \to z \leq \overline{4} \cdot v^2$

Proof. Let $\mathit{Pair}(x, y, z)$ be

$$\overline{2} \cdot z = ((x + y + \overline{1}) \cdot (x + y)) + \overline{2} \cdot x \tag{3.4}$$

to define Cantor's pairing bijection

$$(m, n) \mapsto \Big(\sum_{i=0}^{m+n} i\Big) + m.$$

Reason in $\mathbf{I}(\Sigma_0)$ to formally derive (1)–(6).

- For (1), observe that, for all x, y, either $(x + y)$ or $(x + y + \overline{1})$ must be even. (u is even if and only if $(\exists v \leq u)(\overline{2} \cdot v = u)$.) Thus, for all x, y, we have that $((x + y + \overline{1}) \cdot (x + y)) + \overline{2} \cdot x$ must be even.
- For (2), assume $Pair(x, y, u)$ and $Pair(x, y, v)$. Then we obtain $\overline{2} \cdot u = \overline{2} \cdot v$. Therefore, $u = v$.
- For (3), first establish that, for all z, there is u such that $u \cdot (u + \overline{1}) \leq \overline{2} \cdot z < (u + \overline{1}) \cdot (u + \overline{2})$. Let x be such that $\overline{2} \cdot x = \overline{2} \cdot z - u \cdot (u + \overline{1})$. So, $x > u \rightarrow \overline{2} \cdot z \geq (u + \overline{1}) \cdot (u + \overline{2})$ and, hence, $x \leq u$. Finally, we can take y to be $u - x$.
- For (4), first derive that, for all u, v, we have $u \cdot (u + \overline{1}) = v \cdot (v + \overline{1}) \rightarrow u = v$. Then establish, for all u, v, w, that $w + \overline{2} \cdot u = w + \overline{2} \cdot v \rightarrow u = v$.
- For (5), the claim is immediate if $z = \overline{0}$. So, assume $z > \overline{0}$. Suppose $y > z$. Then $(y + \overline{1}) \cdot y > \overline{2} \cdot y > \overline{2} \cdot z$ and, hence, $\neg Pair(x, y, z)$. Accordingly for $x > z$.
- (6) is immediate by (3.4). ☐

In Section 2.2, we have distinguished complexities of formulas from outside a theory. We now add, generally:

Definition 3.7. Let S be some theory and let Γ be some class of formulas. Formula A is a Γ *formula in* S if and only if there is a formula B in Γ such that $\mathsf{S} \vdash A \leftrightarrow B$.

For complexities of formulas, as seen from inside a theory, we may define, in particular:

Definition 3.8. Let $\mathcal{L}$ be a language and let S be a theory. A formula A is $\Sigma_n^{\mathcal{L}}$ ($\Pi_n^{\mathcal{L}}$) *in* S if and only if there is some $\Sigma_n^{\mathcal{L}}$ ($\Pi_n^{\mathcal{L}}$) formula B such that $\mathsf{S} \vdash A \leftrightarrow B$. Moreover, A is $\Delta_n^{\mathcal{L}}$ *in* S if it is both $\Sigma_n^{\mathcal{L}}$ in S and $\Pi_n^{\mathcal{L}}$ in S.

If theory S extends $\mathbf{I}(\Sigma_0)$ and, so, comprises a (sub-)theory of pairing, then we have in S the subsequent closure conditions on complexity classes.

Proposition 3.9. *Let* S *be an extension of* $\mathbf{I}(\Sigma_0)$ *in language* $\mathcal{L}$. *Then, for all* $\mathcal{L}$ *formulas* B, C*, the following hold.*

- *If* B, C *are* $\Sigma_n^{\mathcal{L}}$ $(\Pi_n^{\mathcal{L}}, \Delta_n^{\mathcal{L}})$ *in* S*, then so is* $(B \vee C)$.
- *If* B, C *are* $\Sigma_n^{\mathcal{L}}$ $(\Pi_n^{\mathcal{L}}, \Delta_n^{\mathcal{L}})$ *in* S*, then so is* $(B \wedge C)$.
- *If* B *is* $\Delta_n^{\mathcal{L}}$ *in* S*, then so is* $\neg B$.
- *If* B *is* $\Sigma_{n+1}^{\mathcal{L}}$ *in* S*, then so is* $\exists x B$.
- *If* B *is* $\Pi_{n+1}^{\mathcal{L}}$ *in* S*, then so is* $\forall x B$.

Proof. We only prove the last two items. We proceed by (meta) induction on n. First, we consider the fourth item, where $n = 0$. Assume B is a $\Sigma_1^{\mathcal{L}}$ formula $\exists y D(x, y, \vec{z})$. Then, by virtue of 3.6 (pairing), we may prove $\exists x B(x, \vec{z})$ to be equivalent in S to $\Sigma_1^{\mathcal{L}}$ formula

$$\exists v (\forall x, y \leq v)\big(Pair(x, y, v) \rightarrow D(x, y, \vec{z})\big).$$

For the last item, if B is a $\Pi_1^{\mathcal{L}}$ formula $\forall y E(x, y, \vec{z})$, then S proves $\forall x B(x, \vec{z})$ to be equivalent to $\Pi_1^{\mathcal{L}}$ formula

$$\forall v (\exists x, y \leq v)\big(Pair(x, y, v) \wedge E(x, y, \vec{z})\big).$$

For the induction step, assume the fourth and last item hold for n. Consider a $\Sigma_{n+2}^{\mathcal{L}}$ formula B of the form $\exists y F(x, y, \vec{z})$. By 3.6, S proves $\exists x B(x, \vec{z})$ and

$$\exists v (\forall x, y)\big(Pair(x, y, v) \rightarrow F(x, y, \vec{z})\big)$$

to be equivalent. Moreover, by the first item and the induction hypothesis that the last item holds for n, $(\forall x, y)(Pair(x, y, v) \rightarrow F(x, y, \vec{z}))$ is a $\Pi_{n+1}^{\mathcal{L}}$ formula in S, which concludes the induction step for the fourth item. The induction step for the last item is again dual. □

Let us also add that $\mathbf{I}(\Sigma_0)$ proves the least number principle (schema) for Σ_0 formulas.

Proposition 3.10. $\mathbf{I}(\Sigma_0)$ *proves, for any* Σ_0 *formula* $B(\vec{x}, y)$, *the following.*

$$\exists y B(\vec{x}, y) \to \exists y \big(B(\vec{x}, y) \wedge (\forall z < y) \neg B(\vec{x}, z)\big)$$

Proof. Let $B(\vec{x}, y)$ be a Σ_0 formula. Assume in $\mathbf{I}(\Sigma_0)$ that

$$\forall y \big((\forall z < y) \neg B(\vec{x}, z) \to \neg B(\vec{x}, y)\big). \tag{3.5}$$

We have $(\forall z < \overline{0}) \neg B(\vec{x}, z)$ and, by (3.5), also

$$(\forall z < y) \neg B(\vec{x}, z) \to (\forall z < y + \overline{1}) \neg B(\vec{x}, z).$$

Hence, by $\mathsf{IS}(\Sigma_0)$, we obtain $(\forall z < y) \neg B(\vec{x}, z)$, that is to say, $\forall y \neg B(\vec{x}, y)$. ☐

3.2 Finite Sequences

We now use formula *Pair* to define in $\mathbf{I}(\Sigma_0)$ that x is a finite sequence, written $\mathit{Seq}(x)$, and that x is a member of the finite sequence y, written $x \in y$. So far, we require these only to define the graphs of certain functions, including the graph of the exponentiation function. The sequence coding below is based on ideas by Nelson [39].

Intuitively, $\mathit{Seq}(x)$ expresses that x is a pair (u, v), where v is odd and bigger or equal to the least power 2^n such that $2^n > u$. Thus, for example, x may be the pair $(762, 1189)$, which reads in binary $(1011111010, 10010100101)$, or more suggestively:

$$\begin{array}{lllll} 101 & 11 & 110 & 10 & \\ 100 & 10 & 100 & 10 & 1 \end{array}$$

The first (upper) binary string represents the sequence members according to the separation markers (in form of 1s) in the second (lower) binary string. So, these members are: $101, 11, 110, 10$. In decimal notation, they are: $5, 3, 6, 2$. Consequently, $(762, 1189)$ is the sequence $(5, 3, 6, 2)$. Notice that $(5, 3, 6, 2)$ has several codes. For example, another code for it would be

$$\begin{array}{lllll} 101 & 11 & 110 & 10 & \\ 10000 & 10 & 100 & 10 & 1 \end{array}$$

even though the second (lower) component is unnecessarily large. Thus, we will not be able to derive in $\mathbf{I}(\Sigma_0)$ that *the* code of $(5, 3, 6, 2)$ is $(762, 1189)$, but only that *some* code for it may be bounded by $(762, 1189)$, which will be sufficient for our purposes.

Intuitively, $x \in y$ expresses that y is a pair (u, v) and x can be suitably extracted from u according to the markers in v.

We now proceed formally. First, we define Σ_0 formulas:

$$\begin{aligned} \mathit{Div}(x, y) &:= (\exists u \leq y)(x \cdot u = y) \\ \mathit{Prim}(x) &:= x > \overline{1} \wedge (\forall u, v \leq x)(x = u \cdot v \rightarrow u = \overline{1} \vee v = \overline{1}) \\ \mathit{Pow}(x) &:= x \geq \overline{1} \wedge (\forall u \leq x)(\mathit{Prim}(u) \wedge \mathit{Div}(u, x) \rightarrow u = \overline{2}) \\ \mathit{Lpg}(x, y) &:= \mathit{Pow}(y) \wedge (y \leq \overline{2} \cdot x + \overline{1}) \wedge x < y \end{aligned}$$

Using these, we may define: x is a sequence.

Definition 3.11. The formula $\mathit{Seq}(x)$ is defined as:

$$\begin{aligned} (\exists u, v \leq x)\big(&\mathit{Pair}(u, v, x) \wedge \neg \mathit{Div}(\overline{2}, v) \wedge \\ &(\forall w \leq \overline{2} \cdot u + \overline{1})(\mathit{Lpg}(u, w) \rightarrow w \leq v)\big) \end{aligned}$$

Sequences may be bounded in $\mathbf{I}(\Sigma_0)$ as follows.

Proposition 3.12. $\mathbf{I}(\Sigma_0)$ *proves the following.*

$$\mathit{Pair}(x, y, z) \wedge \mathit{Seq}(z) \rightarrow z \leq \overline{4} \cdot y^2$$

Proof. By the definition of $\mathit{Seq}(z)$, we may prove in $\mathbf{I}(\Sigma_0)$ that $x < y$. So, by Proposition 3.6 (6), we obtain $z \leq \overline{4} \cdot y^2$. ☐

Furthermore, we may define: x belongs to sequence y.

Definition 3.13. The formula $x \in y$ is defined as follows.

$$
\begin{aligned}
&(\exists z_1, z_2 \leq y)\big(Pair(z_1, z_2, y) \land \\
&(\exists w_1, w_2, u_2, v_2 \leq z_2)(\exists u_1, v_1 < w_1)\big(Pow(w_1) \land Pow(w_2) \land \\
&\qquad\qquad w_2 \geq \overline{2} \land x < w_2 \land \\
&\qquad\qquad z_1 = u_2 \cdot w_2 \cdot w_1 + x \cdot w_1 + u_1 \land \\
&\qquad\qquad z_2 = v_2 \cdot \overline{2} \cdot w_2 \cdot w_1 + w_2 \cdot w_1 + w_1 + v_1\big)\big)
\end{aligned}
$$

Notation: We occasionally use the abbreviations

$$
\begin{aligned}
(\exists x \in t)A &:= (\exists x \leq t)(x \in t \land A) \\
(\forall x \in t)A &:= (\forall x \leq t)(x \in t \to A)
\end{aligned}
$$

where it is assumed that the variable x does not occur inside the term t.

We now derive formally in $\mathbf{I}(\Sigma_0)$ that $(0, 1)$ is the empty sequence and that sequences may be prolonged by adding a number.

Proposition 3.14. $\mathbf{I}(\Sigma_0)$ *proves the following.*

(1) $Pair(\overline{0}, \overline{1}, z) \to x \notin z$

(2) $Pair(u, v, x) \land Seq(x) \to$
$(\exists y \leq t(x, z))\big(Seq(y) \land \forall w(w \in y \leftrightarrow w \in x \lor w = z)\big)$

Here $t(x, z)$ *is the term* $x \cdot (\overline{3} \cdot (z + \overline{1}))^2$.

Proof. Both items just rely on earlier observations.

- (1) is actually immediate: By Definition 3.13, we find that $Pair(u, v, z) \land x \in z \to v > \overline{1}$.
- For (2), we only indicate how to calculate the bound. More details are to be found in [18], proof of Lemma 3.7.

Suppose that $Pair(u_0, v_0, x)$ and that $Seq(x)$. If $z \neq \overline{0}$, let $Lpg(z, w_0)$; if $z = \overline{0}$, let $w_0 = \overline{2}$. Consider u_1, v_1, y so that $Pair(u_1, v_1, y)$ and

$$u_1 = u_0 \cdot w_0 + z, \tag{3.6}$$

$$v_1 = v_0 \cdot w_0 + \overline{1}. \tag{3.7}$$

Since $Seq(x)$, we have $v_0 > \overline{0}$ and

$$v_1 \leq v_0 \cdot \overline{2} \cdot (z + \overline{1}) + \overline{1} \leq v_0 \cdot \overline{3} \cdot (z + \overline{1}). \tag{3.8}$$

Moreover, if $u_0 > \overline{0}$, then also

$$u_1 \leq u_0 \cdot \overline{2} \cdot (z + \overline{1}) + z \leq u_0 \cdot \overline{3} \cdot (z + \overline{1}).$$

Thus, by definition (3.4), we have

$$y \leq x \cdot (\overline{3} \cdot (z + \overline{1}))^2. \tag{3.9}$$

If $u_0 = \overline{0}$, then by (3.6) we have $u_1 < \overline{3} \cdot (z + \overline{1})$. Thus, by definition (3.4) again, we obtain (3.9). ☐

3.3 Function Symbols and Valuation

As we will see in Section 3.4, Proposition 3.14 allows us to define graphs of functions from the (smashed) Grzegorczyk hierarchy.

In the present section, we shall briefly explain what brings us to define these graphs in the first place: namely, our intention to use auxiliary function symbols for working more conveniently in base theories, but also stay flexible enough to get rid of auxiliary function symbols whenever it is desirable. At the same time, we intend to sustain a core intuition behind type-free truth.

To begin with, the levels of the Grzegorczyk hierarchy are defined as follows (for $k > 1$).

Definition 3.15. For $k > 1$, *level k of the Grzegorczyk hierarchy* comprises, for every $i \leq k$, the function $\gamma_i : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ which is defined recursively as follows.

$$\gamma_i(m, n) := \begin{cases} n+1 & : i = 0 \\ m & : i = 1 \text{ and } n = 0 \\ 0 & : i = 2 \text{ and } n = 0 \\ 1 & : i > 2 \text{ and } n = 0 \\ \gamma_{i-1}(m, \gamma_i(m, n-1)) & : i > 0 \text{ and } n \neq 0 \end{cases}$$

Moreover, level $k > 1$ of the Grzegorczyk hierarchy includes the constant 0 function, projection functions (on the i-th coordinate), the modified subtraction function, and is closed under composition and the bounded μ operation.

- Composition: If the functions $f_0, \ldots, f_n, f_{n+1}$ belong to the level k of the Grzegorczyk hierarchy, then the function that maps $(\vec{m})$ to
$$f_{n+1}(f_0(\vec{m}), \ldots, f_n(\vec{m}))$$
belongs to level k also.
- Bounded μ operation: If the function f belongs to the level k of the Grzegorczyk hierarchy, then the function that maps $(\vec{m}, a)$ to
$$\begin{cases} \min\{b < a : f(\vec{m}, b) = 0\} & : \{b < a : f(\vec{m}, b) = 0\} \neq \varnothing \\ a & : \{b < a : f(\vec{m}, b) = 0\} = \varnothing \end{cases}$$
belongs to level k also.

We call γ_i the *i-th Grzegorczyk function.*

Every function that belongs to the Grzegorczyk hierarchy is primitive recursive. Conversely, every primitive recursive function belongs to some level of the Grzegorczyk hierarchy.

In particular, γ_0 is the (binary) successor function, γ_1 is the addition function, γ_2 is the multiplication function, γ_3 is the exponentiation function, γ_4 is the tetration function, and so forth.

Remark 3.16. Practically, we will consider, for $i > 2$, the i-th Grzegorczyk function only with fixed $m = 2$:

$$\gamma_i(2, n) = \begin{cases} 1 & : n = 0 \\ \gamma_{i-1}(2, \gamma_i(2, n-1)) & : n > 0 \end{cases}$$

That is, for $i > 2$, we will consider γ_i as an *unary* function and simply write $\gamma_i(n)$ in place of $\gamma_i(2, n)$.

For later purposes, we will have to consider an intermediate level between the 2nd and 3rd level of the Grzegorczyk hierarchy specifically for Nelson's smash function. So, let us also define a *smashed* hierarchy as follows, where $|n|$ is the number of bits in (least) binary representation of n.

Definition 3.17. For $k > 0$, *level k of the smashed Grzegorczyk hierarchy* comprises, for any $i \leq k$, the function $\sigma_i : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by reference to the i-th Grzegorczyk function γ_i as:

$$\sigma_i(m, n) := 2^{\gamma_i(|m|, |n|)}$$

Moreover, level $k > 0$ of the smashed Grzegorczyk hierarchy includes the constant 0 function, projection functions (on the i-th coordinate), the modified subtraction function, and is closed under composition and the bounded μ operation. We call σ_i the *i-th smashed Grzegorczyk function.*

We have the following (in-)equalities (for $m, n > 0$).

$$\begin{aligned} m \cdot n < 2^{|m|+|n|} &= \sigma_1(m, n) \leq 4 \cdot m \cdot n \\ 2^{|m| \cdot |n|} &= \sigma_2(m, n) \\ m^n < 2^{m \cdot n} < 2^{|2|^{|m \cdot n|}} &= \sigma_3(2, m \cdot n) \end{aligned}$$

$$\sigma_3(m,n) = 2^{|m|^{|n|}} \leq 2^{m^n}$$

and so forth. That is, due to the closure under composition and bounded μ operation, we may simply identify the 1st level of the smashed Grzegorczyk hierarchy with the 2nd level of the (usual) Grzegorczyk hierarchy and, for any $k > 2$, the k-th level of the smashed Grzegorczyk hierarchy with the k-th level of the (usual) Grzegorczyk hierarchy. The only difference, as needed, is that the smashed hierarchy additionally provides some intermediate level 2 for the smash function.

Moreover, since $2^{|m|\cdot|n|} < (2^{|m|+|n|})^{|2^{|m|+|n|}|}$ and (for $m > 2$)

$$\begin{aligned}
t_0(m) &:= m^{|m|} < m^m \\
t_1(m) &:= t_0(m)^{|t_0(m)|} < (m^m)^{|m^m|} < (m^m)^{|(2^{|m|})^m|} < m^{m^3} \\
t_2(m) &:= t_1(m)^{|t_1(m)|} < (m^{m^3})^{|m^{m^3}|} < m^{m^7}
\end{aligned}$$

and, more generally (for $n > 0$),

$$t_n(m) := t_{n-1}(m)^{|t_{n-1}(m)|} < m^{m^{2^{n+1}}}$$

we see that the 3rd level of the (standard) Grzegorczyk hierarchy provides a function to bound the growth rate of functions on the 2nd level of the smashed Grzegorczyk hierarchy.

Later we shall deal, for example, with a language $\mathcal{L}(0,3)$ for elementary arithmetic, corresponding to level 3 of the (smashed) Grzegorczyk hierarchy. In general, if we want to highlight that a language corresponds, via function symbols, to level k of the smashed Grzegorczyk hierarchy, we indicate this by a second argument k in $\mathcal{L}(m,k)$. The first argument m, as we will see later, marks the presence of certain truth predicates.

For complexity classes, in general, we extend superscripts accordingly to:

$$\Sigma_n^{(m,k)}, \Pi_n^{(m,k)}, \Delta_n^{(m,k)}, \mathrm{E}\Sigma_n^{(m,k)}, \mathrm{E}\Pi_n^{(m,k)}$$

Now, let us focus on the issue of auxiliary function symbols. For the sake of illustration, consider language $\mathcal{L}(1,k)$, where the number 1 marks that it comprises only one truth predicate $T(z)$, and consider a truth axiom such as

$$Sent([x=y]) \to \big(T([x=y]) \leftrightarrow val(x)=val(y)\big) \tag{3.10}$$

where the notation is such that, if $\ulcorner s \urcorner, \ulcorner t \urcorner$ encode closed terms s,t, then $[\ulcorner s \urcorner = \ulcorner t \urcorner]$ encodes sentence $(s=t)$. Furthermore, the consequent of (3.10) means that $(s=t)$ is true if and only if s,t have the same value.

Suppose that the target language of $T(z)$ should be $\mathcal{L}(1,k)$. That is, $Sent(z)$ in (3.10) should mean that z is (a code of) an $\mathcal{L}(1,k)$ sentence. Can (3.10) itself be stated in $\mathcal{L}(1,k)$?

To see the problem, let us focus on involved symbols more closely. Assume that the language $\mathcal{L}(1,k)$ corresponds to level k of the smashed hierarchy (provided $k>1$). In particular, assume that $\mathcal{L}(1,k)$ has a function symbol $num(x)$ to denote (a code of) the x-th numeral and the function symbol $sub(x,y)$ to denote (a code of) the term that is obtained by uniformly substituting the $\mathcal{L}(1,k)$ term (coded by) y for the only free variable of the $\mathcal{L}(1,k)$ term (coded by) x.

Again, let $Sent([x=y])$ in (3.10) mean that $[x=y]$ is (a code of) an $\mathcal{L}(1,k)$ sentence, which is to say, x,y are (codes of) closed $\mathcal{L}(1,k)$ terms without free variables.

Then, necessarily, $val(z)$ in (3.10) for the valuation of $\mathcal{L}(1,k)$ terms cannot be a function symbol in $\mathcal{L}(1,k)$. To see why, note that $val(z)$ must be such that we have (may prove) for arbitrary $\mathcal{L}(1,k)$ terms $t(x)$ that

$$t(x) = val(sub(\ulcorner t \urcorner, num(x))). \tag{3.11}$$

Now, assume towards a contradiction that $val(z)$ was a function symbol in $\mathcal{L}(1,k)$. Then we may construct another $\mathcal{L}(1,k)$ term $d(x)$ (for diagonalization) such that

$$d(x) = val(sub(x, num(x))) + \overline{1}. \tag{3.12}$$

However, for $d(\ulcorner d \urcorner)$ in particular, (3.11) and (3.12) would lead to $d(\ulcorner d \urcorner) = d(\ulcorner d \urcorner) + \overline{1}$, a contradiction.

We may draw from this the consequence that axiom (3.10), where the target language of $T(z)$ is $\mathcal{L}(1, k)$, has to be stated in some higher-level language $\mathcal{L}(1, k+1)$ which provides a function symbol $val(z)$ for the valuation of $\mathcal{L}(1, k)$ terms. However, as far as we can see, this consequence would violate a crucial intuition behind type-free – conceived as natural – truth: Namely, that we should be in a position to formulate the statements to which the truth predicate(s) apply and the axioms which involve the truth predicate(s) in one and the same encompassing language.

(Since, in particular, the axioms of *finitist* truth will be restricted to certain logical operations and do not cover an overall first-order language, as we will see in later chapters, we may be more cautious and talk about *partial* type-free truth in the given setting.)

So, we shall not draw the consequence that (3.10) has to be stated in a language different from $\mathcal{L}(1, k)$. Rather, as we have announced, we will show in the remainder of the present chapter that defining graphs of certain functions enables us to formulate axioms like (3.10) in $\mathcal{L}(1, k)$ (or even in $\mathcal{L}(1, \mathrm{Q})$) itself, and that – in a precise sense – auxiliary function symbols serve as mere shorthands for these definitions of graphs.

The concern of auxiliary function symbols is postponed until Sections 3.7–3.8. In the following Sections 3.4–3.6, we first have to set the stage for the application of a common proof-theoretic method. Namely, we want to interchange function symbols and definitions of graphs under the requirement that the latter satisfy totality and uniqueness.

Notice that, more specifically, by showing that truth axioms may be stated in $\mathcal{L}(m, \mathrm{Q})$, we at the same time allow for a neat comparison between theories of truth and the theories from the arithmetical hierarchy. (See Chapter 11.)

Moreover, as the following demonstrations of how to express truth axioms without the auxiliary function symbols bear on the relation of *retractability* between theories, several preparatory results for later investigations will already be obtained in Sections 3.7–3.8.

3.4 Exponentiation

To show that certain function symbols might be eliminated, we have to show how to define the graphs of the denoted functions in a given theory. That is, focussing on the levels $k > 2$ of the (regular) Grzegorczyk hierarchy, we may now use $Pair(x, y, z)$ to define, as announced, the graph of the function (denoted via) γ_k. We start off with (base 2) exponentiation. Once we have defined γ_k for $k > 2$, the definition of σ_k is straightforward. (Note that $|n| = \min\{i > 0 : n < 2^i\}$.)

Definition 3.18. $Exp(x, y)$ (respectively $G_3(x, y)$) is defined to be formula

$$(\exists u \leq \overline{30} \cdot y^{16})(\exists v \in u)\big(G_3Seq(u) \wedge Pair(y, x, v)\big),$$

where $G_3Seq(u)$ is

$$\begin{aligned}
&Seq(u) \wedge (\exists v_1 \in u)\big(Pair(\overline{1}, \overline{0}, v_1)\big) \wedge \\
&\qquad (\forall v_1 \in u)\big(Pair(\overline{1}, \overline{0}, v_1) \vee \\
&(\exists v_0 \in u)(\exists x_0, x_1, y_0, y_1 \leq u)\big(x_1 > \overline{0} \wedge \\
&\qquad\qquad Pair(y_1, x_1, v_1) \wedge Pair(y_0, x_0, v_0) \wedge \\
&\qquad\qquad \big(x_1 = \overline{2} \cdot x_0 \wedge y_1 = y_0 \cdot y_0\big) \vee \\
&\qquad\qquad \big(x_1 = \overline{2} \cdot x_0 + \overline{1} \wedge y_1 = y_0 \cdot y_0 \cdot \overline{2}\big)\big)\big)
\end{aligned}$$

So, for example, $Exp(\overline{6}, \overline{64})$ means that, for $6 = 110$ in binary, we may successively compute: $2^0 = 1$ and $2^1 = 2$ and $2^{11} = 2^2 \cdot 2$ and $2^{110} = (2^2 \cdot 2)^2 = 64$. The bound $\overline{30} \cdot y^{16}$ in formula $Exp(x, y)$ is specific to sequences that stepwise trace such computations, as shown by Lemma 3.19. As the proof of that lemma is parallel to the proof of the subsequent Proposition 3.20, which relies on Proposition 3.14, we omit the proof of Lemma 3.19 and only refer to [18], Chapter V, Section 3 (c).

Lemma 3.19. $\mathbf{I}(\Sigma_0)$ *proves the following.*

$$(\forall v, x, y \leq u)\big(G_3Seq(u) \land Pair(y, x, v) \land v \in u \to \\ (\exists w \leq \overline{30} \cdot y^{16})\big(G_3Seq(w) \land v \in w\big)\big)$$

We shall now derive the properties of $Exp(x, y)$ formally in $\mathbf{I}(\Sigma_0)$. Notice: We have uniqueness, but not totality.

Proposition 3.20. $\mathbf{I}(\Sigma_0)$ *proves the following.*

- $Exp(x, y) \land Exp(x, z) \to y = z$
- $Exp(\overline{0}, \overline{1})$
- $Exp(x + \overline{1}, y) \leftrightarrow \exists z(Exp(x, z) \land y = z \cdot \overline{2})$

Proof. We verify the three items individually.

- To verify the second item of Proposition 3.20, we may derive in $\mathbf{I}(\Sigma_0)$ that 30 codes the sequence $((1, 0))$ with the only member $(1, 0)$ and that 30 satisfies G_3Seq.
- For the third item, we have to consider two directions. We prove only the right-to-left direction here. The proof of the other direction is similar. (For better readability, suggestive notation is also used in the next paragraphs.)

  For the right-to-left direction, we want to derive

  $$\forall x(\forall z \leq y)\big((Exp(x, z) \land y = z \cdot \overline{2}) \to Exp(x + \overline{1}, y)\big),$$

  which we do by deriving a contradiction from its negation. Hence, assume

  $$\exists x(\exists z \leq y)\big(Exp(x, z) \land y = z \cdot \overline{2} \land \neg Exp(x + \overline{1}, y)\big)$$

and observe that, by Proposition 3.10, we may assume for that x that

$$(\forall u < x)(\forall z \leq y)\big((Exp(u,z) \wedge y = z \cdot \overline{2}) \rightarrow Exp(u + \overline{1}, y)\big). \tag{3.13}$$

Moreover, we may reason against the background that, by $\mathsf{IS}(\Sigma_0)$, we have

$$(\forall v, x, y \leq u)\big((G_3Seq(u) \wedge Pair(y,x,v) \wedge v \in u) \rightarrow x \leq y\big).$$

We distinguish two cases: Either x is even or x is odd.

Suppose that x is even: $x = 2 \cdot a$. In this case, by the assumption $Exp(x,z) \wedge y = z \cdot 2$, there exists some sequence s such that $G_3Seq(s)$ and s contains the pair $(b^2, 2 \cdot a)$, where $b^2 = z$. So, by definition of G_3Seq and Lemma 3.19, there exists another sequence $r \leq 30 \cdot b^{16}$ so that $G_3Seq(r)$ and r contains (b, a). If $a = 0$, then $b = 1$ and

$$\begin{aligned}(b^2 \cdot 2, 2 \cdot a + 1) &= (2,1)\\ &\leq 16\end{aligned}$$

If $a > 0$, then $b > 1$. Consequently, since $a \leq b$, we have $2 \cdot a + 1 \leq b^2 \cdot 2$ and, by Proposition 3.6 (6), again that

$$(b^2 \cdot 2, 2 \cdot a + 1) \leq 16 \cdot b^4$$

Thus, by Proposition 3.14, there exists a prolongation t of r by the pair $(b^2 \cdot 2, 2 \cdot a + 1)$ such that

$$\begin{aligned}t &\leq 9 \cdot (16 \cdot b^4 + 1)^2 \cdot r\\ &\leq 9 \cdot (16 \cdot b^4 + 1)^2 \cdot 30 \cdot b^{16}\end{aligned}$$

and, by inequality $9 \cdot (16 \cdot b^4 + 1)^2 \leq (2 \cdot b)^{16}$ for $b > 0$, also

$$t \leq (2 \cdot b)^{16} \cdot 30 \cdot b^{16}$$

$$= 30 \cdot (b^2 \cdot 2)^{16}$$

which is $t \leq 30 \cdot y^{16}$. Therefore, $Exp(x + \overline{1}, y)$, contradicting our assumption.

Assume that x is odd: $x = 2 \cdot a + 1$. Then, by assumption $Exp(x, z) \wedge y = z \cdot 2$, there exists a sequence s such that $G_3Seq(s)$ and s contains $(b^2 \cdot 2, 2 \cdot a + 1)$, where $b^2 \cdot 2 = z$. Thus, by the definition of G_3Seq and Lemma 3.19, there is some sequence $r \leq 30 \cdot b^{16}$ so that $G_3Seq(r)$ and r contains (b, a), that is, $Exp(a, b)$. By (3.13),

$$(Exp(a, b) \wedge b \cdot 2 = b \cdot 2) \to Exp(a + 1, b \cdot 2).$$

That is to say, there is some sequence $q \leq 30 \cdot (b \cdot 2)^{16}$ such that $G_3Seq(q)$ and q contains $(b \cdot 2, a + 1)$. Like in the even case, we may now apply Proposition 3.6 (6) to obtain:

$$(b^2 \cdot 4, 2 \cdot a + 2) \leq 64 \cdot b^4$$

Thus, by Proposition 3.14, there exists a prolongation t of q by the pair $(b^2 \cdot 4, 2 \cdot a + 2)$ such that

$$\begin{aligned} t &\leq 9 \cdot (64 \cdot b^4 + 1)^2 \cdot r \\ &\leq 9 \cdot (64 \cdot b^4 + 1)^2 \cdot 30 \cdot (b \cdot 2)^{16} \end{aligned}$$

and, by inequality $9 \cdot (64 \cdot b^4 + 1)^2 \leq (2 \cdot b)^{16}$ for $b > 0$, also

$$\begin{aligned} t &\leq (2 \cdot b)^{16} \cdot 30 \cdot (b \cdot 2)^{16} \\ &\leq 30 \cdot (b^2 \cdot 4)^{16} \end{aligned}$$

which is $t \leq 30 \cdot y^{16}$. Therefore, $Exp(x + \overline{1}, y)$, contradicting our assumption.

- Finally, to show the first item, we use $\mathsf{IS}(\Sigma_0)$ with induction variable x and induction formula

  $$(\forall y, z < u)\big(Exp(x, y) \wedge Exp(x, z) \to y = z\big),$$

  relying on the second and third item.

This concludes the proof. □

In addition to exponentiation, we can define in $\mathbf{I}(\Sigma_0)$ also the predecessor function.

Proposition 3.21. *There is some Σ_0 formula $Pred(x, y)$ so that $\mathbf{I}(\Sigma_0)$ proves the following.*

- $(\exists y)Pred(x, y)$
- $Pred(x, y) \wedge Pred(x, z) \to y = z$
- $Pred(\overline{0}, \overline{0})$
- $Pred(x + \overline{1}, x)$

Proof. Define $Pred(x, y)$ as

$$(x = \overline{0} \wedge y = \overline{0}) \vee (x > \overline{0} \wedge x = y + \overline{1}),$$

using the symbol for the successor function. □

Once the predecessor function has been defined, we may also define the modified subtraction function.

Proposition 3.22. *There is some Σ_0 formula $Subtr(x, y, z)$ such that $\mathbf{I}(\Sigma_0)$ proves the following.*

- $(\exists z)Subtr(x, y, z)$
- $Subtr(x, y, u) \wedge Subtr(x, y, v) \to u = v$
- $Subtr(x, \overline{0}, x)$
- $Subtr(x, y + \overline{1}, z) \leftrightarrow \exists w(Subtr(x, y, w) \wedge Pred(w, z))$

Proof. Define $Subtr(x, y, z)$ as

$$(y \geq x \wedge z = \overline{0}) \vee (y < x \wedge x = z + y),$$

using the symbol for the addition function. □

3.5 Collection

We now turn to a brief discussion of the collection schema. In particular, we prove (Proposition 3.25) that, given the collection schema, the induction formulas may involve certain definitions of graphs, such as $Exp(x, y)$ that we gave in the foregoing section – provided that we can derive that the respective definition applies to just one unique and existing object.

We will also see (Proposition 3.26) that collection is a finitist principle.

Definition 3.23. Let Γ be some class of formulas. $\mathsf{BS}(\Gamma)$ is the *collection or bounding schema*

$$(\forall x \leq y)\exists z A \to \exists v(\forall x \leq y)(\exists z \leq v)A,$$

where A belongs to Γ.

Proposition 3.24. *Let* S *be an extension of* $\mathbf{I}(\Sigma_0) \cup \mathsf{BS}(\Sigma_n^{\mathcal{L}})$ *in a language* $\mathcal{L}$*. Then, for any* $i \leq n$ *and* $\mathcal{L}$ *formula* B*, the following hold.*

- *If* B *is* $\Sigma_i^{\mathcal{L}}$ $(\Pi_i^{\mathcal{L}})$ *in* S*, then so is* $(\exists x \leq y)B$*.*
- *If* B *is* $\Sigma_i^{\mathcal{L}}$ $(\Pi_i^{\mathcal{L}})$ *in* S*, then so is* $(\forall x \leq y)B$*.*

Proof. By (meta) induction on i. For $i = 0$, the claims are immediate by Definition 2.12. For the induction step, suppose that $\mathsf{S} \vdash B \leftrightarrow \exists z D$, where D is $\Pi_i^{\mathcal{L}}$. Observe that S proves

$$(\forall x \leq y)\exists z D \leftrightarrow \exists v(\forall x \leq y)(\exists z \leq v)D,$$

where $\exists v(\forall x \leq y)(\exists z \leq v)D$ is $\Sigma_{i+1}^{\mathcal{L}}$ in S by induction hypothesis. In other cases, we also use Proposition 3.9. ☐

We now turn to the main result of the present section. As we will see later, to every Σ_0 formula A in which auxiliary function symbols occur, there corresponds a formula A° that involves the definitions of the graphs of the functions denoted in A and has complexity Δ_1 in a certain base theory. By the next proposition, the collection schema allows us to perform, in that base theory, inductions with A°.

Proposition 3.25. *Let $\mathcal{L}$ be some extension of $\mathcal{L}(0, \mathrm{Q})$ and S an extension of $\mathbf{I}(\Sigma_0^{\mathcal{L}}) \cup \mathsf{BS}(\Sigma_1^{\mathcal{L}})$. Then S proves, for all $\Sigma_1^{\mathcal{L}}$ formulas B, C, the following.*

$$\begin{gathered}\forall y\big(B(\vec{x}, y) \leftrightarrow \neg C(\vec{x}, y)\big) \rightarrow \\ \big(B(\vec{x}, \overline{0}) \wedge \forall y\big(B(\vec{x}, y) \rightarrow B(\vec{x}, y + \overline{1})\big) \rightarrow B(\vec{x}, y)\big)\end{gathered}$$

Proof. Let B be $\exists z D(\vec{x}, y, z)$, let C be $\exists z E(\vec{x}, y, z)$, and let D, E be two $\Sigma_0^{\mathcal{L}}$ formulas. Assume in S that

$$\forall y\big(\exists z D(\vec{x}, y, z) \leftrightarrow \neg \exists z E(\vec{x}, y, z)\big), \tag{3.14}$$

that $B(\vec{x}, \overline{0})$, and that $\forall y(B(\vec{x}, y) \rightarrow B(\vec{x}, y + \overline{1}))$, where the latter yield

$$B(\vec{x}, \overline{0}) \wedge (\forall y < u)\big(B(\vec{x}, y) \rightarrow B(\vec{x}, y + \overline{1})\big). \tag{3.15}$$

Observe that (3.14) implies

$$(\forall y \leq u)\exists z\big(D(\vec{x}, y, z) \vee E(\vec{x}, y, z)\big).$$

Hence, $\mathsf{BS}(\Sigma_1^{\mathcal{L}})$ can be applied to obtain

$$\exists v(\forall y \leq u)(\exists z \leq v)\big(D(\vec{x}, y, z) \vee E(\vec{x}, y, z)\big)$$

and, by (3.14), also

$$\exists v(\forall y \leq u)\big(\exists z D(\vec{x}, y, z) \leftrightarrow (\exists z \leq v) D(\vec{x}, y, z)\big). \tag{3.16}$$

By Remark 3.5, $\mathsf{IS}^{<}(\Sigma_0^{\mathcal{L}})$ is available in S and, hence, $B(\vec{x}, y)$ is obtained by (3.15) and (3.16). ☐

We will also require the following well-known results, of which the first shows that collection is a finitist principle.

Proposition 3.26. *Let $\mathcal{L}$ be extension of $\mathcal{L}(0, \mathrm{Q})$. Then $\mathbf{I}(\Sigma^{\mathcal{L}}_{n+1})$ proves all formulas in $\mathsf{BS}(\Sigma^{\mathcal{L}}_{n+1})$.*

Proof. See, for example, [18], proof of Lemma 2.11. □

Proposition 3.27. *Let $\mathcal{L}$ be extension of $\mathcal{L}(0, \mathrm{Q})$. Then $\mathbf{I}(\Sigma^{\mathcal{L}}_{0}) \cup \mathsf{BS}(\Sigma^{\mathcal{L}}_{n+1})$ proves all formulas in $\mathsf{IS}(\Sigma^{\mathcal{L}}_{n})$.*

Proof. See, for example, [18], proof of Lemma 2.15. □

3.6 Base Theories

We now introduce, for all the theories of truth to be investigated below, the base theory of syntax and of term valuation for their respective language. As extensions of $\mathcal{L}(0, \mathrm{Q})$, we first define the following languages.

Definition 3.28. For $m > 0$, $\mathcal{L}(m, \mathrm{Q})$ is the language

$$\mathcal{L}(0, \mathrm{Q}) \cup \{T_i : 0 < i \leq m\},$$

where, for $0 < i \leq m$, T_i is an unary predicate symbol that does not belong to $\mathcal{L}(0, \mathrm{Q})$ and is called the *i-th truth predicate.*

The base theory for language $\mathcal{L}(m, \mathrm{Q})$ is defined as follows.

Definition 3.29. $\mathbf{B}(m, 3)$ is the $\mathcal{L}(m, \mathrm{Q})$ theory

$$\mathbf{I}(\Sigma_0) \cup \mathsf{BS}(\Sigma_1) \cup \{\exists y Exp(x, y)\}.$$

Observe that $\exists y Exp(x, y)$ states the totality of (base 2) exponentiation.

We now address the addition of auxiliary function symbols.

3.7 Function Symbols: Step 1

Since the following is uniform in $m \geq 0$, let us momentarily fix $m \geq 0$ and consider some fixed language $\mathcal{L}(m, \mathrm{Q})$ and $\mathcal{L}(m, \mathrm{Q})$ theory $\mathbf{B}(m, 3)$.

$\mathbf{B}(m, 3)$ qualifies as a base for the axiomatization of type-free truth since, as we will see, we may augment it with further axioms, stated in the language $\mathcal{L}(m, \mathrm{Q})$, which are about the truth of $\mathcal{L}(m, \mathrm{Q})$ statements. However, the formulation of truth axioms is very cumbersome if we cannot make use of auxiliary function symbols as expressive means.

Thus, the purpose of this and the next Section 3.8 is to show that certain extensions S of $\mathbf{B}(m, 3)$ by additional axioms for the auxiliary function symbols satisfy that, for every $\mathcal{L}(m, \mathrm{Q})$ formula B and every formula A in the extended language of S:

- $\mathsf{S} \vdash A$ if and only if $\mathbf{B}(m, 3) \vdash A^\circ$
- $\mathsf{S} \vdash B$ if and only if $\mathbf{B}(m, 3) \vdash B$

where the formula A° is some translation of A in the language $\mathcal{L}(m, \mathrm{Q})$. By Proposition 2.27 and Corollary 2.30, the two items may be achieved by *retracting* S to $\mathbf{B}(m, 3)$. Therefore, the main theorems of Sections 3.7–3.8 may be applied repeatedly in later retractability proofs as well.

The plan is the following. The addition of auxiliary function symbols is divided in two steps: step 1 in this section and step 2 in the next section. The division is only strategic: It allows us to give proofs to some detail in step 1, where only few auxiliary function symbols are introduced, and merely indicate how these proofs ought to be extended so as to treat the function symbols that are introduced in step 2.

Similarly, we focus on base theory $\mathbf{B}(m, 3)$, as it has special status, and indicate how $\mathbf{B}(m, 3)$ may be modified to cover other levels $k > 1$ of the (smashed) Grzegorczyk hierarchy only later.

Our choice of auxiliary function symbols is inspired by a formulation of *elementary arithmetic* by Schwichtenberg and Wainer [43], Section 2.2.3. The choice is not substantial, but often proves convenient due to the use of the bounded μ operator.

Let $\gamma_3(x)$ be some unary function symbol and let $-(x,y)$ be some binary function symbol, both not included in $\mathcal{L}(m,\mathrm{Q})$. We write $x - y$ for $-(x,y)$ and $\overline{2}^x$ for $\gamma_3(x)$ and define, in general:

Definition 3.30. For any language $\mathcal{L}$,

$$\mathcal{L}^* := \mathcal{L} \cup \{-, \overline{2}^{(\ )}\}.$$

In addition, we define:

Definition 3.31. For any $\mathcal{L}$ theory S, S^* is the $\mathcal{L}^*$ theory that extends S by every (missing) axiom in $\mathsf{IS}(\Sigma_0^{\mathcal{L}^*})$ and the following axioms.

(A1) $\overline{0} - \overline{1} = \overline{0}$

(A2) $(x + \overline{1}) - \overline{1} = x$

(A3) $x - \overline{0} = x$

(A4) $x - (y + \overline{1}) = (x - y) - \overline{1}$

(A5) $\overline{2}^{\overline{0}} = \overline{1}$

(A6) $\overline{2}^{x+\overline{1}} = \overline{2}^x \cdot \overline{2}$

Note that (A1) and (A2) axiomatize the term $x - \overline{1}$ for the predecessor function, which is used in (A4).

As said above, our goal is to interpret $\mathcal{L}(m,\mathrm{Q})^*$ formulas in $\mathcal{L}(m,\mathrm{Q})$. Since no new function symbols are added in languages $\mathcal{L}(m,\mathrm{Q})$ with $m > 0$, we only need to define, for every $\mathcal{L}(0,\mathrm{Q})^*$ term $t(\vec{z})$, a specific $\mathcal{L}(0,\mathrm{Q})$ formula $[\![t]\!](x)$, with a distinguished variable x not among $\vec{z}$, to express that x is the value of term t.

Definition 3.32. The $\mathcal{L}(0,\mathrm{Q})$ formula $[\![t]\!](x)$, where $t(\vec{z})$ is some $\mathcal{L}(0,\mathrm{Q})^*$ term and the variable x is not among the variables $\vec{z}$, is defined recursively as follows.

- For variables v and constant $\overline{0}$:

$$[\![v]\!](x) := (v = x) \tag{3.17}$$

$$[\![\overline{0}]\!](x) := (\overline{0} = x) \tag{3.18}$$

- For function symbols in $\mathcal{L}(0,\mathrm{Q})$:

$$[\![t_0 + t_1]\!](x) := (\exists u, v)\big([\![t_0]\!](u) \wedge [\![t_1]\!](v) \wedge u + v = x\big)$$

$$[\![t_0 \cdot t_1]\!](x) := (\exists u, v)\big([\![t_0]\!](u) \wedge [\![t_1]\!](v) \wedge u \cdot v = x\big)$$

  For function symbols in $(\mathcal{L}(0,\mathrm{Q})^* - \mathcal{L}(0,\mathrm{Q}))$:

$$[\![t_0 - t_1]\!](x) := (\exists u, v)\big([\![t_0]\!](u) \wedge [\![t_1]\!](v) \wedge \mathit{Subtr}(u, v, x)\big)$$

$$[\![\overline{2}^{t_0}]\!](x) := (\exists u)\big([\![t_0]\!](u) \wedge \mathit{Exp}(u, x)\big) \tag{3.19}$$

  Here $\mathit{Subtr}(u,v,x)$ and $\mathit{Exp}(u,x)$ are the $\mathcal{L}(0,\mathrm{Q})$ formulas from Proposition 3.22 and Proposition 3.20, respectively.

We verify formally in $\mathbf{B}(0,3)$ that $[\![t]\!](x)$ denotes, for every $\mathcal{L}(0,\mathrm{Q})^*$ term $t(\vec{z})$, a unique object (value) x.

Lemma 3.33. *In* $\mathbf{B}(0,3)$ *we have, for every* $\mathcal{L}(0,\mathrm{Q})^*$ *term* t*, the following.*

- $\exists x([\![t]\!](x))$
- $[\![t]\!](x) \wedge [\![t]\!](y) \rightarrow x = y$

Proof. By a (meta) induction on the rank of t. We focus on two cases to show the strategy.

- t is a variable or $\overline{0}$. In this case, we observe that $\mathbf{B}(0,3)$ proves $\exists x(t = x)$ and
$$t = x \wedge t = y \rightarrow x = y$$
by logic. Thus, we rely on (3.17) and (3.18).
- t is $\overline{2}^{t_0(\vec{z})}$. Then we first have, by the induction hypothesis, that $\mathbf{B}(0,3)$ proves $\exists u([\![t_0(\vec{z})]\!](u))$ and
$$[\![t_0(\vec{z})]\!](u) \wedge [\![t_0(\vec{z})]\!](v) \rightarrow u = v.$$
By Proposition 3.21 and totality of exponentiation, we get $(\exists x, u)\big([\![t_0(\vec{z})]\!](u) \wedge Exp(u,x)\big)$ and
$$\begin{aligned}&\exists u\big([\![t_0(\vec{z})]\!](u) \wedge Exp(u,x)\big) \wedge \\ &\exists v\big([\![t_0(\vec{z})]\!](v) \wedge Exp(v,y)\big) \rightarrow x = y\end{aligned}$$
in $\mathbf{B}(0,3)$. Thus, we may rely on (3.19). □

Next, function $\circledast : \mathcal{L}(m,\mathrm{Q})^* \rightarrow \mathcal{L}(m,\mathrm{Q})$ for step 1 of adding auxiliary function symbols is defined (for $0 < i \leq m$) as follows.

Definition 3.34. $\circledast : \mathcal{L}(m,\mathrm{Q})^* \rightarrow \mathcal{L}(m,\mathrm{Q})$ is defined recursively.

$$\begin{aligned}
(s = t)^\circledast &:= (\exists u, v)\big([\![s]\!](u) \wedge [\![t]\!](v) \wedge u = v\big) && (3.20)\\
(s \neq t)^\circledast &:= (\exists u, v)\big([\![s]\!](u) \wedge [\![t]\!](v) \wedge u \neq v\big) && (3.21)\\
(s \leq t)^\circledast &:= (\exists u, v)\big([\![s]\!](u) \wedge [\![t]\!](v) \wedge u \leq v\big)\\
(s > t)^\circledast &:= (\exists u, v)\big([\![s]\!](u) \wedge [\![t]\!](v) \wedge u > v\big)\\
T_i(t)^\circledast &:= (\exists u)\big([\![t]\!](u) \wedge T_i(u)\big)\\
({\sim}T_i(t))^\circledast &:= (\exists u)\big([\![t]\!](u) \wedge {\sim}T_i(u)\big)\\
(B \vee C)^\circledast &:= (B)^\circledast \vee (C)^\circledast\\
(B \wedge C)^\circledast &:= (B)^\circledast \wedge (C)^\circledast\\
(\exists x B)^\circledast &:= \exists x (B)^\circledast\\
(\forall x B)^\circledast &:= \forall x (B)^\circledast
\end{aligned}$$

We successively verify conditions from Section 2.4.

Lemma 3.35. *In* $\mathbf{B}(m,3)$ *we obtain, for each* $\mathcal{L}(m,\mathrm{Q})^*$ *literal* B*, the following.*

$$({\sim}B)^{\circledast} \leftrightarrow \neg(B)^{\circledast}$$

Proof. We consider only one case to show the general strategy. If B is $(s = t)$, where s, t are $\mathcal{L}(0,\mathrm{Q})^*$ terms, we apply (3.20) and Lemma 3.33 to obtain in $\mathbf{B}(m,3)$ that

$$(s=t)^{\circledast} \leftrightarrow (\forall u, v)\big([\![s]\!](u) \wedge [\![t]\!](v) \to u = v\big). \tag{3.22}$$

So, by (3.21), we obtain that $\mathbf{B}(m,3) \vdash ({\sim}(s=t))^{\circledast} \leftrightarrow \neg(s=t)^{\circledast}$, as required. □

Corollary 3.36. $\circledast$ *commutes with logical operators in* $\mathbf{B}(m,3)$*.*

Proof. By Lemma 3.35, the definition of $\circledast$, and logic. □

Corollary 3.37. $\mathbf{B}(m,3)$ *proves, for all* $\mathcal{L}(m,\mathrm{Q})^*$ *formulas* B, C *and all* $\mathcal{L}(0,\mathrm{Q})^*$ *terms* t*, the following.*

- $(\neg B)^{\circledast} \leftrightarrow \neg(B)^{\circledast}$
- $(B \to C)^{\circledast} \leftrightarrow (B^{\circledast} \to C^{\circledast})$
- $(B \leftrightarrow C)^{\circledast} \leftrightarrow (B^{\circledast} \leftrightarrow C^{\circledast})$
- $((\forall x \le t)B)^{\circledast} \leftrightarrow \forall u([\![t]\!](u) \to (\forall x \le u)B^{\circledast})$
- $((\exists x \le t)B)^{\circledast} \leftrightarrow \forall u([\![t]\!](u) \to (\exists x \le u)B^{\circledast})$

Proof. By applying Proposition 2.22. □

Lemma 3.38. *For every* $\mathcal{L}(m,\mathrm{Q})$ *formula* A *and every* $\mathcal{L}(m,\mathrm{Q})^*$ *formula* B*, the following hold.*

(1) $\mathbf{B}(m,3) \vdash A \leftrightarrow A^{\circledast}$

(2) $\mathbf{B}(m,3)^* \vdash B \leftrightarrow B^{\circledast}$

Proof. We prove (1) and (2) individually.

- For (1), the logical part of $\mathbf{B}(m,3)$ guarantees that, for any $\mathcal{L}(m,\mathrm{Q})$ atom A, we have $\mathbf{B}(m,3) \vdash A \leftrightarrow A^{\circledast}$. Hence, the claim follows by Proposition 2.23 and Corollary 3.36.
- For (2), we rely on Proposition 3.21 and apply axioms (A1) and (A2) with $\mathsf{IS}(\Sigma_0^{\mathcal{L}(m,\mathrm{Q})^*})$ to derive in $\mathbf{B}(m,3)^*$ that

  $$Pred(x,y) \leftrightarrow x - \overline{1} = y.$$

  Similarly for $Subtr(x,y,z)$ and $Exp(x,y)$. So, by (meta) induction on the rank of the $\mathcal{L}(m,\mathrm{Q})^*$ term t, we have that $\mathbf{B}(m,3)^*$ proves $x = t \leftrightarrow [\![t]\!](x)$. For the rest, we continue similarly as in (1). □

We may now proceed to preservation of logical structure.

Lemma 3.39. *For any $\mathcal{L}(m,\mathrm{Q})^*$ formulas B, C and any $\mathcal{L}(0,\mathrm{Q})^*$ term t, the following hold.*

(1) $\mathbf{B}(m,3) \vdash B(t/x)^{\circledast} \rightarrow \exists x(B)^{\circledast}$

(2) $\mathbf{B}(m,3) \vdash \forall x(B)^{\circledast} \rightarrow B(t/x)^{\circledast}$

(3) $\mathbf{B}(m,3) \vdash B^{\circledast} \rightarrow C^{\circledast}$ *implies* $\mathbf{B}(m,3) \vdash \exists x(B)^{\circledast} \rightarrow C^{\circledast}$

(4) $\mathbf{B}(m,3) \vdash C^{\circledast} \rightarrow B^{\circledast}$ *implies* $\mathbf{B}(m,3) \vdash C^{\circledast} \rightarrow \forall x(B)^{\circledast}$

(5) $\mathbf{B}(m,3) \vdash (x = x)^{\circledast}$

(6) $\mathbf{B}(m,3) \vdash (x = y)^{\circledast} \rightarrow (B^{\circledast} \rightarrow B(y/x)^{\circledast})$

Notice: In (3) *and* (4)*, x is the eigenvariable and must not occur freely in C.*

Proof. As for (1), we proceed by (meta) induction on the rank of B. We consider two examples to show the strategy.

- B is $(s_0 \leq s_1)$, where s_0, s_1 are $\mathcal{L}(0,\mathrm{Q})^*$ terms. Assume in $\mathbf{B}(m,3)$ that $B(t/x)^{\circledast}$, which is
$$(\exists u, v)\big(\llbracket s_0(t/x)\rrbracket(u) \wedge \llbracket s_1(t/x)\rrbracket(v) \wedge u \leq v\big). \tag{3.23}$$
By (meta) induction on the rank of $\mathcal{L}(0,\mathrm{Q})^*$ term s, relying on Lemma 3.33, we obtain in $\mathbf{B}(m,3)$ that, for all $\mathcal{L}(0,\mathrm{Q})^*$ terms s, t,
$$\llbracket s(t/x)\rrbracket(u) \to (\exists x, u)\big(\llbracket t\rrbracket(x) \wedge \llbracket s\rrbracket(u)\big).$$
Hence, by Lemma 3.33 and logic, we may derive in $\mathbf{B}(m,3)$ from (3.23) that
$$\exists x\big(\llbracket t\rrbracket(x) \wedge (\exists u, v)\big(\llbracket s_0\rrbracket(u) \wedge \llbracket s_1\rrbracket(v) \wedge u \leq v\big)\big),$$
which is $\exists x\big(\llbracket t\rrbracket(x) \wedge (s_0 \leq s_1)^{\circledast}\big)$, and so $\exists x\big(\llbracket t\rrbracket(x) \wedge B^{\circledast}\big)$.
- B is $(C \wedge D)$. Assume in $\mathbf{B}(m,3)$ that $B(t/x)^{\circledast}$, which is
$$C(t/x)^{\circledast} \wedge D(t/x)^{\circledast}$$
by definition. By induction hypothesis, we have
$$\exists x\big(\llbracket t\rrbracket(x) \wedge C^{\circledast}\big) \wedge \exists x\big(\llbracket t\rrbracket(x) \wedge D^{\circledast}\big)$$
Therefore, by Lemma 3.33, we obtain $\exists x\big(\llbracket t\rrbracket(x) \wedge C^{\circledast} \wedge D^{\circledast}\big)$, which is $\exists x\big(\llbracket t\rrbracket(x) \wedge B^{\circledast}\big)$ by definition.

We omit (2). For (3) and (4), we revert to Remark 2.25.

For (5), only logic is required. As for (6), the proof is by a (meta) induction on the rank of B. Let us consider two cases to show the strategy.

- B is $(s = t)$, where s, t are $\mathcal{L}(0,\mathrm{Q})^*$ terms. Now, assume in $\mathbf{B}(m,3)$ that $(x = y)^{\circledast}$. Then, by Lemma 3.38 (1), we have that $(x = y)$. Our goal is to derive in $\mathbf{B}(m,3)$ that
$$(s = t)^{\circledast} \to (s = t)(y/x)^{\circledast}. \tag{3.24}$$

However, (3.24) is obtained in $\mathbf{B}(m,3)$ by logic and

$$x = y \to \big([\![t]\!](v) \to [\![t(y/x)]\!](v)\big), \tag{3.25}$$

where (3.25) is provable in $\mathbf{B}(m,3)$ for any $\mathcal{L}(0,\mathrm{Q})^*$ term t, as may be shown by (meta) induction on the rank of t.

- B is $(C \vee D)$. As before, we proceed from the assumption $(x = y)$ and want to derive from it in $\mathbf{B}(m,3)$ that

  $$(C \vee D)^{\circledast} \to (C \vee D)(y/x)^{\circledast}.$$

  By induction hypothesis and logic, we have

  $$x = y \to (C^{\circledast} \to C(y/x)^{\circledast}) \tag{3.26}$$

  and

  $$x = y \to (D^{\circledast} \to D(y/x)^{\circledast}). \tag{3.27}$$

  Therefore, from $(x = y)$ and the assumption that $(C \vee D)^{\circledast}$, which is $C^{\circledast} \vee D^{\circledast}$, we have in $\mathbf{B}(m,3)$ by (3.26) and (3.27) that $C(y/x)^{\circledast} \vee D(y/x)^{\circledast}$, which is $(C \vee D)(y/x)^{\circledast}$. ☐

Corollary 3.40. *$\circledast$ preserves logical structure in $\mathbf{B}(m,3)$.*

Proof. By Corollary 3.36 and Lemma 3.39. ☐

Lemma 3.41. *For every $\Sigma_0^{\mathcal{L}(0,\mathrm{Q})^*}$ formula A, formula $A^{\circledast}$ is $\Delta_1^{\mathcal{L}(0,\mathrm{Q})}$ in $\mathbf{B}(m,3)$.*

Proof. By a (meta) induction on the rank of the $\Sigma_0^{\mathcal{L}(0,\mathrm{Q})^*}$ formula A, relying on Lemma 3.33 and reasoning like in example (3.22) above. For the induction step, we also apply Proposition 3.9 and Proposition 3.24. ☐

We are now ready to prove our first retractability result.

Notation: If T contains axioms A, that is, pairs (Γ, A) where Γ is empty, as well as (proper) inferences (Γ, A), where Γ is not empty, we use T° to denote the collection $\{(\Gamma^\circ, A^\circ) : (\Gamma, A) \in \mathsf{T}\}$.

Theorem 3.42. *Let* T *comprise non-logical* $\mathcal{L}(m, \mathrm{Q})^*$ *axioms and non-logical* $\mathcal{L}(m, \mathrm{Q})^*$ *inferences. Then theory* $\mathbf{B}(m, 3)^* \cup \mathsf{T}$ *is retractable to* $\mathbf{B}(m, 3) \cup \mathsf{T}^\circledast$ *by* $\circledast$*.*

Proof. With regard to Corollary 3.40 and Lemma 3.38 (1), it is enough to verify the subsequent claims.

- For all non-logical A in $\mathbf{B}(m, 3)^* \cup \mathsf{T}$, we have

  $$\mathbf{B}(m, 3) \cup \mathsf{T}^\circledast \vdash A^\circledast.$$

  We distinguish possible cases. If A is a member of $\mathbf{B}(m, 3)$, then use the Lemma 3.38 (1). If A is (A1)–(A6), then use Proposition 3.21, Proposition 3.22, and Proposition 3.20. If A is an instance of $\mathsf{IS}(\Sigma_0^{\mathcal{L}(0,\mathrm{Q})^*})$, use Proposition 3.25 and Lemma 3.41. Finally, if A is a member of T, then $A^\circledast$ belongs to $\mathsf{T}^\circledast$ and $\mathbf{B}(m, 3) \cup \mathsf{T}^\circledast \vdash A^\circledast$ is immediate.
- For all non-logical A in $\mathbf{B}(m, 3) \cup \mathsf{T}^\circledast$, we have

  $$\mathbf{B}(m, 3)^* \cup \mathsf{T} \vdash A.$$

  We distinguish possible cases. If A belongs to $\mathbf{B}(m, 3)$, then it belongs to $\mathbf{B}(m, 3)^*$ as well. If A belongs to $\mathsf{T}^\circledast$, then it is $B^\circledast$ for some B from T. So, we can apply Lemma 3.38 (2).
- For all non-logical (Γ, A) in $\mathbf{B}(m, 3)^* \cup \mathsf{T}$, we have that

  $$\mathbf{B}(m, 3) \cup \mathsf{T}^\circledast \vdash \bigwedge \Gamma^\circledast$$

  implies

  $$\mathbf{B}(m, 3) \cup \mathsf{T}^\circledast \vdash A^\circledast.$$

  This is immediate because $(\Gamma^\circledast, A^\circledast)$ belongs to $\mathsf{T}^\circledast$.

- For all non-logical (Γ, A) in $\mathbf{B}(m,3) \cup \mathsf{T}^{\circledast}$, we have that

  $$\mathbf{B}(m,3)^{*} \cup \mathsf{T} \vdash \bigwedge \Gamma$$

  implies

  $$\mathbf{B}(m,3)^{*} \cup \mathsf{T} \vdash A.$$

  This follows by Lemma 3.38 (2) and logic.

So, $\mathbf{B}(m,3)^{*} \cup \mathsf{T}$ is retractable to $\mathbf{B}(m,3) \cup \mathsf{T}^{\circledast}$ and the proof is concluded. ☐

Summing up, we have finally arrived at our conclusion, in the form of the corollary below, for step 1 of adding new auxiliary function symbols

Corollary 3.43. *Let* T *comprise non-logical* $\mathcal{L}(m,\mathrm{Q})^{*}$ *axioms and non-logical* $\mathcal{L}(m,\mathrm{Q})^{*}$ *inferences. Then, for every* $\mathcal{L}(m,\mathrm{Q})^{*}$ *formula* A *and every* $\mathcal{L}(m,\mathrm{Q})$ *formula* B*, the following hold.*

(1) $\mathbf{B}(m,3)^{*} \cup \mathsf{T} \vdash A$ *if and only if* $\mathbf{B}(m,3) \cup \mathsf{T}^{\circledast} \vdash A^{\circledast}$.

(2) $\mathbf{B}(m,3)^{*} \cup \mathsf{T} \vdash B$ *if and only if* $\mathbf{B}(m,3) \cup \mathsf{T}^{\circledast} \vdash B$.

Proof. By Theorem 3.42, Proposition 2.27, Corollary 2.30, with Lemma 3.38 (2). ☐

3.8 Function Symbols: Step 2

Our next aim is to prove a counterpart of Corollary 3.43 for step 2 of adding auxiliary function symbols to $\mathcal{L}(m,\mathrm{Q})$. Using a definition of *elementary functions* by Schwichtenberg and Wainer [43], Section 2.2.3, we further extend $\mathcal{L}(m,\mathrm{Q})^{*}$ to include new symbols and $\mathbf{B}(m,3)^{*}$ to include respective axioms for all the elementary functions in the sense of that definition.

Definition 3.44. For any language $\mathcal{L}$ and $\mathcal{L}$ theory S, $\mathcal{L}^+$ is an extension of $\mathcal{L}^*$ and S^+ an extension of S^*, where the extensions are given simultaneously as follows.

(A7) For some unary function symbol f in $\mathcal{L}^+$, S^+ comprises the axiom:

$$f(x) = \overline{0}$$

(A8) For every k and every $i \leq k$, there is some $(k+1)$-ary function symbol f in $\mathcal{L}^+$ such that S^+ comprises:

$$f(x_0, \ldots, x_k) = x_i$$

(A9) For all k-ary function symbols $f_0, \ldots, f_n$ in $\mathcal{L}^+$ and any $(n+1)$-ary function symbol f_{n+1} in $\mathcal{L}^+$, there is a k-ary function symbol f in $\mathcal{L}^+$ such that S^+ comprises:

$$f(x_1, \ldots, x_k) = f_{n+1}(f_0(x_1, \ldots, x_k), \ldots, f_n(x_1, \ldots, x_k))$$

(A10) For any $(k+1)$-ary function symbol f_0 in $\mathcal{L}^+$, there is some $(k+1)$-ary function symbol f in $\mathcal{L}^+$ such that S^+ comprises:

$$f(x_1, \ldots, x_k, z) = y \leftrightarrow C(x_1, \ldots, x_k, y, z)$$

where $C(x_1, \ldots, x_k, y, z)$ is the formula

$$\begin{gathered} y \leq z \land \\ (\forall v < y)\big(f_0(x_1, \ldots, x_k, v) \neq \overline{0}\big) \land \\ \big(y = z \lor f_0(x_1, \ldots, x_k, y) = \overline{0}\big) \end{gathered}$$

(A11) S^+ comprises every (missing) instance of $\mathsf{IS}(\Sigma_0^{\mathcal{L}^+})$.

In general, we may define:

Definition 3.45. For any language $\mathcal{L}$,

$$\mathcal{L}^+ := \mathcal{L}^* \cup \{f : f \text{ is introduced by 3.44 (A7)–(A10)}\}$$

The formula $C(x_1, \ldots, x_k, y, z)$ in (A10) expresses that, either there is $v < z$ so that $f_0(x_1, \ldots, x_k, v) = 0$, in which case y is the least such v, or else $y = z$. Thus, the formula describes how the *bounded μ operator* works. Indeed, we have the following.

Proposition 3.46. *Let formula $C(x_1, \ldots, x_k, y, z)$ be defined as in* 3.44 (A10)*. Then $\mathbf{B}(m,3)^+$ proves the following.*

- $\exists y C(x_1, \ldots, x_k, y, z)$
- $C(x_1, \ldots, x_k, u, z) \wedge C(x_1, \ldots, x_k, v, z) \to u = v$

Proof. For the first item, apply $\mathsf{IS}(\Sigma_0^{\mathcal{L}(0,\mathrm{Q})^+})$ to induction formula

$$\exists y C(x_1, \ldots, x_k, y, z), \tag{3.28}$$

where z is the induction variable. Since $y \leq z$ is a conjunct of $C(x_1, \ldots, x_k, y, z)$, formula (3.28) is bounded. For the induction step, in particular, assume $C(x_1, \ldots, x_k, y, z)$ and derive from it that

$$C(x_1, \ldots, x_k, y + \overline{1}, z + \overline{1}) \vee C(x_1, \ldots, x_k, y, z + \overline{1}).$$

For the second item, derive in $\mathbf{B}(m,3)^+$ that

$$C(x_1, \ldots, x_k, u, z) \wedge C(x_1, \ldots, x_k, v, z)$$

is not consistent with $u \neq v$, that is, $u < v \vee u > v$. □

We now indicate how the proof of Corollary 3.43 has to be adapted to arrive at a similar result for $\mathbf{B}(m,3)^+$.

First, we have to provide an extended definition of $[\![t]\!](x)$ to account for function symbols in $(\mathcal{L}(0,\mathrm{Q})^+ - \mathcal{L}(0,\mathrm{Q})^*)$. Thus, for function symbols f defined in 3.44 (A7)–(A10), respectively, the definition of $[\![t]\!](x)$ is as follows.

- For (A7), where t is $f(t_0)$:

$$(\exists u)\big(\llbracket t_0 \rrbracket(u) \wedge x = \overline{0}\big)$$

- For (A8), where t is $f(t_0, \ldots, t_k)$:

$$(\exists u_0, \ldots, u_k)\big(\bigwedge\nolimits_{j=0}^{k} \llbracket t_j \rrbracket(u_j) \wedge x = u_i\big)$$

- For (A9), where t is $f_{n+1}(f_0(t_1, \ldots, t_k), \ldots, f_n(t_1, \ldots, t_k))$:

$$\begin{aligned}(\exists u_0, \ldots, u_{n+1})\big(\bigwedge\nolimits_{i=0}^{n} \llbracket f_i(t_1, \ldots, t_k) \rrbracket(u_i) \wedge \\ \llbracket f_{n+1}(u_0, \ldots, u_n) \rrbracket(u_{n+1}) \wedge x = u_{n+1}\big)\end{aligned}$$

- For (A10), where t is $f(t_1, \ldots, t_k, t_{k+1})$:

$$\begin{gathered}(\exists u_1, \ldots, u_{k+1})\big(\bigwedge\nolimits_{i=1}^{k+1} \llbracket t_i \rrbracket(u_i) \wedge x \leq u_{k+1} \wedge \\ (\forall v < x)(\exists u_{k+2})\big(\llbracket f_0(u_1, \ldots, u_k, v) \rrbracket(u_{k+2}) \wedge u_{k+2} \neq \overline{0}\big) \wedge \\ \big(x = u_{k+1} \vee (\exists u_{k+2})\llbracket f_0(u_1, \ldots, u_k, x) \rrbracket(u_{k+2}) \wedge u_{k+2} = \overline{0}\big)\big)\big)\end{gathered}$$

By reasoning like in the proof of Lemma 3.33, we obtain:

Lemma 3.47. $\mathbf{B}(m,3)^*$ *proves, for every* $\mathcal{L}(0,\mathrm{Q})^+$ *term* t*, the following.*

- $\exists x(\llbracket t \rrbracket(x))$
- $\llbracket t \rrbracket(x) \wedge \llbracket t \rrbracket(y) \rightarrow x = y$

Now, function $\oplus : \mathcal{L}(m,\mathrm{Q})^+ \rightarrow \mathcal{L}(m,\mathrm{Q})^*$ for step 2 of adding auxiliary function symbols may be defined like above for step 1.

That $\oplus$ preserves the logical structure in $\mathbf{B}(m,3)^*$ is shown, using Lemma 3.47, as it was shown above that $\circledast$ preserves the logical structure in $\mathbf{B}(m,3)$.

Moreover, we obtain a counterpart of Lemma 3.41 as follows.

Lemma 3.48. *For any* $\Sigma_0^{\mathcal{L}(0,\mathrm{Q})^+}$ *formula* A*, formula* $A^\oplus$ *is* $\Sigma_0^{\mathcal{L}(0,\mathrm{Q})^*}$ *in* $\mathbf{B}(m,3)^*$.

Proof. Using (base 2) exponentiation $\gamma_3(x)$, define $\mathcal{L}(0,\mathrm{Q})^*$ terms $\varphi_n(x)$ recursively as follows.

$$\begin{aligned} \varphi_0(x) &:= x \\ \varphi_{n+1}(x) &:= \gamma_3(\varphi_n(x)) \end{aligned} \tag{3.29}$$

By (meta) induction on the rank, show that, for every $\mathcal{L}(0,\mathrm{Q})^+$ term $t(z_0,\ldots,z_k)$, there is n such that $\mathbf{B}(m,3)^*$ proves

$$[\![t(z_0,\ldots,z_k)]\!](x) \to x \leq \varphi_n\big(\sum_{i\leq k} z_i\big). \tag{3.30}$$

Thus, for all $\mathcal{L}(0,\mathrm{Q})^+$ terms s,t with variables among $z_0,\ldots,z_k$, there is n such that $\mathbf{B}(m,3)^*$ proves the equivalence ($\leftrightarrow$) of the formula $(s=t)^\oplus$ and the formula

$$(\exists u \leq \varphi_n\big(\sum_{i\leq k} z_i\big))(\exists v \leq \varphi_n\big(\sum_{i\leq k} z_i\big))\big([\![s]\!](u) \wedge [\![t]\!](v) \wedge u = v\big).$$

Similarly for other $\mathcal{L}(0,\mathrm{Q})^+$ literals. Thus, by (meta) induction on the rank, relying on the definition of the function $\oplus$, we may establish that $A^\oplus$ is $\Sigma_0^{\mathcal{L}(0,\mathrm{Q})^*}$ in $\mathbf{B}(m,3)^*$. ☐

By reasoning basically like for step 1 of adding new auxiliary function symbols, we finally obtain the following.

Theorem 3.49. *Let* T *comprise non-logical* $\mathcal{L}(m,\mathrm{Q})^+$ *axioms and non-logical* $\mathcal{L}(m,\mathrm{Q})^+$ *inferences. Then theory* $\mathbf{B}(m,3)^+ \cup \mathsf{T}$ *is retractable to* $\mathbf{B}(m,3)^* \cup \mathsf{T}^\oplus$ *by* $\oplus$.

The associated corollary is the following.

Corollary 3.50. *Let* T *comprise non-logical* $\mathcal{L}(m,\mathrm{Q})^+$ *axioms and non-logical* $\mathcal{L}(m,\mathrm{Q})^+$ *inferences. Then, for each* $\mathcal{L}(m,\mathrm{Q})^+$ *formula* A *and each* $\mathcal{L}(m,\mathrm{Q})^*$ *formula* B*, the following hold.*

(1) $\mathbf{B}(m,3)^{+} \cup \mathsf{T} \vdash A$ *if and only if* $\mathbf{B}(m,3)^{*} \cup \mathsf{T}^{\oplus} \vdash A^{\oplus}$.

(2) $\mathbf{B}(m,3)^{+} \cup \mathsf{T} \vdash B$ *if and only if* $\mathbf{B}(m,3)^{*} \cup \mathsf{T}^{\oplus} \vdash B$.

3.9 Conclusions

Since the elementary functions correspond to the level 3 of the Grzegorczyk hierarchy, we define $\mathcal{L}(m,\mathrm{Q})^{+}$ from the last section as $\mathcal{L}(m,3)$.

Definition 3.51. $\mathcal{L}(m,3)$ is the language $\mathcal{L}(m,\mathrm{Q})^{+}$.

Accordingly, theory $\mathbf{B}(m,3)$ got its name due to its correspondence to the level 3 of the Grzegorczyk hierarchy. To distinguish it from $\mathbf{B}(m,3)^{+}$ by simpler notation, let us rename the latter as $\mathbf{A}(m,3)$ to highlight the presence of the [**A**]uxiliary function symbols.

Definition 3.52. $\mathbf{A}(m,3)$ is the theory $\mathbf{B}(m,3)^{+}$.

Definition 3.53. $\mathcal{A} : \mathcal{L}(m,3) \to \mathcal{L}(m,\mathrm{Q})$ (for [$\mathcal{A}$]uxiliary function symbols) is defined to be the composition of $\oplus$ and $\circledast$ such that, for every $\mathcal{L}(m,3)$ formula B, we have:

$$B^{\mathcal{A}} := (B^{\oplus})^{\circledast}$$

Corollary 3.54. *Let* T *be some class of non-logical* $\mathcal{L}(m,3)$ *axioms and non-logical* $\mathcal{L}(m,3)$ *inferences. Then, for any* $\mathcal{L}(m,\mathrm{Q})$ *formula* A *and* $\mathcal{L}(m,3)$ *formula* B*, the following hold.*

(1) $\mathbf{A}(m,3) \cup \mathsf{T} \vdash B$ *if and only if* $\mathbf{B}(m,3) \cup \mathsf{T}^{\mathcal{A}} \vdash B^{\mathcal{A}}$.

(2) $\mathbf{A}(m,3) \cup \mathsf{T} \vdash A$ *if and only if* $\mathbf{B}(m,3) \cup \mathsf{T}^{\mathcal{A}} \vdash A$.

Proof. By Corollary 3.43 and Corollary 3.50. ☐

This concludes our discussion of the addition of the auxiliary function symbols to $\mathcal{L}(m,\mathrm{Q})$. We state a useful consequence for future reference. It relies on the following proposition.

Proposition 3.55. $\mathbf{I}(\Sigma_1) \vdash \exists y Exp(x,y)$

Proof. By applying $\mathsf{IS}(\Sigma_1)$ with induction variable x. □

Corollary 3.56. *Let* T *be some class of non-logical* $\mathcal{L}(m,3)$ *axioms and non-logical* $\mathcal{L}(m,3)$ *inferences. Suppose that* $\mathbf{A}(m,3) \cup \mathsf{T}$ *is retractable to* $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)})$ *by* $\circ$*. Then* $\mathbf{A}(m,3) \cup \mathsf{T}$ *also is retractable to* $\mathbf{I}(\Sigma_1)$ *by* $(\mathcal{A} \cdot \circ)$*.*

Proof. By Proposition 3.26 and Proposition 3.55, $\mathbf{I}(\Sigma_1)$ proves all theorems of $\mathbf{B}(0,3)$. This allows us to use Proposition 2.31 to see that $(\mathcal{A} \cdot \circ)$ preserves logical structure in $\mathbf{I}(\Sigma_1)$. For the rest, we basically only have to note that, by Corollary 3.54 (1),

$$\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)}) \vdash B \text{ implies } \mathbf{I}(\Sigma_1) \vdash B^{\mathcal{A}},$$

where B may be any $\mathcal{L}(0,3)$ formula. □

3.10 Generalizations

We conclude the present chapter by indicating how the foregoing results may be generalized. These generalizations are given only with regard to the issue of auxiliary function symbols as it will reappear in Chapters 6 and 7.

The arguments which led us to Corollary 3.54 were not only uniform in truth index m: They may even be adapted to cover arbitrary levels $k > 2$ of the (smashed) Grzegorczyk hierarchy.

(We may safely neglect the case $k = 2$ here because it is very similar to the case $k = 3$. That close similarity is due to our definition of $\mathbf{A}(m,2)$ in 3.60 below: The distinguished axioms of $\mathbf{A}(m,2)$ for the length function and the smash function are both stated in terms of Exp (exponentiation) from 3.18. Hence, Exp may again be used for retracting $\mathbf{A}(m,2)$ to $\mathbf{B}(m,2)$.)

Let us first extend our definition of languages.

Definition 3.57. Let $\mathcal{L}(m,1)$ be defined as

$$\{\overline{0}, +\overline{1}, +, -, \cdot, =, \leq\} \cup \{T_1, \ldots, T_m\}$$

closed under 3.44 (A7)–(A10) and let $\mathcal{L}(m,2)$ be defined as

$$\mathcal{L}(m,1) \cup \{|\ |, \sigma_2\}$$

closed under 3.44 (A9)–(A10). Extending the Definition 3.51 to $k > 3$, let $\mathcal{L}(m,k)$ be defined as

$$\mathcal{L}(m,k-1) \cup \{\gamma_k\}$$

closed under 3.44 (A9)–(A10).

Definition 3.58. For $k > 2$, $\mathsf{G}(k)$ comprises, for every $2 < i \leq k$, the following axioms.

(A12) $\gamma_i(\overline{0}) = \overline{1}$

(A13) $\gamma_i(x + \overline{1}) = \gamma_{i-1}(\gamma_i(x))$

To obtain the counterpart of Proposition 3.20 for $k > 3$, we define a formula $G_k Seq(u)$ as

$$\begin{aligned} &Seq(u) \wedge (\exists w_1 \in u)\big(Pair(\overline{1}, \overline{0}, w_1)\big) \wedge \\ &\quad (\forall w_1 \in u)\big(Pair(\overline{1}, \overline{0}, w_1) \vee \\ &(\exists w_0 \in u)(\exists x_0, x_1, y_0, y_1 \leq u)\big(x_1 > \overline{0} \wedge \\ &\qquad Pair(y_1, x_1, w_1) \wedge Pair(y_0, x_0, w_0) \wedge \\ &\qquad x_1 = x_0 + \overline{1} \wedge y_1 = \gamma_{k-1}(y_0)\big)\big) \end{aligned}$$

where γ_{k-1} is the $(k-1)$-th Grzegorczyk function symbol. As in the proof of 3.20, we may now define $\Sigma_0^{(0,k-1)}$ formula $G_k(x,y)$ for which the following proposition holds.

Proposition 3.59. *For every $k > 3$, $\mathsf{G}(k-1) \cup \mathbf{I}(\Sigma_0^{(0,k-1)})$ proves the following.*

- $G_k(x, y) \land G_k(x, z) \to y = z$
- $G_k(\overline{0}, \overline{1})$
- $G_k(x + \overline{1}, y) \leftrightarrow \exists z(G_k(x, z) \land y = \gamma_{k-1}(z))$

Let us consider level 4 for the sake of illustration. This time, $G_4(\overline{4}, \overline{65536})$, for example, means that a stepwise computation proceeds like this: ${}^{0}2 = 1$ and ${}^{1}2 = 2^1 = 2$ and ${}^{2}2 = 2^2 = 4$ and ${}^{3}2 = 2^4 = 16$ and ${}^{4}2 = 2^{16} = 65536$. More specifically, G_4 may be defined as follows.

$$(\exists u \leq \overline{30} \cdot \gamma_3(\overline{16} \cdot y))(\exists v \in u)\big(G_4Seq(u) \land Pair(y, x, v)\big)$$

The verification of the third item of Proposition 3.59, for example, is again by means of (a counterpart of Lemma 3.19 and) Proposition 3.14. To show the right-to-left direction, for example, we assume that some sequence $s \leq 30 \cdot \gamma_3(z \cdot 16)$ satisfies $G_4Seq(s)$ and contains (z, x). To prolong it by adding $(\gamma_3(z), x + 1)$ to it, where $\gamma_3(z) = y$, we may apply Proposition 3.6 (6) to obtain

$$(\gamma_3(z), x + 1) \leq 4 \cdot \gamma_3(2 \cdot z)$$

and apply Proposition 3.14 to bound the resulting sequence t as follows.

$$\begin{aligned} t &\leq 9 \cdot (4 \cdot \gamma_3(2 \cdot z) + 1)^2 \cdot s \\ &\leq 9 \cdot (4 \cdot \gamma_3(2 \cdot z) + 1)^2 \cdot 30 \cdot \gamma_3(z \cdot 16) \end{aligned}$$

Hence, by $9 \cdot (4 \cdot \gamma_3(2 \cdot z) + 1)^2 \leq \gamma_3(4 \cdot z + 10)$, we have

$$t \leq \gamma_3(4 \cdot z + 10) \cdot 30 \cdot \gamma_3(z \cdot 16)$$

and by $4 \cdot z + 10 + z \cdot 16 \leq 16 \cdot \gamma_3(z)$ for $z > 0$, we have

$$t \leq \gamma_3(\gamma_3(z) \cdot 16) \cdot 30$$

which means $t \leq \gamma_3(y \cdot 16) \cdot 30$. We omit further details and now turn to extensions of base theories.

For $k > 3$, $\mathbf{C}(m,k)$ will be the [**C**]onnecting theory below.

$$\mathsf{G}(k-1) \cup \mathbf{I}(\Sigma_0^{(0,k-1)}) \cup \mathsf{BS}(\Sigma_1^{(0,k-1)}) \cup \{\exists y G_k(x,y)\}$$

Definition 3.60. Let $\mathbf{A}(m,2)$ be defined as $\mathbf{A}(m,3)$ in 3.52, but with the pair (A5), (A6) of axioms for exponentiation replaced by the subsequent axioms for the length and the smash function.

(A14) $|x| = y \leftrightarrow y > \overline{0} \wedge (\exists z < \overline{2} \cdot x + \overline{2})(Exp(y,z) \wedge z > x)$

(A15) $\sigma_2(x,y) = z \leftrightarrow Exp(|x| \cdot |y|, z)$

For $k > 3$, $\mathbf{A}(m,k)$ is the theory

$$\mathbf{A}(m,k-1) \cup \mathsf{G}(k)$$

closed under 3.44 (A9) and (A10).

To obtain a counterpart of Lemma 3.48 for $k > 3$, definition (3.29) is replaced by

$$\varphi_{n+1}(x) := \gamma_k(\varphi_n(x))$$

and it is proved that, for any $\mathcal{L}(0,k)$ term $t(z_0,\ldots,z_k)$, there is some n such that $\mathsf{G}(k) \cup \mathbf{I}(\Sigma_0^{(0,k)})$ proves (3.30).

Now, for some ([$\mathcal{C}$]onnecting) retractability mapping

$$\mathcal{C}_k : \mathcal{L}(m,k) \to \mathcal{L}(m,k-1)$$

the first generalization (to $k > 3$) that we get is:

Corollary 3.61. *For any $k > 3$, let T be some class of non-logical $\mathcal{L}(m,k)$ axioms and non-logical $\mathcal{L}(m,k)$ inferences. Then, for any $\mathcal{L}(m,k-1)$ formula A and $\mathcal{L}(m,k)$ formula B, the following hold.*

(1) $\mathbf{A}(m,k) \cup \mathsf{T} \vdash B$ *if and only if* $\mathbf{C}(m,k) \cup \mathcal{C}_k(\mathsf{T}) \vdash \mathcal{C}_k(B)$.

(2) $\mathbf{A}(m,k) \cup \mathsf{T} \vdash A$ *if and only if* $\mathbf{C}(m,k) \cup \mathcal{C}_k(\mathsf{T}) \vdash A$.

Definition 3.62. Let $\mathcal{A}$ (which is actually $\mathcal{A}_3$) be defined as in 3.53. For $k > 3$, $\mathcal{A}_k$ is defined recursively as $\mathcal{A}_{k-1} \cdot \mathcal{C}_k$.

Definition 3.63. For $k > 3$, $\mathbf{B}(m, k)$ is defined as

$$\mathbf{B}(m, k-1) \cup \{\mathcal{A}_k(\exists y G_k(x, y))\}.$$

Note: For every $k > 3$, $\mathbf{B}(m, k)$ is an $\mathcal{L}(m, \mathrm{Q})$ theory.

Note: Without going into the details, we may define $\mathbf{B}(m, 2)$ just as $\mathbf{B}(m, 3)$ in 3.29, but with axiom $\exists y G_3(x, y)$ replaced by

$$\exists z(\sigma_2(x, y) = z)^\circ$$

where $\circ$ is the interpretation of $\mathcal{L}(m, 2)$ in terms of *Exp* that is very similar as $\mathcal{A}_3$ and has been omitted above.

Corollary 3.64. *For any $k > 3$, let T be some class of non-logical $\mathcal{L}(m, k)$ axioms and non-logical $\mathcal{L}(m, k)$ inferences. Then, for any $\mathcal{L}(m, \mathrm{Q})$ formula A and $\mathcal{L}(m, k)$ formula B, the following hold.*

(1) $\mathbf{A}(m, k) \cup \mathsf{T} \vdash B$ *if and only if* $\mathbf{B}(m, k) \cup \mathcal{A}_k(\mathsf{T}) \vdash \mathcal{A}_k(B)$.

(2) $\mathbf{A}(m, k) \cup \mathsf{T} \vdash A$ *if and only if* $\mathbf{B}(m, k) \cup \mathcal{A}_k(\mathsf{T}) \vdash A$.

These generalizations already suffice for our later purposes in Chapters 6 and 7 below.

We will give some related generalizations in Section 5.3.

4 Finite Sequences

In Section 3.2, we have reasoned in theory $\mathbf{I}(\Sigma_0)$ and applied Cantor's pairing function to derive two basic properties of finite sequences of numbers: Specifically, in Proposition 3.14, we derived that finite sequences of numbers may be prolonged by adding a number. The main aim of the present chapter is to elaborate on this and to derive, in extended base theory $\mathbf{A}(0,2)$ (respectively $\mathbf{A}(0,3)$), some key properties of finite sequences of numbers. We will rely on these repeatedly in what follows. Importantly, in Section 4.3, we will rely on them to define (for simplicity in $\mathbf{A}(0,3)$) function symbols by a form of bounded recursion and a form of bounded course-of-values recursion.

Concerning the efficiency of the sequence coding below, note that we work in the theory $\mathbf{A}(0,2)$ and estimate the bounds on codes by means of Nelson's smash function

$$m\#n := 2^{|m|\cdot|n|}$$

where $|n|$ is the number of bits in the binary expansion of the number n. That codes may be bounded by means of the smash function will give us enough flexibility for later chapters. As in Chapter 3, the crucial ideas of our coding are taken from Hájek and Pudlák [18], Chapter V, Section 3.

We start off by introducing some general concepts. Notation: We write $\overline{2}^{|x|\cdot|y|}$ for $\sigma_2(x, y)$ for better readability.

4.1 Definitions

Definition 4.1. Let $A(x_1, \ldots, x_k, z)$ be some $\mathcal{L}(0, 2)$ formula and let $f_A(x_1, \ldots, x_k, z)$ be some function symbol in $\mathcal{L}(0, 2)$ such that $\mathbf{A}(0, 2)$ proves the following.

$$\begin{aligned} \neg A(x_1, \ldots, x_k, z) &\leftrightarrow f_A(x_1, \ldots, x_k, z) = \overline{0} \\ A(x_1, \ldots, x_k, z) &\leftrightarrow f_A(x_1, \ldots, x_k, z) = \overline{1} \end{aligned}$$

Then f_A is a *characteristic function symbol* of A.

Let $A(x_1, \ldots, x_k, z)$ be $\mathcal{L}(0, 2)$ formula and $f_{\neg A}(x_1, \ldots, x_k, z)$ a characteristic function symbol for the negation of it. Also, let $f(x_1, \ldots, x_k, z)$ be a function symbol in $\mathcal{L}(0, 2)$ such that $\mathbf{A}(0, 2)$ proves (contains the axiom)

$$f(x_1, \ldots, x_k, z) = y \leftrightarrow C(x_1, \ldots, x_k, y, z),$$

where $C(x_1, \ldots, x_k, y, z)$ is the formula

$$\begin{gathered} y \leq z \wedge \\ (\forall v < y)(f_{\neg A}(x_1, \ldots, x_k, v) = \overline{1}) \wedge \\ (y = z \vee f_{\neg A}(x_1, \ldots, x_k, y) = \overline{0}) \end{gathered}$$

as in Definition 3.44 (A10). Then we write

$$(\mu v < z)A(x_1, \ldots, x_k, z)$$

for the function symbol $f(x_1, \ldots, x_k, z)$ in order to highlight the *bounded μ operator.*

Proposition 4.2. *For every $\mathcal{L}(0, 2)$ term $t(\vec{x})$, there is a function symbol f in $\mathcal{L}(0, 2)$ such that $\mathbf{A}(0, 2) \vdash t(\vec{x}) = f(\vec{x})$.*

Proof. By (meta) induction on the rank of t, relying on closure condition (composition) 3.44 (A9). $\square$

Proposition 4.3. *For every $\Sigma_0^{(0,2)}$ formula A, there is a characteristic function symbol f_A.*

Proof. By (meta) induction on the rank of A. We consider a few selected cases.

- A is $(s \leq t)$, where s, t are $\mathcal{L}(0,2)$ terms. Then consider the $\mathcal{L}(0,2)$ term $\overline{1} - (s - t)$ and apply Proposition 4.2.
- A is $(B \wedge C)$. By induction hypothesis, we get characteristic function symbols f_B, f_C of B, C, respectively. Hence, by Proposition 4.2, we have a function symbol in $\mathcal{L}(0,2)$ for $f_B(\vec{x}) \cdot f_C(\vec{x})$.
- A is $((\forall v \leq t)B)$, where t is an $\mathcal{L}(0,2)$ term. By induction hypothesis, we get a characteristic function symbol f_B of B and a characteristic function symbol $f_=$ of identity. So, by Proposition 4.2, we have a function symbol in $\mathcal{L}(0,2)$ for

$$f_=(t(\vec{x}) + \overline{1}, (\mu v < t(\vec{x}) + \overline{1})(f_B(\vec{x}) = \overline{0})).$$

Similarly for other cases. $\square$

Remark 4.4. Without going into details, let us remark that the counterparts of Proposition 4.2 and Proposition 4.3 can similarly be established for $\mathbf{A}(0,k)$, where $k > 2$.

By Proposition 4.2, we also have function symbols in $\mathcal{L}(0,2)$ such that

$$\begin{aligned} \lfloor x/y \rfloor &= (\mu v < x)(x < (v + \overline{1}) \cdot y) \\ x \bmod y &= x - \lfloor x/y \rfloor \cdot y \\ bit(x,y) &= \lfloor y/\overline{2}^x \rfloor \bmod \overline{2} \end{aligned}$$

are provable in $\mathbf{A}(0,2)$.

Recall that $\mathcal{L}(0,2)$ provides function symbol $|x|$ for the (bit) length of x, which is the least $v > 0$ so that $x < 2^v$. In particular, $|\overline{0}| = |\overline{1}| = \overline{1}$, which will occasionally allow us to indicate bounds on codes more conveniently.

Furthermore, we have a function symbol $\pi(x,y)$ in $\mathcal{L}(0,2)$ so that $\mathbf{A}(0,2) \vdash \pi(x,y) = z \leftrightarrow \mathit{Pair}(x,y,z)$, where $\mathit{Pair}(x,y,z)$ is the formula from Proposition 3.6.

Definition 4.5. The *pairing function* is denoted in $\mathcal{L}(0,2)$ by the symbol $\pi(x,y)$, where $\mathbf{A}(0,2)$ proves that $\pi(x,y)$ equals (=) the following.

$$(\mu v < \overline{4} \cdot (x+y)^2 + \overline{1})(\overline{2} \cdot v = ((x+y+\overline{1}) \cdot (x+y)) + \overline{2} \cdot x)$$

Projection functions are denoted in $\mathcal{L}(0,2)$ by $\pi_0(x)$ and $\pi_1(x)$, respectively, where $\mathbf{A}(0,2)$ proves the following.

$$\pi_0(x) = (\mu v < x + \overline{1})(\exists y \leq x)(\pi(v,y) = x)$$
$$\pi_1(x) = (\mu v < x + \overline{1})(\exists y \leq x)(\pi(y,v) = x)$$

Sequence numbers are defined in accordance with the coding from Section 3.2.

Definition 4.6. The *empty sequence* ($k = 0$) is denoted in $\mathcal{L}(0,2)$ by $\langle\ \rangle$, where $\mathbf{A}(0,2)$ proves

$$\langle\ \rangle = \pi(\overline{0}, \overline{1}).$$

Sequences of length $k+1$ are denoted in $\mathcal{L}(0,2)$ by $\langle x_0, \ldots, x_k \rangle$, where $\mathbf{A}(0,2)$ proves

$$\langle x_0, \ldots, x_k \rangle = \pi(f_0(x_0, \ldots, x_k), f_1(x_0, \ldots, x_k))$$

and equations:

$$f_0(x_0, \ldots, x_k) = \sum_{i<k} (x_i \cdot \overline{2}^{\sum_{j>i}^{k} |x_j|}) + x_k$$

$$f_1(x_0, \ldots, x_k) = \sum_{i \leq k} (\overline{2}^{\sum_{j \geq i}^{k} |x_j|}) + \overline{1}$$

Towards defining the number of occurrences of 1 in a binary expansion of a number x, we first define a formula $N1Seq(x, v, z)$. Intuitively, for example, $N1Seq(\overline{9}, \overline{2}, z)$ expresses that z is a sequence, such as $((0,1), (1,1), (2,1), (3,2))$, where every pair (i, u) contained in it represents that, up to the i-th bit of 9 in binary, the number of occurrences of 1 is u. Thus, for $i = 3 = |9| - 1$, we have $u = v = 2$. So, $N1Seq(x, v, z)$ may be defined as follows.

$$\begin{aligned}
&Seq(z) \wedge \\
&(\forall u \leq z)\big(\pi(u, \overline{0}) \in z \leftrightarrow u = bit(\overline{0}, x)\big) \wedge \\
&\pi(v, |x| - \overline{1}) \in z \wedge \\
&(\forall u < |x| - \overline{1})(\forall w_1 \leq z)\big(\pi(w_1, u + \overline{1}) \in z \leftrightarrow \\
&(\exists w_2 \leq z)(\pi(w_2, u) \in z \wedge w_1 = w_2 + bit(u + \overline{1}, x))\big).
\end{aligned}$$

The number of occurrences of 1 in the binary expansion of x is denoted in $\mathcal{L}(0, 2)$ by $n1(x)$, where $\mathbf{A}(0, 2)$ proves

$$n1(x) = (\mu v < |x| + \overline{1})(\exists z < f(x)) N1Seq(x, v, z)$$

and the following.

$$\begin{aligned}
f(x) &= \overline{4} \cdot (\overline{2}^{p(x)})^2 \\
p(x) &= (|x| \cdot |\pi(|x|, |x|)|) + \overline{1}
\end{aligned}$$

To verify that $n1(x)$ is suitably defined, note that, for $x = 0$ and $v = 0$ and some code z of the sequence $((0,0))$, we may prove in $\mathbf{A}(0, 2)$ that

$$\begin{aligned}
v < |x| + \overline{1} \wedge z < f(x) \wedge N1Seq(x, v, z) \wedge& \\
\pi_0(z) < \overline{2}^{p(x)} \wedge& \\
\pi_1(z) < \overline{2}^{p(x)}&
\end{aligned} \tag{4.1}$$

For $x > 0$, let us suppose in $\mathbf{A}(0,2)$ that, for some v and z, we have (4.1). Take $q = \pi(v + bit(|x + \overline{1}| - \overline{1}, x + \overline{1}), |x + \overline{1}| - \overline{1})$ and $\pi(z_0, z_1)$ with the following equalities.

$$\begin{aligned} z_0 &= \pi_0(z) \cdot \overline{2}^{|q|} + q \\ z_1 &= \pi_1(z) \cdot \overline{2}^{|q|} + \overline{1} \end{aligned}$$

Then $N1Seq(x + \overline{1}, v + bit(|x + \overline{1}| - \overline{1}, x + \overline{1}), \pi(z_0, z_1))$ is derivable in $\mathbf{A}(0,2)$. Also, by (4.1) and $|q| < |\pi(|x + \overline{1}|, |x + \overline{1}|)|$, we obtain in $\mathbf{A}(0,2)$ that

$$\begin{aligned} z_0 &< \overline{2}^{p(x+\overline{1})} \\ z_1 &< \overline{2}^{p(x+\overline{1})}. \end{aligned}$$

By Proposition 3.6 (6), therefore, we have in $\mathbf{A}(0,2)$:

$$\begin{aligned} \pi(z_0, z_1) &< \overline{4} \cdot (\overline{2}^{p(x+\overline{1})})^2 \\ &= f(x + \overline{1}) \end{aligned}$$

Thus, $\mathsf{IS}(\Sigma_0^{(0,2)})$ may be used to derive in $\mathbf{A}(0,2)$ that, for every x, there exist $v < |x| + \overline{1}$ and $z < f(x)$ so that $N1Seq(x, v, z)$ and $n1(x)$ indeed is the number of occurrences of 1 in the binary expansion of x.

Consider a code

$$\begin{array}{lllll} 101 & 11 & 110 & 10 & \\ 100 & 10 & 100 & 10 & 1 \end{array}$$

of sequence $(5, 3, 6, 2)$. In 3.13, we defined membership in a sequence. That 6 is the third member (with index 2) in $(5, 3, 6, 2)$, for example, may accordingly be expressed as: 110 occurs in the upper binary string and the number of markers (1s) to its left in the lower binary string is 3. This is the idea behind the following definition.

Definition 4.7. The *y-th member of a sequence x* is denoted in $\mathcal{L}(0,2)$ by $(x)_y$, where $\mathbf{A}(0,2)$ proves

$$(x)_y = (\mu v < x + \overline{1})(NMem(x,y,v)) \cdot f(x,y),$$

$f(x,y)$ is characteristic function symbol of the bounded formula $(\exists v < x + \overline{1})NMem(x,y,v)$, and $NMem(x,y,v)$ is defined as follows.

$$\begin{aligned}
&Seq(x) \wedge (\exists z_1, z_2 \leq x)\big(Pair(z_1,z_2,x) \wedge \\
&(\exists w_1, w_2, u_2, v_2 \leq z_2)(\exists u_1, v_1 < w_1)\big(Pow(w_1) \wedge Pow(w_2) \wedge \\
&\qquad\qquad w_2 \geq \overline{2} \wedge v < w_2 \wedge \\
&\qquad\qquad z_1 = u_2 \cdot w_2 \cdot w_1 + v \cdot w_1 + u_1 \wedge \\
&\qquad\qquad z_2 = v_2 \cdot \overline{2} \cdot w_2 \cdot w_1 + w_2 \cdot w_1 + w_1 + v_1 \wedge \\
&\qquad\qquad y = n1(v_2)\big)\big)
\end{aligned}$$

The length of a sequence, where a sequence is conceived as a pair of binary strings, is given by subtracting 1 from the number of markers (1s) in the lower (second) binary string.

Definition 4.8. The *length of sequence x* is denoted in $\mathcal{L}(0,2)$ by $lh(x)$, where $\mathbf{A}(0,2)$ proves

$$lh(x) = (\mu v < x + \overline{1})(Lh(x,v)) \cdot f(x),$$

$f(x)$ is characteristic function symbol of $(\exists v < x + \overline{1})Lh(x,v)$, and $Lh(x,v)$ is the following formula.

$$Seq(x) \wedge (\exists u_1, u_2 \leq x)\big(x = \pi(u_1,u_2) \wedge v = n1(u_2) - \overline{1}\big)$$

We may now define concatenation of sequences.

Definition 4.9. The *concatenation of sequences x, y* is denoted in $\mathcal{L}(0,2)$ by $x * y$, where $\mathbf{A}(0,2)$ proves

$$x * y = (\mu v < p(x,y))Conc(x,y,v) \cdot f(x,y),$$

$f(x, y)$ is characteristic function symbol of the bounded formula $(\exists v < p(x, y))Conc(x, y, v)$, and $Conc(x, y, v)$ is defined as

$$\begin{aligned}
&Seq(x) \wedge Seq(y) \wedge \\
&(\exists u_1, u_2 \leq x)(\exists v_1, v_2 \leq y)\big(x = \pi(u_1, u_2) \wedge \\
&\qquad\qquad\qquad\qquad\qquad\quad y = \pi(v_1, v_2) \wedge \\
&\qquad\qquad\qquad\qquad\qquad\quad v = \pi(p_0(u_1, v_1), p_1(u_2, v_2))\big)
\end{aligned}$$

and $\mathbf{A}(0, 2)$ proves the following.

$$\begin{aligned}
q(x, y) &= |\pi_1(x)| + |\pi_1(y)| + \overline{1} \\
p(x, y) &= \overline{4} \cdot (\overline{2}^{q(x,y)})^2 \\
p_0(u_1, v_1) &= u_1 \cdot \overline{2}^{|v_1|} + v_1 \\
p_1(u_2, v_2) &= (u_2 - \overline{1}) \cdot \overline{2}^{|v_2|} + v_2
\end{aligned}$$

To show in $\mathbf{A}(0, 2)$ that $x * y$ is appropriately defined, we may reason as for $n1(x)$ above. However, we omit details here and establish a more general Proposition 4.11 in the next section.

4.2 Basic Properties

We first consider two properties of the empty sequence.

Proposition 4.10. $\mathbf{A}(0, 2)$ *proves the following.*

- $Seq(\langle\ \rangle)$
- $lh(\langle\ \rangle) = \overline{0}$

Proof. Immediate by definition of $\langle\ \rangle$. $\square$

Proposition 4.11. *Let* $f(\vec{x})$ *be some function symbol in* $\mathcal{L}(0, 2)$. *Then* $\mathbf{A}(0, 2)$ *proves*

$$\big(Seq(u) \wedge (\forall w < lh(u))((u)_w \leq f(\vec{x}))\big) \rightarrow$$

$$(\exists v < q(\vec{x}, lh(u)))\big(Seq(v) \wedge \\ (\forall w < lh(u))((v)_w = (u)_w) \wedge lh(v) = lh(u)\big)$$

and $q(\vec{x}, y) = \overline{4} \cdot (\overline{2}^{p(\vec{x},y)})^2$ *and*

$$p(\vec{x}, y) = (y \cdot |f(\vec{x})|) + \overline{1}.$$

Proof. Assume in $\mathbf{A}(0,2)$ that

$$Seq(u) \wedge (\forall w < lh(u))((u)_w \leq f(\vec{x})). \tag{4.2}$$

Apply $\mathsf{IS}(\Sigma_0^{(0,2)})$ to the induction formula

$$y \leq lh(u) \rightarrow (\exists v < q(\vec{x}, y))\big(Seq(v) \wedge \\ (\forall w < y)((v)_w = (u)_w) \wedge lh(v) = y\big), \tag{4.3}$$

where y is induction variable. For $y = \overline{0}$, we get in $\mathbf{A}(0,2)$ that $\langle\ \rangle < q(\vec{x}, \overline{0})$. Consequently, Proposition 4.10 and logical reasoning cover the rest. For the induction step, assume as the induction hypothesis that (4.3) holds for y. Assume that $y + \overline{1} \leq lh(u)$. By (4.3), there exists a sequence $v < q(\vec{x}, y)$ such that $lh(v) = y$ and $(\forall w < y)((v)_w = (u)_w)$. Combining this with (4.2), we have that $(\forall w < y)((v)_w \leq f(\vec{x}))$ and $(u)_y \leq f(\vec{x})$. Thus, for $v * \langle (u)_y \rangle = \pi(v_0, v_1)$ such that

$$v_0 = \pi_0(v) \cdot \overline{2}^{|(u)_y|} + (u)_y$$
$$v_1 = \pi_1(v) \cdot \overline{2}^{|(u)_y|} + \overline{1}$$

we obtain that $v * \langle (u)_y \rangle < q(\vec{x}, y + 1)$, that $v * \langle (u)_y \rangle$ is some sequence of length $y + 1$, and $(\forall w < y + 1)((v * \langle (u)_y \rangle)_w = (u)_w)$. This concludes the induction step. □

Remark 4.12. To see that bound $q(\vec{x}, y)$ is given by means of smash function, observe that $lh(u) < |u|$.

To illustrate the use of Proposition 4.11, we add the following proposition about extracting an initial segment from a given sequence.

Proposition 4.13. $\mathbf{A}(0,2)$ *proves the following.*

$$\begin{aligned}&\big(Seq(u) \wedge y \leq lh(u)\big) \to \\ &\exists v\big(Seq(v) \wedge y = lh(v) \wedge (\forall w < y)((v)_w = (u)_w)\big)\end{aligned} \tag{4.4}$$

Proof. Assume in $\mathbf{A}(0,2)$ that $Seq(u)$ and

$$\begin{aligned}r(u,y) &= (\mu x < y)(\forall w < y)((u)_w \leq (u)_x)\\ f(u,y) &= (u)_{r(u,y)}\end{aligned}$$

to denote the maximum $f(u,y)$ of the first y members of the sequence u. Moreover, let $Seg(u,v,y)$ be defined to be the following formula.

$$Seq(v) \wedge y = lh(v) \wedge (\forall w < y)((v)_w = (u)_w)$$

Assume in $\mathbf{A}(0,2)$ that $y \leq lh(u)$ and $Seg(u,v,y)$. Use $\mathsf{IS}(\Sigma_0^{(0,2)})$ with induction variable z and induction formula

$$z \leq y \to (\forall w < z)((v)_w \leq f(u,y))$$

to derive:

$$(\forall w < y)((v)_w \leq f(u,y))$$

where $y = lh(v)$. Thus, under assumption $Seq(u) \wedge y \leq lh(u)$, we obtain by Proposition 4.11 the following in $\mathbf{A}(0,2)$, where $q(u,y)$ may be defined from $f(u,y)$ as in 4.11.

$$(\exists v)Seg(u,v,y) \leftrightarrow (\exists v < q(u,y))Seg(u,v,y)$$

That is, we may consider (4.4) as induction formula and y as induction variable. Therefore, it suffices to note that the following are derivable in $\mathbf{A}(0,2)$ from the assumption $Seq(u)$: First, that $\overline{0} \leq lh(u) \wedge Seg(u, \langle\ \rangle, \overline{0})$ and, second, that

$$y + \overline{1} \leq lh(u) \wedge Seg(u, v, y) \rightarrow Seg(u, v * \langle (u)_y \rangle, y + \overline{1}).$$

This concludes the proof. $\square$

The proposition below is about prolonging sequences and, as such, is just a (more natural) restatement of Proposition 3.14 (2) by means of concepts from Section 4.1. We omit its proof.

Proposition 4.14. $\mathbf{A}(0,2)$ *proves the following.*

$$\begin{aligned} &Seq(u) \rightarrow \\ &\exists v\big(Seq(v) \wedge lh(v) = lh(u) + \overline{1} \wedge \\ &\quad (\forall w < lh(u))((v)_w = (u)_w) \wedge (v)_{lh(u)} = x\big) \end{aligned}$$

4.3 Bounded Recursion

We may now prove two forms of bounded recursion, as we have announced. In contrast to the sequence coding of the foregoing sections, we will state the two forms of bounded recursion more naturally along values n which are exponential in $|n|$. Since the exponentiation function will almost always be at our disposal in what follows, it is preferable to strive for naturalness rather than for efficiency with regard to bounded recursion. However, in view of some issue in Chapter 6, we also remark on bounded recursion by means of the smash function in 4.16.

Proposition 4.15. *Let* $f_0(\vec{x})$, $f_1(\vec{x}, y, z)$, *and* $f_2(\vec{x}, y)$ *be function symbols in* $\mathcal{L}(0,3)$ *and define the following formulas.*

(1) $(\forall \vec{x})\big(f_0(\vec{x}) \leq f_2(\vec{x}, \overline{0})\big)$

(2) $(\forall \vec{x}, y, z)\big(z \leq f_2(\vec{x}, y) \rightarrow f_1(\vec{x}, y, z) \leq f_2(\vec{x}, y + \overline{1})\big)$

(3) $(\forall \vec{x}, y, z)\big(z \leq y \rightarrow f_2(\vec{x}, z) \leq f_2(\vec{x}, y)\big)$

Then, for some function symbol $f(\vec{x}, y)$ *in* $\mathcal{L}(0,3)$, *the following are derivable in* $\mathbf{A}(0,3)$ *from* (1)–(3).

(4) $f(\vec{x}, \overline{0}) = f_0(\vec{x})$

(5) $f(\vec{x}, y + \overline{1}) = f_1(\vec{x}, y, f(\vec{x}, y))$

Proof. Assume conditions (1)–(3) in $\mathbf{A}(0,3)$. Define the formula $RSeq(\vec{x}, y, z)$ as follows.

$$\begin{aligned} Seq(z) \wedge lh(z) = y + \overline{1} \wedge (z)_{\overline{0}} = f_0(\vec{x}) \wedge \\ (\forall w < y)((z)_{w+\overline{1}} = f_1(\vec{x}, w, (z)_w)) \end{aligned}$$

Use (1)–(3) and $\mathsf{IS}(\Sigma_0^{(0,3)})$ to show in $\mathbf{A}(0,3)$ that, for any $\vec{x}, y, z$, the following holds.

$$RSeq(\vec{x}, y, z) \to (\forall w < y + \overline{1})((z)_w \leq f_2(\vec{x}, w))$$

So, by Proposition 4.11, we have in $\mathbf{A}(0,3)$ that

$$(\exists z) RSeq(\vec{x}, y, z) \leftrightarrow (\exists z < q(\vec{x}, y + \overline{1})) RSeq(\vec{x}, y, z)$$

where q is obtained from f_2 as in Proposition 4.11. Furthermore, $\mathbf{A}(0,3)$ proves that $RSeq(\vec{x}, \overline{0}, \langle f_0(\vec{x}) \rangle)$ and that

$$RSeq(\vec{x}, y, z) \to RSeq(\vec{x}, y + \overline{1}, z * \langle f_1(\vec{x}, y, (z)_y) \rangle).$$

Hence, $(\exists z < q(\vec{x}, y + \overline{1})) RSeq(\vec{x}, y, z)$ may be derived in $\mathbf{A}(0,3)$. So, for

$$f(\vec{x}, y) = ((\mu v < q(\vec{x}, y + \overline{1})) RSeq(\vec{x}, y, v))_y$$

we may use $\mathsf{IS}(\Sigma_0^{(0,3)})$ with induction variable w to conclude the following in $\mathbf{A}(0,3)$.

$$RSeq(\vec{x}, y, z) \to (\forall w < y + \overline{1})((z)_w = f(\vec{x}, w))$$

Consequently, (4) and (5) are verified in $\mathbf{A}(0,3)$. □

Remark 4.16. By inspection of its definition in the proof just given, it becomes apparent that $f(n) = m$ has to be understood as: m is the last member of the least sequence $s < q(n+1)$ such that $RSeq(n, s)$, where it is guaranteed that, for any n, there is such a course-of-values sequence $s < q(n+1)$.

As for the efficiency, the problem with this definition is that, for the computation of $f(n)$ as described by $RSeq(n, s)$, we have to consider a course-of-values sequence s *of length* $n+1$, where $n+1$ is *exponential* in the (bit) length $|n+1|$ of its (least) binary representation.

So, if we want to describe a recursive computation of $f(n)$ by means of the smash function, we have to find a bound $p(|n|)$ of a course-of-values sequence that is polynomial in $|n|$. In view of Proposition 4.11 and as the general recipe, we may do so by a Σ_0 formula $F(n, m)$ which expresses, roughly: There is a course-of-values sequence s, where the length and maximum member of s are both bounded by respective polynomials in $|n|$, and so that $G(s)$ and m is the last member of s. Here $G(s)$ is an analogue of $RSeq(n, s)$, expressing a condition on members of s, such as, suggestively: $(\forall i < lh(s))(\ldots(s)_i\ldots)$.

As a corollary of 4.15, we also have a form of (exponentially) bounded course-of-values recursion.

Corollary 4.17. *Let* $f_1(\vec{x}, y, z)$ *and* $f_2(\vec{x}, y)$ *be function symbols in* $\mathcal{L}(0,3)$ *and define the following formulas.*

(1) $(\forall \vec{x})\big(\langle\ \rangle \leq f_2(\vec{x}, \overline{0})\big)$

(2) $(\forall \vec{x}, y, z)\big(z \leq f_2(\vec{x}, y) \rightarrow z * \langle f_1(\vec{x}, y, z)\rangle \leq f_2(\vec{x}, y + \overline{1})\big)$

(3) $(\forall \vec{x}, y)\big(z \leq y \rightarrow f_2(\vec{x}, z) \leq f_2(\vec{x}, y)\big)$

Then, for some $\mathcal{L}(0,3)$ *function symbols* $f(\vec{x}, y)$ *and* $f^*(\vec{x}, y)$*, the following are derivable in* $\mathbf{A}(0,3)$ *from* (1)–(3).

(4) $f(\vec{x}, y) = f_1(\vec{x}, y, f^*(\vec{x}, y))$

(5) $f^*(\vec{x},\overline{0}) = \langle\ \rangle$

(6) $f^*(\vec{x},y+\overline{1}) = f^*(\vec{x},y) * \langle f(\vec{x},y)\rangle$

Proof. Assume (1)–(3) in $\mathbf{A}(0,3)$. By Proposition 4.15, there is some function symbol $f^*(\vec{x},y)$ in $\mathcal{L}(0,3)$ such that $\mathbf{A}(0,3)$ proves the following.

$$f^*(\vec{x},\overline{0}) = \langle\ \rangle$$
$$f^*(\vec{x},y+\overline{1}) = f^*(\vec{x},y) * \langle f_1(\vec{x},y,f^*(\vec{x},y))\rangle$$

Consequently, for

$$f(\vec{x},y) = (f^*(\vec{x},y+\overline{1}))_y$$

we may verify (4)–(6) in $\mathbf{A}(0,3)$. □

Remark 4.18. Without going into details, let us also remark at this point that the counterparts of Proposition 4.15 and Corollary 4.17 may similarly be proved for $\mathbf{A}(0,k)$, where $k > 3$. This will be required occasionally in Chapter 7 below.

5 Syntax and Valuation

One objective of this chapter is to derive in $\mathbf{A}(0,2)$ (respectively $\mathbf{A}(0,3)$) several theorems about the syntax of languages $\mathcal{L}(m,\mathrm{Q})$ and valuation of $\mathcal{L}(0,\mathrm{Q})$ terms. In Section 5.1, we will define the required notions. In Section 5.2, we will use the notions to derive the respective theorems. For any such theorem, it will follow by Corollary 3.54 (1) that its respective $\mathcal{L}(0,\mathrm{Q})$ interpretation is derivable in $\mathbf{B}(0,2)$ (respectively $\mathbf{B}(0,3)$).

In Section 5.3, we will briefly demonstrate how to obtain in theory $\mathbf{A}(0,k+1)$ respective theorems about syntax of language $\mathcal{L}(m,k)$ and the valuation of $\mathcal{L}(0,k)$ terms, where $k > 1$. These generalizations will play a role later in Chapters 6 and 7.

We shall follow closely the usual meta-theoretic definitions in our standard coding of syntax. For convenience, we will usually make use of the exponentiation function to bound code numbers. This also enables us to encode, for example, a numeral such as $((\overline{0}+\overline{1})+\overline{1})+\overline{1}$ simply as nested sequence

$$(\#(+\overline{1}),(\#(+\overline{1}),(\#(+\overline{1}),\#(\overline{0}))))$$

where the numbers $\#(+\overline{1})$ and $\#(\overline{0})$ encode the function symbol $x+\overline{1}$ and constant $\overline{0}$, respectively. However, in view of certain issues in Chapter 6, we shall occasionally built on the efficiency of the sequence coding from Chapter 4 and also indicate how to bound codes of syntactical objects respectively by means of the smash function.

5.1 Definitions

For $0 < i \leq m$, we have function symbols in $\mathcal{L}(0,2)$ such that $\mathbf{A}(0,2)$ proves the following.

$$\begin{array}{rlrlrl}
\ulcorner +\overline{1} \urcorner &= \langle \overline{3} \rangle & \ulcorner \leq \urcorner &= \langle \overline{11} \rangle & \ulcorner \exists \urcorner &= \langle \overline{19} \rangle \\
\ulcorner + \urcorner &= \langle \overline{5} \rangle & \ulcorner \sim \urcorner &= \langle \overline{13} \rangle & \ulcorner \forall \urcorner &= \langle \overline{21} \rangle \\
\ulcorner \cdot \urcorner &= \langle \overline{7} \rangle & \ulcorner \vee \urcorner &= \langle \overline{15} \rangle & \ulcorner T_i \urcorner &= \langle \overline{29} + \overline{4} \cdot \overline{i} \rangle \\
\ulcorner = \urcorner &= \langle \overline{9} \rangle & \ulcorner \wedge \urcorner &= \langle \overline{17} \rangle & &
\end{array}$$

Recall: T_i is the i-th truth predicate.

Definition 5.1. The $\mathcal{L}(0,2)$ term $\ulcorner t \urcorner$, where t is some $\mathcal{L}(0,\mathrm{Q})$ term, is defined recursively as follows.

$$\begin{aligned}
\ulcorner x_i \urcorner &:= \langle \overline{2} \cdot \overline{i} \rangle \\
\ulcorner \overline{0} \urcorner &:= \langle \overline{1} \rangle \\
\ulcorner t_0 + \overline{1} \urcorner &:= \langle \ulcorner +\overline{1} \urcorner, \ulcorner t_0 \urcorner \rangle \\
\ulcorner t_0 + t_1 \urcorner &:= \langle \ulcorner + \urcorner, \ulcorner t_0 \urcorner, \ulcorner t_1 \urcorner \rangle \\
\ulcorner t_0 \cdot t_1 \urcorner &:= \langle \ulcorner \cdot \urcorner, \ulcorner t_0 \urcorner, \ulcorner t_1 \urcorner \rangle
\end{aligned}$$

We now define a formula $Var(x)$ that expresses that x is a variable.

Definition 5.2. $Var(x)$ is the formula $(\exists u < x)(x = \langle \overline{2} \cdot u \rangle)$.

That x is some $\mathcal{L}(0,\mathrm{Q})$ term may be expressed by a formula $Term(x)$, as shown in the next proposition. To prevent eyestrain in the present and the next Section 5.2, we do not endow meta names, such as $Term(x)$, always with arguments $(0,\mathrm{Q})$.

Proposition 5.3. *There is some* $\Sigma_0^{(0,2)}$ *formula* $Term(x)$ *such that* $\mathbf{A}(0,2)$ *proves*

$$Term(x) \leftrightarrow \bigvee \{(1)\text{–}(5)\},$$

where (1)–(5) *are the following disjuncts.*

(1) $Var(x)$

(2) $x = \ulcorner \overline{0} \urcorner$

(3) $(\exists u < x)\big(Term(u) \wedge x = \langle \ulcorner + \overline{1} \urcorner, u \rangle\big)$

(4) $(\exists u, v < x)\big(Term(u) \wedge Term(v) \wedge x = \langle \ulcorner + \urcorner, u, v \rangle\big)$

(5) $(\exists u, v < x)\big(Term(u) \wedge Term(v) \wedge x = \langle \ulcorner \cdot \urcorner, u, v \rangle\big)$

Proof. If we do not mind to use the exponentiation function, we may just consider a characteristic function symbol $c(z, x)$ of the formula $\bigvee\{(1)\text{–}(5)\}$, where $Term(u)$ and $Term(v)$ are replaced by $(z)_u = \overline{1}$ and $(z)_v = \overline{1}$, respectively, and z is the course-of-values sequence, as indicated below.

$$\begin{array}{cccccccccc} x & & 0 & 1 & 2 & \dots & 5 & 6 & \dots & 11 & \dots \\ z & (& 0 & 0 & 0 & \dots & 0 & 1 & \dots & 1 & \dots) \end{array}$$

Then we let $f_1(z, x) := c(z, x)$ and bound z in terms of x by

$$f_2(x) = \overline{4} \cdot (\overline{2}^{x+\overline{2}})^2.$$

So, (1)–(3) of Corollary 4.17 are satisfied and we have function symbols $f(x)$ and $f^*(x)$ in $\mathcal{L}(0,3)$ such that $\mathbf{A}(0,3)$ proves the following.

$$\begin{aligned} f(x) &= f_1(f^*(x), x) \\ f^*(\overline{0}) &= \langle\ \rangle \\ f^*(x + \overline{1}) &= f^*(x) * \langle f(x) \rangle. \end{aligned}$$

Consequently, we may define $Term(x)$ as $f(x) = \overline{1}$.

If we want to work with the smash function, relying on Remark 4.16, we may define $Term(x)$ to express that there exists a course-of-values sequence

$$z < \overline{4} \cdot (\overline{2}^{|x|^2+\overline{1}})^2$$

such that the following hold: Every member y of z satisfies that $c_{\#}(z, y) = \overline{1}$, where $c_{\#}(z, y)$ is a characteristic function symbol of $\bigvee\{(1)–(5)\}$ with x and $Term(u)$ and $Term(v)$ replaced, respectively, by y and $u \in z$ and $v \in z$, and x is the last member of z. So, this time, z may be depicted as follows.

$$\begin{array}{cccccccccccc} x & & 0 & 1 & 2 & \dots & 5 & 6 & \dots & 11 & \dots & \\ z & (& & & & & & 6 & & 11 & \dots &) \end{array}$$

The bound on z is by the fact that, for the built-up of a term x, we need to take into consideration at most $|x|$ many sub-terms of size at most x. $\square$

That x is a numeral (Num), an atom ($Atom$), literal ($PLit$, $NLit$), or some formula ($Form$), may be expressed by respective $\Sigma_0^{(0,2)}$ formulas in a similar fashion.

Moreover, there is a $\Sigma_0^{(0,2)}$ formula $FVar(x, v)$ to express that v is a free variable of x. Therefore, we may also define formulas to express that x is a closed term or a sentence, respectively, as follows.

$$\begin{aligned} ClTerm(x) &:= Term(x) \wedge (\forall v \leq x)\neg FVar(x, v) \\ Sent(x) &:= Form(x) \wedge (\forall v \leq x)\neg FVar(x, v) \end{aligned}$$

By Remark 4.16, we obtain a function symbol $rk(x)$ in $\mathcal{L}(0, 2)$ for the *rank* of a term or formula, respectively. For simplicity, we shall use just one $\mathcal{L}(0, 2)$ symbol for both.

We also have a function symbol in $\mathcal{L}(0, 3)$ to denote the x-th numeral, as shown in the next proposition. Note that the coding in the next proposition is chosen to match specifically our metatheoretic definition of numerals. To estimate the bounds on such numeral codes, the exponentiation function is required.

Proposition 5.4. $\mathbf{A}(0,3)$ *proves, for some* $\mathcal{L}(0,3)$ *function symbol* $num(x)$*, the following.*

- $num(\overline{0}) = \ulcorner\overline{0}\urcorner$
- $num(x+\overline{1}) = \langle \ulcorner{+}\overline{1}\urcorner, num(x)\rangle$

Proof. As $\ulcorner{+}\overline{1}\urcorner > \ulcorner\overline{0}\urcorner$, we have by Definition 4.6 and Proposition 3.6 (6) the following in $\mathbf{A}(0,3)$.

$$\begin{aligned}
|num(\overline{1})| &\leq |\langle \ulcorner{+}\overline{1}\urcorner, \ulcorner{+}\overline{1}\urcorner\rangle| \\
&\leq |\pi(\ulcorner{+}\overline{1}\urcorner \cdot \overline{2}^{|\ulcorner{+}\overline{1}\urcorner|} + \ulcorner{+}\overline{1}\urcorner, \overline{2}^{\overline{2}\cdot|\ulcorner{+}\overline{1}\urcorner|} + \overline{2}^{|\ulcorner{+}\overline{1}\urcorner|} + \overline{1})| \\
&\leq |\overline{4} \cdot (\overline{2}^{\overline{2}\cdot|\ulcorner{+}\overline{1}\urcorner|} + \overline{2}^{|\ulcorner{+}\overline{1}\urcorner|} + \overline{1})^2| \\
&\leq |\overline{4} \cdot (\overline{2}^{\overline{2}\cdot|\ulcorner{+}\overline{1}\urcorner|} + \overline{2}^{\overline{2}\cdot|\ulcorner{+}\overline{1}\urcorner|})^2| \\
&\leq \overline{4} \cdot |\ulcorner{+}\overline{1}\urcorner| + \overline{6}
\end{aligned}$$

By repeating the argument, we obtain:

$$\begin{aligned}
|num(\overline{2})| &\leq |\langle\langle \ulcorner{+}\overline{1}\urcorner, \ulcorner{+}\overline{1}\urcorner\rangle, \langle \ulcorner{+}\overline{1}\urcorner, \ulcorner{+}\overline{1}\urcorner\rangle\rangle| \\
&\leq \overline{4} \cdot |\langle \ulcorner{+}\overline{1}\urcorner, \ulcorner{+}\overline{1}\urcorner\rangle| + \overline{6} \\
&\leq \overline{4} \cdot (\overline{4} \cdot |\ulcorner{+}\overline{1}\urcorner| + \overline{6}) + \overline{6} \\
&\leq \overline{4}^2 \cdot |\ulcorner{+}\overline{1}\urcorner| + \overline{6}^2
\end{aligned}$$

In general, for $p(x) = \overline{4}^x \cdot |\ulcorner{+}\overline{1}\urcorner| + \overline{6}^x$, we can apply $\mathsf{IS}(\Sigma_0^{(0,3)})$ to derive $|num(x)| \leq p(x)$. Hence, for $f_2(x) = \overline{2}^{p(x)}$, we have the conditions (1)–(3) of Proposition 4.15 in $\mathbf{A}(0,3)$.

$$\begin{aligned}
f_0 = \ulcorner\overline{0}\urcorner &\leq f_2(\overline{0}) \\
z \leq f_2(x) &\rightarrow \langle \ulcorner{+}\overline{1}\urcorner, z\rangle \leq f_2(x+\overline{1}) \\
z \leq y &\rightarrow f_2(z) \leq f_2(y)
\end{aligned}$$

So, by that Proposition 4.15, there is a function symbol $f(x)$ in $\mathcal{L}(0,3)$ such that $\mathbf{A}(0,3)$ proves

$$\begin{aligned} f(\overline{0}) &= \ulcorner\overline{0}\urcorner \\ f(x+\overline{1}) &= \langle \ulcorner{+}\overline{1}\urcorner, f(x)\rangle \end{aligned}$$

and we may define $num(x)$ as $f(x)$. □

Remark 5.5. To match our meta-theoretic definitions, we shall continue to use the coding of numerals from Proposition 5.4. But in view of some issues in Chapter 6, let us indicate very briefly how numerals may alternatively be coded by means of the smash function in the current setting.

We may do so by course-of-values recursion, where the course-of-values sequence for computing $num(x)$ has length at most $|x|^2$ and the sub-numerals which occur in the course of this sequence are nested at most $2 \cdot ||x||$ times. For example, $num(44)$ will take the form

$$\langle \ulcorner\cdot\urcorner, num(4), \langle \ulcorner+\urcorner, num(3), \langle \ulcorner\cdot\urcorner, num(2), num(4)\rangle\rangle\rangle$$

according to the term $((\overline{2}\cdot\overline{4})+\overline{3})\cdot\overline{4}$. Accordingly, we will be able to bound the members of the course-of-values sequence by a polynomial in $|x|$ and rely on Remark 4.16.

Rather than stating the formal details, let us illustrate by an example computation that should make the idea sufficiently clear. Suppose that we want to compute $num(866)$, where 866 reads in binary

$$1101100010.$$

We sketch how to construct a course-of-values sequence, starting from $(\overline{0},\overline{1},\overline{2})$ whose members are some initial constants.

- Step 1: Add the numerals for the numbers (represented in binary)

$$11 \quad 01 \quad 10 \quad 00 \quad 10 \tag{5.1}$$

and the (auxiliary) numeral for 100. Observe that the members just added increase the maximum depth of the nesting by 2, like in all steps to follow. This is so because 100 and every member of (5.1) may be defined to have the form

$$(\overline{2^{2^{k-1}}}\cdot\bar{\imath})+\bar{\jmath}$$

where k (now $k=1$) is the step number and $\bar{\imath},\bar{\jmath}$ are subnumerals that have already been computed. (In the proper formal account, we could define members of the course-of-values sequence as pairs $p=(n,\ulcorner\overline{n}\urcorner)$ and define membership in the course-of-values sequence s as: There are pairs $p_0,p_1,p_2<p$ in s such that p is a pair

$$\big((p_0)_0\cdot(p_1)_0+(p_2)_0,\langle\ulcorner{+}\urcorner,\langle\ulcorner{\cdot}\urcorner,(p_0)_1,(p_1)_1\rangle,(p_2)_1\rangle\big)$$

of required form.) For the rest, we merely have to continue the procedure.

- Step 2: Add the numerals for

  1101 1000 10

  and the (auxiliary) numeral for 10000.
- Step 3: Add the numerals for

  11011000 10

  and, as final step, the numeral for 1101100010.

Importantly, at most $|866|=10$ many members (pairs) are added at each step and at most $|10|=4$ steps have to be taken. So, the length of the course-of-values sequences is bounded by $|866|^2$. Furthermore, numerals added at step k (as components of pairs) are nested at most $2\cdot k$ times and, thus, at most $2\cdot|10|$ times in total. That is, the overall course-of-values sequences for $num(n)$ may be bounded by applying the Proposition 4.11 as usual.

The subsequent proposition is concerned with substitutions in terms.

Proposition 5.6. *For any* $k \geq 0$*, there is some function symbol* $x[u_0, \ldots, u_k/v_0, \ldots, v_k]$ *in* $\mathcal{L}(0,2)$ *such that* $\mathbf{A}(0,2)$ *proves*

$$x[u_0, \ldots, u_k/v_0, \ldots, v_k] = y \leftrightarrow \bigvee \big\{(1)\text{–}(6)\big\}$$

where (1)–(6) *are the following disjuncts.*

(1) $\big(\neg Term(x) \vee \bigvee_{i=0}^{k} (\neg Term(u_i) \vee \neg Var(v_i))\big) \wedge y = \overline{0}$

(2) $\bigwedge_{i=0}^{k} \big(Term(u_i) \wedge Var(v_i)\big) \wedge$
$\big((Var(x) \wedge \bigwedge_{i=0}^{k} x \neq v_i) \vee x = \ulcorner \overline{0} \urcorner\big) \wedge y = x$

(3) $\bigwedge_{i=0}^{k} \big(Term(u_i) \wedge Var(v_i)\big) \wedge \bigvee_{i=0}^{k} \big(x = v_i \wedge y = u_i\big)$

(4) $\bigwedge_{i=0}^{k} \big(Term(u_i) \wedge Var(v_i)\big) \wedge$
$(\exists w < x)\big(Term(w) \wedge$
$x = \langle \ulcorner + \overline{1} \urcorner, w \rangle \wedge$
$y = \langle \ulcorner + \overline{1} \urcorner, w[u_0, \ldots, u_k/v_0, \ldots, v_k] \rangle\big)$

(5) $\bigwedge_{i=0}^{k} \big(Term(u_i) \wedge Var(v_i)\big) \wedge$
$(\exists w_1, w_2 < x)\big(Term(w_1) \wedge Term(w_2) \wedge$
$x = \langle \ulcorner + \urcorner, w_1, w_2 \rangle \wedge$
$y = \langle \ulcorner + \urcorner, w_1[u_0, \ldots, u_k/v_0, \ldots, v_k], w_2[u_0, \ldots, u_k/v_0, \ldots v_k] \rangle\big)$

(6) $\bigwedge_{i=0}^{k} \big(Term(u_i) \wedge Var(v_i)\big) \wedge$
$(\exists w_1, w_2 < x)\big(Term(w_1) \wedge Term(w_2) \wedge$
$x = \langle \ulcorner \cdot \urcorner, w_1, w_2 \rangle \wedge$
$y = \langle \ulcorner \cdot \urcorner, w_1[u_0, \ldots, u_k/v_0, \ldots, v_k], w_2[u_0, \ldots, u_k/v_0, \ldots, v_k] \rangle\big)$

Proof. Like the proof of Proposition 5.3 above and the proof of Proposition 5.13 below. If we want to do without exponentiation function and rather use the smash function, the bounds on codes may be estimated in terms of $|x|, |u_0|, \dots |u_k|$. However, we omit details here. $\square$

Remark 5.7. The function symbol $x[u_0, \dots, u_k/v_0, \dots, v_k]$ is defined with external arity k – as it will prove convenient in later chapters for the presentation of certain proofs. By relying on the results about sequences from Chapter 4, however, we also get a (binary) symbol $x[u]$ in $\mathcal{L}(0,2)$ so that, for example, we have in $\mathbf{A}(0,2)$ in place of 5.6 (3) the following.

$$Seq(u) \wedge (\forall v < lh(u)) Term((u)_v) \wedge Var(x) \wedge x[u] = (u)_{\overline{0}}$$

Notice: Letting $f(x)$ be an $\mathcal{L}(0,2)$ symbol to denote a sequence (without any repetitions) of all the free variables of x, so that $f(x) = \langle x \rangle$ if x is variable, the above condition incorporates the idea that $\overline{0} < lh(u) \wedge x = (f(x))_{\overline{0}} \wedge x[u] = (u)_{\overline{0}}$.

Definition 5.8. The $\mathcal{L}(0,2)$ term $\ulcorner A \urcorner$, where A is some $\mathcal{L}(m, \mathrm{Q})$ formula, is defined recursively as follows.

$$\begin{aligned} \ulcorner s = t \urcorner &:= \langle \ulcorner = \urcorner, \ulcorner s \urcorner, \ulcorner t \urcorner \rangle \\ \ulcorner s \neq t \urcorner &:= \langle \ulcorner \sim \urcorner, \ulcorner s = t \urcorner \rangle \end{aligned}$$

(Similarly for $\leq, >, T_i, \sim T_i$.)

$$\begin{aligned} \ulcorner B \vee C \urcorner &:= \langle \ulcorner \vee \urcorner, \ulcorner B \urcorner, \ulcorner C \urcorner \rangle \\ \ulcorner B \wedge C \urcorner &:= \langle \ulcorner \wedge \urcorner, \ulcorner B \urcorner, \ulcorner C \urcorner \rangle \\ \ulcorner \exists x B \urcorner &:= \langle \ulcorner \exists \urcorner, \ulcorner x \urcorner, \ulcorner B \urcorner \rangle \\ \ulcorner \forall x B \urcorner &:= \langle \ulcorner \forall \urcorner, \ulcorner x \urcorner, \ulcorner B \urcorner \rangle \end{aligned}$$

We add a remark about substitution in formulas.

Remark 5.9. Remember that

$$A(s_0, \dots, s_k/v_0, \dots, v_k)$$

was defined on the meta-level to uniformly replace, for any $i \leq k$, the variable v_i by the term s_i only if v_i is *free* in A, but was *not defined* to rename bound variables if variable clashes make it necessary. Rather, variable clashes are prevented by renaming bound variables *before* a substitution is performed.

We may implement this in our formalization of simultaneous substitution as follows. First note that, by extending 5.6 accordingly, we obtain naive replacement (*rep*) of variables by terms, where variables are replaced irrespectively of whether they occur freely or not in a formula. So, using the characteristic function symbol of $FVar(\ulcorner A \urcorner, \ulcorner x \urcorner)$ (free variable), we obtain substitution $\ulcorner A \urcorner[\ulcorner s_0 \urcorner, \dots, \ulcorner s_k \urcorner / \ulcorner x_0 \urcorner, \dots, \ulcorner x_k \urcorner]$ restricted to free variables such that, for example,

$$\begin{aligned}&\bigwedge\nolimits_{i=0}^{1} FVar(\ulcorner A \urcorner, \ulcorner x_i \urcorner) \wedge \bigwedge\nolimits_{i=2}^{k} \neg FVar(\ulcorner A \urcorner, \ulcorner x_i \urcorner) \to \\ &rep(\ulcorner A \urcorner, \ulcorner s_0 \urcorner, \ulcorner s_1 \urcorner, \ulcorner x_0 \urcorner, \ulcorner x_1 \urcorner)\end{aligned}$$

and similarly for the other $(2^k - 1)$ possible cases.

To prevent variable clashes, we may, for example, first apply Remark 4.16 to obtain a bound b on all p such that, for some $i \leq k$, p is an index of a variable in term s_i, replace each bound variable z_q in A by variable z_{b+q}, and finally apply free-variable substitution $\ulcorner B \urcorner[\ulcorner s_0 \urcorner, \dots, \ulcorner s_k \urcorner / \ulcorner x_0 \urcorner, \dots, \ulcorner x_k \urcorner]$ to the result B of this replacement.

To simplify our notation, we will henceforth use the symbol $x[u_0, \dots, u_k/v_0, \dots, v_k]$ for both: substitutions in terms and substitutions in formulas. To save horizontal space, we also write

$$x\left[\frac{u_0, \dots, u_k}{v_0, \dots, v_k}\right]$$

in place of $x[u_0, \ldots, u_k/v_0, \ldots, v_k]$.

As counterparts of terms that are defined in 5.8, we consider respective function symbols that will simplify the formulation of truth axioms in later chapters: That is, for $0 < i \leq m$, we have function symbols in $\mathcal{L}(0, 2)$ so that $\mathbf{A}(0, 2)$ proves the following.

$$\begin{aligned}[x = y] &= \langle \ulcorner = \urcorner, x, y \rangle \\ [x \neq y] &= \langle \ulcorner \sim \urcorner, [x = y] \rangle\end{aligned}$$

(Similarly for $\leq, >, T_i, \sim T_i$.)

$$\begin{aligned}[x \vee y] &= \langle \ulcorner \vee \urcorner, x, y \rangle \\ [x \wedge y] &= \langle \ulcorner \wedge \urcorner, x, y \rangle \\ [(\exists x) y] &= \langle \ulcorner \exists \urcorner, x, y \rangle \\ [(\forall x) y] &= \langle \ulcorner \forall \urcorner, x, y \rangle \\ [(\exists x \leq y) z] &= [(\exists x)[[x \leq y] \wedge z]] \\ [(\forall x \leq y) z] &= [(\forall x)[[x > y] \vee z]]\end{aligned}$$

For future reference, we also highlight the following (selected) properties.

Proposition 5.10. $\mathbf{A}(0, 2)$ *proves the following.*

(1) $\bigwedge_{i=0}^{k} \big(Term(u_i) \wedge Var(v_i)\big) \wedge Form([x_0 = x_1]) \to$
$[x_0 = x_1] \left[\dfrac{u_0, \ldots u_k}{v_0, \ldots, v_k}\right] =$
$[x_0 \left[\dfrac{u_0, \ldots u_k}{v_0, \ldots, v_k}\right] = x_1 \left[\dfrac{u_0, \ldots u_k}{v_0, \ldots, v_k}\right]]$

(2) $\bigwedge_{i=0}^{k} \big(Term(u_i) \wedge Var(v_i)\big) \wedge Form([x_0 \vee x_1]) \to$
$[x_0 \vee x_1] \left[\dfrac{u_0, \ldots u_k}{v_0, \ldots, v_k}\right] =$
$[x_0 \left[\dfrac{u_0, \ldots u_k}{v_0, \ldots, v_k}\right] \vee x_1 \left[\dfrac{u_0, \ldots u_k}{v_0, \ldots, v_k}\right]]$

(3) $\bigwedge_{i=0}^{k} \big(Term(u_i) \wedge Var(v_i)\big) \wedge Form([x_0 \wedge x_1]) \rightarrow$
$[x_0 \wedge x_1] \left[\dfrac{u_0, \dots u_k}{v_0, \dots, v_k}\right] =$
$[x_0 \left[\dfrac{u_0, \dots u_k}{v_0, \dots, v_k}\right] \wedge\ x_1 \left[\dfrac{u_0, \dots u_k}{v_0, \dots, v_k}\right]]$

(4) $\bigwedge_{i=0}^{k} \big(Term(u_i) \wedge Var(v_i)\big) \wedge Form([(\forall x_0 \leq x_1)x_2]) \rightarrow$
$[(\forall x_0 \leq x_1)x_2] \left[\dfrac{u_0, \dots u_k}{v_0, \dots, v_k}\right] =$
$[(\forall x_0 \leq x_1 \left[\dfrac{u_0, \dots u_k}{v_0, \dots, v_k}\right])\, x_2 \left[\dfrac{u_0, \dots u_k}{v_0, \dots, v_k}\right]]$

Negations may be denoted in $\mathbf{A}(0,2)$ as follows.

Proposition 5.11. *There is a function symbol* $[\neg x]$ *in* $\mathcal{L}(0,2)$ *so that* $\mathbf{A}(0,2)$ *proves*

$$[\neg x] = y \leftrightarrow \bigvee \big\{(1)\text{–}(7)\big\},$$

where (1)–(7) *are the following disjuncts.*

(1) $\neg Form(x) \wedge y = \overline{0}$

(2) $Atom(x) \wedge y = \langle \ulcorner {\sim} \urcorner, x\rangle$

(3) $(\exists u < x)\big(Atom(u) \wedge x = \langle \ulcorner {\sim} \urcorner, u\rangle \wedge y = u\big)$

(4) $(\exists u, v < x)\big(Form(u) \wedge Form(v) \wedge$
$x = \langle \ulcorner \vee \urcorner, u, v\rangle \wedge y = \langle \ulcorner \wedge \urcorner, [\neg u], [\neg v]\rangle\big)$

(5) $(\exists u, v < x)\big(Form(u) \wedge Form(v) \wedge$
$x = \langle \ulcorner \wedge \urcorner, u, v\rangle \wedge y = \langle \ulcorner \vee \urcorner, [\neg u], [\neg v]\rangle\big)$

(6) $(\exists u, v < x)\big(Var(u) \wedge Form(v) \wedge$
$x = \langle \ulcorner \exists \urcorner, u, v\rangle \wedge y = \langle \ulcorner \forall \urcorner, u, [\neg v]\rangle\big)$

(7) $(\exists u, v < x)\big(Var(u) \wedge Form(v) \wedge$
$x = \langle \ulcorner\forall\urcorner, u, v\rangle \wedge y = \langle \ulcorner\exists\urcorner, u, [\neg v]\rangle\big)$

We now move on to valuation of closed $\mathcal{L}(0, \mathrm{Q})$ terms. Let us remark that, in view of later issues in Chapters 6 and 7, we have given already enough indications of how to encode specific syntactical concepts in $\mathbf{A}(0, 2)$ and may continue our work more conveniently in $\mathbf{A}(0, 3)$ for the remainder of this chapter. It also allows us to rely on the more natural coding 5.4 of numerals and the more natural forms 4.15 and 4.17 of bounded recursion.

To compute values of closed $\mathcal{L}(0, \mathrm{Q})$ terms, we may appeal to the following.

Proposition 5.12. *Let $f(\vec{x}, y)$ be some function symbol in $\mathcal{L}(0, 3)$. There also are function symbols $\sum_{v<y} f(\vec{x}, v)$ and $\prod_{v<y} f(\vec{x}, v)$ in $\mathcal{L}(0, 3)$ such that* $\mathbf{A}(0, 3)$ *proves*

$$\sum_{v<0} f(\vec{x}, v) = \overline{0}$$
$$\sum_{v<y+\overline{1}} f(\vec{x}, v) = \sum_{v<y} f(\vec{x}, v) + f(\vec{x}, y)$$

and the following.

$$\prod_{v<0} f(\vec{x}, v) = \overline{1}$$
$$\prod_{v<y+\overline{1}} f(\vec{x}, v) = \prod_{v<y} f(\vec{x}, v) \cdot f(\vec{x}, y)$$

Proof. By Proposition 4.15. For example, for $\sum_{v<y} f(\vec{x}, v)$, it is enough to observe that, for $f_2(\vec{x}, y) = y \cdot \max_{v<y} f(\vec{x}, v)$, where the maximum may be taken as in the proof of Proposition 4.13, conditions (1)–(3) of Proposition 4.15 may be verified in $\mathbf{A}(0, 3)$. Similarly for $\prod_{v<y} f(\vec{x}, v)$. $\square$

Proposition 5.13. *There is a function symbol* $val(x)$ *in* $\mathcal{L}(0,3)$ *so that* $\mathbf{A}(0,3)$ *proves*

$$val(x) = y \leftrightarrow \bigvee \big\{(1)\text{–}(5)\big\},$$

where (1)–(5) *are the following disjuncts.*

(1) $\neg ClTerm(x) \wedge y = \overline{0}$

(2) $x = \ulcorner\overline{0}\urcorner \wedge y = \overline{0}$

(3) $(\exists u < x)\big(ClTerm(u) \wedge$
$x = \langle \ulcorner{+}\overline{1}\urcorner, u\rangle \wedge y = val(u) + \overline{1}\big)$

(4) $(\exists u, v < x)\big(ClTerm(u) \wedge ClTerm(v) \wedge$
$x = \langle \ulcorner{+}\urcorner, u, v\rangle \wedge y = val(u) + val(v)\big)$

(5) $(\exists u, v < x)\big(ClTerm(u) \wedge ClTerm(v) \wedge$
$x = \langle \ulcorner{\cdot}\urcorner, u, v\rangle \wedge y = val(u) \cdot val(v)\big)$

Proof. In accordance with (3), let $c_3(u,x)$ be some characteristic function symbol of formula $ClTerm(u) \wedge x = \langle \ulcorner{+}\overline{1}\urcorner, u\rangle$. Similarly for $c_4(u,v,x)$, which concerns (4), and $c_5(u,v,x)$, which concerns (5). Building on Proposition 5.12, let $f_1(x,z)$ be the subsequent sum

$$\begin{gathered}\sum\nolimits_{u<x} \big(c_3(u,x) \cdot ((z)_u + \overline{1})\big) + \\ \sum\nolimits_{u<x} \sum\nolimits_{v<x} \big(c_4(u,v,x) \cdot ((z)_u + (z)_v)\big) + \\ \sum\nolimits_{u<x} \sum\nolimits_{v<x} \big(c_5(u,v,x) \cdot ((z)_u \cdot (z)_v)\big)\end{gathered}$$

where z is the course-of-values sequence. To bound z, it suffices (see also Remark 5.14 below) to bound the sequence s with

$$\begin{array}{cccccccc} x & 0 & 1 & 2 & 3 & 4 & 5 & \dots \\ s & (\,0 & 1 & 2 & 3 & 4 & 5 & \dots\,) \end{array}$$

which can be done in terms of x as follows.

$$f_2(x) = \overline{4} \cdot (\prod_{u \leq x} \overline{2}^{|u|} \cdot \overline{2})^2 \tag{5.2}$$

We may then verify in $\mathbf{A}(0,3)$ that (1)–(3) of Corollary 4.17 are satisfied for f_1, f_2. $\square$

Remark 5.14. Indeed, that (5.2) may be used as bound for the course-of-values sequence s is due to the fact that, for arbitrary closed $\mathcal{L}(0,\mathrm{Q})$ terms t, we have in $\mathbf{A}(0,3)$ that $t < \ulcorner t \urcorner$. We will return to this issue in the next Chapter 6.

Actually, as $t < \ulcorner t \urcorner$, we are not reliant on the exponentiation function in the proof of Proposition 5.13, for we may use exactly the same strategy as in the proof of Proposition 5.3 to bound the values of closed $\mathcal{L}(0,\mathrm{Q})$ terms in $\mathbf{A}(0,2)$ by means of the smash function. But we may ignore this here.

5.2 Basic Properties

In the present section, we will derive in $\mathbf{A}(0,3)$ several theorems about the syntax of $\mathcal{L}(m,\mathrm{Q})$ and valuation of $\mathcal{L}(0,\mathrm{Q})$ terms. (We reason in $\mathbf{A}(0,3)$ because we want to rely on the more natural coding of numerals, as we have mentioned. However, the reader should keep in mind that the exponentiation function is actually dispensable in what we are going to prove.)

We first derive in $\mathbf{A}(0,3)$ that every numeral is a closed term.

Proposition 5.15. $\mathbf{A}(0,3) \vdash \mathit{ClTerm}(\mathit{num}(x))$

Proof. We have in $\mathbf{A}(0,3)$ that $\mathit{ClTerm}(\mathit{num}(\overline{0}))$ and

$$\mathit{ClTerm}(\mathit{num}(x)) \to \mathit{ClTerm}(\mathit{num}(x + \overline{1})).$$

Thus, we may apply $\mathsf{IS}(\Sigma_0^{(0,3)})$, using $\mathit{ClTerm}(\mathit{num}(x))$ as the induction formula. $\square$

We now prove several $\mathbf{A}(0,3)$ theorems about substitutions in closed $\mathcal{L}(0,\mathrm{Q})$ terms.

Proposition 5.16. *Let* $x, u_0, \ldots, u_k, v_0, \ldots, v_k$ *be variables and let* $t(v_0, \ldots, v_k), s_0, \ldots, s_k$ *be* $\mathcal{L}(0,\mathrm{Q})$ *terms. The following are provable in* $\mathbf{A}(0,3)$*.*

(1) $$\ulcorner t \urcorner \left[\frac{\ulcorner s_0 \urcorner, \ldots, \ulcorner s_k \urcorner}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = \ulcorner t \left(\frac{s_0, \ldots, s_k}{v_0, \ldots, v_k}\right) \urcorner$$

(2) $$\mathit{ClTerm}(x) \to x \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = x$$

(3) $$\bigwedge\nolimits_{i=0}^{k} \mathit{ClTerm}(u_i) \to$$
$$\ulcorner t \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] =$$
$$\ulcorner t \urcorner \left[\frac{u_0, \ldots, u_j}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_j \urcorner}\right] \left[\frac{u_{j+1}, \ldots, u_k}{\ulcorner v_{j+1} \urcorner, \ldots, \ulcorner v_k \urcorner}\right]$$

(4) $$\bigwedge\nolimits_{i=0}^{k} \mathit{ClTerm}(u_i) \to$$
$$\ulcorner t \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = \ulcorner t \urcorner \left[\frac{u_0}{\ulcorner v_0 \urcorner}\right] \cdots \left[\frac{u_k}{\ulcorner v_k \urcorner}\right]$$

(5) $$\bigwedge\nolimits_{i=0}^{k} \mathit{ClTerm}(u_i) \to$$
$$\mathit{ClTerm}(\ulcorner t \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right])$$

Proof. We consider only (2)–(5). For (2), apply $\mathsf{IS}(\Sigma_0^{(0,3)})$ with induction formula

$$(\forall z < x)\big(\mathit{ClTerm}(z) \to z \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = z\big)$$

and induction variable x, by relying on Proposition 5.6 and the counterpart of Proposition 5.3 for closed terms.

For (3), we proceed by (meta) induction on k. If $k = 0$, we get the result by logic. For the induction step (for $k > 0$), we proceed by a sub (meta) induction on the rank of t. We consider two examples to show the strategy.

- t is variable v_i ($i \leq k$). If $i \leq j$, reason in $\mathbf{A}(0,3)$ as follows. Assume that $\bigwedge_{i=0}^{k} ClTerm(u_i)$. Use Proposition 5.6 (3) to obtain

  $$\ulcorner v_i \urcorner \left[\frac{u_0, \ldots, u_j}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_j \urcorner}\right] = u_i$$

  and

  $$\ulcorner v_i \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = u_i.$$

  By (2), we also have

  $$u_i \left[\frac{u_{j+1}, \ldots, u_k}{\ulcorner v_{j+1} \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = u_i.$$

  Hence, we are done. The case for $i > j$ is similar.

- t is $(t_0 + \overline{1})$. By 5.6 (4), we have in $\mathbf{A}(0,3)$ that

  $$\ulcorner t_0 + \overline{1} \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = \langle \ulcorner + \overline{1} \urcorner, \ulcorner t_0 \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] \rangle.$$

  Similarly, we have in $\mathbf{A}(0,3)$ that

  $$\ulcorner t_0 + \overline{1} \urcorner \left[\frac{u_0, \ldots, u_j}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_j \urcorner}\right] \left[\frac{u_{j+1}, \ldots, u_{k+1}}{\ulcorner v_{j+1} \urcorner, \ldots, \ulcorner v_{k+1} \urcorner}\right]$$

  is identical (=) to

  $$\langle \ulcorner + \overline{1} \urcorner, \ulcorner t_0 \urcorner \left[\frac{u_0, \ldots, u_j}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_j \urcorner}\right] \left[\frac{u_{j+1}, \ldots, u_{k+1}}{\ulcorner v_{j+1} \urcorner, \ldots, \ulcorner v_{k+1} \urcorner}\right] \rangle.$$

  Hence, it only remains to apply the induction hypothesis.

(4) is obtained by iterated application of (3).

For (5), we proceed by (meta) induction on the rank of t. We consider only three cases.

- t is v_i $(i \leq k)$. Assume in $\mathbf{A}(0,3)$ that $\bigwedge_{i=0}^{k} \mathit{ClTerm}(u_i)$. By Proposition 5.6 (3), we find
$$\ulcorner v_i \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = u_i.$$
So, we are done.
- t is $\overline{0}$. Assume in $\mathbf{A}(0,3)$ that $\bigwedge_{i=0}^{k} \mathit{ClTerm}(u_i)$. We get in $\mathbf{A}(0,3)$ that $\mathit{ClTerm}(\ulcorner \overline{0} \urcorner)$. By (2), we find
$$\ulcorner \overline{0} \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = \ulcorner \overline{0} \urcorner.$$
So, we are done.
- t is $(t_0 + \overline{1})$. Assume in $\mathbf{A}(0,3)$ that $\bigwedge_{i=0}^{k} \mathit{ClTerm}(u_i)$. By Proposition 5.6 (4), we have in $\mathbf{A}(0,3)$ that
$$\ulcorner t_0 + \overline{1} \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] = \langle \ulcorner + \overline{1} \urcorner, \ulcorner t_0 \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] \rangle.$$
So, we obtain the result by the induction hypothesis and a counterpart of 5.3 (3) for closed terms.

This completes the proof. □

Notice that 5.16 (3) is reminiscent of the famous S_n^m *Theorem* from recursion theory: A further application of 5.15 and 5.16 (3) yields, for example, the subsequent corollary.

Corollary 5.17. $\mathbf{A}(0,3)$ *proves, for any arbitrary function symbol* $f(v_0, \ldots, v_k)$ *from* $\mathcal{L}(0, \mathrm{Q})$*, the following.*

$$\ulcorner f(v_0, \ldots, v_k) \urcorner \left[\frac{\mathit{num}(v_0), \ldots, \mathit{num}(v_k)}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right] =$$

$$\ulcorner f(v_0, \ldots, v_k)\urcorner \left[\frac{num(v_0), \ldots, num(v_j)}{\ulcorner v_0\urcorner, \ldots, \ulcorner v_j\urcorner}\right] \left[\frac{num(v_{j+1}), \ldots, num(v_k)}{\ulcorner v_{j+1}\urcorner, \ldots, \ulcorner v_k\urcorner}\right]$$

The following proposition is the counterpart of 5.16 for sentences. We omit its proof, because it is very similar to the one of 5.16, but remark that it would not go through if we had not defined substitutions in sentences formally as we defined them on the meta-level: Recall Remark 5.9.

Proposition 5.18. *Let $x, u_0, \ldots, u_k, v_0, \ldots, v_k$ all be variables, let $A(v_0, \ldots, v_k)$ be an $\mathcal{L}(m, \mathrm{Q})$ formula, and let $s_0, \ldots, s_k$ be $\mathcal{L}(0, \mathrm{Q})$ terms. Then $\mathbf{A}(0,3)$ proves the following.*

(1) $$\ulcorner A\urcorner \left[\frac{\ulcorner s_0\urcorner, \ldots, \ulcorner s_k\urcorner}{\ulcorner v_0\urcorner, \ldots, \ulcorner v_k\urcorner}\right] = \ulcorner A\left(\frac{s_0, \ldots s_k}{v_0, \ldots, v_k}\right)\urcorner$$

(2) $$Sent(x) \to x \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0\urcorner, \ldots, \ulcorner v_k\urcorner}\right] = x$$

(3) $$\bigwedge\nolimits_{i=0}^{k} ClTerm(u_i) \to$$
$$\ulcorner A\urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0\urcorner, \ldots, \ulcorner v_k\urcorner}\right] =$$
$$\ulcorner A\urcorner \left[\frac{u_0, \ldots, u_j}{\ulcorner v_0\urcorner, \ldots, \ulcorner v_j\urcorner}\right] \left[\frac{u_{j+1}, \ldots, u_k}{\ulcorner v_{j+1}\urcorner, \ldots, \ulcorner v_k\urcorner}\right]$$

(4) $$\bigwedge\nolimits_{i=0}^{k} ClTerm(u_i) \to$$
$$\ulcorner A\urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0\urcorner, \ldots, \ulcorner v_k\urcorner}\right] = \ulcorner A\urcorner \left[\frac{u_0}{\ulcorner v_0\urcorner}\right] \cdots \left[\frac{u_k}{\ulcorner v_k\urcorner}\right]$$

(5) $$\bigwedge\nolimits_{i=0}^{k} ClTerm(u_i) \to$$
$$Sent(\ulcorner A\urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0\urcorner, \ldots, \ulcorner v_k\urcorner}\right])$$

We also add the formal counterpart of Proposition 2.10.

Proposition 5.19. $\mathbf{A}(0,3) \vdash Form(x) \rightarrow [\neg[\neg x]] = x$

Proof. Apply $\mathsf{IS}(\Sigma_0^{(0,3)})$ with induction formula

$$(\forall u < x)(Form(u) \rightarrow [\neg[\neg u]] = u)$$

and induction variable x. □

We now turn to valuation of closed $\mathcal{L}(0, \mathrm{Q})$ terms.

Proposition 5.20. *Let x be variable and let t be a closed $\mathcal{L}(0, \mathrm{Q})$ term. Then $\mathbf{A}(0,3)$ proves the following.*

(1) $val(num(x)) = x$

(2) $val(\ulcorner t \urcorner) = t$

Proof. For (1), we have in $\mathbf{A}(0,3)$ that $val(num(\overline{0})) = \overline{0}$ and

$$val(num(x + \overline{1})) = val(num(x)) + \overline{1}.$$

Thus, we may apply $\mathsf{IS}(\Sigma_0^{(0,3)})$, where $val(num(x)) = x$ is the induction formula.

For (2), we perform a (meta) induction on the rank of t. We consider only two cases.

- t is $\overline{0}$. By Proposition 5.13 (2).
- t is $(t_0 + \overline{1})$, where t_0 is closed. By Definition 5.1 and logic, we have in $\mathbf{A}(0,3)$ that

  $$\ulcorner t_0 + \overline{1} \urcorner = \langle \ulcorner + \overline{1} \urcorner, \ulcorner t_0 \urcorner \rangle.$$

  By Proposition 5.13 (3), $\mathbf{A}(0,3)$ proves that

  $$val(\langle \ulcorner + \overline{1} \urcorner, \ulcorner t_0 \urcorner \rangle) = val(\ulcorner t_0 \urcorner) + \overline{1}.$$

  So, the result is obtained by the induction hypothesis that $\mathbf{A}(0,3) \vdash val(\ulcorner t_0 \urcorner) = t_0$.

Other cases are treated accordingly. □

We may now generalize 5.20 (2) as follows.

Proposition 5.21. $\mathbf{A}(0,3)$ *proves, for any arbitrary* $\mathcal{L}(0,\mathrm{Q})$ *term* $t(v_0,\ldots,v_k)$*, the following.*

$$\bigwedge\nolimits_{i=0}^{k} ClTerm(u_i) \to$$
$$val(\ulcorner t \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right]) = t\left(\frac{val(u_0),\ldots,val(u_k)}{v_0,\ldots,v_k}\right)$$

Proof. By (meta) induction on the rank of t. We consider only three cases to show the strategy.

- t is a variable v_i ($i \leq k$). Then, by Proposition 5.6 (3), we have in $\mathbf{A}(0,3)$ that
$$\ulcorner v_i \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right] = u_i$$
where, by (meta) definition, $val(u_i)$ is $v_i(val(u_i)/v_i)$.
- t is the constant $\overline{0}$. Like the last item, by observing that $val(\ulcorner \overline{0} \urcorner) = \overline{0}$ and
$$\ulcorner \overline{0} \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right] = \ulcorner \overline{0} \urcorner.$$
- t is $(t_0 + \overline{1})$. By Definition 5.1 and Proposition 5.13 (3), we have in $\mathbf{A}(0,3)$ that
$$val(\ulcorner t_0 + \overline{1} \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right])$$
is identical (=) to
$$val(\ulcorner t_0 \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right]) + \overline{1}.$$

So, it only remains to use the induction hypothesis and to note that, by (meta) definition,

$$(t_0 + \overline{1}) \left(\frac{val(u_0), \ldots, val(u_k)}{v_0, \ldots, v_k} \right)$$

is the term

$$t_0 \left(\frac{val(u_0), \ldots, val(u_k)}{v_0, \ldots, v_k} \right) + \overline{1}.$$

Other cases are treated accordingly. □

Observe that 5.21 is reminiscent of the *Kleene Normal Form Theorem* from recursion theory: By 5.15 and 5.20 (1), we obtain, specifically, the following corollary.

Corollary 5.22. $\mathbf{A}(0,3)$ *proves*

$$val(\ulcorner f \urcorner \left[\frac{num(v_0), \ldots, num(v_k)}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner} \right]) = f(v_0, \ldots, v_k)$$

where $f(v_0, \ldots, v_k)$ *may be any function symbol from* $\mathcal{L}(0, \mathrm{Q})$.

5.3 Generalizations

Thus far, we obtained in $\mathbf{A}(0,2)$ (or $\mathbf{A}(0,3)$) theorems about the syntax of language $\mathcal{L}(m, \mathrm{Q})$ and valuation of $\mathcal{L}(0, \mathrm{Q})$ terms. We conclude the present chapter by a few generalizations on how to get in $\mathbf{A}(0, k+1)$ respective theorems for the syntax of language $\mathcal{L}(m, k)$ and valuation of $\mathcal{L}(0, k)$ terms, where $k > 1$. We will require these generalizations later in Chapters 6 and 7.

The only difference between language $\mathcal{L}(m, \mathrm{Q})$ and languages $\mathcal{L}(m, k)$, where $k > 1$, is the built-up of terms by the additional function symbols. In Section 2.3, we already indicated how these additional function symbols may be encoded. Let us adapt our coding formally in $\mathbf{A}(0,2)$ or $\mathbf{A}(0,3)$, depending on which coding of numerals and which forms of bounded recursion are applied in the background.

Continuing from $\ulcorner +\overline{1} \urcorner, \ulcorner + \urcorner, \ldots, \ulcorner \forall \urcorner$, let us consider function symbols such that we may prove

$$\begin{aligned} \ulcorner - \urcorner &= \langle \overline{23} \rangle && \text{(modified subtraction)} \\ \ulcorner 0 \urcorner &= \langle \overline{25} \rangle && \text{(constant function)} \\ \ulcorner \mathrm{P} \urcorner &= \langle \overline{27} \rangle && \text{(projection)} \\ \ulcorner \mathrm{C} \urcorner &= \langle \overline{29} \rangle && \text{(composition)} \\ \ulcorner \mu \urcorner &= \langle \overline{31} \rangle && \text{(μ operation)} \end{aligned}$$

and, for truth predicates and Grzegorczyk function symbols:

$$\begin{aligned} \ulcorner T_i \urcorner &= \langle \overline{29 + 4 \cdot i} \rangle \\ \ulcorner \gamma_i \urcorner &= \langle \overline{31 + 4 \cdot i} \rangle \end{aligned}$$

As indicated in Section 2.3, but formally now, we may apply bounded recursion to derive, for appropriate formula $FSym_k(x)$ (function symbol), that

$$FSym_k(x) \leftrightarrow D(x),$$

where the disjunction $D(x)$ comprises, for example (and among other disjuncts), the following disjuncts.

- $x = \langle \ulcorner 0 \urcorner, \overline{1} \rangle$
- $(\exists v < x)(\exists u < v)\big(x = \langle \ulcorner \mathrm{P} \urcorner, u, v \rangle\big)$
- $(\exists u, v < x)(\exists w < u)\big(FSym_k(u) \wedge lh(u) = w + \overline{1} \wedge$ $Var(v) \wedge x = \langle \ulcorner \mu \urcorner, v, u, (u)_w \rangle\big)$

(Note: The last component of x indicates the arity. Furthermore, v in $\langle \ulcorner \mu \urcorner, v, u, (u)_w \rangle$ specifies the variable of the μ operation.)

Like in Definition 5.1, we accordingly have, for every $\mathcal{L}(0,k)$ function symbol f, a closed term $\ulcorner f \urcorner$ to formally denote f.

As for the definition of $Term_k(x)$ (term), we proceed like in Section 5.1 before to derive

$$Term_k(x) \leftrightarrow E(x),$$

where disjunction $E(x)$ reads as follows.

$$\begin{gathered} Var(x) \vee x = \ulcorner \overline{0} \urcorner \vee \\ (\exists u, v, w < x)\big(FSym_k(u) \wedge lh(u) = v + \overline{1} \wedge (u)_v = lh(w) \wedge \\ (\forall y < lh(w)) Term_k((w)_y) \wedge x = \langle u \rangle * w\big) \end{gathered}$$

Respective terms $\ulcorner t \urcorner$ to formally denote $\mathcal{L}(0,k)$ terms t and the respective extensions for $\mathcal{L}(m,k)$ formulas ($Form_{m,k}$) and for $\mathcal{L}(m,k)$ sentences ($Sent_{m,k}$) may then be defined similarly as in Section 5.1 before.

It only remains to observe that we have, by counterparts of Proposition 5.13, $\mathcal{L}(0,k+1)$ symbols $val_k(x)$ for the valuation of closed $\mathcal{L}(0,k)$ terms x in $\mathbf{A}(0,k+1)$. There are many ways to achieve these counterparts. We sketch an approach in accordance with several other elaborations in this thesis.

(Remember from Section 3.3 that, by diagonalization, $val_k(x)$ cannot be an $\mathcal{L}(0,k)$ term, provided $k > 1$. Indeed, for $k > 2$, consider the k-th Grzegorczyk function symbol $\gamma_k(x)$ and notice that, by iterated substitution

$$\begin{aligned} t_0 &:= \overline{1} \\ t_{n+1} &:= \gamma_k(t_n) \end{aligned}$$

the value of t_n is precisely $\gamma_{k+1}(n)$. Moreover, right after Definition 3.17, we have also indicated that iterated application of the smash function amounts to exponentiation. So, we can see that valuation of closed $\mathcal{L}(0,k)$ terms is an $\mathcal{L}(0,k+1)$ concept.)

Due to the closure conditions (remember 3.44 (A7)–(A10)) on function symbols, we first require some symbol $fval_k(x,y)$ with the subsequent intention: If x is (code of) some function symbol

($FSym_k$) f, and y is (code of) some sequence $(n_0, \ldots, n_r)$, and the arity of f is at most $r+1$, then $fval_k(x,y)$ is the value of $f(\overline{n_0}, \ldots, \overline{n_r})$. Once we have $fval_k(x,y)$ available, the valuation (val_k) of arbitrary closed terms may afterwards be obtained by bounded course-of-values recursion in an expected way: That is, the value of a closed term of form

$$(p, m_0, \ldots, m_r),$$

where p is (a code of) a function symbol ($FSym_k$) and m_i are (codes of) closed sub-terms, equals

$$fval_k(p, (val_k(m_0), \ldots, val_k(m_r))).$$

So, let us briefly focus on $fval_k(x,y)$.

To obtain $fval_k(x,y)$, we may first use a bounded course-of-values recursion to compute the construction tree of x, which we may compress into some list $fsub(x)$ (repetitions possible) of all sub-function symbols in x. For example, if x has the form of bounded μ operation $\langle \ulcorner \mu \urcorner, v, u, z \rangle$, the following would be derivable.

$$fsub(x) = fsub(u) * \langle x \rangle$$

If x has the form of a composition $\langle \ulcorner \mathrm{C} \urcorner, u, z \rangle$, where u is (code of) a sequence $f_0, \ldots, f_n, f_{n+1}$ of composed functions, we have by bounded recursion for a symbol $h(u,w)$ that

$$h(u,w) = \begin{cases} \langle\ \rangle & : w = \overline{0} \\ h(u, w - \overline{1}) * fsub((u)_{w-\overline{1}}) & : w > \overline{0} \end{cases}$$

and may derive

$$fsub(x) = h(u, lh(u)) * \langle x \rangle$$

to prolong the list.

The second to last step is to successively compute, for every single item in the list $fsub(x)$, a respective sequence of various intermediate valuation results, that is, a sequence that gets (by concatenation) successively prolonged and that we just call s for short. If (say) the j-th item of $fsub(x)$ is a code of g and g is neither a μ operation nor a composition, we prolong s (that we have obtained at step $j-1$) by adding, for various sequences

$$(n_0, \dots, n_l),$$

the respective value of

$$g(\overline{n_0}, \dots \overline{n_l}).$$

In particular, the sequences $(n_0, \dots, n_l)$ to be considered at step j are such that l is bounded by the maximum arity of items in $fsub(x)$ (so by x) and any n_i is bounded by the maximum of the numbers in (initial substitution sequence) y and the intermediate values in s (obtained at step $j-1$).

Importantly: To successively bound the n_i to be considered in $\mathbf{A}(0, k+1)$, we have the $(k+1)$-th Grzegorczyk function symbol (γ_{k+1}) available. (Compare Lemma 6.16, where such bounds on values of nested terms are proved in more formal detail. We omit details here.)

Thus, if the j-th item of $fsub(x)$ codes a composition

$$f_{n+1}(f_0(v_0, \dots, v_r), \dots, f_n(v_0, \dots v_r))$$

and y codes a sequence $(n_0, \dots, n_r)$, then s at step $j-1$ already includes the value of

$$f_{n+1}(f_0(\overline{n_0}, \dots, \overline{n_r}), \dots, f_n(\overline{n_0}, \dots \overline{n_r}))$$

since the code of f_{n+1} is the $(j-1)$-th item in $fsub(x)$ and the codes of $f_0, \dots, f_n$ come before the $(j-1)$-th item in $fsub(x)$.

Similarly, if the j-th item of $fsub(x)$ codes a μ operation

$$(\mu u < v_r)(f_0(v_0, \dots, v_{r-1}, v_r) = 0)$$

and y codes a sequence $(n_0, \ldots, n_r)$, then s at step $j-1$ already includes the values of

$$\begin{gathered} f_0(\overline{n_0}, \ldots, \overline{n_{r-1}}, \overline{0}) \\ \vdots \\ f_0(\overline{n_0}, \ldots, \overline{n_{r-1}}, \overline{n_r}) \end{gathered}$$

since the code of f_0 is the $(j-1)$-th item in $fsub(x)$.

By its construction, s at the final step comprises the value of $f(\overline{n_0}, \ldots, \overline{n_r})$. So, the computation of $fval_k(x, y)$ is concluded by extracting this value from s.

Part II

Finitist Truth

6 Finitist Heck

We now turn to the proof-theoretic analysis of theories of finitist truth. Since in this and in the next chapter, we will be concerned exclusively with axiomatizations of just *one single* truth predicate $T_1(x)$, we let

$$T(x) := T_1(x)$$

to simplify the notation.

As indicated in Section 1.4, already the atomic instances

$$T(t) \leftrightarrow T(\ulcorner T(t) \urcorner), \tag{6.1}$$

of the Tarski Schema, where t is some arbitrary closed term, may lead into contradictions if they are carelessly conjoined with other principles of truth. Specifically, in conjunction with the classical negation principle that

$$\neg T(\ulcorner A \urcorner) \leftrightarrow T(\ulcorner \neg A \urcorner), \tag{6.2}$$

where A is some arbitrary sentence, they yield

$$\neg T(t) \leftrightarrow T(\ulcorner \neg T(t) \urcorner). \tag{6.3}$$

So, if for some closed term t, we may prove $\ulcorner \neg T(t) \urcorner = t$, we are landed in a contradiction.

More formally, and to make the target language of the truth axiom visible, we may put (6.1) in the form of a single axiom as

$$T([T(x)]) \leftrightarrow T(val(x)), \tag{6.4}$$

where x ranges over codes of closed terms. Accordingly, we may put (6.2) in the form of a single axiom as

$$T([\neg x]) \leftrightarrow \neg T(x), \tag{6.5}$$

where x ranges over codes of sentences. By the same thought as before, an inconsistency arises if the target language of the pair of truth axioms comprises a closed term t such that $val(\ulcorner t \urcorner) = [\sim T(\ulcorner t \urcorner)]$ is provable.

In the present chapter, we will consider the case in which the target language of the axioms of truth does *not* comprise such a problematic closed term. (The other case will be studied in the following chapter.) Specifically, to show that a given language $\mathcal{L}$ does not comprise a closed term (encoded as) x so that $val(x) = [\sim T(x)]$, it suffices to prove that, for every closed term (encoded as) x, we have

$$val(x) < x. \tag{6.6}$$

For, by (6.6) and $x < [\sim T(x)]$, we get $val(x) < [\sim T(x)]$. Indeed, (6.6) is satisfied by our natural coding of syntax from Chapter 5.

That axioms (6.4) and (6.5) may be consistently conjoined in the absence of certain terms has been already shown – as Volker Halbach kindly brought to our attention – by Heck [24]. Here, however, we are interested not only in the axioms (6.4) and (6.5), but in a consistent theory of *naive type-free* truth for an entire (sub-)finitist fragment more generally.

Specifically, the aim of this chapter is to present a consistent theory of naive type-free truth for the (decidable, bounded) Σ_0 fragment of the language $\mathcal{L} = \mathcal{L}(1,1)$. It will be called the theory $\mathbf{FH}(1,1)$ of Finitist Heck Truth for $\mathcal{L}(1,1)$ in dependence on the work by Heck [24].

As a further refinement of Heck's results in [24], we will also show that $\mathcal{L}(1,1)$ indeed marks a bound for naive type-free truth in the current setting, in the sense that the corresponding theory of naive type-free truth for the Σ_0 fragment of $\mathcal{L}(1,2)$ would be inconsistent (see Proposition 6.6). A boundary will be drawn in Sections 6.1 and 6.2, where we show that a common diagonalization argument undermines the attempt to conjoin (6.4) and (6.5) consistently in cases where the axioms are about (the codes of) $\mathcal{L}(1,k)$ expressions with $k > 1$.

The theory $\mathbf{FH}(1,1)$ will be defined in Section 6.2. In the remainder of the present chapter, it will be shown that $\mathbf{FH}(1,1)$ is consistent. (Since, for $k > 2$, truth axioms of $\mathbf{FH}(1,1)$ do not increase the number of arithmetical theorems of their base $\mathbf{A}(1,k)$ anyways, we put our focus primarily on consistency in the current chapter.)

Later in Chapter 7, we will then put our focus on theories of Decidable Arithmetical Truth for a Σ_0 fragment of $\mathcal{L}(0,k)$, where $k > 1$ and – according to our findings in Sections 6.1 and 6.2 – naivety has to be overcome by postulating certain truth *rules* in place of the axiom (6.4) to obtain consistent extensions.

A major interest of $\mathbf{FH}(1,1)$ relies in its *naivety* as a theory which entails the *untyped* Tarski Schema for a whole sub-finitist (Σ_0) fragment.

6.1 Self-Reference

To unveil contradictions that may be raised by (6.4) and (6.5) in more detail, we use the common diagonalization method. Let us start off with generalizing a few points from Section 5.2.

For $k > 1$, let $val_k(x)$ be $\mathcal{L}(0,k+1)$ symbol for the valuation of closed $\mathcal{L}(0,k)$ terms, indicated in Section 5.3. We obtain the following propositions, where $ClTerm_k(x)$ represents that x is a closed $\mathcal{L}(0,k)$ term.

Proposition 6.1. *Let* $k > 1$. *Then* $\mathbf{A}(0, k+1)$ *proves, for each* $\mathcal{L}(0,k)$ *term* $t(v_0, \ldots, v_n)$, *the following.*

$$\bigwedge\nolimits_{i=0}^{n} ClTerm_k(u_i) \rightarrow$$
$$val_k(\ulcorner t \urcorner \left[\frac{u_0, \ldots, u_n}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_n \urcorner}\right]) = t\left(\frac{val_k(u_0), \ldots, val_k(u_n)}{v_0, \ldots, v_n}\right)$$

Proof. Like the proof of Proposition 5.21. ☐

As a corollary, we obtain a typed version of (Kleene) normal form, as follows.

Corollary 6.2. *Let* $k > 1$. *Then* $\mathbf{A}(0, k+1)$ *proves, for all the function symbols* $f(v_0, \ldots, v_n)$ *from* $\mathcal{L}(0,k)$, *the following.*

$$val_k(\ulcorner f \urcorner \left[\frac{num(v_0), \ldots, num(v_n)}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_n \urcorner}\right]) = f(v_0, \ldots, v_n)$$

Proof. Like the proof of Corollary 5.22. ☐

Self-reference via terms, as we have called it in the informal Section 1.4, may now be achieved by diagonalization as follows.

Proposition 6.3. $\mathbf{A}(0,2)$ *proves, for some* $\mathcal{L}(0,2)$ *term* t, *the following.*

$$t = [\sim T(\ulcorner t \urcorner)]$$

Proof. We have seen in Chapter 5 that substitution (into terms) and numerals are definable in $\mathcal{L}(0,2)$. Thus, we have an $\mathcal{L}(0,2)$ function symbol $d(y)$ (diagonalization) such that $\mathbf{A}(0,2)$ proves:

$$d(y) = \ulcorner \sim T(x) \urcorner [y[num(y)/\ulcorner y \urcorner]/\ulcorner x \urcorner]$$

Furthermore, for a closed $\mathcal{L}(0,2)$ term $\ulcorner d(y) \urcorner$ and a number n, we have $\ulcorner d(y) \urcorner = \overline{n}$. Therefore,

$$d(\overline{n}) = \ulcorner \sim T(x) \urcorner [\overline{n}[num(\overline{n})/\ulcorner y \urcorner]/\ulcorner x \urcorner]$$

and, because $\overline{n}[num(\overline{n})/\ulcorner y\urcorner] = \ulcorner d(\overline{n})\urcorner$, also:

$$d(\overline{n}) = \ulcorner{\sim}T(x)\urcorner[\ulcorner d(\overline{n})\urcorner/\ulcorner x\urcorner]$$

That is, $d(\overline{n}) = [{\sim}T(\ulcorner d(\overline{n})\urcorner)]$ is provable in $\mathbf{A}(0,2)$. $\square$

Informally, t from Proposition 6.3 is a term "which denotes – that is: has as its value – the (code of the) statement that t itself is untrue". Indeed, a formal specification of the informal idea is achieved by Proposition 6.1, according to which $\mathbf{A}(0,3)$ proves $val_2(\ulcorner t\urcorner) = [{\sim}T(\ulcorner t\urcorner)]$.

It may be illustrative to also mention that the diagonalization proof of Proposition 6.3 – computing systematically a code of a self-referential term t from a code of diagonalization function d – can be internalized in $\mathbf{A}(0,4)$ by using formal (typed) versions of Kleene's Second Recursion Theorem.

Proposition 6.4. *Let $k > 1$. Then $\mathcal{L}(0,2)$ comprises a function symbol $\rho_k(x)$ so that $\mathbf{A}(0,k+2)$ proves, for each function symbol $f(v)$ from $\mathcal{L}(0,k)$, the following.*

$$val_{k+1}(\rho_k(\ulcorner f\urcorner)) = f(\rho_k(\ulcorner f\urcorner))$$

Proof. We work in theory $\mathbf{A}(0,k+2)$. To start with, we define $r(x,y)$ for fixed variables x, y as

$$val_k\big(x\left[\frac{num(y[num(x),num(y)/\ulcorner x\urcorner,\ulcorner y\urcorner])}{\ulcorner y\urcorner}\right]\big)$$

and $\rho_k(x)$ correspondingly as

$$\ulcorner r\urcorner\left[\frac{num(x),num(\ulcorner r\urcorner)}{\ulcorner x\urcorner,\ulcorner y\urcorner}\right].$$

By definition, we have, for every $f(v)$ from $\mathcal{L}(0,k)$, that

$$val_{k+1}(\rho_k(\ulcorner f\urcorner)) = val_{k+1}\big(\ulcorner r\urcorner\left[\frac{num(\ulcorner f\urcorner),num(\ulcorner r\urcorner)}{\ulcorner x\urcorner,\ulcorner y\urcorner}\right]\big)$$

By Corollary 6.2:

$$val_{k+1}(\rho_k(\ulcorner f \urcorner)) = r(\ulcorner f \urcorner, \ulcorner r \urcorner)$$

By definition again:

$$val_{k+1}(\rho_k(\ulcorner f \urcorner)) = val_k\big(\ulcorner f \urcorner \left[\frac{num(\rho_k(\ulcorner f \urcorner))}{\ulcorner y \urcorner}\right]\big)$$

By Corollary 6.2:

$$val_{k+1}(\rho_k(\ulcorner f \urcorner)) = f(\rho_k(\ulcorner f \urcorner))$$

This concludes the proof. □

Hence, for some $\mathcal{L}(0,2)$ function symbol $f(v) := [\sim T(v)]$, we have by Proposition 6.4 an $\mathcal{L}(0,2)$ function symbol $\rho_2(x)$ so that $\mathbf{A}(1,4)$ proves:

$$val_3(\rho_2(\ulcorner f \urcorner)) = [\sim T(\rho_2(\ulcorner f \urcorner))]$$

Remark 6.5. Note that the line of reasoning in the above proof is really internalized and would also yield in $\mathbf{A}(0,4)$ that, for any (possibly non-standard) $\mathcal{L}(0,2)$ term x with exactly one free variable,

$$val_3(\rho_2(x)) = val_2(x[\langle num(\rho_2(x))\rangle])$$

where $x[\langle num(\rho_2(x))\rangle]$ is the result of substituting the only free variable of x by the $\rho_2(x)$-th numeral.

Moreover, to extend Proposition 6.4 to $k = 1$, we require the valuation of closed $\mathcal{L}(0,1)$ terms by means of the smash function, as was indicated in Remark 5.14.

6.2 The Theory

We may now make use of Proposition 6.3 to show more precisely why, for every smashed Grzegorczyk level $k > 1$ (or usual level $k > 2$), (6.4) may not be conjoined with (6.5) consistently.

We use $Sent_{m,k}(x)$ to represent that x is an $\mathcal{L}(m,k)$ sentence and obtain the following refinement of an observation by Halbach [19] as formal variant of the Liar Paradox.

Proposition 6.6. *For $k > 1$, $\mathbf{A}(1,k+1)$ is inconsistent with the conjunction of the following.*

$$Sent_{1,k}(x) \to \big(T([\neg x]) \leftrightarrow \neg T(x)\big) \tag{6.7}$$

$$Sent_{1,k}([T(x)]) \to \big(T([T(x)]) \leftrightarrow T(val_k(x))\big). \tag{6.8}$$

Proof. By Proposition 6.1 and Proposition 6.3, we have a closed $\mathcal{L}(0,2)$ term t such that $\mathbf{A}(1,3)$ proves the following.

$$val_2(\ulcorner t \urcorner) = [{\sim} T(\ulcorner t \urcorner)] \tag{6.9}$$

By (6.7), we have:

$$T([{\sim} T(\ulcorner t \urcorner)]) \leftrightarrow {\sim} T([T(\ulcorner t \urcorner)])$$

By (6.8), we have:

$$T([{\sim} T(\ulcorner t \urcorner)]) \leftrightarrow {\sim} T(val_2(\ulcorner t \urcorner))$$

By (6.9), we have

$$T([{\sim} T(\ulcorner t \urcorner)]) \leftrightarrow {\sim} T([{\sim} T(\ulcorner t \urcorner)])$$

and, therefore, we are landed in a contradiction. □

Remark 6.7. By Corollary 3.64 (2), we immediately obtain that, for $k > 1$, the base theory $\mathbf{B}(1,k+1)$ is not consistent with the respective $\mathcal{A}(k+1)$ interpretations of (6.7) and (6.8).

Proposition 6.6 does not generally undermine any attempt to conjoin (6.7) and (6.8) consistently: In particular, we may consistently conjoin them if their range is limited to (codes of) $\mathcal{L}(1,1)$ sentences. Indeed, we will present a consistent theory of Finitist Heck Truth for the Σ_0 fragment of $\mathcal{L}(1,1)$ shortly.

As indicated in Section 1.5, finitist conceptions of truth may be interpreted by respective conceptions of verifiability by finite computations. (In the present chapter, these verifications will be accomplished by computing values of characteristic closed terms.) If $T(x)$ is taken to mean that x is verified by some finite computation (sequence), the presence of the negation axiom gives rise to conceptions of *decidable truth*: By the negation axiom

$$T(\ulcorner \neg A \urcorner) \leftrightarrow \neg T(\ulcorner A \urcorner)$$

we have $T(\ulcorner A \urcorner) \vee T(\ulcorner \neg A \urcorner)$, that is, that either x or the negation of x may be verified by finitely many computation steps.

While decidable problems of arithmetic may be seen to have complexity Δ_1 in theories of arithmetic, the following axioms of Finitist Heck Truth are about the truth of $\Delta_0^{(1,2)}$ sentences in a language with a truth predicate.

Definition 6.8. $\mathsf{FH}(1,1)$ comprises the subsequent axioms of *Finitist Heck Truth* for the Σ_0 fragment of $\mathcal{L}(1,1)$.

(FH1) $Sent_{1,1}([x = y]) \rightarrow$
$\big(T([x = y]) \leftrightarrow val_1(x) = val_1(y)\big)$

(FH2) $Sent_{1,1}([x \leq y]) \rightarrow$
$\big(T([x \leq y]) \leftrightarrow val_1(x) \leq val_1(y)\big)$

(FH3) $Sent_{1,1}([T(x)]) \rightarrow$
$\big(T([T(x)]) \leftrightarrow T(val_1(x))\big)$

(FH4) $Sent_{1,1}(x) \rightarrow$
$\big(T([\neg x]) \leftrightarrow \neg T(x)\big)$

(FH5) $Sent_{1,1}([x \vee y]) \rightarrow$
$\big(T([x \vee y]) \leftrightarrow (T(x) \vee T(y))\big)$

(FH6) $Sent_{1,1}([(\exists x \leq y)z]) \rightarrow$
$\big(T([(\exists x \leq y)z]) \leftrightarrow (\exists u \leq val_1(y))T(z[num(u)/x])\big)$

By applying (FH4), we obtain the following.

Proposition 6.9. $\mathbf{A}(1,3) \cup \mathsf{FH}(1,1)$ *proves the following.*

(1) $Sent_{1,1}([x \neq y]) \rightarrow$
$\big(T([x \neq y]) \leftrightarrow val_1(x) \neq val_1(y)\big)$

(2) $Sent_{1,1}([x > y]) \rightarrow$
$\big(T([x > y]) \leftrightarrow val_1(x) > val_1(y)\big)$

(3) $Sent_{1,1}([\sim T(x)]) \rightarrow$
$\big(T([\sim T(x)]) \leftrightarrow \sim T(val_1(x))\big)$

(4) $Sent_{1,1}([x \wedge y]) \rightarrow$
$\big(T([x \wedge y]) \leftrightarrow (T(x) \wedge T(y))\big)$

(5) $Sent_{1,1}([(\forall x \leq y)z]) \rightarrow$
$\big(T([(\forall x \leq y)z]) \leftrightarrow (\forall u \leq val_1(y))T(z[num(u)/x])\big)$

Proof. Let us consider (3) to show the strategy. Assuming that $Sent_{1,1}([\sim T(x)])$, we have by (FH3) and logic that

$$\sim T([T(x)]) \leftrightarrow \sim T(val_1(x)).$$

Hence, an application of (FH4) yields

$$T([\sim T(x)]) \leftrightarrow \sim T(val_1(x)).$$

Similarly for the other cases. □

As we are interested exclusively in *arithmetical* truth in this study, we shall equip theories of truth accordingly with an axiom of Truth Induction, expressing basically that truth conforms with the principle of induction for the natural numbers.

Definition 6.10. $\mathsf{TI}(1,1)$ contains the following axiom of *Truth Induction.*

$$\begin{aligned}&Sent_{1,1}([(\forall z)x]) \to \big(\big(T(x[num(\overline{0})/z]) \land\\&\forall y(T(x[num(y)/z]) \to T(x[num(y+\overline{1})/z]))\big) \to\\&T(x[num(y)/z]))\big)\end{aligned}$$

Finally, we define theory $\mathbf{FH}(1,1)$ of Finitist Heck Truth by means of the auxiliary function symbols as follows. (But see Remark 6.13.)

Definition 6.11. $\mathbf{FH}(1,1)$ is the *Theory of Finitist Heck Truth* (in auxiliary form) and defined as follows.

$$\mathbf{A}(1,3) \cup \mathsf{FH}(1,1) \cup \mathsf{TI}(1,1)$$

Remark 6.12. In view of Remark 5.14, we may also have based $\mathbf{FH}(1,1)$ on the theory $\mathbf{A}(1,2)$, but as mentioned repeatedly, we will refrain from doing so and let all theories of truth include at least elementary arithmetic for convenience.

Remark 6.13. We reserve the label $\mathbf{FH}(1,1)$ for a theory in a language with auxiliary function symbols because, practically, we will more often have to refer to this theory in auxiliary form in what follows. However, with regard to intuitions behind *type-free truth* that we briefly addressed in Section 3.3, keep in mind that we always have the $\mathcal{L}(1,\mathrm{Q})$ variant

$$\mathbf{B}(1,3) \cup \big(\mathsf{FH}(1,1) \cup \mathsf{TI}(1,1)\big)^{\mathcal{A}}$$

in the background, where $\mathcal{A}$ is the mapping from Definition 3.62 for eliminating auxiliary function symbols.

With regard to type-free (as more natural) truth, the important point is that both the universal closures of truth axioms in $\mathsf{FH}(1,1) \cup \mathsf{TI}(1,1)$ and the sentences from the Σ_0 fragment of the language $\mathcal{L}(1,1)$ as targets of these axioms are $\Pi_2^{(1,\mathrm{Q})}$ in $\mathbf{A}(1,3)$. Hence, in a sense, they are unified in the Π_2 fragment of $\mathcal{L}(1,\mathrm{Q})$.

Similar remarks apply also to all the other theories of truth which will be studied in later chapters.

Our major goal in this chapter is to show the consistency of $\mathbf{FH}(1,1)$. This will be achieved by retracting $\mathbf{FH}(1,1)$ to base $\mathbf{A}(0,3)$. (As an immediate consequence, this will yield that, for $k > 2$ in general,

$$\mathbf{A}(1,k) \cup \mathsf{FH}(1,1) \cup \mathsf{TI}(1,1)$$

is retractable to $\mathbf{A}(0,k)$.) Then, by Proposition 2.27 (2),

$$\mathbf{A}(0,3) \nvdash \overline{0} \neq \overline{0}$$

implies

$$\mathbf{FH}(1,1) \nvdash \overline{0} \neq \overline{0}$$

and, by Corollary 3.54 (2), the latter implies

$$\mathbf{B}(1,3) \cup \big(\mathsf{FH}(1,1) \cup \mathsf{TI}(1,1)\big)^{\mathcal{A}} \nvdash \overline{0} \neq \overline{0},$$

showing the consistency of both Finitist Heck variants.

6.3 Retractability: Valuation

To a large extend, the retraction of $\mathbf{FH}(1,1)$ to $\mathbf{A}(0,3)$ is by formalizing the proof of Proposition 4.3: To any Σ_0 formula in the language $\mathcal{L}(1,1)$, we assign some $\mathcal{L}(0,1)$ term whose value is 1 if the formula is true, and 0 else. For interpreting (FH4), we need to prove by formal induction that, for any sentence A, the term assigned to A has value 0 if and only if the term assigned to $\neg A$ has value 1.

(This is a downside of our initial choice of defining negation as non-primitive; the upside will become clear once we consider computations in the form of finite positive operations.)

We require several preliminaries for interpreting the truth axioms of $\mathbf{FH}(1,1)$: in particular, some preliminaries that allow us to derive the interpretation of negation axiom (FH4) by a formal induction. As we have just stated, the required property is: For every sentence A, the term assigned to A has value 0 precisely if the term assigned to $\neg A$ has value 1.

The difficulty is with bounded quantifications. For example, suppose that we want to derive the required property for a term assigned to $(\forall x \leq t)B$. By the limitation of $\mathsf{IS}(\Sigma_0^{(0,3)})$, we cannot apply a universally quantified induction hypothesis that, for every number n, the required property holds for the term assigned to $B(\overline{n}/x)$. So, we have to iteratively bound the numbers n to be considered according to the rank of the formula B.

The first step is to define a term complexity measure (tcm) on the basis of a complexity measure (fcm) for function symbols occurring in these terms. Both measures are defined in a way to match specifically the proof of Lemma 6.16 below. (For instance, in that proof, we will require that the bound on

$$(\mu v < z)(f_0(\vec{x}, v) = \overline{0})$$

does not depend on f_0, but only on z.)

Notice: In what follows, we apply the extended coding from Section 5.3 for the $\mathcal{L}(1,k)$ syntax with $k > 1$. To simplify our notation, however, we shall not endow tcm and fcm with indices k. These indices are sufficiently clear by context.

Definition 6.14. For every $k > 1$ and every function symbol f from $\mathcal{L}(0,k)$, $fcm(f)$ is defined recursively as follows.

- $fcm(f) := 1$ if f is not composed (as in 3.44 (A9)) and not a bounded μ operation (as in 3.44 (A10)).
- $fcm(f) := \max\{fcm(f_i) : i \leq m\} + fcm(f_{m+1})$ if f is some composition
$$f_{m+1}(f_0(\vec{x}), \ldots, f_m(\vec{x}))$$
as in 3.44 (A9).
- $fcm(f) := fcm(f_0)$ if f is a bounded μ operation
$$(\mu v < z)(f_0(\vec{x}, v) = \overline{0})$$
as in 3.44 (A10).

On the basis of fcm, we define the following.

Definition 6.15. For every $k > 1$ and $\mathcal{L}(0,k)$ term t, $tcm(t)$ is defined recursively as follows.

- $tcm(t) := 0$ if t is a variable or constant $\overline{0}$.
- $tcm(t) := \max\{tcm(t_i) : i \leq n\} + fcm(f)$ if t has form
$$f(t_0, \ldots, t_n).$$

Definition 6.14 and Definition 6.15 have just been stated on the meta level. But notice that, as formal counterparts by Corollary 4.17, we get symbols $fcm(x)$ and $tcm(x)$ in $\mathcal{L}(0,3)$ for the functions (on codes) defined in 6.14 and 6.15, respectively.

We need several further tools that we only sketch here. First, we have in $\mathcal{L}(0,3)$ a function symbol $frvar(x)$ such that, if x is a term or formula, then $frvar(x)$ is a specific sequence, without repetitions, of all free variables involved in x. Similarly, we have in $\mathcal{L}(0,3)$ some function symbol $subtm(x)$ so that, if x is a term,

$subtm(x)$ is a sequence (repetitions possible) of all sub-terms of x. For formulas x, let $subfm(x)$ denote the respective sequence of sub-formulas of x. Finally, we use the binary symbol $x[y]$ for substitution with internal arity, indicated in 5.7: If x is term or formula and y is a sequence of terms, $x[y]$ is the term or formula obtained by replacing, for every $u < lh(y)$, the term $(y)_u$ for the free variable $(frvar(x))_u$ in x (provided $u < lh(frvar(x))$).

(Again, to prevent eyestrain, we do not endow these symbols always with indices k. The indices are clear by context.)

We apply Proposition 4.15 to code in $\mathbf{A}(0,3)$, for each given number n, sequences $(\overline{n}, \overline{n}, \ldots, \overline{n})$ of numerals. That is, we use $numseq(x,y)$ as $\mathcal{L}(0,3)$ function symbol for which the following may be derived in $\mathbf{A}(0,3)$.

$$numseq(x,y) = \begin{cases} \langle\ \rangle & : y = \overline{0} \\ numseq(x, y - \overline{1}) * \langle num(x) \rangle & : y > \overline{0} \end{cases}$$

Moreover, we require nested applications of Grzegorczyk function symbols (γ_i), as in the proof of Lemma 3.48, but formally now: For $k > 1$, $\varphi_{k+1}(x,y)$ is an $\mathcal{L}(0,k+1)$ function symbol so that $\mathbf{A}(0,k+1)$ proves (for $k = 2$)

$$\varphi_3(x,y) = \begin{cases} x & : y = \overline{0} \\ \overline{2} \cdot \varphi_3(x, y - \overline{1}) & : y > \overline{0} \end{cases}$$

and (for $k > 2$) the following.

$$\varphi_{k+1}(x,y) = \begin{cases} x & : y = \overline{0} \\ \gamma_k(\varphi_{k+1}(x, y - \overline{1})) & : y > \overline{0} \end{cases}$$

(Recall (A13) from Section 3.10: $\gamma_k(y + \overline{1}) = \gamma_{k-1}(\gamma_k(y))$.)

We first prove a lemma that, if all the numerals in a given sequence s have their own value bounded by $\varphi_{k+1}(4,i)$, then the value of substitution $\ulcorner t \urcorner[s]$ is bounded by $\varphi_{k+1}(4, i + tcm(\ulcorner t \urcorner))$.

Remember: We have to iteratively put a bound on the values of all members of s because Π_1 induction is not available.

Lemma 6.16. *For $k > 1$, $\mathbf{A}(0, k+1)$ proves the following.*

$$\begin{aligned}&(\forall y \leq numseq(\varphi_{k+1}(\overline{4}, w), z))\\ &\big(\big(Seq(y) \land lh(y) \leq z \land\\ &(\forall u \in y)(\exists v \leq \varphi_{k+1}(\overline{4}, w))(u = num(v))\big) \to\\ &val_k(x[y]) \leq \varphi_{k+1}(\overline{4}, w + tcm(x))\big)\end{aligned}$$

Proof. We reason in $\mathbf{A}(0, k+1)$, but use suggestive notation to improve readability. We use $\mathsf{IS}(\Sigma_0^{(0,k+1)})$ to derive that, for every $a \in subtm(x)$ and every sequence y which consists of at most z many numeral codes among

$$num(0), \ldots, num(\varphi_{k+1}(4, w)) \tag{6.10}$$

we have that

$$val_k(a[y]) \leq \varphi_{k+1}(\varphi_{k+1}(4, w), tcm(a)). \tag{6.11}$$

In particular, if $a = x$, this will finally yield the lemma because clearly

$$\varphi_{k+1}(\varphi_{k+1}(4, w), tcm(a)) = \varphi_{k+1}(4, w + tcm(a)).$$

Since $val_k(a[y]) = 0$ if $a[y]$ is not a closed term, we may presume that $a[y]$ is some closed term. We consider only two cases (two specific function symbols) of the induction step and add a more general case afterwards to show the strategy. (Note: The line of reasoning is actually similar as for valuation in Section 5.3.)

- If a is code of a term

  $$\sigma_2(t_0(v_0, \ldots, v_r), t_1(v_0, \ldots, v_r))$$

  and b_0, b_1 are codes of t_0, t_1, respectively, we must show, for any code p of a sequence $(\overline{p_0}, \ldots, \overline{p_r})$ with $r < z$ and

  $$p_i \leq \varphi_{k+1}(4, w),$$

that the value (val_k) of

$$\sigma_2(t_0(\overline{p_0}, \ldots, \overline{p_r}), t_1(\overline{p_0}, \ldots, \overline{p_r})) \tag{6.12}$$

is bounded by $\varphi_{k+1}(\varphi_{k+1}(4, w), tcm(a))$. Notice that, by induction hypothesis, the value of $t_j(\overline{p_0}, \ldots, \overline{p_r})$ ($j \leq 1$) is yet bounded by $\varphi_{k+1}(\varphi_{k+1}(4, w), tcm(b_j))$. Furthermore, we also have by $tcm(b_j) > 0$ that

$$|\varphi_{k+1}(\varphi_{k+1}(4, w), tcm(b_j))| > 3$$

and, for every $|m| > 3$, the following inequalities.

$$2^{(|m|^2)} \leq 2^{(2^{|m|})} \leq 2^{2 \cdot m}$$

Thus, since $tcm(a) = \max_{j \leq 1}(tcm(b_j)) + 1$, as required, the value of (6.12) is bounded by $\varphi_{k+1}(\varphi_{k+1}(4, w), tcm(a))$.

- If a is code of a term $\gamma_k(t(v_0, \ldots, v_r))$ ($k > 2$) and b is code of t, we must establish, for every code p of some sequence $(\overline{p_0}, \ldots, \overline{p_r})$ with $r < z$ and

  $$p_i \leq \varphi_{k+1}(4, w),$$

  that the value (val_k) of

  $$\gamma_k(t(\overline{p_0}, \ldots, \overline{p_r})) \tag{6.13}$$

  can be bounded by $\varphi_{k+1}(\varphi_{k+1}(4, w), tcm(a))$. By induction hypothesis, the value of $t(\overline{p_0}, \ldots, \overline{p_r})$ is already bounded by $\varphi_{k+1}(\varphi_{k+1}(4, w), tcm(b))$. Therefore, the value of (6.13) can be bounded by

  $$\gamma_k(\varphi_{k+1}(\varphi_{k+1}(4, w), tcm(b)))$$

  and, as $tcm(a) > tcm(b)$ by Definition 6.15 above, also by $\varphi_{k+1}(\varphi_{k+1}(4, w), tcm(a))$.

- More generally, if a is code of
$$f(t_0(v_0,\dots,v_r),\dots,t_n(v_0,\dots,v_r)),$$
where f is an $(n+1)$-ary function symbol, consider again a code p of a sequence $(\overline{p_0},\dots,\overline{p_r})$ with $r<z$ and
$$p_i \leq \varphi_{k+1}(4,w).$$
For $j\leq n$, let b_j be some code of t_j. By the induction hypothesis, we have, for any $j\leq n$, that the value (val_k) of
$$t_j(\overline{p_0},\dots,\overline{p_r})$$
can be bounded by $\varphi_{k+1}(\varphi_{k+1}(4,w),tcm(b_j))$. Furthermore, by a sub-induction on the rank of *function symbols*, along exactly the same lines as the actual main induction on the rank of *terms* (see Remark 6.17), we derive an analogue of Lemma 6.16 for terms x of form $f(v_0,\dots,v_n)$, where any v_i is a variable. This gives us that, for every code q of some sequence $(\overline{q_0},\dots,\overline{q_n})$ with
$$q_j \leq \varphi_{k+1}\big(4,w+\max_{j\leq n}(tcm(b_j))\big),$$
the value (val_k) of
$$f(\overline{q_0},\dots,\overline{q_n})$$
is bounded by
$$\varphi_{k+1}\big(\varphi_{k+1}(4,w),\max_{j\leq n}(tcm(b_j))+fcm(c)\big),$$
where c encodes f. As $tcm(a)=\max_{j\leq n}(tcm(b_j))+fcm(c)$, as required, the value of
$$f(t_0(\overline{p_0},\dots,\overline{p_r}),\dots,t_n(\overline{p_0},\dots,\overline{p_r}))$$
is bounded by $\varphi_{k+1}(\varphi_{k+1}(4,w),tcm(a))$.

This conclues the proof. ☐

Remark 6.17. As a technical aside, let us briefly indicate how composed function symbols (as in 3.44 (A9)) are treated in the sub-induction above. First, let us reformulate the sub-induction statement (where a now is sub-function of given function x) by replacing the range (6.10) of considered numerals by

$$num(0), \ldots, num(\varphi_{k+1}(4, v)),$$

where we now allow $v \leq w + fcm(x) - fcm(a)$, and by replacing (6.11) accordingly by

$$val_k(a[y]) \leq \varphi_{k+1}(\varphi_{k+1}(4, v), fcm(a)).$$

(Since we are now considering only sub-terms $t(v_0, \ldots, v_r)$ of the form $f(v_0, \ldots, v_r)$, we have $tcm(t) = fcm(f)$. Hence, we do not need to distinguish between t and f and may simply speak of $fcm(t)$ for the current purpose of illustration.)

If a is a code of a composition

$$f_{m+1}(f_0(v_0, \ldots, v_r), \ldots, f_m(v_0, \ldots, v_r)),$$

where $b_0, \ldots, b_m, b_{m+1}$ are codes of the functions $f_0, \ldots, f_m, f_{m+1}$, respectively, and we are considering a sequence $(\overline{p_0}, \ldots, \overline{p_r})$ with (let us say)

$$p_i \leq \varphi_{k+1}(4, w + fcm(x) - fcm(a)),$$

then (as $fcm(a) > fcm(b_j)$ for any $j \leq m$) we get by induction hypothesis that the value (val_k) of

$$f_j(\overline{p_0}, \ldots, \overline{p_r})$$

is bounded by q_j given as

$$\varphi_{k+1}\big(\varphi_{k+1}\big(4, w + fcm(x) - fcm(a)\big), fcm(b_j)\big)$$

and, therefore, is bounded also by the following.

$$\varphi_{k+1}\big(4, w + fcm(x) - \big(fcm(a) - \max_{j \leq m}(fcm(b_j))\big)\big)$$

As $fcm(b_{m+1}) = fcm(a) - \max_{j \leq m}(fcm(b_j))$, we may employ the induction hypothesis again to obtain that the value (val_k) of

$$f_{m+1}(\overline{q_0}, \ldots, \overline{q_m})$$

is bounded by $\varphi_{k+1}(4, w + fcm(x))$, which is what we wanted to derive.

In what follows, we make further use of the extended coding from Section 5.3. Specifically, we will need to code the $\mathcal{L}(0,1)$ symbol $\min(x, y)$ for which $\mathbf{A}(0,3)$ proves

$$\min(x, y) = f_{\leq}(x, y) \cdot x + f_{>}(x, y) \cdot y$$

and (see also below) $f_{\leq}$ and $f_{>}$ are characteristic function symbols for $\leq$ and $>$, respectively. Similarly for $\max(x, y)$, bounded minimum

$$\min_{x \leq y}(f(x)) = f\big((\mu v < y + \overline{1})(\forall x < y + \overline{1})(f(v) \leq f(x))\big)$$

and bounded maximum

$$\max_{x \leq y}(f(x)) = f\big((\mu v < y + \overline{1})(\forall x < y + \overline{1})(f(v) \geq f(x))\big)$$

as function symbols in $\mathcal{L}(0,1)$.

(Note again: To keep notation simpler, we do not endow these symbols with indices k. The indices are sufficiently clear by context.)

Below we shall formally assign characteristic terms to $\mathcal{L}(1,1)$ formulas by bounded course-of-values recursion. The underlying idea is to assign to $T(x)$ the same characteristic term that is also assigned to $val_1(x)$, and to $\sim T(x)$ the characteristic term $\overline{1} - t$, where t is the characteristic term assigned to $val_1(x)$.

To make the bounded course-of-values recursion work, $val_1(x)$ must be bounded by (code of) $T(x)$ somehow. The following key lemma ensures precisely this.

Lemma 6.18. $\mathbf{A}(0,3) \vdash x > \overline{0} \to val_1(x) < x$

Proof. By applying $\mathsf{IS}(\Sigma_0^{(0,3)})$ with induction formula

$$(\forall x < y)(x > \overline{0} \to val_1(x) < x)$$

and induction variable y, relying on the counterparts of Proposition 5.13 indicated in Section 5.3.

Let us illustrate a case of the formal induction step. Suppose that y has the form $\langle \ulcorner f \urcorner, \ulcorner t_0 \urcorner, \dots, \ulcorner t_n \urcorner \rangle$ and $f(x_0, \dots, x_n)$ is

$$x_0 \cdot \big(\cdots \big(x_{n-2} \cdot \big(x_{n-1} \cdot x_n\big) \cdots\big)$$

obtained by composition. By induction hypothesis, we have, for every $i \leq n$, that

$$val_1(\ulcorner t_i \urcorner) < \ulcorner t_i \urcorner.$$

Therefore, $val_1(\langle \ulcorner f \urcorner, \ulcorner t_0 \urcorner, \dots, \ulcorner t_n \urcorner \rangle)$ is smaller than

$$\ulcorner t_0 \urcorner \cdot \big(\cdots \big(\ulcorner t_{n-2} \urcorner \cdot \big(\ulcorner t_{n-1} \urcorner \cdot \ulcorner t_n \urcorner\big) \cdots\big)$$

and we obtain the following.

$$val_1(\langle \ulcorner f \urcorner, \ulcorner t_0 \urcorner, \dots, \ulcorner t_n \urcorner \rangle) < \overline{2}^{(\sum_{i \leq n} |\ulcorner t_i \urcorner|)}$$

Consequently, we may conclude by our Definition 4.6 of sequence numbers that

$$val_1(\langle \ulcorner f \urcorner, \ulcorner t_0 \urcorner, \dots, \ulcorner t_n \urcorner \rangle) < \langle \ulcorner f \urcorner, \ulcorner t_0 \urcorner, \dots, \ulcorner t_n \urcorner \rangle$$

which means: $val_1(\langle \ulcorner f \urcorner, \ulcorner t_0 \urcorner, \dots, \ulcorner t_n \urcorner \rangle) < y$. □

We can now formally assign characteristic terms to $\mathcal{L}(1,1)$ formulas. Notice that, in the following lemma, we use $\min(\overline{1},x)$ in most cases only to simplify to proof of Lemma 6.20 below. Moreover, the cases for *unbounded* quantifiers have to be trivialized.

We will require the coding of $\mathcal{L}(0,1)$ symbols $f_{\leq}$, $f_{>}$, and so forth, where

$$\begin{aligned}
f_{\leq}(x,y) &= \overline{1} - (x - y)\\
f_{>}(x,y) &= \overline{1} - f_{\leq}(x,y)\\
f_{=}(x,y) &= \overline{1} - (f_{>}(x,y) + f_{>}(y,x))\\
f_{\neq}(x,y) &= \overline{1} - f_{=}(x,y)
\end{aligned}$$

are derivable in $\mathbf{A}(0,3)$. To simplify the notation, we also use an $\mathcal{L}(0,3)$ symbol $[\max_{x\leq y}(z)]$ for the function that maps the codes of variables v and terms t_1, t_2 to the code of the respective term $\max_{v\leq t_1}(t_2)$. Similarly for $[\min_{x\leq y}(z)]$

Lemma 6.19. $\mathbf{A}(0,3)$ *proves, for function symbol $\iota(x)$ in $\mathcal{L}(0,3)$, the following.*

$$\iota(x) = y \leftrightarrow \bigvee \big\{(1)\text{–}(13)\big\},$$

where (1)–(13) *are the following disjuncts that we sub-divide into cases for better readability.*

Case A: Non-Formulas

(1) $\neg Form_{1,1}(x) \wedge y = \overline{0}$

Case B: Arithmetical Literals

(2) $(\exists x_0, x_1 < x)\big(x = [x_0 = x_1] \wedge Form_{1,1}(x) \wedge$
$y = \ulcorner \min(\overline{1}, f_{=}(u,v))\urcorner[x_0, x_1/\ulcorner u\urcorner, \ulcorner v\urcorner]\big)$

(3) $(\exists x_0, x_1 < x)\big(x = [x_0 \neq x_1] \wedge Form_{1,1}(x) \wedge$
$y = \ulcorner \min(\overline{1}, f_{\neq}(u,v))\urcorner[x_0, x_1/\ulcorner u\urcorner, \ulcorner v\urcorner]\big)$

(4) $(\exists x_0, x_1 < x)\big(x = [x_0 \leq x_1] \wedge \mathit{Form}_{1,1}(x) \wedge$
$y = \ulcorner \min(\overline{1}, f_{\leq}(u,v)) \urcorner [x_0, x_1 / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)$

(5) $(\exists x_0, x_1 < x)\big(x = [x_0 > x_1] \wedge \mathit{Form}_{1,1}(x) \wedge$
$y = \ulcorner \min(\overline{1}, f_{>}(u,v)) \urcorner [x_0, x_1 / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)$

Case C: Truth Literals

(6) $(\exists x_0, x_1 < x)\big(x = [T(x_0)] \wedge \mathit{Form}_{1,1}(x) \wedge$
$x_1 = \mathit{val}_1(x_0) \wedge y = \ulcorner \min(\overline{1}, v) \urcorner [\iota(x_1) / \ulcorner v \urcorner]\big)$

(7) $(\exists x_0, x_1 < x)\big(x = [{\sim}T(x_0)] \wedge \mathit{Form}_{1,1}(x) \wedge$
$x_1 = \mathit{val}_1(x_0) \wedge y = \ulcorner \overline{1} - v \urcorner [\iota(x_1) / \ulcorner v \urcorner]\big)$

Case D: Connectives and Bounded Quantifiers

(8) $(\exists x_0, x_1 < x)\big(x = [x_0 \vee x_1] \wedge \mathit{Form}_{1,1}(x) \wedge$
$y = \ulcorner \min(\overline{1}, \max(u,v)) \urcorner [\iota(x_0), \iota(x_1) / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)$

(9) $(\exists x_0, x_1 < x)\big(x = [x_0 \wedge x_1] \wedge \mathit{Form}_{1,1}(x) \wedge$
$y = \ulcorner \min(\overline{1}, \min(u,v)) \urcorner [\iota(x_0), \iota(x_1) / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)$

(10) $(\exists x_0, x_1, x_2, x_3 < x)\big(x = [(\exists x_0)x_1] \wedge \mathit{Form}_{1,1}(x) \wedge$
$x_1 = [[x_0 \leq x_2] \wedge x_3] \wedge$
$y = \ulcorner \min(\overline{1}, u) \urcorner [[\max_{x_0 \leq x_2}(\iota(x_3))] / \ulcorner u \urcorner]\big)$

(11) $(\exists x_0, x_1, x_2, x_3 < x)\big(x = [(\forall x_0)x_1] \wedge \mathit{Form}_{1,1}(x) \wedge$
$x_1 = [[x_0 \leq x_2] \rightarrow x_3] \wedge$
$y = \ulcorner \min(\overline{1}, u) \urcorner [[\min_{x_0 \leq x_2}(\iota(x_3))] / \ulcorner u \urcorner]\big)$

Case E: Unbounded Quantifiers

(12) $(\exists x_0, x_1 < x)(\forall x_2, x_3 < x)\big(x = [(\exists x_0)x_1] \wedge \mathit{Form}_{1,1}(x) \wedge$
$x_1 \neq [[x_0 \leq x_2] \wedge x_3] \wedge y = \ulcorner \overline{1} \urcorner\big)$

(13) $(\exists x_0, x_1 < x)(\forall x_2, x_3 < x)\big(x = [(\forall x_0)x_1] \wedge \mathit{Form}_{1,1}(x) \wedge$
$x_1 \neq [[x_0 \leq x_2] \rightarrow x_3] \wedge y = \ulcorner \overline{0} \urcorner\big)$

Proof. By Corollary 4.17, where the main Lemma 6.18 guarantees $val_1(x_0) < x$ in (6) and (7). ☐

The following is immediate.

Lemma 6.20. $\mathbf{A}(0,3) \vdash val_1(\iota(x)) \leq \overline{1}$

Proof. In view of the properties of ι, it suffices to note that, by $\mathsf{IS}(\Sigma_0^{(0,3)})$, we have $\overline{0} \leq \min(\overline{1}, x) \leq \overline{1}$. ☐

6.4 Retractability: Upper Bound

We may now define truth of t as

$$val_1(\iota(t)) = \overline{1}$$

and proceed as follows.

Definition 6.21. The mapping $\mathcal{V} : \mathcal{L}(1,3) \to \mathcal{L}(0,3)$ is defined recursively as follows.

$$\begin{aligned}
(s = t)^{\mathcal{V}} &:= (s = t)\\
(s \neq t)^{\mathcal{V}} &:= (s \neq t)\\
(s \leq t)^{\mathcal{V}} &:= (s \leq t)\\
(s > t)^{\mathcal{V}} &:= (s > t)\\
(T(t))^{\mathcal{V}} &:= val_1(\iota(t)) = \overline{1}\\
({\sim}T(t))^{\mathcal{V}} &:= val_1(\iota(t)) = \overline{0}\\
(B \vee C)^{\mathcal{V}} &:= (B^{\mathcal{V}} \vee C^{\mathcal{V}})\\
(B \wedge C)^{\mathcal{V}} &:= (B^{\mathcal{V}} \wedge C^{\mathcal{V}})\\
(\exists x B)^{\mathcal{V}} &:= \exists x (B)^{\mathcal{V}}\\
(\forall x B)^{\mathcal{V}} &:= \forall x (B)^{\mathcal{V}}
\end{aligned}$$

Lemma 6.22. *$\mathcal{V}$ preserves logical structure in* $\mathbf{A}(0,3)$.

Proof. By definition and Lemma 6.20. □

Lemma 6.23. *For every $\mathcal{L}(0,3)$ formula A, $A^{\mathcal{V}}$ is A.*

Proof. By (meta) induction on the rank of A. □

We may now interpret the truth axioms of $\mathsf{FH}(1,1)$.

Lemma 6.24. $\mathbf{A}(0,3)$ *proves the following.*

(1) $Sent_{1,1}([x = y]) \to$
$\big(T([x = y])^{\mathcal{V}} \leftrightarrow val_1(x) = val_1(y)\big)$

(2) $Sent_{1,1}([x \leq y]) \to$
$\big(T([x \leq y])^{\mathcal{V}} \leftrightarrow val_1(x) \leq val_1(y)\big)$

(3) $Sent_{1,1}([T(x)]) \to$
$\big(T([T(x)])^{\mathcal{V}} \leftrightarrow T(val_1(x))^{\mathcal{V}}\big)$

(4) $Sent_{1,1}(x) \to$
$\big(T([\neg x])^{\mathcal{V}} \leftrightarrow \neg(T(x))^{\mathcal{V}}\big)$

(5) $Sent_{1,1}([x \vee y]) \to$
$\big(T([x \vee y])^{\mathcal{V}} \leftrightarrow \big(T(x)^{\mathcal{V}} \vee T(y)^{\mathcal{V}}\big)\big)$

(6) $Sent_{1,1}([(\exists x \leq y)z]) \to$
$\big(T([(\exists x \leq y)z])^{\mathcal{V}} \leftrightarrow (\exists u \leq val_1(y))(T(z[num(u)/x])^{\mathcal{V}}\big)$

Proof. We focus on the derivation of (4) because the remaining derivations basically use the properties of ι. For instance: By Proposition 6.1 (and by noting that $\mathcal{L}(0,2)$ includes $\mathcal{L}(0,1)$), we have, for all closed $\mathcal{L}(0,1)$ terms x, y,

$$val_1(\ulcorner \min(\overline{1}, f_{=}(u,v))\urcorner \left[\frac{x, y}{\ulcorner u\urcorner, \ulcorner v\urcorner}\right]) = \min(\overline{1}, f_{=}(val_1(x), val_1(y)))$$

and, relying on the properties of ι and $f_{=}$,

$$val_1(\iota([x = y])) = \overline{1} \leftrightarrow val_1(x) = val_1(y)$$

which takes care of (1). As for (3), given some $\mathcal{L}(1,1)$ sentence $[T(x)]$, we have by key Lemma 6.18 that $val_1(x) < [T(x)]$. Thus, by Proposition 6.1 and the properties of ι,

$$val_1(\iota([T(x)])) = \min(\overline{1}, val_1(\iota(val_1(x))))$$

and, therefore,

$$val_1(\iota([T(x)])) = \overline{1} \leftrightarrow val_1(\iota(val_1(x))) = \overline{1}$$

which takes care of (3).

So, let us now consider (4). (We use suggestive notation to improve readability.) Observe first that, for any sub-formula $a \in subfml(x)$, terms involved in a have at most x many free variables and a complexity measure (*tcm*) bounded by x. So, we apply $\mathsf{IS}(\Sigma_0^{(0,3)})$ to derive that, for every $a \in subfml(x)$ and (like in Lemma 6.16 above) each sequence y of at most x many numeral codes among

$$num(0), \ldots, num(\varphi_{k+1}(4, (rk(x) - rk(a)) \cdot x)),$$

where rk denotes the rank, we have

$$\begin{aligned}&Sent_{1,1}(a[y]) \rightarrow \\ &\big(val_1(\iota(a[y])) = \overline{0} \leftrightarrow val_1(\iota([\neg a][y])) = \overline{1}\big),\end{aligned}$$

where Lemma 6.20 guarantees

$$val_1(\iota(a[y])) = \overline{0} \leftrightarrow val_1(\iota(a[y])) \neq \overline{1}.$$

Specifically, for $a = x$ and $lh(frvar(a)) = \overline{0}$, this will eventually yield claim (4). We concentrate on selected cases of the formal induction (using suggestive notation) to show the strategy.

- If $a[y]$ is $(s \leq t)$, then the claim is immediate: $\iota([\neg a][y])$ is $\min(\overline{1}, f_{>}(s,t))$ and $\iota(a[y])$ is $\min(\overline{1}, f_{\leq}(s,t))$ and we obtain $f_{>}(s,t) = \overline{1} - f_{\leq}(s,t)$.

- If $a[y]$ is $T(t)$, where t is an $\mathcal{L}(0,1)$ term, then the claim is immediate: By key Lemma 6.18, the value (val_1) of t is bounded by $a[y]$, such that by the Proposition 6.1 and the properties of ι,

  $$val_1(\iota([\neg a][y])) = \overline{1} - val_1(\iota(a[y])).$$

- If $a[y]$ is a conjunction $(B \wedge C)$ and b, c are codes of B, C, respectively, then $\iota(a[y])$ is

  $$\min(\overline{1}, \min(\iota(b), \iota(c)))$$

  and $\iota([\neg a][y])$ is

  $$\min(\overline{1}, \max(\iota([\neg b]), \iota([\neg c]))).$$

  So, we have the claim by induction hypothesis, according to which $val_1(\iota([\neg b])) = 1$ if and only if $val_1(\iota(b)) = 0$, and accordingly for c.

- If a has form $(\forall u)B(u, v_0, \dots, v_r)$, then we distinguish two sub-cases.

  In the first sub-case, $B(u, v_0, \dots, v_r)$ actually has the form $u \leq t(v_0, \dots, v_r) \rightarrow C(u, v_0, \dots, v_r)$. Let p be a substitution sequence $(\overline{p_0}, \dots, \overline{p_r})$ with $r \leq x$ and

  $$p_i \leq \varphi_3(4, (rk(x) - rk(a)) \cdot x).$$

  Let c, d be codes of C, t, respectively. By Lemma 6.16, the value (val_1) of $t(\overline{p_0}, \dots, \overline{p_r})$ is bounded by

  $$\varphi_3(4, (rk(x) - rk(a)) \cdot x + tcm(d))$$

  and, since $tcm(d) < x$, also by

  $$\varphi_3(4, (rk(x) - (rk(a) - 1)) \cdot x)$$

and, since $rk(c) < rk(a) - 1$, also by

$$\varphi_3(4, (rk(x) - rk(c)) \cdot x).$$

So, we obtain by the induction hypothesis that, for every q that is bounded by the value (val_1) of term $t(\overline{p_0}, \ldots, \overline{p_r})$, the equivalence of the following.

$$\begin{aligned} val_1(\iota([\neg c][(\overline{q}, \overline{p_0}, \ldots, \overline{p_r})])) = 1 \\ val_1(\iota(c[(\overline{q}, \overline{p_0}, \ldots, \overline{p_r})])) = 0 \end{aligned}$$

From this, we may finally prove the claim by appeal to the properties of ι, according to which $\iota(a[y])$ is a bounded minimum and $\iota([\neg a][y])$ is the dual bounded maximum.

The second sub-case, where $B(u, v_0, \ldots, v_r)$ does not have form $u \leq t(v_0, \ldots, v_r) \to C(u, v_0, \ldots, v_r)$, is trivial by the properties of ι.

Other (dual) cases are treated accordingly. □

Finally, the interpretation of $\mathsf{TI}(1,1)$ is immediate.

Lemma 6.25. $\mathbf{A}(0,3)$ *proves the following.*

$$\begin{aligned} &Sent_{1,1}([(\forall z)x]) \to \big(\big(T(x[num(\overline{0})/z])^{\mathcal{V}} \wedge \\ &\forall y(T(x[num(y)/z])^{\mathcal{V}} \to T(x[num(y+\overline{1})/z])^{\mathcal{V}})\big) \to \\ &T(x[num(y)/z])^{\mathcal{V}})\big) \end{aligned}$$

Proof. By $\mathsf{IS}(\Sigma_0^{(0,3)})$, noticing that $val_1(\iota(x[num(y)/z])) = \overline{1}$ is a $\Sigma_0^{(0,3)}$ formula. □

6.5 Conclusions

Let us put together our previous findings to conclude the main results of the present chapter.

Theorem 6.26. $\mathbf{FH}(1,1)$ *is retractable to* $\mathbf{A}(0,3)$*.*

Proof. We successively verify the conditions of Definition 2.26 to show that $\mathbf{FH}(1,1)$ is retractable to $\mathbf{A}(0,3)$ by $\mathcal{V}$.

- By Lemma 6.22, $\mathcal{V}$ preserves logical structure in $\mathbf{A}(0,3)$.
- By Lemma 6.23 and logic, we have, for any $\mathcal{L}(0,3)$ atom A, that $\mathbf{A}(0,3) \vdash A \leftrightarrow A^{\mathcal{V}}$.
- By Lemma 6.24 and Lemma 6.25, we have, for any A from $\mathsf{FH}(1,1) \cup \mathsf{TI}(1,1)$, that $\mathbf{A}(0,3) \vdash A^{\mathcal{V}}$. Moreover, by preservation of logical structure, we obtain for (additional logical axioms) B in $\mathbf{A}(1,3)$ that $\mathbf{A}(0,3) \vdash B^{\mathcal{V}}$.
- Finally, every non-logical axiom of $\mathbf{A}(0,3)$ is a non-logical axiom of $\mathbf{A}(1,3)$.

This conclues the proof. ☐

Corollary 6.27. $\mathbf{FH}(1,1)$ *is consistent.*

Proof. Since $\mathbf{A}(0,3) \nvdash \overline{0} \neq \overline{0}$ (as could be shown), we obtain by Theorem 6.26 and Proposition 2.27 (2) that

$$\mathbf{FH}(1,1) \nvdash \overline{0} \neq \overline{0},$$

as required. ☐

Thus, by Corollary 3.54 (2), we immediately obtain

$$\mathbf{B}(1,3) \cup \big(\mathsf{FH}(1,1) \cup \mathsf{TI}(1,1)\big)^{\mathcal{A}} \nvdash \overline{0} \neq \overline{0},$$

the consistency of the $\mathcal{L}(1, \mathrm{Q})$ variant of $\mathbf{FH}(1,1)$.

Indeed, the following theorem clarifies the role of $\mathbf{FH}(1,1)$ as a theory of *naive type-free truth* for the Σ_0 fragment of the language $\mathcal{L}(1,1)$. We postpone the details of the proof to Chapter 8, Lemma 8.14, where we will be concerned with the Tarski Schema in more detail.

Theorem 6.28. $\mathbf{FH}(1,1)$ *proves, for any* $\Sigma_0^{(1,1)}$ *sentence* A, *the following.*

$$A \leftrightarrow T(\ulcorner A \urcorner)$$

That is to say, Corollary 6.27 and Theorem 6.28 jointly entail the consistency of

$$\mathbf{A}(1,3) \cup \left\{A \leftrightarrow T(\ulcorner A \urcorner) : A \text{ is } \Sigma_0^{(1,1)}\right\}$$

where $\ulcorner A \urcorner$ is definable even in the Σ_0 fragment of $\mathcal{L}(0,1)$ – and, therefore, in a considerably weaker base theory $\mathbf{A}(1,1)$ – by employing a coding machinery which is more efficient than the one from Sections 4 and 5. See, for example, Hájek and Pudlák [18], Chapter V, Section 3.

7 Finitist Friedman-Sheard

In the foregoing Section 6, we have established that the classical negation principle that, for arbitrary $\mathcal{L}$ sentences A,

$$\neg T(\ulcorner A \urcorner) \leftrightarrow T(\ulcorner \neg A \urcorner) \tag{7.1}$$

may be conjoined with the iteration principle that, for arbitrary closed $\mathcal{L}$ terms t,

$$T(t) \leftrightarrow T(\ulcorner T(t) \urcorner), \tag{7.2}$$

in a consistent manner in the theory $\mathbf{FH}(1,1)$ of naive type-free truth for the Σ_0 fragment of $\mathcal{L} = \mathcal{L}(1,1)$.

(Notation: Will still adopt the convention that

$$T(x) := T_1(x)$$

because we are still dealing with only a single truth predicate.)

As was shown, the consistency of $\mathbf{FH}(1,1)$ relies on the fact that $\mathcal{L}(0,1)$ does not provide a (diagonalization) term t so that $t = \ulcorner \neg T(t) \urcorner$ is provable in the underlying base theory.

However, by Proposition 6.6, we also know that (with respect to our syntax coding from Chapter 5) we cannot have consistent theories of naive type-free truth for Σ_0 fragment of $\mathcal{L} = \mathcal{L}(1,k)$ with $k > 1$.

Thus, if our concern is type-free truth on the higher levels of the Grzegorczyk hierarchy, either (7.1) or (7.2) has to be dropped.

In the present chapter, we will consider the attempt to stick with the negation axiom (7.1) and neglect iteration axiom (7.2). (In the next chapter, we will endorse (7.2) and drop (7.1).)

The idea which we will consider in the current chapter is to replace iteration axiom (7.2) by a pair of *truth rules* to the effect that, in particular,

$$\mathsf{S} \vdash T(t) \text{ iff } \mathsf{S} \vdash T(\ulcorner T(t) \urcorner)$$

and, for certain A more generally,

$$\mathsf{S} \vdash A \text{ iff } \mathsf{S} \vdash T(\ulcorner A \urcorner), \tag{7.3}$$

where S is one of the theories of truth to be considered in this chapter. By virtue of (7.3), the theory S is also said to ensure *symmetry* (with respect to a fragment).

The theories of truth to be studied in this chapter are finitist variants of the so-called Friedman-Sheard (FS) axiomatization, which in the current form we take from Halbach [19], but which goes back to the work of Friedman and Sheard [14].

FS is another axiomatization of *type-free* truth. Just as for $\mathbf{FH}(1,1)$, the presence of a negation axiom in finitist variants of FS gives rise to notions of *decidable truth*: By the negation axiom, we obtain $T(\ulcorner A \urcorner) \vee T(\ulcorner \neg A \urcorner)$, where in finitist contexts we take $T(\ulcorner A \urcorner)$ to express that A is verified (witnessed) by a finite computation. (Verifications will proceed by computing values of characteristic closed terms, like in the last chapter.)

While decidable problems of arithmetic may be seen to have complexity Δ_1 in arithmetical theories, the finitist variants of FS will be axiomatizations of Δ_0 truth in a language with a truth predicate. Accordingly, the finitist variants of FS will guarantee symmetry with respect to a (sub-)finitist Δ_0 fragment.

Finitist variants of the FS axiomatization, called $\mathsf{FFS}(1,k)$ to denote the level k of the (smashed) Grzegorczyk hierarchy, shall be defined in Section 7.1 below. They comprise the class $\mathsf{FD}(1,k)$ of axioms of arithmetical truth for the *decidable* Σ_0 fragment of $\mathcal{L}(0,k)$, but the axioms still range over codes of sentences in the overall language $\mathcal{L}(1,k)$. In Section 7.2, it will be shown that, for $k>1$, the theory

$$\mathbf{A}(1,k+1)\cup\mathsf{FD}(1,k)\cup\mathsf{TI}(1,k) \tag{7.4}$$

is interpretable in and conservative over $\mathbf{A}(0,k+1)$. This first outcome will not be concerned with truth rules, but verifies that, for any $k>1$, the $(k+1)$-th level of the (smashed) Grzegorczyk hierarchy features – via term valuation – a conception of Σ_0 truth for the k-th level of the (smashed) Grzegorczyk hierarchy.

In Section 7.3, results of 7.2 will be extended. We will show that, for $k>1$, the Finitist FS Theory $\mathbf{FFS}(1,k)$ that extends (7.4) by some (uniform) truth rules for Σ_0 formulas in language $\mathcal{L}(1,k)$ is *locally* interpretable in and conservative over the base theory $\mathbf{A}(0,k+1)$. Here *locality* means that, for each number n, if at most n truth rules are applied in a derivation of A, some interpretation of A *depending on* n can be derived in $\mathbf{A}(0,k+1)$.

An unrestricted FS theory of truth with unrestricted compositional axioms and unrestricted symmetry rules for a whole first-order language of arithmetic (with Peano Arithmetic as its base theory) has been studied by Halbach [19] and its proof-theoretic strength (in terms of provable arithmetical statements) has been determined to coincide with the one of Ramified Analysis $\mathbf{RA}_{<\omega}$ up to level ω. The important peculiarity of this unrestricted FS theory, a peculiarity that is not shared by its finitist variants, is that it is – while consistent in the regular sense – *ω-inconsistent.* (See Corollary 7.33.)

That it would be impossible to *retract* Halbach's unrestricted FS theory to a consistent theory of first-order arithmetic will be (implicitly) highlighted in Remark 8.2, where our concern will be with the Tarski Schema.

From a philosophical point of view, it might also be worth to mention that natural models for (subsystems of unrestricted) FS can be obtained via revision semantics. Revision semantics was introduced by Gupta [17] and Herzberger [25] and was further developed in Belnap and Gupta [1]. For more details, the reader is also referred to Halbach [21], Section 14.

7.1 The Theories

The axioms of Finitist Decidable Truth read as follows. Notice: They range (by the scope of *Sent*) over language $\mathcal{L}(1,k)$ with a truth predicate, but – in contrast to the naive theory – they are silent about the truth of literals $T(t)$ and $\sim T(t)$.

Definition 7.1. For $k > 1$, $\mathsf{FD}(1,k)$ comprises the subsequent axioms of *Finitist Decidable Truth* for Σ_0 fragment of $\mathcal{L}(0,k)$.

(FFS1) $Sent_{1,k}([x = y]) \rightarrow$
$\big(T([x = y]) \leftrightarrow val_k(x) = val_k(y)\big)$

(FFS2) $Sent_{1,k}([x \leq y]) \rightarrow$
$\big(T([x \leq y]) \leftrightarrow val_k(x) \leq val_k(y)\big)$

(FFS3) $Sent_{1,k}(x) \rightarrow$
$\big(T([\neg x]) \leftrightarrow \neg T(x)\big)$

(FFS4) $Sent_{1,k}([x \vee y]) \rightarrow$
$\big(T([x \vee y]) \leftrightarrow (T(x) \vee T(y))\big)$

(FFS5) $Sent_{1,k}([(\exists x \leq y)z]) \rightarrow$
$\big(T([(\exists x \leq y)z]) \leftrightarrow (\exists u \leq val_k(y))T(z[num(u)/x])\big)$

In addition, we should generalize Truth Induction for levels of the (smashed) Grzegorczyk hierarchy as follows.

Definition 7.2. For $k > 1$, $\mathsf{TI}(1,k)$ comprises the subsequent axiom of *Truth Induction.*

$$\begin{aligned}&Sent_{1,k}([(\forall z)x]) \to \big(\big(T(x[num(\overline{0})/z]) \wedge \\ &\forall y(T(x[num(y)/z]) \to T(x[num(y+\overline{1})/z]))\big) \to \\ &T(x[num(y)/z]))\big)\end{aligned}$$

The theories of Decidable Arithmetical Truth may now be defined as follows.

Definition 7.3. For $k > 1$, $\mathbf{FD}(1,k)$ is the *Theory of Finitist Decidable Truth* (in auxiliary form) defined as follows.

$$\mathbf{A}(1,k+1) \cup \mathsf{FD}(1,k) \cup \mathsf{TI}(1,k)$$

Remark 7.4. The very same remarks as in 6.13 also apply for $\mathbf{FD}(1,k)$ and its respective $\mathcal{L}(1,\mathrm{Q})$ variant:

$$\mathbf{B}(1,k) \cup \big(\mathsf{FD}(1,k) \cup \mathsf{TI}(1,k)\big)^{\mathcal{A}}$$

Here $\mathcal{A}$ actually is the mapping

$$\mathcal{A}_k : \mathcal{L}(1,k) \to \mathcal{L}(1,\mathrm{Q})$$

for eliminating auxiliary function symbols, as in Definition 3.62. However, to simplify notation, will shall not always endow them with an index k, as this index k is clear by the respective base theory $\mathbf{B}(1,k)$.

A note on the auxiliary form of $\mathbf{FD}(1,k)$ is in order. Observe that, in contrast to $\mathbf{FD}(1,k)$, the $\mathcal{L}(1,\mathrm{Q})$ variant requires only a base theory for level k of the (smashed) Grzegorczyk hierarchy, where the latter does not provide a total valuation function for closed terms of level k. Indeed, for $k > 1$, the truth axioms

$$\big(\mathsf{FD}(1,k) \cup \mathsf{TI}(1,k)\big)^{\mathcal{A}}$$

have genuine proof-theoretic power, as witnessed by the following proposition, whose formulation requires $G_k(x,y)$ from Proposition 3.59.

Proposition 7.5. *For $k > 1$, totality assertion $(\exists y G_{k+1}(x,y))^{\mathcal{A}}$ is provable in*

$$\mathbf{B}(1,k) \cup \big(\mathsf{FD}(1,k) \cup \mathsf{TI}(1,k)\big)^{\mathcal{A}}.$$

Proof. In view of Corollary 3.64, it is enough to establish that $\mathbf{A}(1,k) \cup \mathsf{FD}(1,k)^{\mathcal{A}} \cup \mathsf{TI}(1,k) \vdash \exists y G_{k+1}(x,y)$. We focus on $k > 2$. (For $k = 2$, we apply the more efficient forms of course-of-values recursion and numeral coding by means of the smash function.)

So, let $f(x)$ be an $\mathcal{L}(0,3)$ function symbol for which $\mathbf{A}(0,3)$ proves

$$f(x) = \begin{cases} num(\overline{1}) & : x = \overline{0} \\ \ulcorner \gamma_k(v) \urcorner [f(x - \overline{1}) / \ulcorner v \urcorner] & : x > \overline{0} \end{cases}$$

where γ_k is the k-th Grzegorczyk function symbol. Then we get $\mathbf{A}(0,3) \vdash ClTerm_k(f(x))$. Therefore, by the $\mathcal{A}$ interpretation of (FFS1), we have in $\mathbf{A}(1,k) \cup \mathsf{FD}(1,k)^{\mathcal{A}}$ the following equivalence.

$$T([f(x) = f(x)]) \leftrightarrow \exists y (val_k(f(x)) = y)^{\mathcal{A}}$$

Hence, by $\mathsf{TI}(1,k)$, we obtain $\exists y (val_k(f(x)) = y)^{\mathcal{A}}$ and via

$$(\forall y < z)\big((val_k(f(x)) = y)^{\mathcal{A}} \to G_{k+1}(x,y)\big)$$

the totality assertion: $\exists y G_{k+1}(x,y)$. ☐

Thus, by Corollary 3.64 (1), we may work in auxiliary theory $\mathbf{A}(1, k+1) \cup \mathsf{FD}(1, k)$ and derive B if we actually want to verify that $B^{\mathcal{A}}$ is derivable in the theory

$$\mathbf{B}(1, k) \cup \big(\mathsf{FD}(1, k) \cup \mathsf{TI}(1, k)\big)^{\mathcal{A}}$$

as, for example, in the next proposition.

The proof of the following is parallel to the proof of Proposition 6.9 and may be omitted.

Proposition 7.6. *For $k > 1$, $\mathbf{A}(1, k+1) \cup \mathsf{FD}(1, k)$ proves every formula below.*

(1) $Sent_{1,k}([x \neq y]) \rightarrow$
$\big(T([x \neq y]) \leftrightarrow val_k(x) \neq val_k(y)\big)$

(2) $Sent_{1,k}([x > y]) \rightarrow$
$\big(T([x > y]) \leftrightarrow val_k(x) > val_k(y)\big)$

(3) $Sent_{1,k}([x \wedge y]) \rightarrow$
$\big(T([x \wedge y]) \leftrightarrow (T(x) \wedge T(y))\big)$

(4) $Sent_{1,k}([(\forall x \leq y)z]) \rightarrow$
$\big(T([(\forall x \leq y)z]) \leftrightarrow (\forall u \leq val_k(y))T(z[num(u)/x])\big)$

Finally, the axioms and inferences of Finitist Friedman-Sheard are defined as follows. While Halbach [21], Section 14, states the truth rules in the form

$$\frac{A}{T(\ulcorner A \urcorner)} \qquad \frac{T(\ulcorner A \urcorner)}{A}$$

where A is any *sentence*, we shall be a bit more liberal here and define truth rules *uniformly* (see below). This also serves an explanatory purpose: Namely, it illustrates that, while our *finitist* theories of truth may not comprise a truth axiom

$$\begin{aligned}&Sent([(\forall x)y]) \to \\ &\quad \big(T([(\forall x)y]) \leftrightarrow \forall u\big(ClTerm(u) \to T(y[u/x])\big)\big)\end{aligned} \tag{7.5}$$

for *unbounded* universal quantification, they may well feature the uniform truth *rules* as counterparts of such an axiom.

Definition 7.7. For $k > 1$, the *Finitist Friedman-Sheard* axioms/inferences $\mathsf{FFS}(1,k)$ comprise the axioms in $\mathsf{FD}(1,k)$ and the two inferences

(FFS6)
$$\frac{A(v_0,\ldots,v_n)}{\bigwedge_{i=0}^{n} ClTerm_k(u_i) \to T(\ulcorner A \urcorner \left[\frac{u_0,\ldots,u_n}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_n \urcorner}\right])}$$

(FFS7)
$$\frac{\bigwedge_{i=0}^{n} ClTerm_k(u_i) \to T(\ulcorner A \urcorner \left[\frac{u_0,\ldots,u_n}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_n \urcorner}\right])}{A(v_0,\ldots,v_n)}$$

where $A(v_0,\ldots,v_n)$ may be any $\Sigma_0^{(1,k)}$ formula.

It is a distinctive feature of the FFS axioms/inferences that, while they do not include truth axioms for the literals $T(t)$ and $\sim T(t)$, they make theory S *symmetric with respect to* Σ_0: If A is some Σ_0 sentence in $\mathcal{L}(1,k)$, then $\mathsf{S} \vdash A$ precisely if $\mathsf{S} \vdash T(\ulcorner A \urcorner)$.

Definition 7.8. For $k > 1$, $\mathbf{FFS}(1,k)$ is the *Finitist Friedman-Sheard Theory* (in auxiliary form) defined as follows.

$$\mathbf{A}(1,k+1) \cup \mathsf{FFS}(1,k) \cup \mathsf{TI}(1,k)$$

We now move to the proof-theoretic analysis.

7.2 Retractability: Valuation

Our first aim is to confirm formally that, for every $k > 1$, the $(k+1)$-th level of the (smashed) Grzegorczyk hierarchy provides – due to term valuation – a notion of Σ_0 truth for level k of the (smashed) Grzegorczyk hierarchy.

Like in Section 6.3, the confirmation is a formalization of the proof of Proposition 4.3: To every Σ_0 formula in the language $\mathcal{L}(1,k)$, we formally assign some $\mathcal{L}(0,k)$ term whose value is 1 if the formula is true, and 0 otherwise. Fortunately, this allows us to build on the work of Section 6.3 and repeat the same line of argument again (by omitting several details).

So, once more, let us assign characteristic terms to $\mathcal{L}(1,k)$ formulas, for $k > 1$. Again, we use $\min(\overline{1}, x)$ in most cases only to simplify the proof of Lemma 7.10 below. Moreover, the cases for $T(t)$ and $\sim T(t)$ now have to be trivialized, just like the cases for unbounded quantifiers.

Recall: We use $\mathcal{L}(0,1)$ symbols for which $\mathbf{A}(0,3)$ proves

$$\begin{aligned}
f_{\leq}(x,y) &= \overline{1} - (x - y)\\
f_{>}(x,y) &= \overline{1} - f_{\leq}(x,y)\\
f_{=}(x,y) &= \overline{1} - (f_{>}(x,y) + f_{>}(y,x))\\
f_{\neq}(x,y) &= \overline{1} - f_{=}(x,y)
\end{aligned}$$

as well as

$$\min_{x \leq y}(f(x)) = f\big((\mu v < y + \overline{1})(\forall x < y + \overline{1})(f(v) \leq f(x))\big)$$

and, accordingly, the following.

$$\max_{x \leq y}(f(x)) = f\big((\mu v < y + \overline{1})(\forall x < y + \overline{1})(f(v) \geq f(x))\big)$$

By bounded course-of-values recursion, we have the following.

Proposition 7.9. $\mathbf{A}(0,3)$ *proves, for the function symbol* $\tau_k(x)$ *in* $\mathcal{L}(0,3)$*, the following.*

$$\tau_k(x) = y \leftrightarrow \bigvee \big\{(1)\text{–}(13)\big\}$$

where (1)–(13) *are the following disjuncts that we sub-divide into cases for better readability.*

Case A: Non-Formulas

(1) $\neg Form_{1,k}(x) \wedge y = \overline{0}$

Case B: Arithmetical Literals

(2) $(\exists x_0, x_1 < x)\big(x = [x_0 = x_1] \wedge Form_{1,k}(x) \wedge$ $y = \ulcorner \min(\overline{1}, f_{=}(u,v))\urcorner[x_0, x_1/\ulcorner u\urcorner, \ulcorner v\urcorner]\big)$

(3) $(\exists x_0, x_1 < x)\big(x = [x_0 \neq x_1] \wedge Form_{1,k}(x) \wedge$ $y = \ulcorner \min(\overline{1}, f_{\neq}(u,v))\urcorner[x_0, x_1/\ulcorner u\urcorner, \ulcorner v\urcorner]\big)$

(4) $(\exists x_0, x_1 < x)\big(x = [x_0 \leq x_1] \wedge Form_{1,k}(x) \wedge$ $y = \ulcorner \min(\overline{1}, f_{\leq}(u,v))\urcorner[x_0, x_1/\ulcorner u\urcorner, \ulcorner v\urcorner]\big)$

(5) $(\exists x_0, x_1 < x)\big(x = [x_0 > x_1] \wedge Form_{1,k}(x) \wedge$ $y = \ulcorner \min(\overline{1}, f_{>}(u,v))\urcorner[x_0, x_1/\ulcorner u\urcorner, \ulcorner v\urcorner]\big)$

Case C: Truth Literals

(6) $(\exists x_0 < x)\big(x = [T(x_0)] \wedge Form_{1,k}(x) \wedge y = \ulcorner \overline{1}\urcorner\big)$

(7) $(\exists x_0 < x)\big(x = [{\sim}T(x_0)] \wedge Form_{1,k}(x) \wedge y = \ulcorner \overline{0}\urcorner\big)$

Case D: Connectives and Bounded Quantifiers

(8) $(\exists x_0, x_1 < x)\big(x = [x_0 \vee x_1] \wedge Form_{1,k}(x) \wedge$ $y = \ulcorner \min(\overline{1}, \max(u,v))\urcorner[\tau_k(x_0), \tau_k(x_1)/\ulcorner u\urcorner, \ulcorner v\urcorner]\big)$

(9) $(\exists x_0, x_1 < x)\big(x = [x_0 \wedge x_1] \wedge Form_{1,k}(x) \wedge$
$y = \ulcorner \min(\overline{1}, \min(u, v))\urcorner[\tau_k(x_0), \tau_k(x_1)/\ulcorner u\urcorner, \ulcorner v\urcorner]\big)$

(10) $(\exists x_0, x_1, x_2, x_3 < x)\big(x = [(\exists x_0)x_1] \wedge Form_{1,k}(x) \wedge$
$x_1 = [[x_0 \leq x_2] \wedge x_3] \wedge$
$y = \ulcorner \min(\overline{1}, u)\urcorner[[\max_{x_0 \leq x_2}(\tau_k(x_3))]/\ulcorner u\urcorner]\big)$

(11) $(\exists x_0, x_1, x_2, x_3 < x)\big(x = [(\forall x_0)x_1] \wedge Form_{1,k}(x) \wedge$
$x_1 = [[x_0 \leq x_2] \rightarrow x_3] \wedge$
$y = \ulcorner \min(\overline{1}, u)\urcorner[[\min_{x_0 \leq x_2}(\tau_k(x_3))]/\ulcorner u\urcorner]\big)$

Case E: Unbounded Quantifiers

(12) $(\exists x_0, x_1 < x)(\forall x_2, x_3 < x)\big(x = [(\exists x_0)x_1] \wedge Form_{1,k}(x) \wedge$
$x_1 \neq [[x_0 \leq x_2] \wedge x_3] \wedge y = \ulcorner \overline{1}\urcorner\big)$

(13) $(\exists x_0, x_1 < x)(\forall x_2, x_3 < x)\big(x = [(\forall x_0)x_1] \wedge Form_{1,k}(x) \wedge$
$x_1 \neq [[x_0 \leq x_2] \rightarrow x_3] \wedge y = \ulcorner \overline{0}\urcorner\big)$

Proof. By applying Corollary 4.17. ☐

The following is immediate.

Lemma 7.10. *For $k > 1$, $\mathbf{A}(0, k+1) \vdash val_k(\tau_k(x)) \leq \overline{1}$.*

Proof. Like the proof of Lemma 6.20. ☐

The following will be the key lemmas in our concluding proof that, for $k > 1$, the $(k+1)$-th level of the (smashed) Grzegorczyk hierarchy provides a notion of Σ_0 truth for level k of the (smashed) Grzegorczyk hierarchy: Truth of t might be expressed as $val_k(\tau_k(t)) = \overline{1}$ and untruth of t as $val_k(\tau_k(t)) = \overline{0}$. Just to simplify the notation for the following lemma, we define

$$\mathcal{V}_k(T(t)) := (val_k(\tau_k(t)) = \overline{1})$$

in accordance with notations in Section 7.3 below.

Lemma 7.11. *For $k > 1$, $\mathbf{A}(0, k+1)$ proves the following.*

(1) $Sent_{1,k}([x = y]) \to$
$\big(\mathcal{V}_k(T([x = y])) \leftrightarrow val_k(x) = val_k(y)\big)$

(2) $Sent_{1,k}([x \leq y]) \to$
$\big(\mathcal{V}_k(T([x \leq y])) \leftrightarrow val_k(x) \leq val_k(y)\big)$

(3) $Sent_{1,k}(x) \to$
$\big(\mathcal{V}_k(T([\neg x])) \leftrightarrow \neg\mathcal{V}_k(T(x))\big)$

(4) $Sent_{1,k}([x \vee y]) \to$
$\big(\mathcal{V}_k(T([x \vee y])) \leftrightarrow (\mathcal{V}_k(T(x)) \vee \mathcal{V}_k(T(y)))\big)$

(5) $Sent_{1,k}([(\exists x \leq y)z]) \to$
$\big(\mathcal{V}_k(T([(\exists x \leq y)z])) \leftrightarrow (\exists u \leq val_k(y))\mathcal{V}_k(T(z[num(u)/x]))\big)$

Proof. Like the proof of Lemma 6.24, but without taking care of truth axiom for atoms $T(t)$. □

As for Truth Induction, we have the following.

Lemma 7.12. *For $k > 1$, $\mathbf{A}(0, k+1)$ proves the following.*

$$\begin{aligned}&Sent_{1,k}([(\forall z)x]) \to \Big(\big(\mathcal{V}_k(T(x[num(\overline{0})/z])) \wedge \\ &\forall y(\mathcal{V}_k(T(x[num(y)/z])) \to \mathcal{V}_k(T(x[num(y+\overline{1})/z])))\big) \to \\ &\mathcal{V}_k(T(x[num(y)/z]))\Big)\end{aligned}$$

Proof. By $\mathsf{IS}(\Sigma_0^{(0,k+1)})$, noting that $val_k(\tau_k(x[num(y)/z])) = \overline{1}$ is a $\Sigma_0^{(0,k+1)}$ formula. □

We will define retractability mappings on the basis of the key lemmas 7.11 and 7.12 in more general terms later in Section 7.3. We conclude the present section by two brief remarks.

Remark 7.13. Since, for $k > 1$, $\mathcal{V}_k(T(t))$ and $\neg\mathcal{V}_k(T(t))$ both are $\Sigma_0^{(0,k+1)}$ formulas, we could define $\mathbf{FD}(1,k)$ as

$$\mathbf{A}(1,k+1) \cup \mathsf{FD}(1,k) \cup \mathsf{IS}(\Sigma_0^{(1,k+1)})$$

without affecting the validity of the main theorems of the current chapter.

Remark 7.14. By following the indications that led us to (2.5) from Section 2.20, we have some $\Sigma_1^{(0,3)}$ formula Val (valuation of primitive recursive function symbols) such that, for every $k > 1$, the formula $val_k(x) = y$ is equivalent in $\mathbf{I}(\Sigma_1^{(0,k+1)})$ to the formula $Val(\ulcorner val_k(v)\urcorner[num(x)/\ulcorner v\urcorner], y)$. Moreover, we have

$$\mathbf{I}(\Sigma_1^{(0,3)}) \vdash (\exists y)Val(\ulcorner val_k(v)\urcorner[num(x)/\ulcorner v\urcorner], y)$$

and

$$\begin{aligned}\mathbf{I}(\Sigma_1^{(0,3)}) \vdash\ & Val(\ulcorner val_k(v)\urcorner[num(x)/\ulcorner v\urcorner], y) \wedge \\ & Val(\ulcorner val_k(v)\urcorner[num(x)/\ulcorner v\urcorner], z) \to y = z.\end{aligned}$$

So, $Val(\ulcorner val_k(v)\urcorner[num(x)/\ulcorner v\urcorner], y)$ is $\Delta_1^{(0,3)}$ in $\mathbf{I}(\Sigma_1^{(0,3)})$. (Similarly for all the other primitive recursive function symbols.) We will revert to this observation later in Remark 7.32.

7.3 Retractability: Levels

In the Proposition 7.9, we trivialized the cases for literals $T(t)$ and $\sim T(t)$, respectively. Our next aim is to treat these cases in a non-trivial manner. As we shall see shortly, this gives rise to *leveled* interpretations of truth, where individual levels of interpretation correspond to the number of applications of the rules (FFS6) or (FFS7).

For any $k > 1$, let $\tau_k^1(x)$ be the $\mathcal{L}(0,3)$ function symbol from Proposition 7.9. We now consider $\tau_k^m(x)$, where $m > 1$.

Proposition 7.15. *Let $k > 1$ and $m > 1$. $\mathbf{A}(0, k+1)$ proves, for some function symbol $\tau_k^m(x)$ in $\mathcal{L}(0, k+1)$, the following.*

$$\tau_k^m(x) = y \leftrightarrow \bigvee \big\{(1)\text{–}(13)\big\}$$

where (1)–(13) *are the subsequent disjuncts in which the function symbol $\tau_k^{m-1}(x)$ from $\mathcal{L}(0, k+1)$ is used. We sub-divide these disjuncts into cases for better readability.*

Case A: Non-Formulas

(1) $\neg Form_{1,k}(x) \wedge y = \overline{0}$

Case B: Arithmetical Literals

(2) $(\exists x_0, x_1 < x)\big(x = [x_0 = x_1] \wedge Form_{1,k}(x) \wedge$
$y = \ulcorner \min(\overline{1}, f_{=}(u, v)) \urcorner [x_0, x_1 / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)$

(3) $(\exists x_0, x_1 < x)\big(x = [x_0 \neq x_1] \wedge Form_{1,k}(x) \wedge$
$y = \ulcorner \min(\overline{1}, f_{\neq}(u, v)) \urcorner [x_0, x_1 / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)$

(4) $(\exists x_0, x_1 < x)\big(x = [x_0 \leq x_1] \wedge Form_{1,k}(x) \wedge$
$y = \ulcorner \min(\overline{1}, f_{\leq}(u, v)) \urcorner [x_0, x_1 / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)$

(5) $(\exists x_0, x_1 < x)\big(x = [x_0 > x_1] \wedge Form_{1,k}(x) \wedge$
$y = \ulcorner \min(\overline{1}, f_{>}(u, v)) \urcorner [x_0, x_1 / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)$

Case C: Truth Literals

(6) $(\exists x_0 < x)\big(x = [T(x_0)] \wedge Form_{1,k}(x) \wedge$
$y = \ulcorner \min(\overline{1}, v) \urcorner [\tau_k^{m-1}(val_k(x_0)) / \ulcorner v \urcorner]\big)$

(7) $(\exists x_0 < x)\big(x = [{\sim} T(x_0)] \wedge Form_{1,k}(x) \wedge$
$y = \ulcorner \overline{1} - v \urcorner [\tau_k^{m-1}(val_k(x_0)) / \ulcorner v \urcorner]\big)$

Case D: Connectives and Bounded Quantifiers

(8)
$$\begin{aligned}&(\exists x_0, x_1 < x)\big(x = [x_0 \vee x_1] \wedge Form_{1,k}(x) \wedge \\ &y = \ulcorner \min(\overline{1}, \max(u, v)) \urcorner [\tau_k^m(x_0), \tau_k^m(x_1) / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)\end{aligned}$$

(9)
$$\begin{aligned}&(\exists x_0, x_1 < x)\big(x = [x_0 \wedge x_1] \wedge Form_{1,k}(x) \wedge \\ &y = \ulcorner \min(\overline{1}, \min(u, v)) \urcorner [\tau_k^m(x_0), \tau_k^m(x_1) / \ulcorner u \urcorner, \ulcorner v \urcorner]\big)\end{aligned}$$

(10)
$$\begin{aligned}&(\exists x_0, x_1, x_2, x_3 < x)\big(x = [(\exists x_0) x_1] \wedge Form_{1,k}(x) \wedge \\ &x_1 = [[x_0 \leq x_2] \wedge x_3] \wedge \\ &y = \ulcorner \min(\overline{1}, u) \urcorner [[\max_{x_0 \leq x_2}(\tau_k^m(x_3))] / \ulcorner u \urcorner]\big)\end{aligned}$$

(11)
$$\begin{aligned}&(\exists x_0, x_1, x_2, x_3 < x)\big(x = [(\forall x_0) x_1] \wedge Form_{1,k}(x) \wedge \\ &x_1 = [[x_0 \leq x_2] \rightarrow x_3] \wedge \\ &y = \ulcorner \min(\overline{1}, u) \urcorner [[\min_{x_0 \leq x_2}(\tau_k^m(x_3))] / \ulcorner u \urcorner]\big)\end{aligned}$$

Case E: Unbounded Quantifiers

(12)
$$\begin{aligned}&(\exists x_0, x_1 < x)(\forall x_2, x_3 < x)\big(x = [(\exists x_0) x_1] \wedge Form_{1,k}(x) \wedge \\ &x_1 \neq [[x_0 \leq x_2] \wedge x_3] \wedge y = \ulcorner \overline{1} \urcorner\big)\end{aligned}$$

(13)
$$\begin{aligned}&(\exists x_0, x_1 < x)(\forall x_2, x_3 < x)\big(x = [(\forall x_0) x_1] \wedge Form_{1,k}(x) \wedge \\ &x_1 \neq [[x_0 \leq x_2] \rightarrow x_3] \wedge y = \ulcorner \overline{0} \urcorner\big)\end{aligned}$$

Proof. By applying Corollary 4.17 (for $\mathbf{A}(0, k+1)$). See also Remark 4.18. □

Lemma 7.16. *For $k > 1$ and $m > 1$, the following holds.*

$$\mathbf{A}(0, k+1) \vdash val_k(\tau_k^m(x)) \leq \overline{1}$$

Proof. Like the proof of Lemma 7.10. □

We may now define leveled retractability mappings.

Definition 7.17. The mapping $\mathcal{V}_k^m : \mathcal{L}(1, k+1) \rightarrow \mathcal{L}(0, k+1)$, where $k > 1$ and $m > 0$, is defined recursively as follows.

$$\begin{aligned}
\mathcal{V}_k^m(s = t) &:= (s = t)\\
\mathcal{V}_k^m(s \neq t) &:= (s \neq t)\\
\mathcal{V}_k^m(s \leq t) &:= (s \leq t)\\
\mathcal{V}_k^m(s > t) &:= (s > t)\\
\mathcal{V}_k^m(T(t)) &:= val_k(\tau_k^m(t)) = \overline{1}\\
\mathcal{V}_k^m({\sim}T(t)) &:= val_k(\tau_k^m(t)) = \overline{0}\\
\mathcal{V}_k^m(B \vee C) &:= (\mathcal{V}_k^m(B) \vee \mathcal{V}_k^m(C))\\
\mathcal{V}_k^m(B \wedge C) &:= (\mathcal{V}_k^m(B) \wedge \mathcal{V}_k^m(C))\\
\mathcal{V}_k^m(\exists x B) &:= \exists x(\mathcal{V}_k^m(B))\\
\mathcal{V}_k^m(\forall x B) &:= \forall x(\mathcal{V}_k^m(B))
\end{aligned}$$

Lemma 7.18. *For every $k > 1$ and every $m > 0$, $\mathcal{V}_k^m$ preserves logical structure in $\mathbf{A}(0, k+1)$.*

Proof. By definition, Lemma 7.10, and Lemma 7.16. □

Lemma 7.19. *For every $k > 1$ and every $m > 0$, $\mathcal{V}_k^m(A)$ is the formula A whenever A is an $\mathcal{L}(0, k+1)$ formula.*

Proof. By (meta) induction on the rank of A. □

By using these leveled retractability mappings, we obtain the following as extension of Lemma 7.11. Notice the key difference between 7.20 (3) and the interpretation 6.24 (3) in the context of naive truth.

Lemma 7.20. *For $k > 1$ and $m > 1$, the subsequent formulas are provable in $\mathbf{A}(0, k+1)$.*

(1) $Sent_{1,k}([x = y]) \rightarrow$
$\big(\mathcal{V}_k^m(T([x = y])) \leftrightarrow val_k(x) = val_k(y)\big)$

(2) $Sent_{1,k}([x \leq y]) \rightarrow$
$\big(\mathcal{V}_k^m(T([x \leq y])) \leftrightarrow val_k(x) \leq val_k(y)\big)$

(3) $Sent_{1,k}([T(x)]) \to$
$\big(\mathcal{V}_k^m(T([T(x)])) \leftrightarrow \mathcal{V}_k^{m-1}(T(val_k(x)))\big)$

(4) $Sent_{1,k}(x) \to$
$\big(\mathcal{V}_k^m(T([\neg x])) \leftrightarrow \neg\mathcal{V}_k^m(T(x))\big)$

(5) $Sent_{1,k}([x \vee y]) \to$
$\big(\mathcal{V}_k^m(T([x \vee y])) \leftrightarrow (\mathcal{V}_k^m(T(x)) \vee \mathcal{V}_k^m(T(y)))\big)$

(6) $Sent_{1,k}([(\exists x \leq y)z]) \to$
$\big(\mathcal{V}_k^m(T([(\exists x \leq y)z])) \leftrightarrow (\exists u \leq val_k(y))\mathcal{V}_k^m(T(z[num(u)/x]))\big)$

Proof. Like the proof of Lemma 7.11, but considering also item (3). For this item (3), assume that $Sent_{1,k}([T(x)])$. Notice that $\mathcal{V}_k^m(T([T(x)]))$ is defined as

$$\overline{1} = val_k(\tau_k^m([T(x)])) \tag{7.6}$$

and that $\mathbf{A}(0, k+1)$ proves

$$\tau_k^m([T(x)]) = \ulcorner \min(\overline{1}, v)\urcorner[\tau_k^{m-1}(val_k(x))/\ulcorner v \urcorner].$$

Hence, by Proposition 6.1, we have

$$val_k(\tau_k^m([T(x)])) = \min(\overline{1}, val_k(\tau_k^{m-1}(val_k(x)))).$$

In sum, $\mathbf{A}(0, k+1)$ proves that (7.6) is equivalent to

$$\overline{1} = val_k(\tau_k^{m-1}(val_k(x))),$$

where the latter is $\mathcal{V}_k^{m-1}(T(val_k(x)))$ by definition. □

As for Truth Induction, we reapply the proof of Lemma 7.12.

Lemma 7.21. *For $k > 1$ and $m > 1$, the subsequent formula is provable in $\mathbf{A}(0, k+1)$.*

$$Sent_{1,k}([(\forall z)x]) \to \Big(\big(\mathcal{V}_k^m(T(x[num(\overline{0})/z])) \wedge$$
$$\forall y\big(\mathcal{V}_k^m(T(x[num(y)/z])) \to \mathcal{V}_k^m(T(x[num(y+\overline{1})/z]))\big)\big) \to$$
$$\mathcal{V}_k^m(T(x[num(y)/z]))\Big)$$

Proof. Like the proof of Lemma 7.12. □

We are now ready to estimate proof-theoretic upper bounds.

7.4 Retractability: Upper Bounds

The next lemma may now be proved exactly like in the case of $\mathsf{FD}(1, k)$ above. We omit details, but add two remarks about the derivation of item 7.22 (3): First, Lemma 7.10 and Lemma 7.16 jointly ensure $\mathcal{V}_k^{m-1}({\sim}T(t)) \leftrightarrow {\sim}\mathcal{V}_k^{m-1}(T(t))$; second, joined with equality

$$\tau_k^m([{\sim}T(t)]) = \ulcorner\overline{1} - v\urcorner[\tau_k^{m-1}(val_k(t))/\ulcorner v\urcorner],$$

they also ensure

$$\mathcal{V}_k^m(T([{\sim}T(t)])) \leftrightarrow {\sim}\mathcal{V}_k^m(T([T(t)])).$$

So, we have the following.

Lemma 7.22. *For $k > 1$ and $m > 1$, the subsequent formulas are provable in $\mathbf{A}(0, k+1)$.*

(1) $Sent_{1,k}([x \neq y]) \rightarrow$
$\big(\mathcal{V}_k^m(T([x \neq y])) \leftrightarrow val_k(x) \neq val_k(y)\big)$

(2) $Sent_{1,k}([x > y]) \rightarrow$
$\big(\mathcal{V}_k^m(T([x > y])) \leftrightarrow val_k(x) > val_k(y)\big)$

(3) $Sent_{1,k}([{\sim}T(x)]) \rightarrow$
$\big(\mathcal{V}_k^m(T([{\sim}T(x)])) \leftrightarrow \mathcal{V}_k^{m-1}({\sim}T(val_k(x)))\big)$

(4) $Sent_{1,k}([x \wedge y]) \rightarrow$
$\big(\mathcal{V}_k^m(T([x \wedge y])) \leftrightarrow (\mathcal{V}_k^m(T(x)) \wedge \mathcal{V}_k^m(T(y)))\big)$

(5) $Sent_{1,k}([(\forall x \leq y)z]) \rightarrow$
$\big(\mathcal{V}_k^m(T([(\forall x \leq y)z])) \leftrightarrow (\forall u \leq val_k(y))\mathcal{V}_k^m(T(z[num(u)/x]))\big)$

Lemma 7.20 and Lemma 7.22 may now be employed to show the following. (To allow for a neat presentation, we stipulate that $\leftrightarrow$ binds stronger than $\rightarrow$ in Lemma 7.23.)

Lemma 7.23. *Let $k > 1$ and let $m > 1$. Then, for every $\Sigma_0^{(1,k)}$ formula $A(v_0, \ldots, v_n)$, the following is provable in $\mathbf{A}(0, k+1)$.*

$$\bigwedge\nolimits_{i=0}^{n} ClTerm_k(u_i) \rightarrow$$
$$\mathcal{V}_k^m\big(T(\ulcorner A \urcorner \left[\frac{u_0, \ldots, u_n}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_n \urcorner}\right])\big) \leftrightarrow \mathcal{V}_k^{m-1}\big(A\left(\frac{val_k(u_0), \ldots, val_k(u_n)}{v_0, \ldots, v_n}\right)\big)$$

Proof. By (meta) induction on the rank of A. (We postpone the details of the proof to Chapter 8, Lemma 8.14, where we will be concerned with the so-called Uniform Tarski Schema.)

We consider only one example here: Suppose $A(v_0, \ldots, v_n)$ is a (negative) literal $\sim T(t(v_0, \ldots, v_n))$. Furthermore, assume that $\bigwedge_{i=0}^{n} ClTerm_k(u_i)$. By Proposition 6.1, we have that $\mathbf{A}(0, k+1)$ proves the following.

$$val_k\big(\ulcorner t \urcorner \left[\frac{u_0, \ldots, u_n}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_n \urcorner}\right]\big) = t\left(\frac{val_k(u_0), \ldots, val_k(u_n)}{v_0, \ldots, v_n}\right)$$

Hence, it only remains to apply Lemma 7.22 (3). $\square$

Lemma 7.23 may now be applied to take (locally) care of the inferences (FFS6) and (FFS7) as follows.

Corollary 7.24. *Let $k > 1$ and $m > 1$. Then, for every $\Sigma_0^{(1,k)}$ formula $A(v_0, \ldots, v_n)$, the subsequent inferences are admissible in $\mathbf{A}(0, k+1)$.*

(1)
$$\frac{\mathcal{V}_k^{m-1}\big(A(v_0, \ldots, v_n)\big)}{\bigwedge_{i=0}^{n} ClTerm_k(u_i) \rightarrow \mathcal{V}_k^m\big(T(\ulcorner A \urcorner \left[\frac{u_0, \ldots, u_n}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_n \urcorner}\right])\big)}$$

$$\textbf{(2)}\quad \frac{\bigwedge_{i=0}^{n} ClTerm_k(u_i) \to \mathcal{V}_k^m\big(T(\ulcorner A \urcorner \left[\frac{u_0, \dots, u_n}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_n \urcorner}\right]))}{\mathcal{V}_k^{m-1}\big(A(v_0, \dots, v_n)\big)}$$

Proof. The proof is by means of Lemma 7.23.

- For (1), suppose $\mathbf{A}(0, k+1) \vdash \mathcal{V}_k^{m-1}\big(A(v_0, \dots, v_n)\big)$, that is, we have the following.

  $$\mathbf{A}(0, k+1) \vdash (\forall v_0, \dots v_n)\mathcal{V}_k^{m-1}\big(A(v_0, \dots, v_n)\big) \tag{7.7}$$

  Assume in $\mathbf{A}(0, k+1)$ that $\bigwedge_{i=0}^{n} ClTerm_k(u_i)$. By (7.7), we obtain $\mathcal{V}_k^{m-1}\big(A(val_k(u_0), \dots, val_k(u_n))\big)$. Thus, Lemma 7.23 yields

  $$\mathcal{V}_k^m\big(T(\ulcorner A \urcorner \left[\frac{u_0, \dots, u_n}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_n \urcorner}\right]))$$

  as required.

- For (2), suppose $\mathbf{A}(0, k+1)$ proves the following.

  $$\bigwedge_{i=0}^{n} ClTerm_k(u_i) \to \mathcal{V}_k^m\big(T(\ulcorner A \urcorner \left[\frac{u_0, \dots, u_n}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_n \urcorner}\right])) \tag{7.8}$$

  Since $\mathbf{A}(0, k+1) \vdash \bigwedge_{i=0}^{n} ClTerm_k(num(v_i))$ (reconsider also Proposition 5.15), we obtain by (7.8) the following.

  $$\mathcal{V}_k^m\big(T(\ulcorner A \urcorner \left[\frac{num(v_0), \dots, num(v_n)}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_n \urcorner}\right]))$$

  Since $\mathbf{A}(0, k+1) \vdash \bigwedge_{i=0}^{n} val_k(num(v_i)) = v_i$ (reconsider also Proposition 5.20 (1)), we obtain by Lemma 7.23 that

  $$\mathcal{V}_k^{m-1}\big(A(v_0, \dots, v_n)\big)$$

  as required. □

For $k > 1$ and $n > 0$, write

$$\mathbf{FFS}(1,k) \overset{n}{\vdash} A$$

to mean that A is a theorem of $\mathbf{FFS}(1,k)$, where derivations may involve at most $n-1$ many applications of truth rules, that is, (FFS6) or (FFS7). (For example, if (FFS6) is applied once and (FFS7) is applied once, we count two applications of truth rules and put $\overset{3}{\vdash}$.)

In a nutshell, we have the following key lemma of the present section. (See also a similar argument for the non-finitist case in Halbach [19].)

Lemma 7.25. *For $k > 1$ and $n > 0$, suppose that $\mathbf{FFS}(1,k) \overset{n}{\vdash} A$. Then $\mathbf{A}(0,k+1) \vdash \mathcal{V}_k^n(A)$.*

Proof. By Lemmas 7.11, 7.12, 7.20, and 7.21, we have, for any $n > 0$: If A is some axiom from $\mathsf{FD}(1,k)$ and also $0 < m \leq 2n$, then

$$\mathbf{A}(0,k+1) \vdash \mathcal{V}_k^m(A).$$

Lemma 7.18 guarantees that, for each m such that $0 < m \leq 2n$, $\mathcal{V}_k^m$ preserves logical structure. Let $0 < i < n$. By Corollary 7.24, Lemma 7.19, and preservation of logical structure, we obtain: If, for every m so that $i \leq m \leq (n-i)+n$, the theory $\mathbf{A}(0,k+1)$ proves

$$\mathcal{V}_k^m\big(B(v_0,\ldots,v_r)\big),$$

then, for any m such that $i+1 \leq m \leq (n-(i+1))+n$, it also holds that $\mathbf{A}(0,k+1)$ proves the following.

$$\mathcal{V}_k^m\Big(\bigwedge_{i=0}^{r} ClTerm_k(u_i) \to T(\ulcorner B \urcorner \left[\frac{u_0,\ldots,u_r}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_r \urcorner}\right])\Big)$$

Conversely: If, for every m such that $i \leq m \leq (n-i)+n$, the theory $\mathbf{A}(0,k+1)$ proves that

$$\mathcal{V}_k^m\Big(\bigwedge_{i=0}^{r} ClTerm_k(u_i) \to T(\ulcorner B \urcorner \left[\frac{u_0,\ldots,u_r}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_r \urcorner}\right])\Big)$$

then, for any m such that $i+1 \leq m \leq (n-(i+1))+n$, it also holds that $\mathbf{A}(0,k+1)$ proves the following.

$$\mathcal{V}_k^m\big(B(v_0,\ldots,v_r)\big)$$

Hence, for every i such that $0 < i < n$, if we consider the i-th application of (FFS6) or (FFS7) to the conclusion A, we obtain $\mathbf{A}(0,k+1) \vdash \mathcal{V}_k^n(A)$ in any case. $\square$

7.5 Conclusions

We are now prepared to conclude the main results of the present chapter.

Theorem 7.26. *Let* $k > 1$. *Then* $\mathbf{FD}(1,k)$ *is retractable to* $\mathbf{A}(0,k+1)$ *by* $\mathcal{V}_k$.

Proof. We verify the conditions from Definition 2.26.

- By Lemma 7.18, mapping $\mathcal{V}_k$ preserves logical structure in $\mathbf{A}(0,k+1)$.
- By Lemma 7.19, we have for each $\mathcal{L}(0,k+1)$ atom A that $\mathcal{V}_k(A)$ is the formula A.
- By Lemma 7.11, Lemma 7.12, and preservation of logical structure, we have for every B in $\mathsf{FD}(1,k) \cup \mathsf{TI}(1,k)$ that

$\mathbf{A}(0, k+1) \vdash \mathcal{V}_k(B)$. Furthermore, by preservation of logical structure, we obtain for (additional logical axioms) B in $\mathbf{A}(1, k+1)$ that $\mathbf{A}(0, k+1) \vdash \mathcal{V}_k(B)$.

- $\mathbf{A}(0, k+1)$ is a sub-theory of $\mathbf{FD}(1, k)$.

This concludes the proof. $\square$

Remark 7.27. For $k > 1$, we also mention the following.

- By Proposition 2.31, the mapping $(\mathcal{A} \cdot \mathcal{V}_k)$ preserves logical structure in $\mathbf{B}(0, k+1)$.
- By Corollary 3.64 (1), we obtain for every axiom B from $\mathsf{FD}(1, k) \cup \mathsf{TI}(1, k)$ that $\mathbf{B}(0, k+1) \vdash \mathcal{A}(\mathcal{V}_k(B))$.

Hence, for $k > 1$, $\mathbf{FD}(1, k)$ is retractable also to $\mathbf{B}(0, k+1)$ by $(\mathcal{A} \cdot \mathcal{V}_k)$.

By Proposition 2.27, the first corollary is the following.

Corollary 7.28. *Let $k > 1$. Then, for any $\mathcal{L}(0, k+1)$ formula A and $\mathcal{L}(1, k+1)$ formula B, the following hold.*

(1) $\mathbf{FD}(1, k) \vdash B$ *implies* $\mathbf{A}(0, k+1) \vdash \mathcal{V}_k(B)$.

(2) $\mathbf{FD}(1, k) \vdash A$ *if and only if* $\mathbf{A}(0, k+1) \vdash A$.

Remark 7.29. For $k > 1$, we also mention the following.

- For every $\mathcal{L}(1, \mathrm{Q})$ formula B, we have by Corollary 3.64 (2) that
$$\mathbf{B}(1, k) \cup \big(\mathsf{FD}(1, k) \cup \mathsf{TI}(1, k)\big)^{\mathcal{A}} \vdash B$$
implies $\mathbf{FD}(1, k) \vdash B$ and, by Remark 7.27, implies
$$\mathbf{B}(0, k+1) \vdash \mathcal{A}(\mathcal{V}_k(B)).$$

- For every $\mathcal{L}(0,\mathrm{Q})$ formula A, we have by Corollary 3.64 (2) and Proposition 7.5 that
$$\mathbf{B}(1,k) \cup \big(\mathsf{FD}(1,k) \cup \mathsf{TI}(1,k)\big)^{\mathcal{A}} \vdash A$$
precisely if $\mathbf{FD}(1,k) \vdash A$ and, by Remark 7.27, precisely if
$$\mathbf{B}(0,k+1) \vdash A.$$

As for the Finitist Friedman-Sheard Theories, we have finally arrived at the following.

Theorem 7.30. *Let $k > 1$ and let $n > 0$. Then, for every $\mathcal{L}(0,k+1)$ formula A and $\mathcal{L}(1,k+1)$ formula B, the following hold.*

(1) $\mathbf{FFS}(1,k) \,\overset{n}{\mid\!\!-}\, B$ *implies* $\mathbf{A}(0,k+1) \vdash \mathcal{V}^n_k(B)$.

(2) $\mathbf{FFS}(1,k) \,\overset{n}{\mid\!\!-}\, A$ *if and only if* $\mathbf{A}(0,k+1) \vdash A$.

Proof. (1) is covered by Lemma 7.25. (2) results as in the proof of Theorem 7.26. □

Remark 7.31. Clearly, a counterpart of Remark 7.29 also applies to Theorem 7.30.

Remark 7.32. By Remark 7.14, we obtain a limiting result for some (local) retractability mapping $\mathcal{W}^n : \mathcal{L}(1,\omega) \to \mathcal{L}(0,\mathrm{Q})$ that we do not specify in more detail here. Namely, using the unified language and base theory

$$\mathcal{L}(1,\omega) := \bigcup_{1<k<\omega} \mathcal{L}(1,k+1)$$
$$\mathbf{A}(1,\omega) := \bigcup_{1<k<\omega} \mathbf{A}(1,k+1)$$

that comprise, respectively, symbols and axioms for all primitive recursive functions, we may consider the extension $\mathbf{FFS}(1,\omega)$ of $\mathbf{I}(\Sigma_1)$ defined as follows.

$$\mathbf{A}(1,\omega) \cup \mathsf{IS}(\Sigma_1^{(1,\omega)}) \cup \bigcup_{1<k<\omega} \mathsf{FFS}(1,k)$$

Assuming $n > 0$, we then have, for any $\mathcal{L}(0,\mathrm{Q})$ formula A and $\mathcal{L}(1,\omega)$ formula B, the following.

(1) $\mathbf{FFS}(1,\omega) \overset{n}{\models} B$ implies $\mathbf{I}(\Sigma_1) \vdash \mathcal{W}^n(B)$.

(2) $\mathbf{FFS}(1,\omega) \overset{n}{\models} A$ if and only if $\mathbf{I}(\Sigma_1) \vdash A$.

Observe that Theorem 7.30 (2) also yields the consistency of $\mathbf{FFS}(1,k)$. Suppose (towards absurdity) that $\mathbf{FFS}(1,k) \vdash \overline{0} \neq \overline{0}$. Then, for some $n > 0$, we have $\mathbf{FFS}(1,k) \overset{n}{\models} \overline{0} \neq \overline{0}$ and, by the Theorem 7.30 (2), also $\mathbf{A}(0,k+1) \vdash \overline{0} \neq \overline{0}$, where we know that actually $\mathbf{A}(0,k+1) \nvdash \overline{0} \neq \overline{0}$. Therefore, $\mathbf{FFS}(1,k) \nvdash \overline{0} \neq \overline{0}$.

That is, we see that $\mathbf{FFS}(1,k)$ has a model $(\mathfrak{M}, S)$, where S serves as the semantic extension of the truth predicate. For the present purpose, we may conceive all the members (Γ, A) of $\mathsf{FFS}(1,k)$ as axioms (of sort): If $\Gamma = \varnothing$ is empty, we may regard A as an *unconditional* axiom that may be added to a derivation without restriction. On the other hand, if $\Gamma \neq \varnothing$, we may regard A as *conditional* axiom that may be added to a derivation only if the latter already comprises all the formulas in Γ. Given this understanding, it makes sense to say that all axioms in

$$\mathsf{FFS}(1,k) \cup \mathsf{TI}(1,k)$$

are Σ_0 formulas in language $\mathcal{L}(1,k+1)$. (As for $\mathsf{TI}(1,k)$, just use the equivalent bounded version by reasoning as in Remark 3.5.)

Therefore, we are in a position to rely on the model-theoretic fact that the Σ_0 fragment of $\mathcal{L}(1,k+1)$ is *absolute* (see, for example, Kaye [30]) in the following sense: For every Σ_0 formula

$A(x_0, \dots, x_r)$ in the language $\mathcal{L}(1, k+1)$ and any given standard numbers $n_0, \dots, n_r$ in the domain ω of the standard model $\mathbb{N}$ of arithmetic (where $\mathbb{N}$ is a sub-model of $\mathfrak{M}$), we have

$$(\mathfrak{M}, S) \vDash A(n_0, \dots, n_r)$$

if and only if

$$(\mathbb{N}^+, S \cap \omega) \vDash A(n_0, \dots, n_r)$$

where $\mathbb{N}^+$ is the extension of $\mathbb{N}$ that also interprets the auxiliary function symbols and $S \cap \omega$ is the restriction of S to ω.

Therefore, by considering the sub-model $(\mathbb{N}^+, S \cap \omega)$ of $(\mathfrak{M}, S)$, we see that $\mathbf{FFS}(1, k)$ has a standard model. (We do not have to worry about the more complex collection schema $\mathsf{BS}(\Sigma_1)$ from $\mathbf{A}(1, k+1)$ because it is known that it holds in $\mathbb{N}$.) But this establishes that $\mathbf{FFS}(1, k)$ is *ω-consistent*: There is no $\mathcal{L}(1, k+1)$ formula $A(x)$ such that, for every $n \in \omega$, we have

$$\mathbf{FFS}(1, k) \vdash A(\overline{n})$$

but also the following.

$$\mathbf{FFS}(1, k) \vdash \exists x \neg A(x)$$

So, we may conclude the following.

Corollary 7.33. *For $k > 1$, $\mathbf{FFS}(1, k)$ is ω-consistent.*

Hence, the situation in the finitist case is drastically different as for the *ω-inconsistent* Friedman-Sheard theory (in unrestricted form) investigated by Halbach in [19], where the ω-inconsistency result is attributed to Cantini and related to a theorem of McGee

[36] which, in turn, bears a strong resemblance to the paradox of Yablo [51].

The ω-inconsistency of the unrestricted FS theory is basically due to the fact that, while the theory includes the truth axiom (7.5) for (theory internal) unbounded universal quantification, the (unrestricted) truth rules can be (from a theory external point of view) applied only a standard number of times in a derivation. Thus, for a diagonalization statement such that, provably,

$$D \leftrightarrow \exists x \neg T(f(x, \ulcorner D \urcorner)),$$

where $f(\overline{n}, \ulcorner D \urcorner)$ denotes a code of sentence

$$\underbrace{T([T(\cdots[T}_{n \text{ times}}(\ulcorner D \urcorner)]\cdots)]),$$

one may derive, for every $n \in \omega$, that $T(f(\overline{n}, \ulcorner D \urcorner))$, but also that $\exists x \neg T(f(x, \ulcorner D \urcorner))$. For a detailed exposition, the reader is also referred to Halbach [21], Theorem 13.9.

8 Finitist Kripke-Feferman

In this chapter, we investigate a finitist variant of the so-called Kripke-Feferman (KF) axiomatization of truth. In original form, the KF axiomatization was given by Reinhardt [42] and Feferman [12]. A list of references for the philosophical motivation of the KF axiomatization is given in Halbach [21]: Halbach and Horsten [22] and Burgess [3], for example, offer detailed discussions.

Notation: Since the variant of KF that we will study in this chapter is an axiomatization of type-free truth in the language $\mathcal{L}(1, \mathrm{Q})$ with only *one single* truth predicate, we (still) let

$$T(x) := T_1(x)$$

and stipulate that $Sent(x)$, $Term(x)$, and so on, are the $\mathcal{L}(0, 3)$ formulas from Section 5.1 that express, respectively, that x is an $\mathcal{L}(1, \mathrm{Q})$ sentence, that x is an $\mathcal{L}(0, \mathrm{Q})$ term, and so on.

Now, as indicated in Section 1.4, and in contrast to the FH axiomatization from Chapter 6 and the FFS axiomatization from Chapter 7, the KF axiomatization relies on the idea that, in the presence of arbitrarily many symbols (via the Δ_1 definitions) for primitive recursive functions, one may reason about truth consistently by neglecting the classical negation principle that, for any arbitrary statement A:

$$\neg T(\ulcorner A \urcorner) \leftrightarrow T(\ulcorner \neg A \urcorner)$$

We will highlight how this negation principle, against the background of KF, relates to Tarski's Theorem and Liar Paradox in more detail later in Section 10.1. At this point, we only want to see how an axiomatization of truth is affected by the lack of the negation principle.

If the negation axiom is missing, *untruth* ($\neg T(x)$) and *falsity* ($T([\neg x])$) come apart, which is why one may have to axiomatize them individually. For example, the original formulations of KF were given in a language comprising an unary predicate symbol $T(x)$ for truth and $F(x)$ for falsity, with an axiom so that

$$F(\ulcorner A \urcorner) \leftrightarrow T(\ulcorner \neg A \urcorner)$$

to fix their relationship. Accordingly, axioms came in pairs: for example, an axiom so that

$$T(\ulcorner A \land B \urcorner) \leftrightarrow (T(\ulcorner A \urcorner) \land T(\ulcorner B \urcorner))$$

for the truth of conjunctions and an axiom so that

$$F(\ulcorner A \land B \urcorner) \leftrightarrow (F(\ulcorner A \urcorner) \lor F(\ulcorner B \urcorner))$$

for the falsity of conjunctions.

In our presentation, we will follow Halbach [21] and avoid the need of the symbol $F(x)$ by denoting falsity of x as $T([\neg x])$. In addition, by our treatment of negation as a non-primitiv concept, certain axioms for falsity do not have to be stated explicitly, but may rather be derived from the respective axioms for truth. This also has the advantage that our KF axioms can nicely reflect our recursive definition of formulas given on the meta-theoretic level. Furthermore, by Proposition 5.19, we do not have to take care specifically of double negations.

(While our decision of defining negation as non-primitive was occasionally uncomfortable in the two preceding chapters, it will turn out to be rather convenient in the following chapters.)

While KF does not feature the negation principle, it includes the iteration principle that, for arbitrary closed terms t:

$$T(t) \leftrightarrow T(\ulcorner T(t) \urcorner)$$

However, in the absence of the negation principle, the iteration principle does not entail (the naive theorem 6.9 (3) from Chapter 6) that:

$$\sim T(t) \leftrightarrow T(\ulcorner \sim T(t) \urcorner) \tag{8.1}$$

Rather, the respective KF axiom takes the negation into the scope of the truth axiom so that only

$$(T([\neg t]) \vee \neg Sent(t)) \leftrightarrow T(\ulcorner \sim T(t) \urcorner) \tag{8.2}$$

where the idea is that, if t is not a sentence ($\neg Sent(t)$), it may not be true after all. It is crucial for axiom (8.2) that (besides $\neg Sent(t)$) negation symbols appear only inside the scope of the truth predicate and that the truth predicate itself appears only unnegated (positively) on both sides of the biconditional. Indeed, we shall see in Section 10.1 that (8.1) is not consistent with the finitist KF theory to be defined and studied in this chapter. (See Remark 10.7.)

Finally, let us mention that the semantics of KF conforms to (the strong Kleene variant of) a well-known outline of a theory of truth by Kripke [32]. We shall have the occasion to say a little more on the semantics of KF in Section 8.3 below.

8.1 The Theory

We first express the (unrestricted) KF axioms according to how negation is treated in this study. It is a peculiarity of these KF axioms that they involve the truth predicate only positively (unnegated) on both sides of the biconditional ($\leftrightarrow$). We shall turn back to this peculiarity in more detail later.

Definition 8.1. $\mathsf{KF}(1, \mathrm{Q})$ comprises the subsequent (unrestricted) *Kripke-Feferman* axioms for the whole language $\mathcal{L}(1, \mathrm{Q})$. We subdivide the axioms into axiom sets for better readability.

Axiom Set A: Arithmetical Literals

(KF1) $Sent([x = y]) \rightarrow$
$\big(T([x = y]) \leftrightarrow val(x) = val(y)\big)$

(KF2) $Sent([x \neq y]) \rightarrow$
$\big(T([x \neq y]) \leftrightarrow val(x) \neq val(y)\big)$

(KF3) $Sent([x \leq y]) \rightarrow$
$\big(T([x \leq y]) \leftrightarrow val(x) \leq val(y)\big)$

(KF4) $Sent([x > y]) \rightarrow$
$\big(T([x > y]) \leftrightarrow val(x) > val(y)\big)$

Axiom Set B: Truth Literals

(KF5) $Sent([T(x)]) \rightarrow$
$\big(T([T(x)]) \leftrightarrow T(val(x))\big)$

(KF6) $Sent([{\sim}T(x)]) \rightarrow$
$\big(T([{\sim}T(x)]) \leftrightarrow (T([\neg val(x)]) \vee \neg Sent(val(x)))\big)$

Axiom Set C: Connectives and Quantifiers

(KF7) $Sent([x \vee y]) \rightarrow$
$\big(T([x \vee y]) \leftrightarrow (T(x) \vee T(y))\big)$

(KF8) $Sent([x \wedge y]) \rightarrow$
$\big(T([x \wedge y]) \leftrightarrow (T(x) \wedge T(y))\big)$

(KF9) $Sent([(\exists x)y]) \rightarrow$
$\big(T([(\exists x)y]) \leftrightarrow \exists u T(y[num(u)/x])\big)$

(KF10) $Sent([(\forall x)y]) \rightarrow$
$\big(T([(\forall x)y]) \leftrightarrow \forall u T(y[num(u)/x])\big)$

We will show how to derive respective falsity axioms only in connection with the finitist restriction of $\mathsf{KF}(1, \mathrm{Q})$ later.

Remark 8.2. We first sketch why (unrestricted) $\mathsf{KF}(1, \mathrm{Q})$ cannot be retractable to a consistent theory of first-order arithmetic at all. Suppose towards an absurdity that $\mathsf{KF}(1, \mathrm{Q})$ was retractable to some theory S of first-order arithmetic. Then, by extension of Lemma 8.14 below and by using also Proposition 2.27 (1) under preservation of logical structure, S may prove, for every $\mathcal{L}(0, \mathrm{Q})$ sentence A, the instance of the Tarski Schema that

$$A \leftrightarrow T(\ulcorner A \urcorner)^\circ,$$

where $T(x)^\circ$ is $\mathcal{L}(0, \mathrm{Q})$ formula to interpret the truth predicate. However, by diagonalization via unbounded existential quantification, there is some $\mathcal{L}(0, \mathrm{Q})$ Liar D such that S proves

$$D \leftrightarrow \neg T(\ulcorner D \urcorner)^\circ.$$

Consequently, S must be inconsistent.

Actually, if $\mathbf{A}(1, 3)$ is augmented by the (unrestricted) axioms from $\mathsf{KF}(1, \mathrm{Q})$ and Truth Induction (defined in 8.6), the resulting system exceeds first-order arithmetic as it reaches the proof-theoretic strength (in terms of provable arithmetical theorems) of Ramified Analysis $\mathbf{RA}_{<\omega^\omega}$ up to level ω^ω. Specifically, this unrestricted KF theory proves the consistency of Peano Arithmetic. (Consult Cantini [6], Paragraph 3.)

To get a finitist variant of the KF axiomatization, we remove – as a tempting move – axiom (KF10) for *universal* quantification and replace it by an axiom for *bounded* universal quantification.

Observe that, in the presence of (KF10), we may establish a dual of Proposition 8.4 below, saying that we do not have to add an axiom for bounded universal quantification as well, as it may be derived. However, if (KF10) is missing, the axiom for bounded universal quantification no longer is redundant, but does neither violate finitist standards, which is why we add it. This leaves us with the type-free KF axiomatization of *essential* Σ_1 *truth in the language of truth.*

While FH from Chapter 6 and FFS from Chapter 7 both are axiomatizations of decidable truth, the finitist KF axiomatization rather is in line with the *asymmetric* concept of *semi-decidability* (or algorithmic verification): Semi-decidability of a problem does not imply the semi-decidability of the problem's complement.

Since finitist KF may be seen as an axiomatization of semi-decidable (recursively enumerable) truth, it must also be closely tied to the conception of *derivability* in a formal system. For a short note on such interplay between truth and derivability, the reader may also consider – as kindly suggested to us by Andrea Cantini – Myhill's "system which can define its own truth" [38].

Definition 8.3. $\mathsf{FKF}(1, \mathrm{Q})$ comprises the subsequent *Finitist Kripke-Feferman* axioms for the $\mathrm{E}\Sigma_1$ fragment of $\mathcal{L}(1, \mathrm{Q})$. The axioms are sub-divided into axiom sets for better readability.

Axiom Set A: Arithmetical Literals

(FKF1) $Sent([x = y]) \rightarrow \big(T([x = y]) \leftrightarrow val(x) = val(y)\big)$

(FKF2) $Sent([x \neq y]) \rightarrow \big(T([x \neq y]) \leftrightarrow val(x) \neq val(y)\big)$

(FKF3) $Sent([x \leq y]) \rightarrow \big(T([x \leq y]) \leftrightarrow val(x) \leq val(y)\big)$

(FKF4) $\mathit{Sent}([x > y]) \rightarrow$
$\big(T([x > y]) \leftrightarrow \mathit{val}(x) > \mathit{val}(y)\big)$

Axiom Set B: Truth Literals

(FKF5) $\mathit{Sent}([T(x)]) \rightarrow$
$\big(T([T(x)]) \leftrightarrow T(\mathit{val}(x))\big)$

(FKF6) $\mathit{Sent}([{\sim}T(x)]) \rightarrow$
$\big(T([{\sim}T(x)]) \leftrightarrow (T([\neg \mathit{val}(x)]) \vee \neg \mathit{Sent}(\mathit{val}(x)))\big)$

Axiom Set C: Connectives

(FKF7) $\mathit{Sent}([x \vee y]) \rightarrow$
$\big(T([x \vee y]) \leftrightarrow (T(x) \vee T(y))\big)$

(FKF8) $\mathit{Sent}([x \wedge y]) \rightarrow$
$\big(T([x \wedge y]) \leftrightarrow (T(x) \wedge T(y))\big)$

Axiom Set D: Bounded Universal Quantification

(FKF9) $\mathit{Sent}([(\forall x \leq y)z]) \rightarrow$
$\big(T([(\forall x \leq y)z]) \leftrightarrow (\forall u \leq \mathit{val}(y))T(z[\mathit{num}(u)/x])\big)$

Axiom Set E: Unbounded Existential Quantification

(FKF10) $\mathit{Sent}([(\exists x)y]) \rightarrow$
$\big(T([(\exists x)y]) \leftrightarrow \exists u T(y[\mathit{num}(u)/x])\big)$

To see that the dual of (FKF9) does not have to be added as axiom, let us show how to derive it.

Proposition 8.4. $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ *proves the following.*

$$\mathit{Sent}([(\exists x \leq y)z]) \rightarrow$$

$$\big(T([(\exists x \leq y)z]) \leftrightarrow (\exists u \leq val(y))T(z[num(u)/x])\big)$$

Proof. Assume in $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ that $Sent([(\exists x \leq y)z])$. By axiom (FKF10), we have the following.

$$T([(\exists x \leq y)z]) \leftrightarrow \exists u T([[x \leq y] \wedge z][num(u)/x])$$

where $T([(\exists x \leq y)z])$ actually is $T([(\exists x)[[x \leq y] \wedge z]])$. Also, by Proposition 5.10, we have in $\mathbf{A}(1,3)$ that

$$[[x \leq y] \wedge z][num(u)/x] = [[num(u) \leq y] \wedge z[num(u)/x]].$$

Hence, by (FKF8), (FKF3), and Proposition 5.20 (1), we may derive that

$$\begin{aligned} T([(\exists x \leq y)z]) &\leftrightarrow \exists u(T([num(u) \leq y]) \wedge T(z[num(u)/x])) \\ &\leftrightarrow (\exists u \leq val(y))T(z[num(u)/x])), \end{aligned}$$

as required. □

As announced, we now derive usual statements about falsity ($T([\neg x])$) and double negation.

Proposition 8.5. $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ *proves the following.*

(1) $Sent(x) \rightarrow$
$\big(T([\neg[\neg x]]) \leftrightarrow T(x)\big)$

(2) $Sent([x \vee y]) \rightarrow$
$\big(T([\neg[x \vee y]]) \leftrightarrow (T([\neg x]) \wedge T([\neg y]))\big)$

(3) $Sent([x \wedge y]) \rightarrow$
$\big(T([\neg[x \wedge y]]) \leftrightarrow (T([\neg x]) \vee T([\neg y]))\big)$

(4) $Sent([(\exists x \leq y)z]) \rightarrow$
$\big(T([\neg[(\exists x \leq y)z]]) \leftrightarrow (\forall u \leq val(y))T([\neg z[num(u)/x]])\big)$

(5) $Sent([(\forall x \leq y)z]) \rightarrow$
$\big(T([\neg[(\forall x \leq y)z]]) \leftrightarrow (\exists u \leq val(y))T([\neg z[num(u)/x]])\big)$

(6) $Sent([(\forall x)y]) \to$
$\big(T([\neg[(\forall x)y]]) \leftrightarrow \exists u T([\neg y[num(u)/x]])\big)$

Proof. (1) is by Proposition 5.19 and logic. For (2), we may use Proposition 5.11 (4) and (FKF8). For (3)–(6), we may proceed accordingly. ☐

Like all theories of truth considered before, finitist KF theory includes some axiom of Truth Induction, but this time only for language $\mathcal{L}(1, \mathrm{Q})$.

Definition 8.6. $\mathsf{TI}(1, \mathrm{Q})$ contains the following (single) axiom of *Truth Induction*.

$$\begin{aligned}&Sent([(\forall z)x]) \to \Big(\big(T(x[num(\overline{0})/z]) \land \\ &\forall y(T(x[num(y)/z]) \to T(x[num(y+\overline{1})/z]))\big) \to \\ &T(x[num(y)/z])\Big)\end{aligned}$$

The Finitist Kripke-Feferman Theory may now be defined as follows.

Definition 8.7. $\mathbf{FKF}(1, \mathrm{Q})$ is *Finitist Kripke-Feferman Theory* (in auxiliary form) defined as follows.

$$\mathbf{A}(1,3) \cup \mathsf{FKF}(1, \mathrm{Q}) \cup \mathsf{TI}(1, \mathrm{Q})$$

Remark 8.8. Similar remarks about the intuitions behind type-free truth as in 6.13 and 7.4 also apply for $\mathbf{FKF}(1, \mathrm{Q})$ and its respective $\mathcal{L}(1, \mathrm{Q})$ variant:

$$\mathbf{B}(1,3) \cup \big(\mathsf{FKF}(1, \mathrm{Q}) \cup \mathsf{TI}(1, \mathrm{Q})\big)^{\mathcal{A}}$$

Here $\mathcal{A}$ is the mapping

$$\mathcal{A} : \mathcal{L}(1,3) \to \mathcal{L}(1,\mathrm{Q})$$

defined in 3.53 for eliminating auxiliary function symbols.

8.2 Retractability: Lower Bound

As in earlier chapters, our goal is to prove that $\mathbf{FKF}(1,\mathrm{Q})$ is retractable to (an auxiliary extension of) $\mathbf{I}(\Sigma_1)$. This requires that we define (as preliminary step) some mapping

$$\circ : \mathcal{L}(1,3) \to \mathcal{L}(0,3)$$

which satisfies all conditions from Definition 2.26. We consider 2.26 (3) first. That is, we first unveil the proof-theoretic power of the truth axioms of $\mathbf{FKF}(1,\mathrm{Q})$.

We start off by defining a new complexity class in accordance with our definition of $\mathsf{FKF}(1,\mathrm{Q})$.

Definition 8.9. For $m > 0$, class $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$ of $\mathrm{E}\Sigma_1^{(m,\mathrm{Q})}$ formulas that are *positive in the truth predicate* T_m is defined recursively, where t is $\mathcal{L}(0,\mathrm{Q})$ term.

- $\mathcal{L}(m-1,\mathrm{Q})$ literals are $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$.
- Positive $\mathcal{L}(m,\mathrm{Q})$ literals $T_m(t)$ are $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$.
- If B, C are $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$, then so is $(B \vee C)$.
- If B, C are $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$, then so is $(B \wedge C)$.
- If B is $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$, then so is $(\forall x \leq t)B$.
- If B is $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$, then so is $\exists x B$.

In accordance with Proposition 8.4, we have:

Proposition 8.10. *If $m > 0$ and B is $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$, also $(\exists x \leq t)B$ is $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$, where t is $\mathcal{L}(0,\mathrm{Q})$ term.*

Proof. As $(x \leq t)$ is $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$ by first item, we may rely on the respective items for conjunction and existential quantification. □

Proposition 8.11 below is proved by (meta) induction on the rank of A. The following Corollary 8.12 is a consequence of 8.11. We omit further details, but will sometimes tacitly rely on such observations in what follows.

Proposition 8.11. *For $m > 0$, the following hold.*

- *If A is $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$, then A is $\mathrm{PE}\Sigma_1^{(m+1,\mathrm{Q})}$.*
- *If A is $\mathrm{E}\Sigma_1^{(m,\mathrm{Q})}$, then A is $\mathrm{PE}\Sigma_1^{(m+1,\mathrm{Q})}$.*

Corollary 8.12. *For $m > 0$, the following hold.*

- *If $0 < i \leq m$ and A is $\mathrm{PE}\Sigma_1^{(i,\mathrm{Q})}$, then A is $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$.*
- *If $i < m$ and A is $\mathrm{E}\Sigma_1^{(i,\mathrm{Q})}$, then A is $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$.*

We now turn to the verification of condition 2.26 (3). For the next lemma, we define a uniform version of the Tarski Schema in accordance with Proposition 5.21 for closed terms.

Definition 8.13. Let Γ be a class of formulas. The *Uniform Tarski Schema* $\mathsf{UTS}(\Gamma)$ is

$$\bigwedge_{i=0}^{k} ClTerm(u_i) \rightarrow$$

$$\left(T(\ulcorner A \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right]) \leftrightarrow A\left(\frac{val(u_0), \ldots, val(u_k)}{v_0, \ldots, v_k}\right)\right)$$

where $A(v_0, \ldots, v_k)$ belongs to Γ.

The following Lemma 8.14 is some modification of Cantini [6], Lemma 3.2.

Lemma 8.14. *For every formula A from* $\mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})})$, *we have that* $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ *proves A.*

Proof. Assume in $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ that $\bigwedge_{i=0}^{k} ClTerm(u_i)$. By (meta) induction on the rank of the formula A, we show that the biconditional

$$T(\ulcorner A \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right]) \leftrightarrow A\left(\frac{val(u_0),\ldots,val(u_k)}{v_0,\ldots,v_k}\right)$$

may be proved in $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$. To keep our proof at a reasonable length, we focus on a few selected cases, since many cases are just dual to each other.

- A is $(s = t)(v_0,\ldots,v_k)$. By Proposition 5.10, we obtain in $\mathbf{A}(1,3)$ that

  $$\ulcorner s = t \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right] \tag{8.3}$$

  is identical (=) to

  $$[\ulcorner s \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right] = \ulcorner t \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right]].$$

  Therefore, we have by logic that

  $$T(\ulcorner s = t \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right])$$

  is equivalent ($\leftrightarrow$) to

  $$T([\ulcorner s \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right] = \ulcorner t \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right]]). \tag{8.4}$$

By Proposition 5.18 (5), we have in $\mathbf{A}(1,3)$ that (8.3) is a sentence. Thus, by (FKF1), we have in $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ that (8.4) is equivalent ($\leftrightarrow$) to

$$val(\ulcorner s \urcorner \left[\frac{u_0, \dots, u_k}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner}\right]) = val(\ulcorner t \urcorner \left[\frac{u_0, \dots, u_k}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner}\right]). \quad (8.5)$$

By Proposition 5.21, (8.5) is equivalent ($\leftrightarrow$) to

$$s\left(\frac{val(u_0), \dots, val(u_k)}{v_0, \dots, v_k}\right) = t\left(\frac{val(u_0), \dots, val(u_k)}{v_0, \dots, v_k}\right)$$

which by (meta) definition is the formula

$$(s = t)\left(\frac{val(u_0), \dots, val(u_k)}{v_0, \dots, v_k}\right).$$

- A is $T(t)(v_0, \dots, v_k)$. Similarly, by using (FKF5).
- A is $(B \vee C)(v_0, \dots, v_k)$. By the induction hypothesis and logical reasoning, $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ proves

$$T(\ulcorner B \urcorner \left[\frac{u_0, \dots, u_k}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner}\right]) \vee T(\ulcorner C \urcorner \left[\frac{u_0, \dots, u_k}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner}\right]) \quad (8.6)$$

to be equivalent ($\leftrightarrow$) to

$$B\left(\frac{val(u_0), \dots, val(u_k)}{v_0, \dots, v_k}\right) \vee C\left(\frac{val(u_0), \dots, val(u_k)}{v_0, \dots, v_k}\right)$$

which is formula

$$(B \vee C)\left(\frac{val(u_0), \dots, val(u_k)}{v_0, \dots, v_k}\right)$$

by (meta) definition. By Proposition 5.18 (5), we can derive in $\mathbf{A}(1,3)$ that

$$\ulcorner B \vee C \urcorner \left[\frac{u_0, \dots, u_k}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner}\right] \quad (8.7)$$

is a sentence. Moreover, by Proposition 5.10, $\mathbf{A}(1,3)$ proves (8.7) to be identical ($=$) to

$$[\ulcorner B\urcorner \left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right] \vee \ulcorner C\urcorner \left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right]].$$

Therefore, it only remains to note that, by (FKF7), we can show in $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ that formula (8.6) is equivalent ($\leftrightarrow$) to formula

$$T([\ulcorner B\urcorner \left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right] \vee \ulcorner C\urcorner \left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right]]).$$

- A is $((\forall x \leq t)B)(v_0,\dots,v_k)$, which is to say, A is formula $(\forall x \leq t(v_0,\dots,v_k))B(v_0,\dots,v_k,x)$. By 5.15 and 5.20 (1), we get in $\mathbf{A}(1,3)$ that $ClTerm(num(x))$ and $val(num(x)) = x$, for arbitrary x. So, by induction hypothesis and logic, we have in $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ that

$$(\forall x \leq t\left(\frac{val(u_0),\dots,val(u_k)}{v_0,\dots,v_k}\right))T(\ulcorner B\urcorner \left[\frac{u_0,\dots,u_k,num(x)}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner,\ulcorner x\urcorner}\right])$$

is equivalent ($\leftrightarrow$) to

$$(\forall x \leq t\left(\frac{val(u_0),\dots,val(u_k)}{v_0,\dots,v_k}\right))B\left(\frac{val(u_0),\dots,val(u_k)}{v_0,\dots,v_k}\right)$$

which is formula

$$((\forall x \leq t)B)\left(\frac{val(u_0),\dots,val(u_k)}{v_0,\dots,v_k}\right)$$

by (meta) definition. By Proposition 5.21, we get in $\mathbf{A}(1,3)$ that

$$t\left(\frac{val(u_0),\dots,val(u_k)}{v_0,\dots,v_k}\right) = val(\ulcorner t\urcorner \left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right]).$$

Moreover, by Proposition 5.18 (5), we can derive in $\mathbf{A}(1,3)$ that

$$\ulcorner(\forall x \leq t)B\urcorner\left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right]$$

is a sentence which, by Proposition 5.10, is identical $(=)$ to

$$[(\forall x \leq \ulcorner t\urcorner\left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right])\ulcorner B\urcorner\left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right]].$$

Consequently, it only remains to note that, by (FKF9), we obtain in $\mathbf{A}(1,3)\cup\mathsf{FKF}(1,\mathrm{Q})$ that

$$T([(\forall x \leq \ulcorner t\urcorner\left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right])\ulcorner B\urcorner\left[\frac{u_0,\dots,u_k}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right]])$$

is equivalent $(\leftrightarrow)$ to

$$(\forall x \leq t\left(\frac{val(u_0),\dots,val(u_k)}{v_0,\dots,v_k}\right))T(\ulcorner B\urcorner\left[\frac{u_0,\dots,u_k,num(x)}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner,\ulcorner x\urcorner}\right])$$

from above.

Other (dual) cases are treated accordingly. □

Corollary 8.15. *For any* $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ *formula* $A(v_0,\dots,v_k)$*, we have that* $\mathbf{A}(1,3)\cup\mathsf{FKF}(1,\mathrm{Q})$ *proves the following.*

$$T(\ulcorner A\urcorner\left[\frac{num(v_0),\dots,num(v_k)}{\ulcorner v_0\urcorner,\dots,\ulcorner v_k\urcorner}\right]) \leftrightarrow A(v_0,\dots,v_k)$$

Proof. By Lemma 8.14 in combination with Proposition 5.15 and Proposition 5.20 (1). □

By the next lemma, condition 2.26 (3) is verified.

Lemma 8.16. *For every formula* A *in* $\mathsf{IS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})})$*, we have that* $\mathbf{FKF}(1,\mathrm{Q})$ *proves* A*.*

Proof. Assume in $\mathbf{FKF}(1,\mathrm{Q})$ that

$$\begin{aligned}&A(x_0,\ldots,x_k,\overline{0})\wedge\\ &\forall y\big(A(x_0,\ldots,x_k,y)\to A(x_0,\ldots,x_k,y+\overline{1})\big)\end{aligned}\tag{8.8}$$

where $A(x_0,\ldots,x_k,y)$ is a $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ formula and y is not among the variables $x_0,\ldots,x_k$. So, by relying on Proposition 5.15 and Proposition 5.18 (5), we obtain

$$Sent(\ulcorner(\forall y)A\urcorner\left[\frac{num(x_0),\ldots,num(x_k)}{\ulcorner x_0\urcorner,\ldots,\ulcorner x_k\urcorner}\right])$$

in $\mathbf{A}(1,3)$ and, therefore, also

$$Sent([(\forall\ulcorner y\urcorner)\ulcorner A\urcorner\left[\frac{num(x_0),\ldots,num(x_k)}{\ulcorner x_0\urcorner,\ldots,\ulcorner x_k\urcorner}\right]]).\tag{8.9}$$

Furthermore, (8.8), Corollary 8.15, and Proposition 5.18 (3) imply that we can derive in $\mathbf{A}(1,3)\cup\mathsf{FKF}(1,\mathrm{Q})$ that

$$T(\ulcorner A\urcorner\left[\frac{num(x_0),\ldots,num(x_k)}{\ulcorner x_0\urcorner,\ldots,\ulcorner x_k\urcorner}\right]\left[\frac{num(\overline{0})}{\ulcorner y\urcorner}\right])$$

and, in addition, that

$$T(\ulcorner A\urcorner\left[\frac{num(x_0),\ldots,num(x_k)}{\ulcorner x_0\urcorner,\ldots,\ulcorner x_k\urcorner}\right]\left[\frac{num(y)}{\ulcorner y\urcorner}\right])$$

implies ($\to$) that

$$T(\ulcorner A\urcorner\left[\frac{num(x_0),\ldots,num(x_k)}{\ulcorner x_0\urcorner,\ldots,\ulcorner x_k\urcorner}\right]\left[\frac{num(y+\overline{1})}{\ulcorner y\urcorner}\right]).$$

Therefore, (8.9) and $\mathsf{TI}(1,\mathrm{Q})$ yield that

$$T(\ulcorner A\urcorner\left[\frac{num(x_0),\ldots,num(x_k)}{\ulcorner x_0\urcorner,\ldots,\ulcorner x_k\urcorner}\right]\left[\frac{num(y)}{\ulcorner y\urcorner}\right]).$$

Finally, by applying Proposition 5.18 (3) and Corollary 8.15, we obtain $A(x_0,\ldots,x_k,y)$ in $\mathbf{FKF}(1,\mathrm{Q})$. □

Against the background of Lemma 8.16, we might, whenever convenient, regard $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ formulas in **FKF**$(1, \mathrm{Q})$ also simply as $\Sigma_1^{(1,\mathrm{Q})}$ formulas which are positive in the truth predicate T_1, as shown in Proposition 8.19 below. We first define:

Definition 8.17. For $m > 0$, the class $\mathrm{P}\Delta_0^{(m,\mathrm{Q})}$ is defined recursively as follows, where t is $\mathcal{L}(0, \mathrm{Q})$ term.

- $\mathcal{L}(m-1, \mathrm{Q})$ literals are $\mathrm{P}\Delta_0^{(m,\mathrm{Q})}$.
- Positive $\mathcal{L}(m, \mathrm{Q})$ literals $T_m(t)$ are $\mathrm{P}\Delta_0^{(m,\mathrm{Q})}$.
- If B, C are $\mathrm{P}\Delta_0^{(m,\mathrm{Q})}$, then so is $(B \vee C)$.
- If B, C are $\mathrm{P}\Delta_0^{(m,\mathrm{Q})}$, then so is $(B \wedge C)$.
- If B is $\mathrm{P}\Delta_0^{(m,\mathrm{Q})}$, then so is $(\exists x \leq t)B$.
- If B is $\mathrm{P}\Delta_0^{(m,\mathrm{Q})}$, then so is $(\forall x \leq t)B$.

Definition 8.18. For $m > 0$, a formula A is $\mathrm{P}\Sigma_1^{(m,\mathrm{Q})}$ if A has the form $\exists x B$, where B is $\mathrm{P}\Delta_0^{(m,\mathrm{Q})}$.

Proposition 8.19. *In the theory* **FKF**$(1, \mathrm{Q})$*, every* $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ *formula is* $\mathrm{P}\Sigma_1^{(1,\mathrm{Q})}$.

Proof. By a (meta) induction on the rank of $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$ formulas, by following the proofs of Proposition 3.9, Proposition 3.24, and Proposition 3.26 (for $n = 0$), respectively. □

8.3 Retractability: Fixed Points

Our next goal is to specify function $\circ : \mathcal{L}(1, 3) \to \mathcal{L}(0, 3)$ which satisfies condition 2.26 (2). Before we do this, we briefly look at the semantics of KF and explain the main strategy.

As said in the introduction to this chapter, untruth $(\neg T(x))$ and falsity $(T([\neg x]))$ may have to be kept apart in the absence of the negation axiom. This is mirrored semantically as follows. A standard model of *classical* arithmetical truth is a pair $(\mathbb{N}, S)$, where $\mathbb{N}$ is the standard model of arithmetic, S a set of natural numbers, serving as the (semantic) extension (or meaning) of the truth predicate T, and the negation axiom is true in the model: For any $\mathcal{L}(1, \mathrm{Q})$ sentence A, we have that

$$(\mathbb{N}, S) \models T(\overline{\#(\neg A)}) \leftrightarrow \neg T(\overline{\#(A)}),$$

where $\#(A)$ is a code number of A. This is the case if

$$(\omega - S) = \{\#(A) : \#(\neg A) \in S\} \cup (\omega - Sent^{\mathbb{N}}),$$

where ω is the domain of $\mathbb{N}$ and $Sent^{\mathbb{N}}$ is the set of all natural numbers that encode an $\mathcal{L}(1, \mathrm{Q})$ sentence.

On the other hand, the semantics of the KF axiomatization is best introduced by reference to the so-called (generalized) Strong Kleene Schema, as in Halbach [21], Definition 15.10, for example. We do not give all the details, but focus on interpreting the truth predicate. For this purpose, we conceive a model of arithmetical truth more generally as triple $(\mathbb{N}, S_1, S_2)$, where S_1, S_2 are sets of natural numbers, called the (semantic) extension and (semantic) anti-extension of T, respectively. The idea is that

$$(\mathbb{N}, S_1, S_2) \models T(\overline{n}) \text{ iff } n \in S_1$$

and that

$$(\mathbb{N}, S_1, S_2) \models \neg T(\overline{n}) \text{ iff } n \in S_2.$$

Like above, we let

$$S_2 := \{\#(A) : \#(\neg A) \in S_1\} \cup (\omega - Sent^{\mathbb{N}}).$$

The crucial difference is this: While the model of classical truth above is a particular model $(\mathbb{N}, S_1, S_2)$ in which S_2 is the complement $(\omega - S_1)$ of S_1 and, hence,

$$S_1 \cap S_2 = \varnothing \qquad \text{(no gluts)}$$

$$S_1 \cup S_2 = \omega \qquad\qquad \text{(no gaps)}$$

a model $(\mathbb{N}, S_1, S_2)$ of the KF axiomatization may allow

$$S_1 \cap S_2 \neq \varnothing$$
$$S_1 \cup S_2 \neq \omega$$

because KF does not feature the negation axiom.

As said in the introduction to this chapter, due to the lack of the negation axiom, the truth predicate does not appear negated on either side of the KF axioms' biconditionals ($\leftrightarrow$). This is why the semantics of KF can be formulated by means of *positive inductive definitions*: Using our syntax coding, and noticing that S_2 was indeed defined on the basis of S_1, we may code the relation $(\mathbb{N}, S_1, S_2) \vDash A$ in the form $\mathbb{P}(n, S_1)$, where S_1 occurs only positively. Let $\wp(\omega)$ be the set of all subsets of ω. By what we have just said, we may define a positive (monotone) operator

$$\mathcal{P} : \wp(\omega) \to \wp(\omega)$$

(for $\mathcal{P}$ositive) as

$$\mathcal{P}(S_1) := \{n : \mathbb{P}(n, S_1)\}.$$

For more details, see also Halbach [21], Lemma 15.6.

Finally, by relying on the general theory of positive inductive definitions (as, for example, in Moschovakis [37]), we obtain that there exist fixed points S_1 such that $S_1 = \mathcal{P}(S_1)$. The upshot is that one will finally be able to establish the following: S_1 is a fixed point of $\mathcal{P}$ if and only if the expansion $(\mathbb{N}, S_1)$ of $\mathbb{N}$ by the (semantic) extension S_1 of the truth predicate is a model of the (unrestricted) KF axiomatization plus the axiom $T(x) \to Sent(x)$. For more details, see Halbach [21], Theorem 15.15.

Let us turn to *Finitist* KF theory $\mathbf{FKF}(1, \mathrm{Q})$. Here, importantly, we have: If $\#(A)$ is included in the (semantic) extension of the truth predicate in any **FKF**$(1, \mathrm{Q})$ model whatsoever, then *a finite number of iterated applications* of a $\mathcal{P}$-like operator already suffice to show that $\#(A)$ belongs to that (semantic) extension – just as a finite number of applications of the operator $\mathcal{V}$ from Section 2.3 suffice to compute the value of a term in the language of primitive recursive arithmetic (PRA). Thus, membership in fixed points of the $\mathcal{P}$-like operator under consideration can be defined by the concept of finite sequences of computation (or verification) steps.

The main idea behind the retraction of **FKF**$(1, \mathrm{Q})$, and the definition of mapping $\circ : \mathcal{L}(1, 3) \to \mathcal{L}(0, 3)$ that fulfills condition 2.26 (2) in particular, is to express in $\mathbf{A}(0, 3) \cup \mathsf{IS}(\Sigma_1^{(0,3)})$ by some interpreting formula $T(x)^\circ$ basically that there exists some finite computation sequence that verifies x.

Let us next proceed formally. For convenience, we shall make use of auxiliary function symbols and work in the extensions of $\mathbf{A}(0, 3)$ rather than in $\mathbf{I}(\Sigma_1)$ directly.

We first define a specific operator form.

Definition 8.20. Let A be a formula with at least x free. The operator form $\mathbb{T}$ (for [$\mathbb{T}$]ruth) is defined as follows.

$$\mathbb{T}(\{x : A(x)\}, y) := \bigvee \big\{(\mathrm{T1})\text{–}(\mathrm{T10})\big\}$$

where (T1)–(T10) are the following disjuncts that we sub-divide into cases for better readability.

Case A: Arithmetical Literals

(T1) $(\exists x_0, x_1 < y)\big(Sent([x_0 = x_1]) \wedge y = [x_0 = x_1] \wedge val(x_0) = val(x_1)\big)$

(T2) $(\exists x_0, x_1 < y)\big(Sent([x_0 \neq x_1]) \wedge y = [x_0 \neq x_1] \wedge val(x_0) \neq val(x_1)\big)$

(T3) $(\exists x_0, x_1 < y)\big(Sent([x_0 \leq x_1]) \wedge y = [x_0 \leq x_1] \wedge val(x_0) \leq val(x_1)\big)$

(T4) $(\exists x_0, x_1 < y)\big(Sent([x_0 > x_1]) \wedge$
$y = [x_0 > x_1] \wedge val(x_0) > val(x_1)\big)$

Case B: Truth Literals

(T5) $(\exists x_0 < y)\big(Sent([T(x_0)]) \wedge$
$y = [T(x_0)] \wedge A(val(x_0))\big)$

(T6) $(\exists x_0 < y)\big(Sent([{\sim}T(x_0)]) \wedge$
$y = [{\sim}T(x_0)] \wedge \big(A([\neg val(x_0)]) \vee \neg Sent(val(x_0))\big)\big)$

Case C: Connectives

(T7) $(\exists x_0, x_1 < y)\big(Sent([x_0 \vee x_1]) \wedge$
$y = [x_0 \vee x_1] \wedge \big(A(x_0) \vee A(x_1)\big)\big)$

(T8) $(\exists x_0, x_1 < y)\big(Sent([x_0 \wedge x_1]) \wedge$
$y = [x_0 \wedge x_1] \wedge \big(A(x_0) \wedge A(x_1)\big)\big)$

Case D: Bounded Universal Quantification

(T9) $(\exists x_0, x_1, x_2 < y)\big(Sent([(\forall x_0 \leq x_1)x_2]) \wedge$
$y = [(\forall x_0 \leq x_1)x_2] \wedge$
$(\forall x_3 \leq val(x_1))A(x_2[num(x_3)/x_0])\big)$

Case E: Unbounded Existential Quantification

(T10) $(\exists x_0, x_1 < y)\big(Sent([(\exists x_0)x_1]) \wedge$
$y = [(\exists x_0)x_1] \wedge \exists x_2 A(x_1[num(x_2)/x_0])\big)$

To prevent confusion, we have to clarify how we measure the complexity of *forms* as opposed to complexity of formulas: The complexity of a form is the complexity of the respective formula which results by replacing schematic letters (in our case: A) by atomic formulas.

Moreover, that A *occurs only positively* in a given form means that A occurs only unnegated therein.

By the results from Chapter 3, we obtain the subsequent Σ_1 normal form.

Lemma 8.21. *Let* $\mathbb{O}(\{x : A(x)\}, y)$ *be an* $\mathrm{E}\Sigma_1^{(0,3)}$ *form in which* A *occurs only positively and let* $B^<(x, z)$ *be a* $\Sigma_0^{(0,3)}$ *formula, where* $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_1^{(0,3)})$ *proves the following.*

$$B(x) \leftrightarrow \exists z B^<(x, z)$$

Then there also is a $\Sigma_0^{(0,3)}$ *form* $\mathbb{O}^<(\{x : A(x)\}, y, z)$ *in which* A *occurs only positively and such that*

$$\mathbb{O}(\{x : B(x)\}, y) \leftrightarrow \exists z \big(\mathbb{O}^<(\{x : B^<(x, z)\}, y, z)\big)$$

is provable in $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_1^{(0,3)})$.

Proof. The direction from right to left is (basically) by logic. For the other direction, we follow the proofs of Proposition 3.9 and Proposition 3.24, respectively. $\square$

Definition 8.22. For every form $\mathbb{O}(\{x : A(x)\}, y)$ in which A occurs only positively, the formula $\mathbb{O}Seq(u, v)$ is defined as follows.

$$\begin{aligned}
&Seq(u) \wedge lh(u) = v + \overline{1} \wedge \\
&(\forall w \leq v) Seq((u)_w) \wedge (u)_{\overline{0}} = \langle\ \rangle \wedge \\
&(\forall w < v)(\forall y < u)\big(\big(y \in (u)_w \to y \in (u)_{w+\overline{1}}\big) \wedge \\
&\qquad\qquad\qquad\qquad \big(y \in (u)_{w+\overline{1}} \to \mathbb{O}(\{x : x \in (u)_w\}, y)\big)\big)
\end{aligned}$$

In particular, $\mathbb{T}Seq$ describes a successively growing extension of the truth predicate. The sub-formula

$$y \in (u)_w \to y \in (u)_{w+\overline{1}}$$

expresses that the extension is growing (monotonicity), while

$$y \in (u)_{w+\overline{1}} \to \mathbb{T}(\{x : x \in (u)_w\}, y)$$

expresses that the extension is growing according to the operator defined by $\mathbb{T}(\{x : x \in (u)_w\}, y)$.

Lemma 8.23. *Let* $\mathbb{O}(\{x : A(x)\}, y)$ *be an* $\mathrm{E}\Sigma_1^{(0,3)}$ *form in which* A *occurs only positively. Then* $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_1^{(0,3)})$ *proves, for some* $\Sigma_0^{(0,3)}$ *formula* $\mathbb{O}Seq^{<}(u,v,z)$*, the following.*

$$\mathbb{O}Seq(u,v) \leftrightarrow \exists z\big(u < z \wedge v < z \wedge \mathbb{O}Seq^{<}(u,v,z)\big)$$

Proof. By Lemma 8.21 and by following the respective proofs of Proposition 3.9 and Proposition 3.24 again. $\square$

We require some preliminaries for unifying the components of bounded operator sequences ($\mathbb{O}Seq^{<}$). First of all, we require the union $r(u,w,v)$ of the initial v components of given sequences u and w.

Proposition 8.24. $\mathbf{A}(0,3)$ *proves, for an* $\mathcal{L}(0,3)$ *function symbol* $r(u,w,v)$*, the following.*

$$r(u,w,\overline{0}) = \Big\{\langle\ \rangle$$

$$r(u,w,v+\overline{1}) = \begin{cases} r(u,w,v) * \langle (u)_v * (w)_v\rangle & : (1)\\ r(u,w,v) * \langle (u)_v\rangle & : (2)\\ r(u,w,v) * \langle (w)_v\rangle & : (3)\\ r(u,w,v) & : (4)\end{cases}$$

Here (1)–(4) *are, respectively, the following conditions.*

$$\begin{aligned} &v < lh(u) \wedge v < lh(w)\\ &v < lh(u) \wedge v \geq lh(w)\\ &v \geq lh(u) \wedge v < lh(w)\\ &v \geq lh(u) \wedge v \geq lh(w)\end{aligned}$$

Proof. In view of Proposition 8.26 below, we only mention that $r(u,w,v)$ may actually be bounded by a term in $\mathcal{L}(0,2)$. This is so because, for the computation of $r(u,w,v)$, it suffices to use a course-of-values sequence of length $\max(|u|,|w|)$ in which every member (sequence) is bounded by $u * w$. Thus, we may revert to Remark 4.16, as usual. □

Proposition 8.25. *Let* $r(u, w, v)$ *be the function symbol in* 8.24. *Then* $\mathbf{A}(0, 3)$ *proves the following.*

(1) $v \leq lh(u) \vee v \leq lh(w) \to lh(r(u, w, v)) = v$

(2) $(\forall v < lh(u))(\forall x < z)\big(x \in (u)_v \to x \in (r(u, w, v + \overline{1}))_v\big)$

(3) $(\forall v < lh(w))(\forall x < z)\big(x \in (w)_v \to x \in (r(u, w, v + \overline{1}))_v\big)$

Proof. We may first derive (1) and obtain (2) and (3) afterwards on the basis of (1).

- As for (1), apply $\mathsf{IS}(\Sigma_0^{(0,3)})$ with induction variable v.
- As for (2), suppose $v < lh(u)$ and $x < z$ and $x \in (u)_v$. By (1), we know that $lh(r(u, w, v + \overline{1})) = v + \overline{1}$. Thus, by the definition of r, we have

  $$(r(u, w, v + \overline{1}))_v = (u)_v * (w)_v$$

  or

  $$(r(u, w, v + \overline{1}))_v = (u)_v$$

  In either case, $x \in (r(u, w, v + \overline{1}))_v$ by assumption.

(3) is similar. □

Next, we may unify ($\mathbb{O}un^{<}$) the components of operator sequences ($\mathbb{O}Seq^{<}$) that are bounded by some number, as shown by the next proposition.

Proposition 8.26. *Let* $r(u, w, v)$ *be the function symbol in* 8.24. *Moreover, let* $\mathbb{O}(\{x : A(x)\}, y)$ *be an* $\mathrm{E}\Sigma_1^{(0,3)}$ *form in which* A *occurs only positively. Then* $\mathbf{A}(0, 3)$ *proves, for an* $\mathcal{L}(0, 3)$ *function symbol* $\mathbb{O}un^{<}(u, z)$*, the following.*

$$\mathbb{O}un^{<}(\overline{0}, z) = \Big\{ \langle\langle\ \rangle\rangle$$

$$\mathbb{O}un^{<}(u+\overline{1},z) = \begin{cases} r(u,\mathbb{O}un^{<}(u,z),f(u,z)) & :(1)\\ \mathbb{O}un^{<}(u,z) & :(2)\\ \mathbb{O}un^{<}(u,z) & :(3)\end{cases}$$

Here (1)–(3) *are, respectively, the conditions*

$$\begin{aligned}&u < z \wedge (\exists v < z)\mathbb{O}Seq^{<}(u,v,z)\\ &u < z \wedge (\forall v < z)\neg\mathbb{O}Seq^{<}(u,v,z)\\ &u \geq z\end{aligned}$$

and $\mathbf{A}(0,3)$ *proves the following equality.*

$$f(u,z) = \max\big(lh(u), lh(\mathbb{O}un^{<}(u,z))\big)$$

Proof. By applying Proposition 4.15 and Proposition 4.3. As for the bound which is required for applying Proposition 4.15, let us recall that r may be bounded by a term in $\mathcal{L}(0,2)$. (This will do the job because iterated application of the smash function can be bounded by exponentiation.) □

Lemma 8.27. *Let* $\mathbb{O}un^{<}(u,z)$ *be as in* 8.26. *Then* $\mathbf{A}(0,3)$ *proves, for any* $\mathrm{E}\Sigma_1^{(0,3)}$ *form* $\mathbb{O}(\{x : A(x)\},y)$ *in which* A *occurs only positively, the following.*

(1) $(\forall u,v < z)\big(\mathbb{O}Seq^{<}(u,v,z) \to$
$(\forall x < z)\big(x \in (u)_v \to x \in (\mathbb{O}un^{<}(u+\overline{1},z))_v\big)\big)$

(2) $u \leq w \to$
$(\forall v,x < z)\big(x \in (\mathbb{O}un^{<}(u,z))_v \to x \in (\mathbb{O}un^{<}(w,z))_v\big)$

(3) $(\exists v < z)\big(\mathbb{O}Seq^{<}(\mathbb{O}un^{<}(u,z),v,z)\big)$

Proof. We consider only (1). Assume that $u < z$ and $v < z$ and $\mathbb{O}Seq^{<}(u,v,z)$. Then we have

$$\mathbb{O}un^{<}(u+\overline{1},z) = r(u,\mathbb{O}un^{<}(u,z),f(u,z))$$

where $f(u, z) = \max(lh(u), lh(\mathbb{O}un^{<}(u, z)))$. Assume further that $x < z$ and $x \in (u)_v$. Then we may rely on Proposition 8.25 (2) to conclude that $x \in (\mathbb{O}un^{<}(u + \overline{1}, z))_v$. ☐

We now define formula $\mathbb{T}Fip(x)$ to express basically that the truth of x is witnessed by a finite computation sequence ($\mathbb{T}Seq$). The definition of $\mathbb{T}Fip(x)$ is as follows.

$$(\exists u, v)\big(\mathbb{T}Seq(u, v) \wedge x \in (u)_v\big) \tag{8.10}$$

In the proof of Lemma 8.29, we will first show formally that $\mathbb{T}Fip(x)$ defines a fixed point of an operator defined by $\mathbb{T}$. That $\mathbb{T}Fip(x)$ defines the least of some (definable) fixed points will be established below in Chapter 9.

In general, we define the following.

Definition 8.28. For every form $\mathbb{O}(\{x : A(x)\}, y)$ in which A occurs only positively, formula $\mathbb{O}Fip(y)$ is defined as follows.

$$(\exists u, v)\big(\mathbb{O}Seq(u, v) \wedge y \in (u)_v\big).$$

Lemma 8.29. $\mathbf{A}(0, 3) \cup \mathsf{BS}(\Sigma_1^{(0,3)})$ *proves, for every* $\mathrm{E}\Sigma_1^{(0,3)}$ *form* $\mathbb{O}(\{x : A(x)\}, y)$ *in which* A *occurs only positively, the following.*

$$\mathbb{O}Fip(y) \leftrightarrow \mathbb{O}(\{x : \mathbb{O}Fip(x)\}, y)$$

Proof. The left-to-right direction holds by monotonicity: From the assumption $\mathbb{O}Fip(y)$, we obtain by definition of $\mathbb{O}Fip$ and $\mathbb{O}Seq$ that $\mathbb{O}(\{x : x \in (u)_{v-\overline{1}}\}, y)$. By monotonicity, we have for every $x < u$ that

$$x \in (u)_{v-\overline{1}} \to x \in (u)_v$$

and obtain $\mathbb{O}(\{x : x \in (u)_v\}, y)$ and, hence, $\mathbb{O}(\{x : \mathbb{O}Fip(x)\}, y)$, as required.

For the right-to-left direction, suppose $\mathbb{O}(\{x : \mathbb{O}Fip(x)\}, y)$. By Lemma 8.23, we have

$$\mathbb{O}Fip(x) \leftrightarrow \exists z(\exists u, v < z)\big(\mathbb{O}Seq^{<}(u, v, z) \wedge x \in (u)_v\big).$$

By Lemma 8.21 and the assumption, we get

$$\exists z\big(\mathbb{O}^{<}\big(\big\{x : (\exists u, v < z)\big(\mathbb{O}Seq^{<}(u, v, z) \wedge x \in (u)_v\big)\big\}, y, z\big)\big).$$

By Lemma 8.27 (3), furthermore, we have for some $v < z$ that $\mathbb{O}Seq^{<}(\mathbb{O}un^{<}(z, z), v, z)$. By Lemma 8.27 (1) and (2), we get

$$\exists z\big(\mathbb{O}^{<}\big(\big\{x : x \in (\mathbb{O}un^{<}(z, z))_v\big\}, y, z\big)\big).$$

Therefore, by Proposition 4.14, we may consider the prolongation $\mathbb{O}un^{<}(z, z) * \langle(\mathbb{O}un^{<}(z, z))_v * \langle y\rangle\rangle$ to finally conclude $\mathbb{O}Fip(y)$, as required. □

Corollary 8.30. $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_1^{(0,3)})$ *proves the following.*

$$\mathbb{T}Fip(y) \leftrightarrow \mathbb{T}(\{x : \mathbb{T}Fip(x)\}, y)$$

Proof. By instantiating Lemma 8.29. □

Remark 8.31. It might be illustrative to describe the proof of Corollary 8.30 in some more detail as shortening and prolonging, respectively, of truth (or computation) sequences ($\mathbb{T}Seq$). Let us consider two brief examples.

- As for the left-to-right direction, assume $\mathbb{T}Fip(y)$. That is to say, there is a truth sequence that contains y in its last component. Call this sequence p. Suppose that y has form $T(t)$. Then, by definition of truth sequences, the value of t is included in the second to last component of p. Hence, using Proposition 4.13, we cut off the last component of p and obtain a truth sequence q which contains the value of t in its last component. This covers the respective case of the verification of $\mathbb{T}(\{x : \mathbb{T}Fip(x)\}, y)$.
- Similarly for the right-to-left direction. Let us assume that $\mathbb{T}(\{x : \mathbb{T}Fip(x)\}, y)$ and that y has form $T(t)$. This means that a truth sequence q contains the value of t in its last

component. By Proposition 4.14, we may prolong q to get a truth sequence p that contains $T(t)$ in its last component. This verifies $\mathbb{T}Fip(y)$.

Similarly for other cases.

8.4 Retractability: Upper Bound

We are now ready to approach the definition of a retractability mapping that satisfies condition 2.26 (2).

As first step, we define a mapping $\mathcal{F} : \mathcal{L}(1,3) \to \mathcal{L}(0,3)$ as follows.

Definition 8.32. $\mathcal{F} : \mathcal{L}(1,3) \to \mathcal{L}(0,3)$ is defined recursively.

$$\begin{aligned}
(s = t)^{\mathcal{F}} &:= (s = t) \\
(s \neq t)^{\mathcal{F}} &:= (s \neq t) \\
(s \leq t)^{\mathcal{F}} &:= (s \leq t) \\
(s > t)^{\mathcal{F}} &:= (s > t) \\
T(t)^{\mathcal{F}} &:= \mathbb{T}Fip(t) \\
({\sim}T(t))^{\mathcal{F}} &:= \neg\mathbb{T}Fip(t) \\
(B \vee C)^{\mathcal{F}} &:= (B^{\mathcal{F}} \vee C^{\mathcal{F}}) \\
(B \wedge C)^{\mathcal{F}} &:= (B^{\mathcal{F}} \wedge C^{\mathcal{F}}) \\
(\exists x B)^{\mathcal{F}} &:= \exists x (B)^{\mathcal{F}} \\
(\forall x B)^{\mathcal{F}} &:= \forall x (B)^{\mathcal{F}}
\end{aligned}$$

The following is immediate for $\mathcal{F}$.

Lemma 8.33. *$\mathcal{F}$ preserves logical structure in* $\mathbf{A}(0,3)$.

Proof. By definition of $\mathcal{F}$ and logic. □

Moreover, we also have the following.

Lemma 8.34. *For every $\mathcal{L}(0,3)$ formula A, $A^{\mathcal{F}}$ is A.*

Proof. By (meta) induction on the rank of A. ☐

We now show that condition 2.26 (2) is fulfilled by $\mathcal{F}$ and get our final retractability mapping afterwards by composing $\mathcal{F}$ with the mapping $\mathcal{A}$ for eliminating auxiliary function symbols.

Lemma 8.35. *For every axiom A from $\mathsf{FKF}(1, \mathrm{Q})$, we have that $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_1^{(0,3)})$ proves $A^{\mathcal{F}}$.*

Proof. We work in theory $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_1^{(0,3)})$ without repeatedly mentioning it. We consider selected axioms in $\mathsf{FKF}(1, \mathrm{Q})$ to show the strategy. Note that the following material tacitly relies on Lemma 8.33, Lemma 8.34, and Proposition 2.22.

- For (FKF5), assume that $Sent([T(x)])$. Then we have

  $$\mathbb{T}Fip([T(x)]) \leftrightarrow \mathbb{T}(\{v : \mathbb{T}Fip(v)\}, [T(x)])$$

  by Lemma 8.29 and

  $$\mathbb{T}Fip(val(x)) \leftrightarrow \mathbb{T}(\{v : \mathbb{T}Fip(v)\}, [T(x)])$$

  because, if $[T(x)]$ satisfies one of the disjuncts (T1)–(T10), then it must be (T5) in particular. So, we obtain:

  $$\mathbb{T}Fip([T(x)]) \leftrightarrow \mathbb{T}Fip(val(x))$$

  As noted, we may finally apply 8.33, 8.34, and 2.22 to conclude that the $\mathcal{F}$ interpretation of the axiom (FKF5) can be derived.

- For (FKF7), assume that $Sent([x \vee y])$. Rely on Lemma 8.29 to obtain

  $$\mathbb{T}Fip([x \vee y]) \leftrightarrow \mathbb{T}(\{v : \mathbb{T}Fip(v)\}, [x \vee y])$$

and reason like before to obtain

$$\big(\mathbb{T}Fip(x) \lor \mathbb{T}Fip(y)\big) \leftrightarrow \mathbb{T}(\{v : \mathbb{T}Fip(v)\}, [x \lor y]).$$

The rest is covered by 8.33, 8.34, and 2.22.

- For (FKF9), assume that $Sent([(\forall x \leq y)z])$. Then we may again rely on Lemma 8.29 and the definition of form $\mathbb{T}$ to derive

  $$\mathbb{T}Fip([(\forall x \leq y)z]) \leftrightarrow \mathbb{T}(\{v : \mathbb{T}Fip(v)\}, [(\forall x \leq y)z])$$

  and that $\mathbb{T}(\{v : \mathbb{T}Fip(v)\}, [(\forall x \leq y)z])$ is equivalent ($\leftrightarrow$) to

  $$(\forall u \leq val(y))\mathbb{T}Fip(z[num(u)/x]),$$

  as required.

Other cases are treated accordingly. □

Remark 8.36. If, in addition, $\mathsf{FKF}(1, \mathrm{Q})$ would include the *one-directional* axiom

(FKF11) $Sent([(\forall x)y]) \land (\forall z_0, z_1 < y)(y \neq [[x \leq z_0] \to z_1]) \to \big(T([(\forall x)y]) \to \forall u T(y[num(u)/x])\big)$

Lemma 8.35 would still hold because the $\mathcal{F}$ interpretation of the axiom (FKF11) could be derived as follows: Assuming

$$Sent([(\forall x)y]) \land (\forall z_0, z_1 < y)(y \neq [[x \leq z_0] \to z_1]),$$

we may derive $\neg\mathbb{T}(\{v : \mathbb{T}Fip(v)\}, [(\forall x)y])$ by using definition of form $\mathbb{T}$ and conclude $\neg\mathbb{T}Fip([(\forall x)y])$ by Lemma 8.29.

Let us now address the derivation of the $\mathcal{F}$ interpretation of Truth Induction $\mathsf{TI}(1, \mathrm{Q})$.

Lemma 8.37. *For the only axiom A from $\mathsf{TI}(1, \mathrm{Q})$, we have that $\mathbf{A}(0, 3) \cup \mathsf{IS}(\Sigma_1^{(0,3)})$ proves $A^{\mathcal{F}}$.*

Proof. By 3.26, 3.9, and 3.24, the formula $\mathbb{T}Fip(x)$ is $\Sigma_1^{(0,3)}$ in the theory $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)})$. Hence, $\mathsf{IS}(\Sigma_1^{(0,3)})$ includes

$$\begin{aligned}&\big(\mathbb{T}Fip(x[num(\overline{0})/z]) \wedge \\ &\forall y\big(\mathbb{T}Fip(x[num(y)/z]) \rightarrow \mathbb{T}Fip(x[num(y+\overline{1})/z])\big)\big) \rightarrow \\ &\mathbb{T}Fip(x[num(y)/z])\end{aligned}$$

So, we may apply 8.33, 8.34, and 2.22 to conclude that the $\mathcal{F}$ interpretation of Truth Induction is derivable. □

8.5 Conclusions

We now piece together our earlier findings to conclude the main results of the present chapter.

Theorem 8.38. $\mathbf{FKF}(1, \mathrm{Q})$ *is retractable to* $\mathbf{I}(\Sigma_1)$ *by* $(\mathcal{A} \cdot \mathcal{F})$.

Proof. In view of Corollary 3.56, it is enough to establish that $\mathbf{FKF}(1, \mathrm{Q})$ is retractable to $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)})$ by $\mathcal{F}$. We verify the conditions from Definition 2.26.

- By Lemma 8.33, mapping $\mathcal{F}$ preserves logical structure in $\mathbf{A}(0,3)$.
- By Lemma 8.34 and logic, we have for all $\mathcal{L}(0,3)$ atoms B that $\mathbf{A}(0,3) \vdash B \leftrightarrow B^{\mathcal{F}}$.
- By preservation of logical structure, we have for (additional logical axioms) B in $\mathbf{A}(1,3)$ that $\mathbf{A}(0,3) \vdash B^{\mathcal{F}}$. Also, by Lemma 8.35, Lemma 8.37, and Proposition 3.26, we get for every B in $\mathsf{FKF}(1, \mathrm{Q}) \cup \mathsf{TI}(1, \mathrm{Q})$ that $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)}) \vdash B^{\mathcal{F}}$.

- By Lemma 8.16, we get for any axiom B in $\mathsf{IS}(\Sigma_1^{(0,3)})$ that $\mathbf{FKF}(1,\mathrm{Q}) \vdash B$. (Recall that any $\Sigma_1^{(0,3)}$ formula is $\Sigma_1^{(0,\mathrm{Q})}$ in $\mathbf{A}(0,3)$.)

This concludes the proof. □

Thus, we have finally arrived at the following.

Corollary 8.39. *For every $\mathcal{L}(0,\mathrm{Q})$ formula A and $\mathcal{L}(1,3)$ formula B, the following hold.*

(1) $\mathbf{FKF}(1,\mathrm{Q}) \vdash B$ *implies* $\mathbf{I}(\Sigma_1) \vdash (B^{\mathcal{F}})^{\mathcal{A}}$.

(2) $\mathbf{FKF}(1,\mathrm{Q}) \vdash A$ *if and only if* $\mathbf{I}(\Sigma_1) \vdash A$.

Proof. By Proposition 2.27. □

Remark 8.40. Let us also add the following.

- For every $\mathcal{L}(1,\mathrm{Q})$ formula B, we have by Corollary 3.54 (2) that
$$\mathbf{B}(1,3) \cup \big(\mathsf{FKF}(1,\mathrm{Q}) \cup \mathsf{TI}(1,\mathrm{Q})\big)^{\mathcal{A}} \vdash B$$
implies $\mathbf{FKF}(1,\mathrm{Q}) \vdash B$ and, by Corollary 8.39 (1), implies
$$\mathbf{I}(\Sigma_1) \vdash (B^{\mathcal{F}})^{\mathcal{A}}.$$

- For every $\mathcal{L}(0,\mathrm{Q})$ formula A, we have by Corollary 3.54 (2) that
$$\mathbf{B}(1,3) \cup \big(\mathsf{FKF}(1,\mathrm{Q}) \cup \mathsf{TI}(1,\mathrm{Q})\big)^{\mathcal{A}} \vdash A$$
precisely if $\mathbf{FKF}(1,\mathrm{Q}) \vdash A$ and, by Corollary 8.39 (2), precisely if
$$\mathbf{I}(\Sigma_1) \vdash A.$$

Remark 8.41. By inspecting the proof of Lemma 8.35 above, we see that $\Sigma_1^{(0,3)}$ induction was not needed. (Moreover, every $\Sigma_1^{(0,3)}$ formula is $\Sigma_1^{(0,\mathrm{Q})}$ in $\mathbf{A}(0,3)$.) Therefore, we also have:

For every $\mathcal{L}(0,3)$ formula A and every $\mathcal{L}(1,3)$ formula B, the following hold.

- $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q}) \vdash B$ implies $\mathbf{A}(0,3) \vdash B^{\mathcal{F}}$.
- $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q}) \vdash A$ if and only if $\mathbf{A}(0,3) \vdash A$.

Remark 8.42. We also notice that the Finitist Kripke-Feferman Theory may alternatively be seen as *theory of* $\mathrm{P}\Sigma_1^{(1,\mathrm{Q})}$ *induction*. Namely, the following *theories are identical*.

- $\mathbf{FKF}(1,\mathrm{Q})$
- $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q}) \cup \mathsf{IS}(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})})$

To see this, we may revert to Corollary 2.29 and retract the former theory to the latter by the identity function. Notice:

- As $T(x[num(y)/z])$ is $\mathrm{P}\Sigma_1^{(1,\mathrm{Q})}$ in $\mathbf{A}(1,3)$, the axiom in the singleton $\mathsf{TI}(1,\mathrm{Q})$ is derivable by means of $\mathsf{IS}(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})})$.
- By Lemma 8.16 and Proposition 8.19, every formula from $\mathsf{IS}(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})})$ is derivable in $\mathbf{FKF}(1,\mathrm{Q})$.

Remark 8.43. We conclude the present chapter by the remark that, since $\mathbf{I}(\Sigma_1)$ proves every instance of $\mathsf{IS}(\Pi_1)$ (see, for example, Hájek and Pudlák [18], Lemma 2.12), we could equally have defined $\mathbf{FKF}(1,\mathrm{Q})$ as a theory of *negative untruth* by rephrasing truth axioms from $\mathsf{FKF}(1,\mathrm{Q})$ (in logically equivalent form) as axioms for (negative) $\sim T(x)$ and by replacing $\mathsf{TI}(1,\mathrm{Q})$ by the axiom of Untruth Induction and we would still have obtained a theory which is retractable to $\mathbf{I}(\Sigma_1)$.

9 Groundedness

According to Kripke [32], pp. 693f., if a sentence in the language of truth involves the notion of truth, its

> truth value in turn must be ascertained by looking at *other* sentences, and so on. If ultimately this process terminates in sentences not mentioning the concept of truth, so that the truth value of the original statement can be ascertained, [then] we call the original sentence *grounded*; otherwise, *ungrounded.*

In our case of *arithmetical* truth: If ultimately this process terminates in sentences with a determinate truth value in the model $\mathbb{N}$, the original sentence may be called *grounded*; otherwise, *ungrounded.*

For instance, let us consider sentence $T([\neg[T(\ulcorner\overline{0} > \overline{1}\urcorner)]])$ and recall the ($\mathcal{P}$ositive) operator

$$\mathcal{P} : \wp(\omega) \to \wp(\omega)$$

from Section 8.3 which has the form (of some positive inductive definition)

$$\mathcal{P}(S) := \{n : \mathbb{P}(n, S)\}$$

as in Halbach [21], Lemma 15.6. To be grounded in the sense of Kripke [32] means to be contained in the *least* fixed point of the operator $\mathcal{P}$: that is, the fixed point that we obtain by iterated application of $\mathcal{P}$ to the empty set $S = \varnothing$. So, by definition of condition $\mathbb{P}$, we have the following.

1. $\#(\neg(\overline{0} > \overline{1})) \in \mathcal{P}(\varnothing)$ because $\mathbb{N} \vDash \neg(\overline{0} > \overline{1})$.
2. $\#(\neg T(\ulcorner\overline{0} > \overline{1}\urcorner)) \in \mathcal{P}(\mathcal{P}(\varnothing))$ because, by the foregoing step, $\#(\neg(\overline{0} > \overline{1})) \in \mathcal{P}(\varnothing)$.
3. $\#(T([\neg[T(\ulcorner\overline{0} > \overline{1}\urcorner)]])) \in \mathcal{P}(\mathcal{P}(\mathcal{P}(\varnothing)))$ because, by the foregoing step, $\#(\neg T(\ulcorner\overline{0} > \overline{1}\urcorner)) \in \mathcal{P}(\mathcal{P}(\varnothing))$.

Thus, 3 iterations of the operator $\mathcal{P}$ ascertain that the considered sentence $T([\neg[T(\ulcorner\overline{0} > \overline{1}\urcorner)]])$ is in the least fixed point.

Similarly, we may verify the following examples.

- ω many iterations of $\mathcal{P}$ ascertain that the sentence

$$T(\ulcorner\forall x(\overline{0} < x + \overline{1})\urcorner)$$

is contained in the least fixed point of $\mathcal{P}$.

- A (Truth Teller) sentence D such that

$$(\mathbb{N}, \varnothing) \vDash D \leftrightarrow T(\ulcorner D\urcorner)$$

is not contained in the least fixed point of $\mathcal{P}$.

- A (Liar) sentence D such that

$$(\mathbb{N}, \varnothing) \vDash D \leftrightarrow \neg T(\ulcorner D\urcorner)$$

is not contained in the least fixed point of $\mathcal{P}$.

The groundedness of a statement may be seen as a guarantee for the statement's naturalness or non-pathological nature, in the sense that the statement's truth or falsity may be ascertained. In contrast, a Truth Teller or a paradoxical Liar, which seem unnatural and do not seem to have a determinate truth value, both are ungrounded. For more details, see Yablo [50] and Leitgeb [34].

On the proof-theoretic side, the addition of the groundedness requirement may increase the strength of a system considerably. Most notably, Cantini [6], Paragraph 3, discusses two theories of truth that both incorporate the unrestricted axioms in KF for a whole first-order language and that differ only in that the second (in contrast to the first one) has axioms to the effect that truth is *grounded* (and consistent). (In contrast to the unrestricted KF theory that we have discussed right after Remark 8.2, the two KF theories now under consideration contain not only internal Truth Induction, but the induction *schema* for an overall first-order language of truth.) Cantini [6] proves that, while the second theory has the full power (in terms of provable arithmetical statements) of theory $\mathbf{ID}_1$ that asserts the existence of a *least* fixed point for any positive arithmetical inductive operator, the first theory only reaches the power of proper restriction $\widehat{\mathbf{ID}}_1$ of $\mathbf{ID}_1$ which asserts only the existence of some *arbitrary* fixed point for the positive arithmetical inductive operators that are elementary. (Note: $\widehat{\mathbf{ID}}_1$ is strictly stronger than $\mathbf{RA}_{<\omega^\omega}$. Consult Feferman [11] for more details on systems of inductive definitions.)

The main purpose of this chapter is to make clear that the situation is different in the setting of *finitist* truth. Let us briefly recall a few points.

In particular, recall from Section 8.3 that the KF axioms may be interpreted as positive inductive truth operation $\mathcal{P}$ which, by restricting the KF axioms as in the *finitist* **FKF**, may take the form of a specific *finite* computation (verification), as defined by $\mathbb{T}Fip(x)$. Indeed, Lemma 8.29 revealed that verifiability by such finite computation defines a fixed point.

In the present chapter, we shall elaborate on this and prove formally by finitist means that $\mathbb{T}Fip(x)$ actually defines the *least* one of certain definable fixed points. The informal idea is that, since computations defined by $\mathbb{T}Fip(x)$ are finite, we can check by some finite procedure that certain properties are preserved by objects generated in the course of these computations.

Hence, in a loose sense, the computations defined by $\mathbb{T}Fip(x)$ have groundeness built-in and, in contrast to the above sketched non-finitist case, we do not have to go beyond the framework of Chapter 8 in order to deal with grounded truth. In more proof-theoretic terms, this clearly reflects that *least* fixed points of Σ_1 operators are Σ_1 definable themselves.

9.1 The Theory

In general, if S is a fixed point of an operator $\mathcal{P}$, then $\mathcal{P}(S) \subseteq S$ holds. We may define the least fixed point L of $\mathcal{P}$ as

$$L := \bigcap \{S : \mathcal{P}(S) \subseteq S\}.$$

(By monotonicity: $L \subseteq \mathcal{P}(L)$. Furthermore, if $\mathcal{P}(S) \subseteq S$, then $L \subseteq S$ and, by monotonicity, $\mathcal{P}(L) \subseteq \mathcal{P}(S)$ and so $\mathcal{P}(L) \subseteq S$. Hence, $\mathcal{P}(L) \subseteq L$. Minimality of L is immediate by definition.)

Precisely this is the idea behind the following schema – which defines the least (Γ definable) fixed point of the operator that we have defined by the form $\mathbb{T}$ in 8.20. Burgess [3] presents a very similar axiomatization of grounded truth.

Note: In what follows, notational conventions from Chapter 8 apply. In particular, we write $T(x)$ for $T_1(x)$, and so on.

Definition 9.1. Let Γ be a given class of formulas. $\mathsf{GS}(\Gamma)$ is called the *Groundedness Schema* and defined as

$$\forall y\big(\mathbb{T}(\{x : A(x)\}, y) \to A(y)\big) \to \big(T(x) \to A(x)\big)$$

where $A(x)$ belongs to Γ.

For convenience, we first change the logical form of the antecedent of $\mathsf{GS}(\Gamma)$ instances by logical reasoning.

Proposition 9.2. $\mathbf{A}(1,3)$ *proves, for every* $\mathcal{L}(1,3)$ *formula* $A(x)$, *the following.*

$$\forall y\big(\mathbb{T}(\{x : A(x)\}, y) \to A(y)\big) \leftrightarrow \forall y\big(\bigwedge\big\{(\text{G1})\text{–}(\text{G10})\big\}\big)$$

where (G1)–(G10) *are the subsequent conjuncts that we sub-divide into cases for better readability.*

Case A: Arithmetical Literals

(G1) $(\forall x_0, x_1 < y)\big(\mathit{Sent}([x_0 = x_1]) \to \big(\mathit{val}(x_0) = \mathit{val}(x_1) \to A([x_0 = x_1])\big)\big)$

(G2) $(\forall x_0, x_1 < y)\big(\mathit{Sent}([x_0 \neq x_1]) \to \big(\mathit{val}(x_0) \neq \mathit{val}(x_1) \to A([x_0 \neq x_1])\big)\big)$

(G3) $(\forall x_0, x_1 < y)\big(\mathit{Sent}([x_0 \leq x_1]) \to \big(\mathit{val}(x_0) \leq \mathit{val}(x_1) \to A([x_0 \leq x_1])\big)\big)$

(G4) $(\forall x_0, x_1 < y)\big(\mathit{Sent}([x_0 > x_1]) \to \big(\mathit{val}(x_0) > \mathit{val}(x_1) \to A([x_0 > x_1])\big)\big)$

Case B: Truth Literals

(G5) $(\forall x_0 < y)\big(\mathit{Sent}([T(x_0)]) \to \big(A(\mathit{val}(x_0)) \to A([T(x_0)])\big)\big)$

(G6) $(\forall x_0 < y)\big(\mathit{Sent}([{\sim}T(x_0)]) \to \big(A([\neg \mathit{val}(x_0)]) \vee \neg \mathit{Sent}(\mathit{val}(x_0)) \to A([{\sim}T(x_0)])\big)\big)$

Case C: Connectives

(G7) $(\forall x_0, x_1 < y)\big(\mathit{Sent}([x_0 \vee x_1]) \to \big(A(x_0) \vee A(x_1) \to A([x_0 \vee x_1])\big)\big)$

(G8) $(\forall x_0, x_1 < y)\big(Sent([x_0 \wedge x_1]) \rightarrow$
$\big(A(x_0) \wedge A(x_1) \rightarrow A([x_0 \wedge x_1])\big)\big)$

Case D: Bounded Universal Quantification

(G9) $(\forall x_0, x_1, x_2 < y)\big(Sent([(\forall x_0 \leq x_1)x_2]) \rightarrow$
$\big((\forall x_3 \leq val(x_1))A(x_2[num(x_3)/x_0]) \rightarrow$
$A([(\forall x_0 \leq x_1)x_2])\big)\big)$

Case E: Unbounded Existential Quantification

(G10) $(\forall x_0, x_1 < y)\big(Sent([(\exists x_0)x_1]) \rightarrow$
$\big(\exists x_2 A(x_1[num(x_2)/x_0]) \rightarrow A([(\exists x_0)x_1])\big)\big)$

Let us now put the groundedness schema to work by deriving several theorems. For this, however, we first have to introduce a new complexity class.

Definition 9.3. Let Γ be some class of $\mathcal{L}$ formulas. The formula class $\Sigma_0(\Gamma)$ (pronounce: Σ_0 in Γ) is defined recursively as follows, where t is an $\mathcal{L}$ term.

- If B belongs to Γ, then B is $\Sigma_0(\Gamma)$.
- If B belongs to Γ, then $\neg B$ is $\Sigma_0(\Gamma)$.
- If B, C are $\Sigma_0(\Gamma)$, then so is $(B \vee C)$.
- If B, C are $\Sigma_0(\Gamma)$, then so is $(B \wedge C)$.
- If B is $\Sigma_0(\Gamma)$, then so is $(\exists x \leq t)B$.
- If B is $\Sigma_0(\Gamma)$, then so is $(\forall x \leq t)B$.

Remember that, by Remark 8.36, we may extend $\mathsf{FKF}(1, \mathrm{Q})$ by a *one-directional* axiom

(FKF11) $Sent([(\forall x)y]) \wedge (\forall z_0, z_1 < y)(y \neq [[x \leq z_0] \rightarrow z_1]) \rightarrow$
$\big(T([(\forall x)y]) \rightarrow \forall u T(y[num(u)/x])\big)$

without changing the validity of Theorem 8.38. Indeed, we shall derive several groundedness results in the following theory.

Definition 9.4. $\mathbf{FG}(1,\mathrm{Q})$ is the *Theory of Finitist Grounded Truth* (in auxiliary form) defined as follows.

$$\mathbf{FKF}(1,\mathrm{Q}) \cup \{(\mathrm{FKF11})\} \cup \mathsf{GS}(\Sigma_0(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})}))$$

In the present section, we focus on the Consistency Theorem and the Truth Teller. In the following chapter, we will turn back to the Liar.

We need a preliminary result:

Proposition 9.5. $\mathbf{FG}(1,\mathrm{Q}) \vdash T(x) \to Sent(x)$

Proof. In view of Proposition 9.2, it suffices to show (G1)–(G10), where $A(x)$ is $Sent(x)$, which is $\Sigma_0(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})})$ formula in $\mathbf{A}(1,3)$. This only requires logical reasoning. □

Thus, the Consistency Theorem may be derived.

Proposition 9.6. $\mathbf{FG}(1,\mathrm{Q}) \vdash T([\neg x]) \to \neg T(x)$

Proof. In view of Proposition 9.2, it suffices to show (G1)–(G10), where $A(x)$ is $\neg T([\neg x])$, which is $\Sigma_0(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})})$ formula in $\mathbf{A}(1,3)$. Let us consider a few selected examples.

- For (G1), suppose $Sent([x_0 = x_1])$ and $val(x_0) = val(x_1)$. Moreover, suppose towards contradiction that $T([x_0 \neq x_1])$. By (FKF2), this implies $val(x_0) \neq val(x_1)$, which yields the contradiction.
- For (G2), proceed accordingly, using Proposition 8.5 (1).

- For (G6), suppose (for contraposition) $Sent([{\sim}T(x_0)])$ and $T([\neg[{\sim}T(x_0)]])$, which by Proposition 8.5 (1) gives $T([T(x_0)])$. Thus, by (FKF5), we obtain $T(val(x_0))$, which by Proposition 8.5 (1) implies $T([\neg[\neg val(x)_0]])$ and by the foregoing Proposition 9.5 implies $Sent(val(x_0))$, as required.
- For (G9), assume that $Sent([(\forall x_0 \leq x_1)x_2])$ and also that $T([\neg[(\forall x_0 \leq x_1)x_2]])$. Then, as required, we directly obtain $(\exists x_3 \leq val(x_1))T([\neg x_2[num(x_3)/x_0]])$ by Proposition 8.5 (5).
- For (G10), assume that $Sent([(\exists x_0)x_1])$ and $T([\neg[(\exists x_0)x_1]])$. If $[(\exists x_0)x_1]$ is bounded, we may apply Proposition 8.5 (4). Otherwise, we obtain by (FKF11) that

  $$T([\neg[(\exists x_0)x_1]]) \to \forall x_2 T([\neg x_1[num(x_2)/x_0]]).$$

  Hence, $\forall x_2 T([\neg x_1[num(x_2)/x_0]])$, as required.

Other cases are treated accordingly. □

Remark 9.7. Let D be a Truth Teller so that

$$(\mathbb{N}, \varnothing) \vDash D \leftrightarrow T(\ulcorner D \urcorner),$$

where D is of the form $\exists v(C \wedge T(v))$ and C is a Σ_1 formula in $\mathcal{L}(0, \mathrm{Q})$. Then $\mathbf{FG}(1, \mathrm{Q}) \vdash \neg T(\ulcorner D \urcorner)$ holds. In view of Proposition 9.2, it is enough to show (G1)–(G10), where $A(x)$ is the following conjunction.

$$\begin{aligned} & x \neq [T(num(\ulcorner D \urcorner))] \wedge \\ & x \neq \ulcorner C \wedge T(v) \urcorner [num(\ulcorner D \urcorner)/\ulcorner v \urcorner] \wedge \\ & x \neq [(\exists \ulcorner v \urcorner)\ulcorner C \wedge T(v) \urcorner] \end{aligned} \tag{9.1}$$

Since we have $\neg A(\ulcorner D \urcorner)$, we will eventually conclude $\neg T(\ulcorner D \urcorner)$ by $\mathsf{GS}(\Sigma_0(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})}))$. We omit further details.

9.2 Retractability: Boosted Induction

Our aim is to show that, by adding the groundedness schema for $\Sigma_0(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})})$ formulas to $\mathbf{FKF}(1,\mathrm{Q})$, we still get a theory that is retractable to $\mathbf{I}(\Sigma_1)$. For this, we first apply our earlier results from Chapter 4 to *boost* inductions: The aim is to show, mainly, that Γ induction can (under some requirements on Γ) be boosted to $\Sigma_0(\Gamma)$ induction.

In the present section, we rely on ideas by Hájek and Pudlák [18], Chapter I, Section 2. Like in Section 8.3, we shall work in extensions of $\mathbf{A}(0,3)$ rather than in $\mathbf{I}(\Sigma_1)$ directly.

For $k \geq 0$, we let the formula $Fun^{k+1}(u,z)$ express that u is a (total) function $\{0,\dots,z\}^{k+1} \to \{0,1\}$. Thus, the definition of $Fun^{k+1}(u,z)$ is as follows.

$$
\begin{aligned}
&Seq(u) \wedge lh(u) = (z+\overline{1})^{k+1} \wedge \\
&(\forall v_0,\dots,v_k \leq z)(\exists v_{k+1} \leq \overline{1})(\exists w \in u) \\
&\big(w = \langle v_0,\dots,v_k,v_{k+1}\rangle\big) \wedge \\
&(\forall w \in u)(\exists v_0,\dots,v_k \leq z)(\exists v_{k+1} \leq \overline{1}) \\
&\big(w = \langle v_0,\dots,v_k,v_{k+1}\rangle\big) \wedge \\
&(\forall v_0,\dots,v_k \leq z)(\forall v_{k+1},v_{k+2} \leq \overline{1}) \\
&\big(\langle v_0,\dots,v_k,v_{k+1}\rangle \in u \wedge \langle v_0,\dots,v_k,v_{k+2}\rangle \in u \to v_{k+1} = v_{k+2}\big)
\end{aligned}
$$

We then have the following.

Proposition 9.8. $\mathbf{A}(0,3)$ *proves, for some function symbols* $p(z)$ *and* $q(z)$ *in* $\mathcal{L}(0,3)$,

$$(\exists u < q(z))\big(Fun^{k+1}(u,z) \wedge (\forall v_0,\dots,v_k \leq z)(\langle v_0,\dots,v_k,\overline{1}\rangle \in u)\big)$$

and $q(z) = \overline{4} \cdot (\overline{2}^{p(z)})^2$ *and*

$$p(z) = (z+\overline{1})^{k+1} \cdot |\langle \underbrace{z,\dots,z}_{k+1 \text{ times}}, \overline{1}\rangle| + \overline{1}.$$

Proof. By applying $\mathsf{IS}(\Sigma_0^{(0,3)})$, we first show

$$(\forall v_0, \ldots, v_k \le z)(\forall v_{k+1} \le \overline{1})\big(\langle v_0, \ldots, v_k, v_{k+1}\rangle \le \langle \underbrace{z, \ldots, z}_{k+1 \text{ times}}, \overline{1}\rangle\big)$$

and then use the definition of $Fun^{k+1}(u, z)$ to obtain

$$(\forall u, z)\big(Fun^{k+1}(u, z) \to (\forall w < lh(u))((u)_w \le \langle \underbrace{z, \ldots, z}_{k+1 \text{ times}}, \overline{1}\rangle)\big).$$

So, we may apply Proposition 4.11 to derive the following. First, that for some $u < q(\overline{0})$, we have $Fun^{k+1}(u, \overline{0}) \wedge \langle \overline{0}, \ldots, \overline{0}, \overline{1}\rangle \in u$. Second, if there is $u < q(z)$ with

$$Fun^{k+1}(u, z) \wedge (\forall v_0, \ldots, v_k \le z)(\langle v_0, \ldots, v_k, \overline{1}\rangle \in u),$$

there is $u < q(z + \overline{1})$ with

$$Fun^{k+1}(u, z + \overline{1}) \wedge (\forall v_0, \ldots, v_k \le z + \overline{1})(\langle v_0, \ldots, v_k, \overline{1}\rangle \in u).$$

(For instance: Apply Corollary 4.17 to obtain a function symbol $f(u, x, z)$ such that $\mathbf{A}(0, 3)$ proves $f(u, \overline{0}, z) = u$ and, for every $x > \overline{0}$,

$$f(u, x, z) = \begin{cases} f(u, x - \overline{1}, z) * \langle x\rangle & : (1) \\ f(u, x - \overline{1}, z) & : (2) \end{cases}$$

where (1) and (2) are, in that given order, the subsequent two conditions.

$$(\exists v_0, \ldots, v_k \le z + \overline{1})\big(x = \langle v_0, \ldots, v_k, \overline{1}\rangle \wedge x \notin f(u, x - \overline{1}, z)\big)$$
$$(\forall v_0, \ldots, v_k \le z + \overline{1})\big(x = \langle v_0, \ldots, v_k, \overline{1}\rangle \to x \in f(u, x - \overline{1}, z)\big)$$

Then consider

$$f(u, \langle \underbrace{z + \overline{1}, \ldots, z + \overline{1}}_{k+1 \text{ times}}, \overline{1}\rangle, z)$$

to conclude the induction step.) By application of $\mathsf{IS}(\Sigma_0^{(0,3)})$, we finally obtain that, for arbitrary z,

$$(\exists u < q(z))(Fun^{k+1}(u,z) \wedge (\forall v_0,\dots,v_k \leq z)(\langle v_0,\dots,v_k,\overline{1}\rangle \in u)),$$

as required. □

For the next step, we have to recall Definition 3.7.

Proposition 9.9. *Let* S *be an extension of* $\mathbf{A}(0,3) \cup \mathbf{I}(\Gamma)$ *and* Γ *a given class of formulas that is closed in* S *as follows: If* B *is a* Γ *formula in* S*, then so is* $(\forall x \leq t)B$*, where* t *is an arbitrary* $\mathcal{L}(0,3)$ *term. Then* S *proves*

$$\exists y \neg B(\vec{x},y) \rightarrow \exists y(\neg B(\vec{x},y) \wedge (\forall z < y)B(\vec{x},z))$$

provided $B(\vec{x},y)$ *is a* Γ *formula in* S *and* y,z *are not among* $\vec{x}$*.*

Proof. Like the proof of Proposition 3.10, but by observing that $(\forall z < y)B(\vec{x},z)$ is a Γ formula in S by assumption. □

For any formula $A(v_0,\dots,v_k)$ with the indicated free variables, we now define a form $\mathbb{C}(A,u,z)$ to express that, for every $v_i \leq z$, the extension (meaning) of A is coded in the sequence u.

Definition 9.10. For given formula $A(v_0,\dots,v_k)$ with the indicated free variables, form $\mathbb{C}(A,u,z)$ is defined as follows.

$$\begin{aligned}&Fun^{k+1}(u,z) \wedge \\ &(\forall v_0,\dots,v_k \leq z)\big(\langle v_0,\dots,v_k,\overline{1}\rangle \in u \leftrightarrow A(v_0,\dots,v_k)\big)\end{aligned}$$

Proposition 9.11. *Let* S *be an extension of* $\mathbf{A}(0,3) \cup \mathbf{I}(\Gamma)$ *and* Γ *some class of formulas that is closed in* S *as follows, where* t *are* $\mathcal{L}(0,3)$ *terms.*

- *If* B *is* $\Sigma_0^{(0,3)}$*, then* B *is a* Γ *formula in* S*.*
- *If* B,C *are* Γ *formulas in* S*, then so is* $(B \vee C)$*.*

- *If B, C are Γ formulas in S, then so is $(B \wedge C)$.*
- *If B is a Γ formula in S, then so is $(\exists x \leq t)B$.*
- *If B is a Γ formula in S, then so is $(\forall x \leq t)B$.*

Then, for every $\Sigma_0(\Gamma)$ formula A, $\mathsf{S} \vdash \exists u \mathbb{C}(A, u, z)$.

Proof. By (meta) induction on the rank of $A(v_0, \dots, v_k)$. For the base case, we distinguish two sub-cases. (Recall the definition of $\Sigma_0(\Gamma)$.)

- $A(v_0, \dots, v_k)$ belongs to Γ. Let z be arbitrary. By Proposition 9.8 and logic, we have for some u that

  $$\begin{aligned}&Fun^{k+1}(u, z) \wedge \\ &(\forall v_0, \dots, v_k \leq z)\big(A(v_0, \dots, v_k) \rightarrow \langle v_0, \dots, v_k, \overline{1}\rangle \in u\big),\end{aligned}$$

  which we abbreviate as $D(u, z)$. By the closure conditions of Γ, $\neg D(u, z)$ is a Γ formula in S. So, by Proposition 9.9, we obtain

  $$\exists u\big(D(u, z) \wedge (\forall v < u)\neg D(v, z)\big).$$

  Now $\exists u \mathbb{C}(A, u, z)$ may be concluded as follows: Suppose towards contradiction that, for some $v_0, \dots, v_k \leq z$, we had $\neg A(v_0, \dots, v_k) \wedge \langle v_0, \dots, v_k, \overline{1}\rangle \in u$. But then, by replacing $\langle v_0, \dots, v_k, \overline{1}\rangle$ by $\langle v_0, \dots, v_k, \overline{0}\rangle$, we would get a $v < u$ such that $D(v, z)$. This contradicts $(\forall v < u)\neg D(v, z)$.

- A is $\neg B(v_0, \dots, v_k)$, where B belongs to Γ. By the first sub-case, we obtain $\exists u \mathbb{C}(B, u, z)$. By Proposition 4.15, there is some function symbol $f(u, y)$ in $\mathcal{L}(0, 3)$ such that $\mathbf{A}(0, 3)$ proves (for $w < (z + \overline{1})^{k+1}$) that

  $$(f(u, (z + \overline{1})^{k+1}))_w = \begin{cases} \langle v_0, \dots, v_k, \overline{0}\rangle & : (1) \\ \langle v_0, \dots, v_k, \overline{1}\rangle & : (2) \end{cases}$$

where (1) and (2) are, in that given order, the subsequent conditions.

$$(u)_w = \langle v_0, \ldots, v_k, \overline{1}\rangle$$
$$(u)_w = \langle v_0, \ldots, v_k, \overline{0}\rangle$$

Thus, $\mathbb{C}(A, f(u, (z+\overline{1})^{k+1}), z)$.

For the induction step, we consider only one sub-case to show the strategy.

- A is $(\exists x \leq t(v_0, \ldots, v_k))B(x, v_0, \ldots, v_k)$, where B is a $\Sigma_0(\Gamma)$ formula. Let $h(z)$ be a function symbol in $\mathcal{L}(0,3)$ to denote the maximum

  $$h(z) = \max(\{z\} \cup \{t(v_0, \ldots, v_k) : v_0, \ldots, v_k \leq z\})$$

  in $\mathbf{A}(0,3)$. By the induction hypothesis, it is assumed that $\exists u \mathbb{C}(B, u, h(z))$. By Proposition 4.15, there is some function symbol $f(u,y)$ in $\mathcal{L}(0,3)$ such that $\mathbf{A}(0,3)$ proves (for any $w < (z+\overline{1})^{k+1}$) that

  $$(f(u, (z+\overline{1})^{k+1}))_w = \begin{cases} \langle v_0, \ldots, v_k, \overline{0}\rangle & : (1) \\ \langle v_0, \ldots, v_k, \overline{1}\rangle & : (2) \end{cases}$$

  where (1) and (2) are, respectively, the conditions:

  $$(\forall x \leq t(v_0, \ldots, v_k))(\langle x, v_0, \ldots, v_k, \overline{0}\rangle \in u)$$
  $$(\exists x \leq t(v_0, \ldots, v_k))(\langle x, v_0, \ldots, v_k, \overline{1}\rangle \in u)$$

  Hence, $\mathbb{C}(A, f(u, (z+\overline{1})^{k+1}), z)$.

Other cases are treated accordingly. □

Before we conclude the present section, we need to recall the following (instance of a) well-known result.

Proposition 9.12. *Let* Γ *be a class of formulas and let* S *be an extension of* $\mathbf{A}(0,3)$ *by the schema*

$$\forall y\big((\forall z < y)A(\vec{x},z) \to A(\vec{x},y)\big) \to A(\vec{x},y), \tag{9.2}$$

where $A(\vec{x},y)$ *belongs to* Γ *and* y, z *are not among* $\vec{x}$*. Then* S *can prove all formulas in* $\mathsf{IS}(\Gamma)$*.*

Proof. Let $A(\vec{x},y)$ from Γ be arbitrary. Assume in S that

$$A(\vec{x},\overline{0}) \wedge \forall y\big(A(\vec{x},y) \to A(\vec{x},y+\overline{1})\big). \tag{9.3}$$

Let v be arbitrary. If $v = \overline{0}$, we have by (9.3) that $A(\vec{x},v)$ and

$$(\forall z < v)A(\vec{x},z) \to A(\vec{x},v). \tag{9.4}$$

Else, if $v > \overline{0}$, we have $v = u + \overline{1}$ for some u. Suppose in S that $(\forall z < v)A(\vec{x},z)$. Then $A(\vec{x},u)$ and, by (9.3), we obtain $A(\vec{x},v)$. Hence, we have (9.4) again. Finally, (9.2) yields $A(\vec{x},y)$. ☐

We eventually obtain the following proposition about *boosted* $\Sigma_0(\Gamma)$ *induction.*

Proposition 9.13. *Let* S *be an extension of* $\mathbf{A}(0,3) \cup \mathbf{I}(\Gamma)$ *and let* Γ *be some class of formulas that is closed in* S *as in* 9.11*. Then* S *proves all formulas in* $\mathsf{IS}(\Sigma_0(\Gamma))$*.*

Proof. Let $A(x_0,\dots,x_k,y)$ be a $\Sigma_0(\Gamma)$ formula. Assume in S that $\exists y \neg A(x_0,\dots,x_k,y)$. By Proposition 9.11, we obtain

$$\exists u \mathbb{C}(\neg A, u, \max(x_0,\dots,x_k,y))$$

and, therefore, may consider

$$w = (\mu v < y + \overline{1})(\langle x_0,\dots,x_k,v,\overline{1}\rangle \in u).$$

Then we have in S that

$$\neg A(x_0,\dots,x_k,w) \wedge (\forall v < w)A(x_0,\dots,x_k,v).$$

In sum:

$$\begin{aligned}&\exists y\big(\neg A(x_0,\ldots,x_k,y)\big) \rightarrow\\&\exists y\big(\neg A(x_0,\ldots,x_k,y) \wedge (\forall v < y)A(x_0,\ldots,x_k,v)\big)\end{aligned}$$

Therefore, it only remains to apply Proposition 9.12. □

9.3 Retractability: Least Fixed Point

We are now in a position to establish formally that $\mathbb{T}Fip(z)$ defines the *least* $\Sigma_0(\Sigma_1^{(0,3)})$ definable fixed point of the finitist truth operator.

The key lemma is the following.

Lemma 9.14. $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)})$ *proves, for every* $\Sigma_0(\Sigma_1^{(0,3)})$ *formula* $A(x)$*, the following.*

$$\forall y\big(\mathbb{T}(\{x : A(x)\},y) \rightarrow A(y)\big) \rightarrow \big(\mathbb{T}Fip(x) \rightarrow A(x)\big)$$

Proof. Let $A(x)$ be some $\Sigma_0(\Sigma_1^{(0,3)})$ formula. Assume in theory $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)})$ that

$$\forall y\big(\mathbb{T}(\{x : A(x)\},y) \rightarrow A(y)\big) \wedge \mathbb{T}Fip(x), \tag{9.5}$$

where $\mathbb{T}Fip(x)$ is the formula $(\exists u,v)\big(\mathbb{T}Seq(u,v) \wedge x \in (u)_v\big)$ and $\mathbb{T}Seq(u,v)$, remember, is the following formula.

$$\begin{aligned}&Seq(u) \wedge lh(u) = v + \overline{1} \wedge\\&(\forall w \leq v)Seq((u)_w) \wedge (u)_{\overline{0}} = \langle\ \rangle \wedge\\&(\forall w < v)(\forall y < u)\big(\big(y \in (u)_w \rightarrow y \in (u)_{w+\overline{1}}\big) \wedge\\&\qquad\qquad\big(y \in (u)_{w+\overline{1}} \rightarrow \mathbb{T}(\{x : x \in (u)_w\},y)\big)\big)\end{aligned} \tag{9.6}$$

By Proposition 9.13, we may apply $\mathsf{IS}(\Sigma_0(\Sigma_1^{(0,3)}))$ with induction formula

$$w \leq v \rightarrow (\forall y < u)\big(y \in (u)_w \rightarrow A(y)\big) \tag{9.7}$$

and induction variable w. For $w = \overline{0}$, the claim is immediate because $(u)_{\overline{0}} = \langle\ \rangle$ by (9.6). For induction step, we have by (9.6) that

$$(\forall y < u)\bigl(y \in (u)_{w+\overline{1}} \rightarrow \mathbb{T}(\{x : x \in (u)_w\}, y)\bigr)$$

and by induction hypothesis that

$$(\forall y < u)\bigl(y \in (u)_{w+\overline{1}} \rightarrow \mathbb{T}(\{x : x \in A(x)\}, y)\bigr).$$

Consequently, by (9.5), we obtain

$$(\forall y < u)\bigl(y \in (u)_{w+\overline{1}} \rightarrow A(y)\bigr).$$

This establishes (9.7) and, in particular,

$$(\forall y < u)\bigl(y \in (u)_v \rightarrow A(y)\bigr).$$

Since $x < u$ and $x \in (u)_v$, we conclude $A(x)$. ☐

9.4 Conclusions

Summing up, we have arrived at the following.

Theorem 9.15. $\mathbf{FG}(1, \mathrm{Q})$ *is retractable to* $\mathbf{I}(\Sigma_1)$ *by* $(\mathcal{A} \cdot \mathcal{F})$.

Proof. It suffices to extend the proof of 8.38: That is, we need to show that, for each formula B in $\mathsf{GS}(\Sigma_0(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})}))$, we have that $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)}) \vdash B^{\mathcal{F}}$. Thus, let A be a $\Sigma_0(\mathrm{P}\Sigma_1^{(1,\mathrm{Q})})$ formula. Then $A^{\mathcal{F}}$ is $\Sigma_0(\Sigma_1^{(0,3)})$ in $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)})$. By Lemma 9.14, the latter proves

$$\forall y\bigl(\mathbb{T}(\{x : A(x)^{\mathcal{F}}\}, y) \rightarrow A(y)^{\mathcal{F}}\bigr) \rightarrow \bigl(\mathbb{T}Fip(x) \rightarrow A(x)^{\mathcal{F}}\bigr).$$

Hence, by preservation of logical structure, we obtain

$$\big(\forall y\big(\mathbb{T}(\{x : A(x)\}, y) \to A(y)\big) \to \big(T(x) \to A(x)\big)\big)^{\mathcal{F}}$$

in $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_1^{(0,3)})$, as required. $\square$

By reasoning as in the proof of Corollary 8.39, we may finally conclude the following.

Corollary 9.16. *For each* $\mathcal{L}(0,\mathrm{Q})$ *formula* A *and each* $\mathcal{L}(1,\mathrm{Q})$ *formula* B*, the following hold.*

(1) $\mathbf{FG}(1,\mathrm{Q}) \vdash B$ *implies* $\mathbf{I}(\Sigma_1) \vdash (B^{\mathcal{F}})^{\mathcal{A}}$.

(2) $\mathbf{FG}(1,\mathrm{Q}) \vdash A$ *if and only if* $\mathbf{I}(\Sigma_1) \vdash A$.

Since FKF axioms may (by adding basically the negation of (9.1) as a further disjunct to $\mathbb{T}$) alternatively be interpreted by finite computation sequences that include a Truth Teller in their initial component, we see by Remark 9.7 above that $\mathbf{FKF}(1,\mathrm{Q})$ and $\mathbf{FG}(1,\mathrm{Q})$ are not *the same theories* after all. However, the following is now also immediate.

Corollary 9.17. *For every* $\mathcal{L}(0,3)$ *formula* A *and* $\mathcal{L}(1,3)$ *formula* B*, the following hold.*

(1) $\mathbf{FG}(1,\mathrm{Q}) \vdash B$ *implies* $\mathbf{FKF}(1,\mathrm{Q}) \vdash B^{\mathcal{F}}$.

(2) $\mathbf{FG}(1,\mathrm{Q}) \vdash A$ *if and only if* $\mathbf{FKF}(1,\mathrm{Q}) \vdash A$.

Proof. We only consider (1) because (2) can be immediately obtained from (1). But, for (1), we only have to observe that $\mathbf{I}(\Sigma_1)$ is a sub-theory of $\mathbf{FKF}(1,\mathrm{Q})$ and that $\mathbf{A}(1,3) \vdash B^{\mathcal{F}} \leftrightarrow (B^{\mathcal{F}})^{\mathcal{A}}$. Hence, Corollary 9.16 (1) applies. $\square$

10 Disquotation

In Section 8.2, we have already encountered the so-called *Uniform Tarski Schema*. Let us recall it.

Definition 10.1. Let Γ be a class of formulas. The *Uniform Tarski Schema* $\mathsf{UTS}(\Gamma)$ is

$$\bigwedge\nolimits_{i=0}^{k} ClTerm(u_i) \rightarrow$$
$$\left(T(\ulcorner A \urcorner \left[\frac{u_0, \dots, u_k}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner}\right]) \leftrightarrow A\left(\frac{val(u_0), \dots, val(u_k)}{v_0, \dots, v_k}\right)\right)$$

where $A(v_0, \dots, v_k)$ belongs to Γ.

As the name suggests, this schema goes back to Tarski [49]. Many philosophers have advocated *deflationism about truth*: the view that, very roughly, all there is about the concept of truth is already captured by the Tarski Schema. One of the prominent supporters of this view is Horwich [28].

In Section 8.2, Lemma 8.14, it was shown that the Finitist Kripke-Feferman Theory $\mathbf{FKF}(1, \mathrm{Q})$ proves (every axiom of) the

Uniform Tarski Schema for all the *positive essential* Σ_1 *formulas*: $\mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})})$. We have interpreted $\mathsf{FKF}(1,\mathrm{Q})$ axioms by a formula $\mathbb{T}Fip(x)$ that expresses, informally, that x is witnessed by some finite computation (sequence). We relied on the fixed point property that

$$\mathbb{T}Fip(y) \leftrightarrow \mathbb{T}(\{x : \mathbb{T}Fip(x)\}, y),$$

where the form $\mathbb{T}$ defines such finite computations, to derive the interpretations of the truth axioms.

The major goal of this brief chapter is to illustrate that, conversely, the schema $\mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})})$ may be used to interpret the $\mathsf{FKF}(1,\mathrm{Q})$ axioms (not in language of arithmetic but) in the language of truth. For this, we only have to mimic the fixed point property: By a diagonalization lemma (Lemma 10.2), we may derive for a $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ formula $D(x)$ that

$$D(y) \leftrightarrow \mathbb{T}(\{x : T(\ulcorner D(z) \urcorner [num(x)/\ulcorner z \urcorner])\}, y)$$

and obtain (the fixed point property)

$$D(y) \leftrightarrow \mathbb{T}(\{x : D(x)\}, y)$$

by means of $\mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})})$. The basic idea is very similar as in Cantini [6] or in Halbach [20].

10.1 Self-Reference

Note: We have introduced formal machinery for self-reference *via terms* in Section 6.1. In the present section, our concern shall be self-reference *via unbounded existential quantification.* (Note: Unbounded existential quantification was needed for defining graphs of primitive recursive functions.)

As usual, we start off by proving a diagonalization lemma, but one which is adapted to our context.

Lemma 10.2. *Let* $A(v_0, \ldots, v_k)$ *be a* $\mathrm{PE}\Sigma_1^{(1,3)}$ *formula. Then there is some* $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ *formula* $B(v_1, \ldots, v_k)$ *so that* $\mathbf{A}(1,3)$ *proves the following.*

$$A(\ulcorner B(v_1, \ldots, v_k)\urcorner, v_1, \ldots, v_k) \leftrightarrow B(v_1, \ldots, v_k)$$

Proof. Let $A(v_0, \ldots, v_k)$ be some $\mathrm{PE}\Sigma_1^{(1,3)}$ formula. Consider the formula $C(v_0, \ldots, v_k)$ defined as

$$A(v_0[num(v_0)/\ulcorner v_0\urcorner], v_1, \ldots, v_k).$$

Note that, by using $\mathcal{A}$ for eliminating auxiliary function symbols, as defined in 3.53, formula $C^{\mathcal{A}}$ is $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ and $\mathbf{A}(1,3)$ proves:

$$C \leftrightarrow C^{\mathcal{A}}$$

Moreover, there is some number n such that we have in $\mathbf{A}(1,3)$ that $\ulcorner C^{\mathcal{A}}\urcorner = \overline{n}$. So, let

$$B(v_1, \ldots, v_k) := C^{\mathcal{A}}(\overline{n}/v_0)$$

and note that, by Proposition 5.18 (1), $\mathbf{A}(1,3)$ proves

$$\ulcorner B(v_1, \ldots, v_k)\urcorner = \overline{n}[num(\overline{n})/\ulcorner v_0\urcorner].$$

Thus, we have

$$A(\ulcorner B(v_1, \ldots, v_k)\urcorner, v_1, \ldots, v_k) \leftrightarrow B(v_1, \ldots, v_k)$$

in $\mathbf{A}(1,3)$, as required. □

Turning back to the Liar Paradox, we get the following.

Corollary 10.3. *There exists some* $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ *sentence* D *such that* $\mathbf{A}(1,3)$ *proves the following.*

$$T(\ulcorner \neg D\urcorner) \leftrightarrow D$$

Proof. Let $A(v)$ be $T([\neg v])$ and apply Lemma 10.2. $\square$

We have considered the putative inconsistencies raised by the classical negation principle that, for arbitrary statements A,

$$T(\ulcorner \neg A \urcorner) \leftrightarrow \neg T(\ulcorner A \urcorner)$$

in conjunction with iteration principle that, for arbitrary closed terms t,

$$T(t) \leftrightarrow T(\ulcorner T(t) \urcorner)$$

in more detail in Section 6.2. As we have seen there, if A is restricted to be contained in the Σ_0 fragment of $\mathcal{L}(1, \mathrm{Q})$ and t is restricted to be an $\mathcal{L}(0, \mathrm{Q})$ term, the two principles are mutually consistent (with respect to our syntax coding from Chapter 5).

However, now that we are considering an axiomatization that features the truth axiom (FKF10) for the *unbounded* existential quantification, the situation is different. Indeed, as shown in the next proposition, the negation principle is not consistent with the Finitist Kripke-Feferman Theory.

Proposition 10.4. $\mathbf{A}(1, 3) \cup \mathsf{FKF}(1, \mathrm{Q})$ *is not consistent with*

$$Sent(x) \to \big(T([\neg x]) \leftrightarrow \neg T(x)\big).$$

Proof. By Corollary 10.3, there is a $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ sentence D so that $\mathbf{A}(1, 3) \cup \mathsf{FKF}(1, \mathrm{Q})$ joined with the negation axiom proves

$$\neg T(\ulcorner D \urcorner) \leftrightarrow D.$$

However, by Lemma 8.14, $\mathbf{A}(1, 3) \cup \mathsf{FKF}(1, \mathrm{Q})$ also proves

$$T(\ulcorner D \urcorner) \leftrightarrow D,$$

and, therefore, we obtain $D \leftrightarrow \neg D$. $\square$

Turning back to groundedness briefly, recall that, by Proposition 9.6, the Groundedness Schema allows us to prove the Consistency Theorem that, for arbitrary statements A,

$$T(\ulcorner \neg A \urcorner) \rightarrow \neg T(\ulcorner A \urcorner).$$

With respect to the Uniform Tarski Schema, this implies the following one-directional variant.

(Consult also Cantini [7], Theorem 8.8., for a similar result in the non-finitist setting.)

Proposition 10.5. *Let* $A(v_0, \ldots, v_k)$ *be an arbitrary* $\mathcal{L}(1, \mathrm{Q})$ *formula. Then* $\mathbf{FG}(1, \mathrm{Q})$ *proves the following.*

$$\begin{gathered} \bigwedge\nolimits_{i=0}^{k} ClTerm(u_i) \rightarrow \\ \left(T(\ulcorner A \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right]) \rightarrow A\left(\frac{val(u_0), \ldots, val(u_k)}{v_0, \ldots, v_k}\right)\right) \end{gathered}$$

Proof. By extending the proof of Lemma 8.14. Most notably, if $A(v_0, \ldots, v_k)$ is negative literal $\neg T(t)(v_0, \ldots, v_k)$, then we can by assumption $\bigwedge_{i=0}^{k} ClTerm(u_i)$ and the Consistency Theorem (see Proposition 9.6) derive that

$$T(\ulcorner \neg T(t) \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right]) \rightarrow \neg T(\ulcorner T(t) \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right])$$

and reason exactly as in the proof of Lemma 8.14 to obtain the equivalence

$$\neg T(\ulcorner T(t) \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right]) \leftrightarrow \neg T(t)\left(\frac{val(u_0), \ldots, val(u_k)}{v_0, \ldots, v_k}\right)$$

as required.

(Keep also in mind that, if $A(v_0, \ldots, v_k)$ has the form of an unbounded universal quantification, we require the one-directional axiom (FKF11).) □

Hence, this time, we may apply Corollary 10.3 to obtain the following proposition related to our earlier remarks from Chapter 9: The Liar is not grounded, that is, not contained in the least fixed point of the finitist truth operator.

Proposition 10.6. *Let D be an $\mathcal{L}(1,\mathrm{Q})$ Liar such that*

$$\mathbf{A}(1,3) \vdash \neg T(\ulcorner D \urcorner) \leftrightarrow D. \tag{10.1}$$

Then $\mathbf{FG}(1,\mathrm{Q}) \vdash \neg T(\ulcorner D \urcorner)$.

Proof. Suppose towards a contradiction that $T(\ulcorner D \urcorner)$. Then, by Proposition 10.5, we have D and, by (10.1), we obtain $\neg T(\ulcorner D \urcorner)$. Hence, $T(\ulcorner D \urcorner) \wedge \neg T(\ulcorner D \urcorner)$, and we conclude $\neg T(\ulcorner D \urcorner)$. ☐

Remark 10.7. To see that $\mathbf{A}(1,3) \cup \mathsf{FKF}(1,\mathrm{Q})$ is not consistent with the axiom that

$$\sim T(val(x)) \leftrightarrow T([\sim T(x)]), \tag{10.2}$$

where x is a code of a closed $\mathcal{L}(1,\mathrm{Q})$ term, we just follow a dual of the argumentation above. That is, we follow the same line of reasoning as in the proof of Lemma 10.2 to get in $\mathbf{A}(1,3)$ that

$$\neg T(\ulcorner D \urcorner) \leftrightarrow D,$$

where D is an $\mathrm{E}\Sigma_1^{(1,\mathrm{Q})}$ sentence (with negated truth predicate) in $\mathbf{A}(1,3)$, and we extend the proof of Lemma 8.14 by using (10.2) to prove (any instance of the Uniform Tarski Schema for $\mathrm{E}\Sigma_1^{(1,\mathrm{Q})}$ formulas and, therefore, in particular)

$$T(\ulcorner D \urcorner) \leftrightarrow D,$$

which results in a contradiction $D \leftrightarrow \neg D$ again.

10.2 Interpretation

Let us now apply Lemma 10.2 to interpret axioms of $\mathsf{FKF}(1, \mathrm{Q})$. For this, recall the operator form $\mathbb{T}$ from Definition 8.20.

The key lemma is the following.

Lemma 10.8. $\mathbf{A}(1,3) \cup \mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})})$ *proves, for some* $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ *formula* $DQ(z)$*, the following.*

$$DQ(z) \leftrightarrow \mathbb{T}(\{x : DQ(x)\}, z)$$

Proof. Define $D(y, z)$ as

$$\mathbb{T}(\{x : T(y[num(x)/\ulcorner z \urcorner])\}, z).$$

By Lemma 10.2, there is a $\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}$ formula $DQ(z)$ such that $\mathbf{A}(1,3)$ proves

$$DQ(z) \leftrightarrow D(\ulcorner DQ(z) \urcorner, z)$$

where $D(\ulcorner DQ(z) \urcorner, z)$ actually is the following.

$$\mathbb{T}(\{x : T(\ulcorner DQ(z) \urcorner[num(x)/\ulcorner z \urcorner])\}, z)$$

Therefore, it only remains to employ $\mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})})$ and Proposition 5.20 (1). ☐

We may now use the formula $DQ(x)$ in exactly the same way as we have used $\mathbb{T}Fip(x)$ (truth fixed point) in Section 8.4.

Thus, the anticipated result is the following.

Lemma 10.9. $\mathbf{A}(1,3) \cup \mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})})$ *proves the following theorems, sub-divided into cases for better readability.*

Case A: Arithmetical Literals

(DQ1) $Sent([x = y]) \rightarrow$
$\big(DQ([x = y]) \leftrightarrow val(x) = val(y)\big)$

(DQ2) $Sent([x \neq y]) \to$
$\big(DQ([x \neq y]) \leftrightarrow val(x) \neq val(y)\big)$

(DQ3) $Sent([x \leq y]) \to$
$\big(DQ([x \leq y]) \leftrightarrow val(x) \leq val(y)\big)$

(DQ4) $Sent([x > y]) \to$
$\big(DQ([x > y]) \leftrightarrow val(x) > val(y)\big)$

Case B: Truth Literals

(DQ5) $Sent([T(x)]) \to$
$\big(DQ([T(x)]) \leftrightarrow DQ(val(x))\big)$

(DQ6) $Sent([{\sim}T(x)]) \to$
$\big(DQ([{\sim}T(x)]) \leftrightarrow (DQ([\neg val(x)]) \vee \neg Sent(val(x)))\big)$

Case C: Connectives

(DQ7) $Sent([x \vee y]) \to$
$\big(DQ([x \vee y]) \leftrightarrow (DQ(x) \vee DQ(y))\big)$

(DQ8) $Sent([x \wedge y]) \to$
$\big(DQ([x \wedge y]) \leftrightarrow (DQ(x) \wedge DQ(y))\big)$

Case D: Bounded Universal Quantification

(DQ9) $Sent([(\forall x \leq y)z]) \to$
$\big(DQ([(\forall x \leq y)z]) \leftrightarrow (\forall u \leq val(y))DQ(z[num(u)/x])\big)$

Case E: Unbounded Existential Quantification

(DQ10) $Sent([(\exists x)y]) \to$
$\big(DQ([(\exists x)y]) \leftrightarrow (\exists u)DQ(y[num(u)/x])\big)$

Proof. We reason in $\mathbf{A}(1,3) \cup \mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})})$ without repeatedly mentioning it. We focus on three examples to show the strategy.

- For (DQ5), assume that $Sent([T(x)])$. Then

$$DQ([T(x)]) \leftrightarrow \mathbb{T}(\{v : DQ(v)\}, [T(x)])$$

by Lemma 10.8 and

$$DQ(val(x)) \leftrightarrow \mathbb{T}(\{v : DQ(v)\}, [T(x)])$$

by definition of $\mathbb{T}$ and logic.

- For (DQ8), assume that $Sent([x \wedge y])$. Then

$$DQ([x \wedge y]) \leftrightarrow \mathbb{T}(\{v : DQ(v)\}, [x \wedge y])$$

by Lemma 10.8 and

$$(DQ(x) \wedge DQ(y)) \leftrightarrow \mathbb{T}(\{v : DQ(v)\}, [x \wedge y])$$

by definition of $\mathbb{T}$ and logic

- For (DQ10), assume that $Sent([(\exists x)y])$. Then

$$DQ([(\exists x)y]) \leftrightarrow \mathbb{T}(\{v : DQ(v)\}, [(\exists x)y])$$

by Lemma 10.8 and

$$(\exists u)DQ(y[num(u)/x]) \leftrightarrow \mathbb{T}(\{v : DQ(v)\}, [(\exists x)y])$$

by definition of $\mathbb{T}$ and logic

Other cases are treated accordingly. □

Finally, we may translate formulas as follows.

Definition 10.10. $\mathcal{D} : \mathcal{L}(1,3) \to \mathcal{L}(1,3)$ is defined recursively.

$$\begin{aligned}(s = t)^{\mathcal{D}} &:= (s = t)\\ (s \neq t)^{\mathcal{D}} &:= (s \neq t)\\ (s \leq t)^{\mathcal{D}} &:= (s \leq t)\end{aligned}$$

$$
\begin{aligned}
(s > t)^{\mathcal{D}} &:= (s > t)\\
T(t)^{\mathcal{D}} &:= DQ(t)\\
({\sim}T(t))^{\mathcal{D}} &:= \neg DQ(t)\\
(B \vee C)^{\mathcal{D}} &:= (B^{\mathcal{D}} \vee C^{\mathcal{D}})\\
(B \wedge C)^{\mathcal{D}} &:= (B^{\mathcal{D}} \wedge C^{\mathcal{D}})\\
(\exists x B)^{\mathcal{D}} &:= \exists x (B)^{\mathcal{D}}\\
(\forall x B)^{\mathcal{D}} &:= \forall x (B)^{\mathcal{D}}
\end{aligned}
$$

The following two observations are immediate.

Lemma 10.11. *$\mathcal{D}$ preserves logical structure in $\mathbf{A}(1,3)$.*

Proof. By definition of $\mathcal{D}$ and logic. ☐

Lemma 10.12. *For every $\mathcal{L}(0,3)$ formula A, $A^{\mathcal{D}}$ is A.*

Proof. By (meta) induction on the rank of A. ☐

10.3 Conclusions

We conclude the present chapter by drawing a few consequences from earlier observations. For simplicity, we add a definition.

Definition 10.13. $\mathbf{FDQ}(1,\mathrm{Q})$ is the *Finitist Disquotation Theory* (in auxiliary form) defined as follows.

$$\mathbf{A}(1,3) \cup \mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}) \cup \mathsf{TI}(1,\mathrm{Q})$$

The first consequence is:

Theorem 10.14. *For every $\mathcal{L}(0,3)$ formula A and $\mathcal{L}(1,3)$ formula B, the following hold.*

(1) $\mathbf{FKF}(1,\mathrm{Q}) \vdash B$ *implies* $\mathbf{FDQ}(1,\mathrm{Q}) \vdash B^{\mathcal{D}}$.

(2) $\mathbf{FKF}(1,\mathrm{Q}) \vdash A$ *if and only if* $\mathbf{FDQ}(1,\mathrm{Q}) \vdash A$.

Proof. We first consider (1) and then infer part of (2) from it.

- For (1), we notice that, for 2.27 (1), we do not have to establish 2.26 (1). So, in view of Lemma 10.11, it suffices to establish that, for every axiom B in $\mathsf{FKF}(1,\mathrm{Q})$, we have

  $$\mathbf{FDQ}(1,\mathrm{Q}) \vdash B^{\mathcal{D}}.$$

  This is covered by Lemma 10.9.
- The left-to-right direction of (2) is yet covered by (1) and Lemma 10.12. As for the right-to-left direction of (2), we merely have to note that, by Lemma 8.14, $\mathbf{FDQ}(1,\mathrm{Q})$ is a sub-theory of $\mathbf{FKF}(1,\mathrm{Q})$. □

Remark 10.15. We add that Halbach [21], Lemma 19.20, shows implicitly that compositional axioms from $\mathsf{FKF}(1,\mathrm{Q})$ may not be derived in $\mathbf{FDQ}(1,\mathrm{Q})$.

Thus, $\mathbf{FDQ}(1,\mathrm{Q})$ and $\mathbf{FKF}(1,\mathrm{Q})$ may not be the *same theory* after all.

In Chapter 8, we implicitly have established the following as well, where $\mathcal{F}$ is defined in 8.32.

Corollary 10.16. *For any $\mathcal{L}(0,\mathrm{Q})$ formula A and any $\mathcal{L}(1,3)$ formula B, the following hold.*

(1) $\mathbf{FDQ}(1,\mathrm{Q}) \vdash B$ *implies* $\mathbf{I}(\Sigma_1) \vdash (B^{\mathcal{F}})^{\mathcal{A}}$.

(2) $\mathbf{FDQ}(1,\mathrm{Q}) \vdash A$ *if and only if* $\mathbf{I}(\Sigma_1) \vdash A$.

Proof. Since $\mathbf{FDQ}(1,\mathrm{Q})$ indeed is sub-theory of $\mathbf{FKF}(1,\mathrm{Q})$, (1) holds by Corollary 8.39 (1). For (2), we merely have to apply Corollary 8.39 (2) and Theorem 10.14 (2). □

Remark 10.17. We finally add the following.

- For every $\mathcal{L}(1,\mathrm{Q})$ formula B, we have by Corollary 3.54 (2) that
$$\mathbf{B}(1,3) \cup \big(\mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}) \cup \mathsf{TI}(1,\mathrm{Q})\big)^{\mathcal{A}} \vdash B$$
implies $\mathbf{FDQ}(1,\mathrm{Q}) \vdash B$ and, by Corollary 10.16 (1), implies
$$\mathbf{I}(\Sigma_1) \vdash (B^{\mathcal{F}})^{\mathcal{A}}.$$
- For every $\mathcal{L}(0,\mathrm{Q})$ formula A, we have by Corollary 3.54 (2) that
$$\mathbf{B}(1,3) \cup \big(\mathsf{UTS}(\mathrm{PE}\Sigma_1^{(1,\mathrm{Q})}) \cup \mathsf{TI}(1,\mathrm{Q})\big)^{\mathcal{A}} \vdash A$$
if and only if $\mathbf{FDQ}(1,\mathrm{Q}) \vdash A$ and, by Corollary 10.16 (2), if and only if
$$\mathbf{I}(\Sigma_1) \vdash A.$$

11 Iteration

In Chapter 7, we have introduced Finitist FS theories $\mathbf{FFS}(1, k)$ which correspond, respectively, to levels $k > 1$ of the Grzegorczyk hierarchy. In the present chapter, we study *iterated* Finitist KF theories $\mathbf{FKF}(m, \mathrm{Q})$ which correspond, respectively, to the levels $m > 1$ of the *arithmetical* hierarchy.

As was highlighted in Remark 10.7, the principle that

$$T([\sim T(x)]) \leftrightarrow \sim T(val(x)), \tag{11.1}$$

where x may be a code of any closed $\mathcal{L}(0, \mathrm{Q})$ term, may not be added consistently to $\mathbf{A}(1, 3) \cup \mathsf{FKF}(1, \mathrm{Q})$ and, thus, may not be added consistently to $\mathbf{FKF}(1, \mathrm{Q})$ either.

In an iterated Finitist KF axiomatization, not only one, but several truth predicates

$$T_1, T_2, \ldots, T_m$$

(where $m > 1$) are employed and (11.1) is replaced by respective *iteration* principles which state, *only for lower* $i < m$, that

$$T_m([\sim T_i(x)]) \leftrightarrow \sim T_i(val(x)), \tag{11.2}$$

where x may be a code of any closed $\mathcal{L}(0, \mathrm{Q})$ term.

The theories $\mathbf{FKF}(n, \mathrm{Q})$ $(n > 1)$ of Iterated Finitist KF truth will be defined in Section 11.1 below as restrictions of transfinite iterations applied by Jäger et al. [29] and studied, for alternative axiomatizations of type-free truth, also by Fujimoto [16].

While Jäger et al. [29], for example, are interested in transfinitely iterated fixed point theories $\widehat{\mathbf{ID}}_\alpha$ $(\alpha \geq \omega)$, our interest is in Iterated Finitist KF truth as an analogue of the Turing jump. That is, we consider only iterations up to level ω.

The key result of the present chapter will be that, for every $n > 1$, $\mathbf{FKF}(n, \mathrm{Q})$ is retractable to $\mathbf{I}(\Sigma_n)$. While the retraction of $\mathbf{FKF}(1, \mathrm{Q})$ to $\mathbf{I}(\Sigma_1)$ used the notion of finite computation (or verification) sequence, retractions of $\mathbf{FKF}(n, \mathrm{Q})$ (with $n > 1$) to $\mathbf{I}(\Sigma_n)$ will make use of respective notions of higher-level (but still finite) computations in which oracles about results of lower-level computations are employed.

Sections 11.2–11.4 below will be devoted entirely to the proof-theoretic analysis of theories $\mathbf{FKF}(n, \mathrm{Q})$ $(n > 1)$. This proof-theoretic analysis will be structured very similarly as the one in Chapter 8.

11.1 The Theories

Let $\mathsf{FKF}(1, \mathrm{Q})$ be defined as in 8.3. For $m > 1$, we proceed as follows.

Notation: Since, in what follows, we will be concerned exclusively with languages $\mathcal{L}(m, \mathrm{Q})$, where $m > 0$, we will not endow $val(x)$ and related symbols always with a subscript Q, which is clear by the context of the present chapter.

Definition 11.1. Let $m > 1$. $\mathsf{FKF}(m, \mathrm{Q})$ comprises, for every $0 < i < m$, the subsequent *Iterated Finitist Kripke-Feferman* axioms for the $\mathrm{E}\Sigma_1$ fragment of $\mathcal{L}(m, \mathrm{Q})$. We sub-divide the axioms into axiom sets for better readability. Furthermore, we write $Sent(x)$ to mean that x is code of an $\mathcal{L}(m, \mathrm{Q})$ sentence.

Axiom Set A: Arithmetical Literals

(I1) $Sent([x = y]) \rightarrow$
$\big(T_m([x = y]) \leftrightarrow val(x) = val(y)\big)$

(I2) $Sent([x \neq y]) \rightarrow$
$\big(T_m([x \neq y]) \leftrightarrow val(x) \neq val(y)\big)$

(I3) $Sent([x \leq y]) \rightarrow$
$\big(T_m([x \leq y]) \leftrightarrow val(x) \leq val(y)\big)$

(I4) $Sent([x > y]) \rightarrow$
$\big(T_m([x > y]) \leftrightarrow val(x) > val(y)\big)$

Axiom Set B: Lower-Level $(i < m)$ Truth Literals

(I5) $Sent([T_i(x)]) \rightarrow$
$\big(T_m([T_i(x)]) \leftrightarrow T_i(val(x))\big)$

(I6) $Sent([{\sim}T_i(x)]) \rightarrow$
$\big(T_m([{\sim}T_i(x)]) \leftrightarrow {\sim}T_i(val(x))\big)$

Axiom Set C: Self-Applicable Truth Literals

(I7) $Sent([T_m(x)]) \rightarrow$
$\big(T_m([T_m(x)]) \leftrightarrow T_m(val(x))\big)$

(I8) $Sent([{\sim}T_m(x)]) \rightarrow$
$\big(T_m([{\sim}T_m(x)]) \leftrightarrow (T_m([\neg val(x)]) \vee \neg Sent(val(x)))\big)$

Axiom Set D: Connectives

(I9) $Sent([x \vee y]) \rightarrow \big(T_m([x \vee y]) \leftrightarrow (T_m(x) \vee T_m(y))\big)$

(I10) $Sent([x \wedge y]) \rightarrow \big(T_m([x \wedge y]) \leftrightarrow (T_m(x) \wedge T_m(y))\big)$

Axiom Set E: Bounded Universal Quantification

(I11) $Sent([(\forall x \leq y)z]) \rightarrow \big(T_m([(\forall x \leq y)z]) \leftrightarrow (\forall u \leq val(y))T_m(z[num(u)/x])\big)$

Axiom Set F: Unbounded Existential Quantification

(I12) $Sent([(\exists x)y]) \rightarrow \big(T_m([(\exists x)y]) \leftrightarrow \exists u T_m(y[num(u)/x])\big)$

As for Truth Induction, we define the following.

Definition 11.2. For $m > 1$, $\mathsf{TI}(m, \mathrm{Q})$ contains the following (single) axiom of *Truth Induction*, where $Sent(x)$ means that x is code of an $\mathcal{L}(m, \mathrm{Q})$ sentence.

$$Sent([(\forall z)x]) \rightarrow \Big(\big(T_m(x[num(\overline{0})/z]) \wedge \forall y(T_m(x[num(y)/z]) \rightarrow T_m(x[num(y+\overline{1})/z]))\big) \rightarrow T_m(x[num(y)/z])\Big)$$

To allow for a simpler comparison of the iterated Finitist KF theories with certain other theories in the next chapter, we will base the former on a common theory $\mathbf{A}(\omega, 3)$ which is defined as follows.

Definition 11.3. As common language, we define:

$$\mathcal{L}(\omega, 3) := \bigcup_{n<\omega} \mathcal{L}(n, 3)$$

Definition 11.4. As common base theory, we define:

$$\mathbf{A}(\omega, 3) := \bigcup_{n<\omega} \mathbf{A}(n, 3)$$

Observe: $\mathbf{A}(\omega, 3)$ is an $\mathcal{L}(\omega, 3)$ theory that extends $\mathbf{A}(0, 3)$ by the logical axioms for all $\mathcal{L}(\omega, 3)$ formulas.

For every $n > 1$, a theory $\mathbf{FKF}(n, \mathrm{Q})$ may now be defined as follows.

Definition 11.5. For every $n > 1$, $\mathbf{FKF}(n, \mathrm{Q})$ is the *Iterated Finitist Kripke-Feferman Theory* (in auxiliary form) defined as follows.

$$\mathbf{A}(\omega, 3) \cup \bigcup_{m=1}^{n} \big(\mathsf{FKF}(m, \mathrm{Q}) \cup \mathsf{TI}(m, \mathrm{Q})\big)$$

Observe: For any $n > 1$, we may have based $\mathbf{FKF}(n, \mathrm{Q})$ just on the $\mathcal{L}(n, 3)$ theory $\mathbf{A}(n, 3)$. However, as we mentioned before, we reserve the names $\mathbf{FKF}(n, \mathrm{Q})$ for theories with a unified base $\mathbf{A}(\omega, 3)$ only to simplify a comparison in the next chapter.

11.2 Retractability: Lower Bounds

By proceeding like in Section 8.2, we want to establish that, for every $n > 1$, $\mathbf{FKF}(n, \mathrm{Q})$ is retractable to $\mathbf{I}(\Sigma_n)$. Thus, we need to define some mapping $\circ : \mathcal{L}(n,3) \to \mathcal{L}(0,\mathrm{Q})$ that satisfies, in particular, condition 2.26 (3) from the Section 2. That is to say, we first unveil the proof-theoretic power of the truth axioms of $\mathbf{FKF}(n,\mathrm{Q})$.

We first define a more general version of the Uniform Tarski Schema. Let $\mathsf{UTS}(1,\Gamma)$ be defined as in 8.13. For $m > 1$, we proceed as follows.

Definition 11.6. Let $m > 1$ and let Γ be any class of formulas. The *Uniform Tarski Schema* $\mathsf{UTS}(m,\Gamma)$ is defined as

$$\bigwedge\nolimits_{i=0}^{k} ClTerm(u_i) \to$$
$$\big(T_m(\ulcorner A \urcorner \left[\frac{u_0,\ldots,u_k}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right]) \leftrightarrow A\left(\frac{val(u_0),\ldots,val(u_k)}{v_0,\ldots,v_k}\right)\big)$$

where $A(v_0,\ldots,v_k)$ belongs to Γ.

The following may be proved very similarly as in Section 8.2, which is why we can safely omit details.

Lemma 11.7. *Let A be a formula from $\mathsf{UTS}(m,\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})})$, where $m > 1$. Then $\mathbf{A}(m,3) \cup \mathsf{FKF}(m,\mathrm{Q})$ proves A.*

Proof. Like the proof of Lemma 8.14. ☐

Corollary 11.8. *Let $A(v_0,\ldots,v_k)$ be arbitrary $\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})}$ formula, where $m > 1$. Then $\mathbf{A}(m,3) \cup \mathsf{FKF}(m,\mathrm{Q})$ proves the following.*

$$T_m(\ulcorner A \urcorner \left[\frac{num(v_0),\ldots,num(v_k)}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner}\right]) \leftrightarrow A(v_0,\ldots,v_k)$$

Proof. Like the proof of Corollary 8.15. ☐

Lemma 11.9. *Let A be a formula in $\mathsf{IS}(\mathrm{PE}\Sigma_1^{(m,\mathrm{Q})})$, where $m > 1$. Then $\mathbf{A}(m,3) \cup \mathsf{FKF}(m,\mathrm{Q}) \cup \mathsf{TI}(m,\mathrm{Q})$ proves A.*

Proof. Like the proof of Lemma 8.16. ☐

In addition to the above counterparts of results from Section 8.2, we also require the following.

Lemma 11.10. *Every* $\Sigma_{n+1}^{(m,\mathrm{Q})}$ *formula is* $\Sigma_1^{(m+n,\mathrm{Q})}$ *in*

$$\mathbf{A}(m+n,3) \cup \bigcup_{i=1}^{m+n} \mathsf{FKF}(i,\mathrm{Q}). \tag{11.3}$$

Proof. By (meta) induction on n, using m as a parameter. The claim is immediate if n is 0. For the induction step, assume as induction hypothesis: Every $\Sigma_{n+1}^{(m,\mathrm{Q})}$ formula is $\Sigma_1^{(m+n,\mathrm{Q})}$ in (11.3). Let $A(v_0,\ldots,v_k)$ be an $\Sigma_{n+2}^{(m,\mathrm{Q})}$ formula. That is, $A(v_0,\ldots,v_k)$ is $\exists x B(v_0,\ldots,v_k,x)$, where $\neg B(v_0,\ldots,v_k,x)$ is $\Sigma_{n+1}^{(m,\mathrm{Q})}$. By the induction hypothesis, there is a $\Sigma_1^{(m+n,\mathrm{Q})}$ formula $C(v_0,\ldots,v_k,x)$ such that (11.3) proves

$$\neg B(v_0,\ldots,v_k,x) \leftrightarrow C(v_0,\ldots,v_k,x). \tag{11.4}$$

Let p be $m+n+1$. By Corollary 11.8, we may prove in

$$\mathbf{A}(p,3) \cup \bigcup_{i=1}^{p} \mathsf{FKF}(i,\mathrm{Q}) \tag{11.5}$$

that $C(v_0,\ldots,v_k,x)$ is equivalent ($\leftrightarrow$) to

$$T_p(\ulcorner C \urcorner \left[\frac{num(v_0),\ldots,num(v_k),num(x)}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner,\ulcorner x \urcorner}\right])$$

and, by (11.4), that $B(v_0,\ldots,v_k,x)$ is equivalent ($\leftrightarrow$) to

$$\neg T_p(\ulcorner C \urcorner \left[\frac{num(v_0),\ldots,num(v_k),num(x)}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner,\ulcorner x \urcorner}\right]).$$

Hence, $A(v_0,\ldots,v_k)$ is equivalent ($\leftrightarrow$) in (11.5) to

$$\exists x(\neg T_p(\ulcorner C \urcorner \left[\frac{num(v_0),\ldots,num(v_k),num(x)}{\ulcorner v_0 \urcorner,\ldots,\ulcorner v_k \urcorner,\ulcorner x \urcorner}\right]))$$

and, consequently, is $\Sigma_1^{(p,\mathrm{Q})}$ in (11.5), as required. This concludes the proof. $\square$

In sum, the following corollary yields condition 2.26 (3) from Section 2.

Corollary 11.11. *Let* A *be a formula from* $\mathsf{IS}(\Sigma_n)$*, where* $n > 1$*. Then* $\mathbf{FKF}(n, \mathrm{Q})$ *proves* A*.*

Proof. By Lemma 11.9 and Lemma 11.10. $\square$

11.3 Retractability: Fixed Points

The goal of this section is to generalize the truth operation ($\mathbb{T}$) which we defined in 8.20 and now rename as $\mathbb{T}_1(\{x : A(x)\}, y)$. For $m > 1$, the form $\mathbb{T}_m(\{x : A(x)\}, y)$ will be defined so that it may be used to describe finite sequences ($\mathbb{T}_m Seq$) of higher-level computation steps on the basis of an oracle

$$\{x : \mathbb{T}_i Fip(x)\} \quad \text{for } 0 < i < m$$

which reveals the outcomes of computations at some lower level $i < m$.

The definition of these forms is as follows.

Definition 11.12. Let A be a formula with at least x free. For $m > 1$, we define

$$\mathbb{T}_m(\{x : A(x)\}, y) := \bigvee \{(\mathrm{O1})\text{–}(\mathrm{O12})\}$$

where (O1)–(O12) are the following disjuncts in which formulas $\mathbb{T}_i Fip(x)$ with $i < m$ are used. We sub-divide the disjuncts into cases for better readability. Furthermore, let us write $Sent(x)$ to mean that x is code of an $\mathcal{L}(m, \mathrm{Q})$ sentence.

Case A: Arithmetical Literals

(O1) $(\exists x_0, x_1 < y)\big(Sent([x_0 = x_1]) \wedge$
$y = [x_0 = x_1] \wedge val(x_0) = val(x_1)\big)$

(O2) $(\exists x_0, x_1 < y)\big(Sent([x_0 \neq x_1]) \wedge$
$y = [x_0 \neq x_1] \wedge val(x_0) \neq val(x_1)\big)$

(O3) $(\exists x_0, x_1 < y)\big(Sent([x_0 \leq x_1]) \wedge$
$y = [x_0 \leq x_1] \wedge val(x_0) \leq val(x_1)\big)$

(O4) $(\exists x_0, x_1 < y)\big(Sent([x_0 > x_1]) \wedge$
$y = [x_0 > x_1] \wedge val(x_0) > val(x_1)\big)$

Case B: Lower-Level $(i < m)$ Truth Literals

(O5) $(\exists x_0 < y)\bigvee_{i=1}^{m-1}\big(Sent([T_i(x_0)]) \wedge$
$y = [T_i(x_0)] \wedge \mathbb{T}_i Fip(val(x_0))\big)$

(O6) $(\exists x_0 < y)\bigvee_{i=1}^{m-1}\big(Sent([{\sim}T_i(x_0)]) \wedge$
$y = [{\sim}T_i(x_0)] \wedge \neg\mathbb{T}_i Fip(val(x_0))\big)$

Case C: Self-Applicable Truth Literals

(O7) $(\exists x_0 < y)\big(Sent([T_m(x_0)]) \wedge$
$y = [T(x_0)] \wedge A(val(x_0))\big)$

(O8) $(\exists x_0 < y)\big(Sent([{\sim}T_m(x_0)]) \wedge$
$y = [{\sim}T_m(x_0)] \wedge \big(A([\neg val(x_0)]) \vee \neg Sent(val(x_0))\big)\big)$

Case D: Connectives

(O9) $(\exists x_0, x_1 < y)\big(Sent([x_0 \vee x_1]) \wedge$
$y = [x_0 \vee x_1] \wedge \big(A(x_0) \vee A(x_1)\big)\big)$

(O10) $(\exists x_0, x_1 < y)\big(Sent([x_0 \wedge x_1]) \wedge$
$y = [x_0 \wedge x_1] \wedge \big(A(x_0) \wedge A(x_1)\big)\big)$

Case E: Bounded Universal Quantification

(O11) $(\exists x_0, x_1, x_2 < y)\big(\mathit{Sent}([(\forall x_0 \leq x_1)x_2]) \land$
$y = [(\forall x_0 \leq x_1)x_2] \land$
$(\forall x_3 \leq \mathit{val}(x_1))A(x_2[\mathit{num}(x_3)/x_0])\big)$

Case F: Unbounded Existential Quantification

(O12) $(\exists x_0, x_1 < y)\big(\mathit{Sent}([(\exists x_0)x_1]) \land$
$y = [(\exists x_0)x_1] \land \exists x_2 A(x_1[\mathit{num}(x_2)/x_0])\big)$

We first need to clarify the complexity of $\mathbb{T}_m$, where $m > 1$. This requires new notation. First, let us define

$$\underset{0<i<m}{\{x : J_i(x)\}} := \{x : J_1(x)\}, \ldots, \{x : J_{m-1}(x)\}$$

as an abbreviation. Moreover, to highlight negative occurrences $\neg J_i$ of schematic letters in operator forms, we can replace them by respective unnegated occurrences $\overline{J_i}$ of schematic letters. We may now state the definition of iterated (indexed) operator forms as follows.

Definition 11.13. Let $\mathbb{O}_1(\{x : B(x)\}, y)$ be an $\mathrm{E}\Sigma_1^{(0,3)}$ form in which B occurs only positively. Moreover, for $m > 1$, let

$$\mathbb{O}\big(\{x : A(x)\}, \underset{0<i<m}{\{x : J_i(x)\}}, y\big)$$

be an $\mathrm{E}\Sigma_1^{(0,3)}$ form in which A occurs only positively. (Thus, J_i is allowed to occur negatively.) Then, for any $m > 1$, we define $\mathbb{O}_m(\{x : A(x)\}, y)$ to be the $\mathrm{E}\Sigma_m^{(0,3)}$ form

$$\mathbb{O}\big(\{x : A(x)\}, \underset{0<i<m}{\{x : \mathbb{O}_i \mathit{Fip}(x)\}}, y\big).$$

That is, $\mathbb{T}_m$ (where $m > 1$) is substitution instance of some $\mathrm{E}\Sigma_1^{(0,3)}$ operator form and the instance $\mathbb{T}_m$ itself has complexity $\mathrm{E}\Sigma_m^{(0,3)}$ according to our terminology from Section 8.3.

Having clarified this issue, we may now formally establish the existence of fixed points along the lines of Section 8.3. (Notice: $\overline{J_i}$ is schematic for $\neg J_i$ in the tedious lemma below.)

Lemma 11.14. *Let* $m > 1$ *and let* S *be an extension of theory* $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_m^{(0,3)})$. *Suppose that*

$$\mathbb{O}\big(\{x : A(x)\}, \underset{0<i<m}{\{x : J_i(x)\}}, \underset{0<i<m}{\{x : \overline{J_i}(x)\}}, y\big)$$

is an $\mathrm{E}\Sigma_1^{(0,3)}$ *form where* $A, J_i, \overline{J_i}$ *occur only positively. Also, let* $B^<(x,z)$ *be an* $\mathrm{E}\Pi_{m-1}^{(0,3)}$ *formula so that* S *proves*

$$B(x) \leftrightarrow \exists z B^<(x,z)$$

and, for any $0 \leq i < m-1$, *let* $K_i^<(x,z)$ *be an* $\mathrm{E}\Pi_i^{(0,3)}$ *formula so that* S *proves the following.*

$$K_i(x) \leftrightarrow \exists z K_i^<(x,z)$$

Then there is a $\Sigma_0^{(0,3)}$ *form* $\mathbb{O}^<$ *so that*

$$\mathbb{O}\big(\{x : B(x)\}, \underset{0<i<m}{\{x : K_i(x)\}}, \underset{0<i<m}{\{x : \neg K_i(x)\}}, y\big)$$

is provably equivalent ($\leftrightarrow$) *to*

$$\exists z\big(\mathbb{O}^<(\{x : B^<(x,z)\}, \underset{0<i<m}{\{x : K_i^<(x,z)\}}, \underset{0<i<m}{\{x : \neg K_i(x)\}}, y, z)\big)$$

in S *also.*

Proof. Like the proof of Lemma 8.21. □

On the basis of Lemma 11.14, we get the following. (Recall Definition 8.22, according to which $\mathbb{O}_m Seq(u,v)$ is the formula

$$Seq(u) \wedge lh(u) = v + \overline{1} \wedge$$

$$\begin{aligned}&(\forall w \le v) Seq((u)_w) \land (u)_{\overline{0}} = \langle\,\rangle \land \\ &(\forall w < v)(\forall y < u)\big(\big(y \in (u)_w \to y \in (u)_{w+\overline{1}}\big) \land \\ &\qquad\qquad\qquad \big(y \in (u)_{w+\overline{1}} \to \mathbb{O}_m(\{x : x \in (u)_w\}, y)\big)\big)\end{aligned}$$

and $\mathbb{O}_m$ is defined as in 11.13.)

Lemma 11.15. *Let $m > 1$ and $\mathbb{O}_1(\{x : B(x)\}, y)$ a given $\mathrm{E}\Sigma_1^{(0,3)}$ form in which B occurs only positively. Then $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_m^{(0,3)})$ proves, for an $\mathrm{E}\Pi_{m-1}^{(0,3)}$ formula $\mathbb{O}_m Seq^<(u, v, z)$, the following.*

$$\mathbb{O}_m Seq(u, v) \leftrightarrow \exists z\big(u < z \land v < z \land \mathbb{O}_m Seq^<(u, v, z)\big)$$

Proof. Like the proof of Lemma 8.23. □

In contrast to Chapter 8, we require some additional coding machinery for the bounded oracle computations.

As in Section 9.2, let $Fun^1(w, z)$ express that w is a (total) function $\{0, \dots, z\} \to \{0, 1\}$. For every $\mathrm{E}\Sigma_m^{(0,3)}$ form $\mathbb{O}_m$, where index $m > 1$, define $\mathbb{O}_m S^<(u, z)$ as

$$(\exists v < z)(\mathbb{O}_m Seq^<(u, v, z))$$

which is a $\Sigma_0(\mathrm{E}\Sigma_{m-1}^{(0,3)})$ formula. Like in Section 9.2, let formula $\mathbb{C}(\mathbb{O}_m S^<, w, z)$ be defined as follows.

$$Fun^1(w, z) \land (\forall u \le z)\big(\langle u, \overline{1}\rangle \in w \leftrightarrow \mathbb{O}_m S^<(u, z)\big)$$

As first main lemma, we have the following.

Lemma 11.16. *Let $m > 1$ and $\mathbb{O}_1(\{x : B(x)\}, y)$ a given $\mathrm{E}\Sigma_1^{(0,3)}$ form in which B occurs only positively. Then $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_m^{(0,3)})$ proves the following.*

$$\exists w \mathbb{C}(\mathbb{O}_m S^<, w, z)$$

Proof. By Proposition 3.27 and by $\mathsf{BS}(\Sigma_m^{(0,3)})$, we have $\mathsf{IS}(\mathrm{E}\Sigma_{m-1}^{(0,3)})$ available. Thus, we may apply Proposition 9.11. □

Having coded $\mathbb{O}_m S^{<}(x,z)$ successively as finite sequence, we shall apply the following counterpart of Proposition 8.26.

Proposition 11.17. *Let* $r(u,w,v)$ *be the function symbol in* 8.24. *Then* $\mathbf{A}(0,3)$ *proves, for a function symbol* $un^{<}(u,w,z)$ *in* $\mathcal{L}(0,3)$, *the following.*

$$un^{<}(\overline{0},w,z) = \begin{cases} \langle\langle\ \rangle\rangle \end{cases}$$

$$un^{<}(u+\overline{1},w,z) = \begin{cases} r(u, un^{<}(u,w,z), f(u,w,z)) & : (1) \\ un^{<}(u,w,z) & : (2) \\ un^{<}(u,w,z) & : (3) \end{cases}$$

Here (1)–(3) *are, respectively, the conditions*

$$\begin{aligned} & u < z \wedge \langle u, \overline{1}\rangle \in w \\ & u < z \wedge \langle u, \overline{0}\rangle \in w \\ & u \geq z \end{aligned}$$

and $\mathbf{A}(0,3)$ *proves the following equality.*

$$f(u,w,z) = \max(lh(u), lh(un^{<}(u,w,z)))$$

Proof. Like the proof of Proposition 8.26. □

Properties (1)–(3) of Lemma 8.27 under assumptions

$$(\forall u < z)\big(\langle u,\overline{1}\rangle \in w \leftrightarrow (\exists v < z)(\mathbb{O}_m Seq^{<}(u,v,z))\big) \tag{11.6}$$

$$(\forall u < z)\big(\langle u,\overline{0}\rangle \in w \leftrightarrow (\forall v < z)\neg(\mathbb{O}_m Seq^{<}(u,v,z))\big) \tag{11.7}$$

may now be derived for $un^{<}(u,w,z)$ exactly as in 8.27. We omit details.

Finally, we get the second main lemma below. (Recall Definition 8.28.)

Lemma 11.18. *Let $m > 1$ and $\mathbb{O}_1(\{x : B(x)\}, y)$ a given $\mathrm{E}\Sigma_1^{(0,3)}$ form where B occurs only positively. Then.*

$$\mathbb{O}_m Fip(y) \leftrightarrow \mathbb{O}_m(\{x : \mathbb{O}_m Fip(x)\}, y)$$

is provable in $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_m^{(0,3)})$.

Proof. It suffices to modify the right-to-left direction of the proof of Lemma 8.29. Thus, assume that $\mathbb{O}_m(\{x : \mathbb{O}_m Fip(x)\}, y)$. By Lemma 11.15, we have

$$\mathbb{O}_m Fip(x) \leftrightarrow \exists z(\exists u, v < z)\big(\mathbb{O}_m Seq^<(u, v, z) \wedge x \in (u)_v\big).$$

By Lemma 11.14 and the assumption, we get

$$\exists z\big(\mathbb{O}_m^<\big(\big\{x : (\exists u, v < z)\big(\mathbb{O}_m Seq^<(u, v, z) \wedge x \in (u)_v\big)\big\}, y, z\big)\big).$$

By Lemma 11.16, there exists w such that (11.6) and (11.7) are satisfied. Therefore, we may proceed precisely as in the proof of Lemma 8.29, using the function symbol $un^<(u, w, z)$, to conclude $\mathbb{O}_m Fip(y)$. $\square$

Corollary 11.19. *For any $m > 1$, $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_m^{(0,3)})$ proves the following.*

$$\mathbb{T}_m Fip(y) \leftrightarrow \mathbb{T}_m(\{x : \mathbb{T}_m Fip(x)\}, y)$$

Proof. By instantiating Lemma 11.18. $\square$

11.4 Retractability: Upper Bounds

Following the procedure of Section 8.4, we may now define some retractability mapping that satisfy condition 2.26 (2).

Definition 11.20. The mapping $\mathcal{I} : \mathcal{L}(\omega, 3) \to \mathcal{L}(0, 3)$ is defined recursively as follows, where $0 < m < \omega$.

$$(s = t)^{\mathcal{I}} := (s = t)$$

$$
\begin{aligned}
(s \neq t)^{\mathcal{I}} &:= (s \neq t)\\
(s \leq t)^{\mathcal{I}} &:= (s \leq t)\\
(s > t)^{\mathcal{I}} &:= (s > t)\\
(T_m(t))^{\mathcal{I}} &:= \mathbb{T}_m Fip(t)\\
({\sim}T_m(t))^{\mathcal{I}} &:= \neg \mathbb{T}_m Fip(t)\\
(B \vee C)^{\mathcal{I}} &:= (B^{\mathcal{I}} \vee C^{\mathcal{I}})\\
(B \wedge C)^{\mathcal{I}} &:= (B^{\mathcal{I}} \wedge C^{\mathcal{I}})\\
(\exists x B)^{\mathcal{I}} &:= \exists x (B)^{\mathcal{I}}\\
(\forall x B)^{\mathcal{I}} &:= \forall x (B)^{\mathcal{I}}
\end{aligned}
$$

The following lemmas may now be proved very similarly as in Section 8.4, which is why we can safely omit details.

Lemma 11.21. *$\mathcal{I}$ preserves logical structure in $\mathbf{A}(0,3)$.*

Lemma 11.22. *For every $\mathcal{L}(0,3)$ formula A, $A^{\mathcal{I}}$ is A.*

Lemma 11.23. *Let $n > 1$ and suppose that A is an axiom from $\bigcup_{m=1}^{n} \mathsf{FKF}(m, \mathrm{Q})$. Then $\mathbf{A}(0,3) \cup \mathsf{BS}(\Sigma_n^{(0,3)})$ proves $A^{\mathcal{I}}$.*

Lemma 11.24. *Let $n > 1$ and suppose that A is an axiom from $\bigcup_{m=1}^{n} \mathsf{TI}(m, \mathrm{Q})$. Then $\mathbf{A}(0,3) \cup \mathsf{IS}(\Sigma_n^{(0,3)})$ proves $A^{\mathcal{I}}$.*

11.5 Conclusions

We now piece together our earlier findings to conclude the main results of the present chapter.

Theorem 11.25. *For any $n > 1$, $\mathbf{FKF}(n, \mathrm{Q})$ is retractable to $\mathbf{I}(\Sigma_n)$ by $(\mathcal{A} \cdot \mathcal{I})$.*

Proof. Like the proof of Theorem 8.38. ☐

Corollary 11.26. *Let $n > 1$. For any $\mathcal{L}(0, \mathrm{Q})$ formula A and $\mathcal{L}(\omega, 3)$ formula B, the following hold.*

(1) $\mathbf{FKF}(n, \mathrm{Q}) \vdash B$ *implies* $\mathbf{I}(\Sigma_n) \vdash (B^{\mathcal{I}})^{\mathcal{A}}$.

(2) $\mathbf{FKF}(n, \mathrm{Q}) \vdash A$ *if and only if* $\mathbf{I}(\Sigma_n) \vdash A$.

Proof. Like the proof of Corollary 8.39. ☐

Remark 11.27. For $n > 1$, let us also add the following.

- For every $\mathcal{L}(n, \mathrm{Q})$ formula B, we get by Corollary 3.54 (2) that
$$\mathbf{B}(n,3) \cup \bigcup_{m=1}^{n} \big(\mathsf{FKF}(m, \mathrm{Q}) \cup \mathsf{TI}(m, \mathrm{Q})\big)^{\mathcal{A}} \vdash B$$
implies $\mathbf{FKF}(n, \mathrm{Q}) \vdash B$ and, by Corollary 11.26 (1), also
$$\mathbf{I}(\Sigma_n) \vdash (B^{\mathcal{I}})^{\mathcal{A}}.$$

- For every $\mathcal{L}(0, \mathrm{Q})$ formula A, we get by Corollary 3.54 (2) that
$$\mathbf{B}(n,3) \cup \bigcup_{m=1}^{n} \big(\mathsf{FKF}(m, \mathrm{Q}) \cup \mathsf{TI}(m, \mathrm{Q})\big)^{\mathcal{A}} \vdash A$$
precisely if $\mathbf{FKF}(n, \mathrm{Q}) \vdash A$ and, by Corollary 11.26 (2), precisely if
$$\mathbf{I}(\Sigma_n) \vdash A.$$

Remark 11.28. Like in Remark 8.42, we also observe that, for every $n > 1$, the n-th iterate of the Finitist KF theory may alternatively be seen as a *theory of* $\mathrm{P}\Sigma_1^{(n,\mathrm{Q})}$ *induction.* Namely, the following *theories are identical* also for $n > 1$.

- $\mathbf{FKF}(n, \mathrm{Q})$
- $\mathbf{A}(\omega, 3) \cup \bigcup_{m=1}^{n} \mathsf{FKF}(m, \mathrm{Q}) \cup \mathsf{IS}(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$

Finally, we define the theory of Peano Arithmetic as follows.

Definition 11.29. PA is the theory

$$\mathrm{Q} \cup \bigcup_{n<\omega} \mathsf{IS}(\Sigma_n).$$

Moreover, an entire hierarchy of iterated Finitist KF truth is defined as follows.

Definition 11.30. $\mathbf{FKF}(\omega, \mathrm{Q})$ is the theory

$$\bigcup_{n<\omega} \mathbf{FKF}(n, \mathrm{Q}).$$

At the limit, we immediately obtain the following.

Corollary 11.31. *For every $\mathcal{L}(0, \mathrm{Q})$ formula A and for every $\mathcal{L}(\omega, 3)$ formula B, the following hold.*

(1) $\mathbf{FKF}(\omega, \mathrm{Q}) \vdash B$ *implies* $\mathsf{PA} \vdash (B^{\mathcal{I}})^{\mathcal{A}}$.

(2) $\mathbf{FKF}(\omega, \mathrm{Q}) \vdash A$ *if and only if* $\mathsf{PA} \vdash A$.

Proof. For (1), let B be some $\mathcal{L}(\omega, 3)$ formula and suppose that $\mathbf{FKF}(\omega, \mathrm{Q}) \vdash B$. Then, for some $n < \omega$, we get $\mathbf{FKF}(n, \mathrm{Q}) \vdash B$. So, by Corollary 11.26, we have $\mathbf{I}(\Sigma_n) \vdash (B^{\mathcal{I}})^{\mathcal{A}}$. Therefore, we obtain $\mathsf{PA} \vdash (B^{\mathcal{I}})^{\mathcal{A}}$. Similarly for (2). $\square$

12 Reflection

In the last Chapter 11, we have studied theories $\mathbf{FKF}(n, \mathrm{Q})$ that comprise the hierarchy of $n > 1$ many levels of Iterated Finitist Kripke-Feferman Truth. The purpose of the present chapter is to state the same theories differently, not as hierarchies, but in the form of *progressions*

$$\mathsf{S}_1, \mathsf{S}_2, \mathsf{S}_3, \ldots$$

of theories, where for $n > 1$, the theory S_n *reflects* certain statements which are *provable in the predecessor theory* S_{n-1}. Namely, S_n reflects $\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ statements that are, inside the theory S_n, actually equivalent to $\Delta_0^{(n,\mathrm{Q})}$ statements: For any such statement, S_n comprises the reflection axiom that, if the statement may be proved in S_{n-1}, then it holds in S_n as well. The principal result of the present chapter will be that, from the point of view of $\mathbf{A}(\omega, 3) \cup \mathsf{FKF}(n, \mathrm{Q})$, reflection of the $\Delta_0^{(n,\mathrm{Q})}$ fragment of S_{n-1} precisely amounts to having the n-level hierarchy

$$\bigcup_{m=1}^{n} \big(\mathsf{FKF}(m, \mathrm{Q}) \cup \mathsf{TI}(m, \mathrm{Q})\big)$$

that we have seen in Chapter 11.

As was indicated before, reflection will be axiomatized in S_n by a *schema*, or a so-called *local* principle below: For each given

$\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ formula $A(v_0, \ldots, v_k)$, we have the respective axiom that, for any (possibly non-standard) codes $u_0, \ldots, u_k$ of closed $\mathcal{L}(0, \mathrm{Q})$ terms, it holds that

$$Prov_{\mathsf{S}_{n-1}}(\ulcorner A \urcorner \left[\frac{u_0, \ldots, u_k}{\ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner}\right]) \to A\left(\frac{val(u_0), \ldots, val(u_k)}{v_0, \ldots, v_k}\right)$$

where $Prov_{\mathsf{S}_{n-1}}(x)$ represents that x is provable in S_{n-1}.

It is illustrative to compare the main theorem of this chapter with a theorem of Leivant [35], adapted to the context of Finitist Kripke-Feferman Truth. While our work is in the setting of axiomatic truth, Leivant [35] works in a completely arithmetical setting. He shows that, for $n > 0$, the induction schema for Σ_n formulas is equivalent over (a first-order formulation of) Primitive Recursive Arithmetic to a slight adaption of the above reflection schema for Σ_{n+1} formulas: Namely, in the adapted schema, S_{n-1} is replaced by (first-order) Primitive Recursive Arithmetic and is *fixed.*

Thus, while Leivant [35] establishes a relationship between Σ_n induction and Σ_{n+1} reflection, we will – in the light of the Remark 11.28 – reveal the relationship between $\mathrm{P}\Sigma_1$ induction and $\Sigma_0(\mathrm{P}\Sigma_1)$ (or, from theory-internal perspective, Δ_0) reflection *in a language of* $n > 1$ *times iterated finitist truth.* Moreover, in sharp contrast to Leivant [35], we will study reflection not over a fixed system S, but along a progression $\mathsf{S}_1, \mathsf{S}_2, \mathsf{S}_3, \ldots$ of systems. This requires an argumentation different from the one of Leivant [35]. Indeed, in our proof, we shall rely heavily on the material from Section 9.2.

12.1 The Theories

Before we define some progression of theories, we should mention that, to keep the current chapter at reasonable length, we present it in a more sketchy manner than preceding chapters. The main omission is that, while we have indicated to some detail how to encode syntactical concepts in preceding chapters, we do not go into the details of how to encode the main syntactical concept of the present chapter: namely, *provability.*

That is, we shall just presume that we have at our disposal a Σ_0 formula $Proof_{\mathsf{S}}(x, y)$ to represent in $\mathcal{L}(0, 3)$ that x is (a code of) a derivation of formula y in a given calculus of theory S. (We will add a few more details about the presumed calculus below.) Accordingly, $Prov_{\mathsf{S}}(y)$ is defined as the respective unbounded Σ_1 formula $\exists x Proof_{\mathsf{S}}(x, y)$. For more details about the encoding of provability, we refer the reader to, for example, Boolos et al. [2], Section 15.3, or Schwichtenberg and Wainer [43], Section 3.2.2, or any other preferred introduction to Gödel's Incompleteness Theorems.

We may now approach the definition of the intended progression of theories by first introducing a so-called Uniform Reflection Schema.

Note: Notational conventions from the foregoing chapters still apply. That is, we write $val(x)$ for $val_{\mathrm{Q}}(x)$, and so forth.

Definition 12.1. Let S be a theory and Γ a collection of formulas. $\mathsf{URS}(\mathsf{S}, \Gamma)$ is the *Uniform Reflection Schema*

$$\bigwedge\nolimits_{i=0}^{k} ClTerm(u_i) \to$$
$$\Big(Prov_{\mathsf{S}}(\ulcorner A \urcorner \left[\frac{u_0, \dots, u_k}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner}\right]) \to A\left(\frac{val(u_0), \dots, val(u_k)}{v_0, \dots, v_k}\right)\Big),$$

where $A(v_0, \dots, v_k)$ belongs to Γ.

To simplify notation in what follows, let us introduce, for the progression

$$\mathsf{S}_1, \mathsf{S}_2, \mathsf{S}_3, \ldots$$

of $\mathcal{L}(\omega, \mathrm{Q})$ theories $\mathsf{S}_n = \mathbf{FR}(n, \mathrm{Q})$ $(n > 0)$ defined below, the following abbreviation.

Definition 12.2. For $n > 1$, $\mathsf{FR}(n, \mathrm{Q})$ is the schema

$$\mathsf{URS}\big(\mathbf{FR}(n-1, \mathrm{Q}), \Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})\big)$$

where $\mathbf{FR}(n-1, \mathrm{Q})$ is defined simultaneously as in 12.4.

Notice: As we are now dealing with a *progression* of systems, axioms of the systems must not be stated by means of auxiliary function symbols, but must be stated by means of the function symbols from $\mathcal{L}(0, \mathrm{Q})$. This is why we apply the mapping $\mathcal{A}$ for eliminating auxiliary function symbols, as defined in 3.53.

According to $\mathbf{A}(\omega, 3)$, we define $\mathbf{B}(\omega, 3)$ as follows.

Definition 12.3. $\mathbf{B}(\omega, 3)$ is the theory

$$\bigcup_{n<\omega} \mathbf{B}(n, 3).$$

Theories $\mathbf{FR}(n, \mathrm{Q})$ $(n > 0)$ may now be defined as follows.

Definition 12.4. For $n > 0$, $\mathbf{FR}(n, \mathrm{Q})$ is the *Theory of Finitist Reflection* defined recursively as follows.

$$\mathbf{FR}(n, \mathrm{Q}) := \begin{cases} \mathbf{B}(\omega, 3) \cup \big(\mathsf{FKF}(1, \mathrm{Q}) \cup \mathsf{TI}(1, \mathrm{Q})\big)^{\mathcal{A}} & : n = 1 \\ \mathbf{B}(\omega, 3) \cup \big(\mathsf{FKF}(n, \mathrm{Q}) \cup \mathsf{FR}(n, \mathrm{Q})\big)^{\mathcal{A}} & : n > 1 \end{cases}$$

12.2 Retractability: Lower Bounds

As we have stressed above, we will usually suppress details about the coding of provability in the proofs to follow. However, if some proof step makes it necessary, we will highlight – well satisfiable – requirements on that coding.

To start with, the following two lemmas are insensitive to the proof system according to which provability is encoded.

The initial lemma is a refinement of a classic result by Kreisel and Lévy [31].

Lemma 12.5. *Let* $n > 1$ *and suppose that* A *is a formula from* $\mathsf{IS}(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$*. Then* $\mathbf{A}(n,3) \cup \mathsf{FR}(n,\mathrm{Q})$ *proves* A*.*

Proof. Let $n > 1$. We work in $\mathbf{A}(n,3) \cup \mathsf{FR}(n,\mathrm{Q})$ throughout the proof.

Let $A(x_0,\dots,x_k,y)$ be $\mathrm{P}\Sigma_1^{(n,\mathrm{Q})}$ formula and let $B(x_0,\dots,x_k,z)$ be the corresponding $\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ formula below.

$$A(\overline{0}/y) \wedge (\forall y < z)(A \to A(y+\overline{1}/y)) \to A(z/y)$$

There are (by logic) proofs of $B(\overline{n_0},\dots,\overline{n_k},\overline{0})$ in $\mathbf{B}(n,3)$, where $n_0,\dots,n_k$ are arbitrary. Also, there is a fixed pattern of how to prove in $\mathbf{B}(n,3)$ the conclusion $B(\overline{n_0},\dots,\overline{n_k},\overline{m+1})$ from the assumption $B(\overline{n_0},\dots,\overline{n_k},\overline{m})$, where $n_0,\dots,n_k$ are arbitrary. So, for some respective function symbol $f(x_0,\dots,x_k,z)$ from $\mathcal{L}(0,3)$, we may derive in $\mathbf{A}(0,3)$ the following. First, that for a number $p < f(x_0,\dots,x_k,\overline{0})$, we have

$$Proof_{\mathsf{S}}(p, \ulcorner B \urcorner \left[\frac{num(x_0),\dots,num(x_k),num(\overline{0})}{\ulcorner x_0 \urcorner,\dots,\ulcorner x_k \urcorner,\ulcorner z \urcorner}\right]).$$

Second, if there is $p < f(x_0,\dots,x_k,z)$ with

$$Proof_{\mathsf{S}}(p, \ulcorner B \urcorner \left[\frac{num(x_0),\dots,num(x_k),num(z)}{\ulcorner x_0 \urcorner,\dots,\ulcorner x_k \urcorner,\ulcorner z \urcorner}\right]),$$

there also is $q < f(x_0, \ldots, x_k, z + \overline{1})$ with

$$Proof_{\mathsf{S}}(q, \ulcorner B \urcorner \left[\frac{num(x_0), \ldots, num(x_k), num(z + \overline{1})}{\ulcorner x_0 \urcorner, \ldots, \ulcorner x_k \urcorner, \ulcorner z \urcorner}\right]),$$

where S is $\mathbf{FR}(n-1, \mathrm{Q})$. So, we obtain in $\mathbf{A}(0,3)$ that

$$Prov_{\mathsf{S}}(\ulcorner B \urcorner \left[\frac{num(x_0), \ldots, num(x_k), num(z)}{\ulcorner x_0 \urcorner, \ldots, \ulcorner x_k \urcorner, \ulcorner z \urcorner}\right]).$$

By Proposition 5.15, we have:

$$\bigwedge\nolimits_{i=0}^{k} ClTerm(num(x_i)) \wedge ClTerm(num(z))$$

Hence, $\mathsf{FR}(n, \mathrm{Q})$ yields that:

$$B(val(num(x_0)), \ldots, val(num(x_k)), val(num(z)))$$

By Proposition 5.20 (1), we also have:

$$\bigwedge\nolimits_{i=0}^{k} val(num(x_i)) = x_i \wedge val(num(z)) = z$$

Therefore, $B(x_0, \ldots, x_k, z)$ and, by Remark 3.5, the claim. □

Lemma 12.6. *Let $n > 1$ and suppose that A is a formula from $\bigcup_{m=1}^{n-1} \mathsf{FKF}(m, \mathrm{Q})$. Then $\mathbf{A}(n, 3) \cup \mathsf{FR}(n, \mathrm{Q})$ proves A.*

Proof. By (meta) induction on $n > 1$. We treat only two cases of the induction step to show the strategy. (The induction base is similar.) We will reason (for $n > 1$) in

$$\mathbf{A}(n+1, 3) \cup \mathsf{FR}(n+1, \mathrm{Q}) \tag{12.1}$$

without repeatedly mentioning it. We first show how to derive formulas in $\mathsf{FKF}(n, \mathrm{Q})$. The key observation is that, in view of Lemma 12.5 (and Proposition 8.19 extended for $n > 1$), all the formulas from $\mathsf{FKF}(n, \mathrm{Q})$ are $\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ in $\mathbf{FR}(n, \mathrm{Q})$.

- For (I1), we find a $\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ formula $B(x,y)$ such that $\mathbf{B}(n,3) \cup \mathsf{IS}(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ proves:

$$(\mathrm{I1})^{\mathcal{A}} \leftrightarrow B(x,y) \tag{12.2}$$

  There exists a particular proof (coded as) p of $B(x,y)$ in $\mathbf{FR}(n,\mathrm{Q})$. Thus, a respective proof (coded as) q of an instance $B(\bar{i},\bar{j})$ merely depends on the terms $\bar{i},\bar{j}$ and can be obtained from p by instantiation of the free variables x,y. Therefore, for a function symbol $f(x,y)$ in $\mathcal{L}(0,3)$, we may derive in $\mathbf{A}(0,3)$ that

$$(\exists u < f(v,w))\mathit{Proof}_{\mathsf{S}}(u, \ulcorner B(x,y)\urcorner \left[\frac{\mathit{num}(v), \mathit{num}(w)}{\ulcorner x\urcorner, \ulcorner y\urcorner}\right]),$$

  where S is $\mathbf{FR}(n,\mathrm{Q})$, and obtain from that

$$\mathit{Prov}_{\mathsf{S}}(\ulcorner B(x,y)\urcorner \left[\frac{\mathit{num}(v), \mathit{num}(w)}{\ulcorner x\urcorner, \ulcorner y\urcorner}\right]).$$

  By Proposition 5.15, we obtain that $\mathit{ClTerm}(\mathit{num}(v))$ and $\mathit{ClTerm}(\mathit{num}(w))$. So, by $\mathsf{FR}(n+1,\mathrm{Q})$, we may derive

$$B(\mathit{val}(\mathit{num}(v)), \mathit{val}(\mathit{num}(w))).$$

  By Proposition 5.20 (1), we get $B(v,w)$. Finally, (12.2) and $\mathbf{A}(n,3) \vdash (\mathrm{I1}) \leftrightarrow (\mathrm{I1})^{\mathcal{A}}$ yield the result.

- For (I12), we find a $\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ formula $C(x,y)$ such that $\mathbf{B}(n,3) \cup \mathsf{IS}(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ proves

$$(\mathrm{I12})^{\mathcal{A}} \leftrightarrow C(x,y).$$

  So, we can proceed as before to derive (I12) by means of $\mathsf{FR}(n+1,\mathrm{Q})$.

To conclude the induction step, we still have to show that all formulas in $\bigcup_{m=1}^{n-1} \mathsf{FKF}(m, \mathrm{Q})$ are provable in (12.1). We consider (I1) to show the strategy. Suppose $0 < m < n$. By induction hypothesis,

$$\mathbf{A}(n,3) \cup \mathsf{FR}(n, \mathrm{Q}) \tag{12.3}$$

proves (I1) with index $m < n$:

$$Sent([x = y]) \to \big(T_m([x = y]) \leftrightarrow val(x) = val(y)\big) \tag{12.4}$$

Moreover, since $\mathbf{A}(n,3) \vdash (12.4) \leftrightarrow (12.4)^{\mathcal{A}}$, we also get that (12.3) proves $(12.4)^{\mathcal{A}}$. By Corollary 3.54 (2) and the definition of $\mathbf{FR}(n, \mathrm{Q})$, we get that $\mathbf{FR}(n, \mathrm{Q})$ proves $(12.4)^{\mathcal{A}}$. Also, we find a $\Sigma_0(\mathrm{P}\Sigma_1^{(m,\mathrm{Q})})$ formula $D(x, y)$ so that $\mathbf{B}(n,3) \cup \mathsf{IS}(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ proves $(12.4)^{\mathcal{A}} \leftrightarrow D(x, y)$. For the rest, we proceed formally in (12.1) precisely as before. □

12.3 Sequent Calculi and Provability

In order to show, conversely, how to derive Uniform Reflection Schema $\mathsf{FR}(n, \mathrm{Q})$ (for $n > 1$) in the theory $\mathbf{FKF}(n, \mathrm{Q})$ which, by Remark 11.28, is the same as the theory

$$\mathbf{A}(n,3) \cup \bigcup_{m=1}^{n} \mathsf{FKF}(m, \mathrm{Q}) \cup \mathsf{IS}(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})}),$$

let us first elaborate somewhat on satisfiable requirements on the coding of provability ($Prov_{\mathrm{S}}$). We shall do this by a series of remarks.

Remark 12.7. For definiteness, we presume in what follows that provability ($Prov_{\mathrm{S}}$) has been encoded according to an underlying Tait-style calculus, which is a one-sided sequent calculus. So, any line of a proof is a (finite) set Γ of formulas with the intended meaning of Γ being the disjunction $\bigvee \Gamma$ of the formulas in Γ. As usual, we write Γ, Δ for $\Gamma \cup \Delta$.

We do not present a Tait-style calculus in detail here, but we refer the reader to, for example, Buss [5], 1.2.14, and the original work by Tait [47].

We may tacitly presuppose that our Tait-style calculus in the background has the substitution property that, for every variable x and every term t, if there is a derivation of a finite set Φ of formulas, then there is a derivation (of same length) of the result $\Phi(t/x)$ of substituting t for x in these formulas.

Remark 12.8. We also require that, in the sequent calculus on which the coding of provability ($\mathit{Prov}_{\mathsf{S}}$) is based, induction and collection are not formalized by axioms, but by respective *rules* with side formulas.

We write $\Delta(\vec{x})$ to denote finite sets of formulas such that all variables occurring freely in the formulas are among $\vec{x}$. For any given class Γ of formulas, the *induction rule* corresponding to axioms $\mathsf{IS}(\Gamma)$ is

$$\frac{\Delta(\vec{x}), \neg A(\vec{x}, y), A(\vec{x}, y + \overline{1})}{\Delta(\vec{x}), \neg A(\vec{x}, \overline{0}), A(\vec{x}, t)}$$

where y is not among $\vec{x}$, and t is an arbitrary term, and $A(\vec{x}, y)$ belongs to Γ.

For any class Γ of formulas, the *boundedness or collection rule* corresponding to axioms $\mathsf{BS}(\Gamma)$ is

$$\frac{\Delta(\vec{x}), (\forall y \leq t)\exists z A(\vec{x}, y, z)}{\Delta(\vec{x}), \exists v (\forall y \leq t)(\exists z \leq v) A(\vec{x}, y, z)}$$

where v, y, z are not among $\vec{x}$, and t is an arbitrary term, and $A(\vec{x}, y, z)$ belongs to Γ.

Since side formulas (in Δ) are allowed, the instances of $\mathsf{IS}(\Gamma)$ may be proved by respective instances of the induction rule, and the induction rule is admissible in the presence of the axioms in $\mathsf{IS}(\Gamma)$. Similarly for $\mathsf{BS}(\Gamma)$ and collection rule. See, for example, Sieg [45], Proposition 1.5.

As usual, we use rule variants for induction and collection in order to reduce the complexity of formulas in derivations. Most notably, we will rely on the following observation: *If any formula in Γ is Σ_1 and we conceive the assumption and conclusion of an induction rule or collection rule in the form $\Delta(\vec{x}), \Phi$, the universal closure of $\bigvee \Phi$ is $\mathrm{E}\Pi_2$ in any case.*

Remark 12.9. Like for induction and collection, we require that, in the sequent calculus on which the coding of provability ($Prov_{\mathsf{S}}$) relies, certain *truth rules* are employed in place of the respective truth axioms.

Let Δ be some finite set of formulas. The rules corresponding to the two directions of the biconditional of the axiom (I1) are, respectively, the following. (For simplicity, we omit subscripts of $T(x)$ and $Sent(x)$. We also let s, t, possibly with some subscript, range over $\mathcal{L}(0, \mathrm{Q})$ terms.)

(IR1)
$$\frac{\Delta, T([s = t])}{\Delta, \neg Sent([s = t]), val(s) = val(t)}$$

(IR2)
$$\frac{\Delta, val(s) = val(t)}{\Delta, \neg Sent([s = t]), T([s = t])}$$

Accordingly for (I2)–(I7). Similarly, the rules for (I8) are the following.

(IR15)
$$\frac{\Delta, T([{\sim}T(t)])}{\Delta, \neg Sent([{\sim}T(t)]), T([\neg val(t)]), \neg Sent(val(t))}$$

(IR16)
$$\frac{\Delta, T([\neg val(t)]), \neg Sent(val(t))}{\Delta, \neg Sent([{\sim}T(t)]), T([{\sim}T(t)])}$$

As for the connectives, let us state the rules for (I9):

(IR17)
$$\frac{\Delta, T([s \vee t])}{\Delta, \neg Sent([s \vee t]), T(s) \vee T(t)}$$

(IR18)
$$\frac{\Delta, T(s) \vee T(t)}{\Delta, \neg Sent([s \vee t]), T([s \vee t])}$$

Accordingly for (I10). As for axiom (I11) on bounded universal quantification, we have the following rules.

(IR21)
$$\frac{\Delta, s \leq val(t_1) \wedge {\sim}T(t_2[num(s)/t_0])}{\Delta, \neg Sent([(\forall t_0 \leq t_1)t_2]), {\sim}T([(\forall t_0 \leq t_1)t_2])}$$

(IR22)
$$\frac{\Delta, x \leq val(t_1) \rightarrow T(t_2[num(x)/t_0])}{\Delta, \neg Sent([(\forall t_0 \leq t_1)t_2]), T([(\forall t_0 \leq t_1)t_2])}$$

Here x is the eigenvariable that must not occur freely in formulas from Δ and terms t_0, t_1, t_2. Similarly in the last two rules that correspond to axiom (I12) on unbounded existential quantification.

(IR23)
$$\frac{\Delta, T(t_1[num(s)/t_0])}{\Delta, \neg Sent([(\exists t_0)t_1]), T([(\exists t_0)t_1])}$$

(IR24)
$$\frac{\Delta, {\sim}T(t_1[num(x)/t_0])}{\Delta, \neg Sent([(\exists t_0)t_1]), {\sim}T([(\exists t_0)t_1])}$$

According to Remark 12.8, we have (for $\mathbf{FR}(1, \mathrm{Q})$) a rule for Truth Induction with eigenvariable x.

$$\frac{\Delta, {\sim}T(t[num(x)/s]), T(t[num(x+\overline{1})/s])}{\Delta, \neg Sent([(\forall s)t]), {\sim}T(t[num(\overline{0})/s]), T(t[num(t_0)/s])}$$

Like for induction and collection before, by allowing the side formulas (in Δ), truth axioms may be derived by respective truth rules, and truth rules are admissible in the presence of respective truth axioms.

Once again, we use rule variants for the truth axioms in order to reduce the complexity of formulas in derivations. Importantly, we will rely on the following observation: *If the assumption and conclusion of a truth rule is conceived in the form* Δ, Φ*, then the universal closure of* $\bigvee \Phi$ *is equivalent inside* $\mathbf{A}(n,3)$ *to some* $\mathrm{E}\Pi_1$ *sentence in language* $\mathcal{L}(n,\mathrm{Q})$*, where any truth predicate in a truth rule has index* $m \leq n$*.*

Remark 12.10. As further ingredient for reducing complexity of formulas in derivations, let us notice that formulas of the form $\exists x(x \leq t \land A)$ and $\forall x(x \leq t \to A)$, where t is a *closed* term to bound a respective quantifier, are equivalent to disjunctions and conjunctions, respectively, over extensions of $\mathbf{I}(\Sigma_0)$.

More precisely, if a closed $\mathcal{L}(0,\mathrm{Q})$ term t has value m in the standard model of the natural numbers, then

$$\mathbf{I}(\Sigma_0^{(0,\mathrm{Q})}) \vdash t = \overline{m},$$

where it also holds that, for every $\mathcal{L}(n,\mathrm{Q})$ formula A and every number m:

$$\begin{aligned}
\mathbf{I}(\Sigma_0^{(n,\mathrm{Q})}) &\vdash (\exists x \leq \overline{m})A \leftrightarrow \bigvee\nolimits_{i=0}^{m} A(\bar{i}/x) \\
\mathbf{I}(\Sigma_0^{(n,\mathrm{Q})}) &\vdash (\forall x \leq \overline{m})A \leftrightarrow \bigwedge\nolimits_{i=0}^{m} A(\bar{i}/x)
\end{aligned}$$

We will require only that this relationship between bounded quantifiers and conjunctions/disjunctions may be formally stated (without caring too much about the efficiency of the coding) in $\mathbf{I}(\Sigma_1^{(0,3)})$.

Let us sketch the idea. Notice first that, by Proposition 4.15, we have a function symbol $f_\vee(x,y,z)$ in $\mathcal{L}(0,3)$ such that

$$f_\vee(x,y,z) = \begin{cases} z[num(\overline{0})/x] & : y = \overline{0} \\ \langle \ulcorner \vee \urcorner, f_\vee(x, y - \overline{1}, z), z[num(y)/x]\rangle & : y > \overline{0} \end{cases}$$

is provable in $\mathbf{A}(0,3)$. Similarly for conjunction:

$$f_\wedge(x,y,z) = \begin{cases} z[num(\overline{0})/x] & : y = \overline{0} \\ \langle \ulcorner \wedge \urcorner, f_\wedge(x, y - \overline{1}, z), z[num(y)/x]\rangle & : y > \overline{0} \end{cases}$$

Now, the formalization we need is that, for extensions S of the theory $\mathbf{I}(\Sigma_0^{(n,\mathrm{Q})})$, we may derive in $\mathbf{I}(\Sigma_1^{(0,3)})$ from the assumption $ClTerm(y)$ the following.

$$\begin{gathered} Prov_{\mathsf{S}}([(\exists \ulcorner x \urcorner \leq y)\ulcorner A \urcorner]) \to Prov_{\mathsf{S}}(f_\vee(\ulcorner x \urcorner, val(y), \ulcorner A \urcorner)) \\ Prov_{\mathsf{S}}([(\forall \ulcorner x \urcorner \leq y)\ulcorner A \urcorner]) \to Prov_{\mathsf{S}}(f_\wedge(\ulcorner x \urcorner, val(y), \ulcorner A \urcorner)) \end{gathered}$$

For definiteness, we fix the form of $\bigvee_{i=0}^{m} A(\bar{i}/x)$ in accordance with $f_\vee(x,y,z)$ as

$$\big(\cdots\big(A(\overline{0}/x) \vee A(\overline{1}/x)\big) \vee \cdots \vee A(\overline{m-1}/x)\big) \vee A(\overline{m}/x)$$

Similarly for $\bigwedge_{i=0}^{m} A(\bar{i}/x)$. Note: By fixing these forms, we also make respective assumptions about the structure of (partial cut-free) derivations of formulas with these forms. Although we will not consider derivations in any detail, we will tacitly rely on the structure of derivations, for example, in Definition 12.12 further below. We do not have to give details, but the reader should keep in mind that, for example, Definition 12.12 will have to be given in dependence on the structure of derivations in a given sequent calculus.

This concludes our remarks on the coding of provability.

12.4 Retractability: Upper Bounds

In derivations of instances of $\mathsf{FR}(n, \mathrm{Q})$, that is,

$$\mathsf{URS}\big(\mathbf{FR}(n-1, \mathrm{Q}), \Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})\big)$$

(for $n > 1$) we are given some (standard) $\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ formula $A(v_0, \ldots, v_k)$ denoted via its code as $\ulcorner A(v_0, \ldots, v_k)\urcorner$. First of all, let us remove bounded quantifiers and transform each respective substitution instance denoted as

$$\ulcorner A\urcorner[u_0, \ldots, u_k/\ulcorner v_0\urcorner, \ldots, \ulcorner v_k\urcorner],$$

where $u_0, \ldots u_k$ are (codes of) some closed terms, into a Boolean combination of $\mathrm{P}\Sigma_1^{(n,\mathrm{Q})}$ formulas. Term β_A to denote the result of this transformation is defined recursively on the rank of A as follows.

Definition 12.11. Let Γ be a class of formulas. Furthermore, let $A(v_0, \ldots, v_k)$ be some $\Sigma_0(\Gamma)$ formula. The term $\beta_A(u_0, \ldots, u_k)$ is defined recursively as follows. (Notation: $\vec{u}$ means $u_0, \ldots, u_k$.)

(∗) If A or $\neg A$ belongs to Γ, then β_A is:

$$\ulcorner A\urcorner[u_0, \ldots, u_k/\ulcorner v_0\urcorner, \ldots, \ulcorner v_k\urcorner]$$

- If not (∗) and A has form $(B_0 \vee B_1)$, then β_A is:

$$[\beta_{B_0}(u_0, \ldots, u_k) \vee \beta_{B_1}(u_0, \ldots, u_k)]$$

- If not (∗) and A has form $(B_0 \wedge B_1)$, then β_A is:

$$[\beta_{B_0}(u_0, \ldots, u_k) \wedge \beta_{B_1}(u_0, \ldots, u_k)]$$

- If not (∗) and A has form $(\exists v_{k+1} \leq t)B$, then β_A is:

$$f_{\vee}\big(\ulcorner v_{k+1}\urcorner, val(\ulcorner t\urcorner[\vec{u}/\ulcorner v_0\urcorner, \ldots, \ulcorner v_k\urcorner]), \beta_B(\vec{u}, \ulcorner v_{k+1}\urcorner)\big)$$

- If not $(*)$ and A has form $(\forall v_{k+1} \leq t)B$, then β_A is:

$$f_\wedge\big(\ulcorner v_{k+1}\urcorner, val(\ulcorner t\urcorner[\vec{u}/\ulcorner v_0\urcorner, \ldots, \ulcorner v_k\urcorner]), \beta_B(\vec{u}, \ulcorner v_{k+1}\urcorner)\big)$$

According to the definition of β_A, we might next give some (pseudo) truth definition in dependence on a particular Boolean combination of sentences from Γ.

Given such a Boolean combination B denoted (by coding) by β_A, the rough intuition behind "a sentence D is true in the sense of B" is that D is some sub-sentence of B which may occur in a (partial cut-free) derivation of B and that D is true.

For instance, if B has the form (recall Remark 12.10)

$$\bigvee_{i=0}^{m} C(\bar{i}/x)$$

corresponding to some bounded existential quantifier, then "D is true in the sense of B" is to say that, for some $p \leq m$, either D has the form $C(\overline{p}/x)$ and is true or (for $p > 0$) has the form

$$\bigvee_{j=0}^{p} C(\bar{j}/x),$$

and, for some j such that $j \leq p$, $C(\bar{j}/x)$ is true. Similarly for bounded universal quantifiers.

For the definition below, let $Sent_{\mathrm{P\Sigma1}}(x)$ represent that x is a $\mathrm{P\Sigma_1}$ sentence in $\mathcal{L}(n, \mathrm{Q})$.

Definition 12.12. For $n > 1$, let $A(v_0, \ldots, v_k)$ be a $\Sigma_0(\mathrm{P\Sigma}_1^{(n,\mathrm{Q})})$ formula. The formula $T_A(u_0, \ldots, u_k, x)$ is defined recursively.

$(*)$ If A or $\neg A$ is $\mathrm{P\Sigma}_1^{(n,\mathrm{Q})}$, then T_A is:

$$\begin{gathered} x = \beta_A(u_0, \ldots, u_k) \wedge \\ \big((Sent_{\mathrm{P\Sigma1}}(x) \wedge T_n(x)\big) \vee \big(Sent_{\mathrm{P\Sigma1}}([\neg x]) \wedge \neg T_n([\neg x])\big)\big) \end{gathered}$$

- If not $(*)$ and A has form $(B_0 \vee B_1)$, then T_A is:

$$\begin{aligned}&\textstyle\bigvee_{j \leq 1} \big(x = \beta_{B_j}(u_0, \ldots, u_k) \wedge T_{B_j}(u_0, \ldots, u_k, x)\big) \vee \\ &\qquad \big(x = \beta_A(u_0, \ldots, u_k) \wedge \\ &\textstyle\bigvee_{j \leq 1} T_{B_j}(u_0, \ldots, u_k, \beta_{B_j}(u_0, \ldots, u_k))\big)\end{aligned}$$

- If not $(*)$ and A has form $(B_0 \wedge B_1)$, then T_A is:

$$\begin{aligned}&\textstyle\bigvee_{j \leq 1} \big(x = \beta_{B_j}(u_0, \ldots, u_k) \wedge T_{B_j}(u_0, \ldots, u_k, x)\big) \vee \\ &\qquad \big(x = \beta_A(u_0, \ldots, u_k) \wedge \\ &\textstyle\bigwedge_{j \leq 1} T_{B_j}(u_0, \ldots, u_k, \beta_{B_j}(u_0, \ldots, u_k))\big)\end{aligned}$$

- If not $(*)$ and A has form $(\exists v_{k+1} \leq t)B$, where C is defined as $(\exists v_{k+1} \leq v_{k+2})B$, then T_A is:

$$\begin{aligned}&(\exists y \leq val(\ulcorner t \urcorner[u_0, \ldots, u_k / \ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner])) \\ &\quad \big(x = \beta_B(u_0, \ldots, u_k, num(y)) \wedge T_B(u_0, \ldots, u_k, num(y), x)\big) \vee \\ &\quad \big(x = \beta_C(u_0, \ldots, u_k, num(y)) \wedge y > \overline{0} \wedge \\ &(\exists z \leq y)(T_B(u_0, \ldots, u_k, num(z), \beta_B(u_0, \ldots, u_k, num(z))))\big)\end{aligned}$$

- If not $(*)$ and A has form $(\forall v_{k+1} \leq t)B$, where C is defined as $(\forall v_{k+1} \leq v_{k+2})B$, then T_A is:

$$\begin{aligned}&(\exists y \leq val(\ulcorner t \urcorner[u_0, \ldots, u_k / \ulcorner v_0 \urcorner, \ldots, \ulcorner v_k \urcorner])) \\ &\quad \big(x = \beta_B(u_0, \ldots, u_k, num(y)) \wedge T_B(u_0, \ldots, u_k, num(y), x)\big) \vee \\ &\quad \big(x = \beta_C(u_0, \ldots, u_k, num(y)) \wedge y > \overline{0} \wedge \\ &(\forall z \leq y)(T_B(u_0, \ldots, u_k, num(z), \beta_B(u_0, \ldots, u_k, num(z))))\big)\end{aligned}$$

The following lemma shows that T_A may serve as a (pseudo) truth definition of sort.

Lemma 12.13. *Let $n > 1$. Then $\mathbf{A}(n, 3) \cup \mathsf{FKF}(n, \mathrm{Q})$ proves, for every $\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ formula $A(v_0, \dots, v_k)$, the following.*

$$\bigwedge\nolimits_{i \le k} ClTerm(u_i) \to \\ \big(T_A(u_0, \dots, u_k, \beta_A(u_0, \dots, u_k)) \leftrightarrow A(val(u_0), \dots, val(u_k))\big)$$

Proof. By (meta) induction on the rank of $A(v_0, \dots, v_k)$. Let us consider two cases to show the strategy.

- If A or $\neg A$ is $\mathrm{P}\Sigma_1^{(n,\mathrm{Q})}$, then we apply Lemma 11.7.
- Otherwise, if A has form $(\exists v_{k+1} \le t)B$, then we get by the induction hypothesis (in combination with Proposition 5.15 and Proposition 5.20 (1)) a provable equivalence of formula $A(val(u_0), \dots, val(u_k))$ and the following.

  $$(\exists z \le t(val(u_0), \dots, val(u_k))) \\ T_B(u_0, \dots, u_k, num(z), \beta_B(u_0, \dots, u_k, num(z)))$$

  Note also that, by Proposition 5.21, we have, provably:

  $$t(val(u_0), \dots, val(u_k)) = val(\ulcorner t \urcorner [u_0, \dots, u_k / \ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner])$$

  Thus, it only remains to observe that $\beta_A(u_0, \dots, u_k)$ is also provably identical to

  $$\beta_C(u_0, \dots, u_k, num(t(val(u_0), \dots, val(u_k)))),$$

  where C is $(\exists v_{k+1} \le v_{k+2})B$, and apply definition of T_A.

Other cases are treated accordingly. $\square$

Before we turn to the main lemma of the current section, we add a preliminary explanation.

Remark 12.14. Recall from Lemma 11.10 that, given some $\mathrm{E}\Pi_2$ sentence A as an example, the axioms in $\bigcup_{m=1}^{2} \mathsf{FKF}(m, \mathrm{Q})$ do not enable us to simply derive $\neg A \leftrightarrow T_2(\ulcorner \neg A \urcorner)$ – recall: $\neg A$ is a Σ sentence and we have the Tarski Schema only for Σ sentences – that is, do not enable us to derive $A \leftrightarrow \neg T_2(\ulcorner \neg A \urcorner)$. Rather, by (meta) recursion on rank of A (or, dually, rank of $\neg A$), we may systematically assign to A (to $\neg A$) some $\mathrm{E}\Sigma_1$ sentence $f_\Pi(A)$ (or $f_\Sigma(\neg A)$) in the language $\mathcal{L}(1, \mathrm{Q})$ of truth such that we may then employ the axioms from $\bigcup_{m=1}^{2} \mathsf{FKF}(m, \mathrm{Q})$ to derive, for example, $A \leftrightarrow \neg T_2(\ulcorner f_\Pi(A) \urcorner)$ and $\neg A \leftrightarrow T_2(\ulcorner f_\Sigma(\neg A) \urcorner)$.

Without going into further details, let us point out that, by respective bounded course-of-values recursions, we get respective function symbols in $\mathcal{L}(0, 3)$, operating on (possibly non-standard) codes x of sentences, as formal counterparts of f_Σ, f_Π in $\mathbf{A}(0, 3)$. A similar remark applies, of course, to other complexity classes as well, also if they comprise formulas from $\mathcal{L}(m, \mathrm{Q})$ with $m > 0$.

We now turn to the main lemma of the current section.

Lemma 12.15. *Let $n > 1$ and suppose that A is a formula from $\mathsf{FR}(n, \mathrm{Q})$. Then A is provable in*

$$\mathbf{A}(n, 3) \cup \bigcup_{m=1}^{n} \mathsf{FKF}(m, \mathrm{Q}) \cup \mathsf{IS}(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})}). \tag{12.5}$$

Proof. We sketch the proof. To simplify notation, we fix S to be $\mathbf{FR}(n - 1, \mathrm{Q})$. We work in (12.5) without repeatedly mentioning this.

Let $A(v_0, \dots, v_k)$ be a $\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ formula. Suppose that

$$\bigwedge_{i=0}^{k} ClTerm(u_i) \wedge Prov_{\mathsf{S}}(\ulcorner A \urcorner \left[\frac{u_0, \dots, u_k}{\ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner}\right]). \tag{12.6}$$

By the considerations from Remark 12.10 above, we obtain that $Prov_{\mathsf{S}}(\beta_A(u_0, \dots, u_k))$. Also, every given proof of $\beta_A(u_0, \dots, u_k)$ may be transformed formally in (12.5) into a partial cut free (or free-cut free) proof of $\beta_A(u_0, \dots, u_k)$. (The transformation only requires tetration. See, for example, Buss [4] and the reference therein.) So, from the initial assumption (12.6), we obtain that there is some partial cut free proof of $\beta_A(u_0,\dots,u_k)$ in S, where cuts are applied only to formulas that were introduced by non-logical axioms/rules.

The (formalized) partial cut free theorem allows us to bound the complexity of the formulas appearing in a proof. In view of the Remarks 12.8 and 12.9, we may suppose that any sequent in a proof of $\beta_A(u_0,\dots,u_k)$ has the form $\Delta_0,\Delta_1,\Delta_2,\Delta_3$, where the universal closures of the formulas in these sequents have at most the following complexities in (12.5).

Set	Origin	Complexity
Δ_0	$\mathbf{B}(\omega,3)$	$\Pi_2^{(0,\mathrm{Q})}$
Δ_1	$\mathsf{TI}(1,\mathrm{Q})$	$\Pi_1^{(1,\mathrm{Q})}$
Δ_2	$\mathsf{FKF}(n-1,\mathrm{Q})\cup\mathsf{FR}(n-1,\mathrm{Q})$	$\Pi_1^{(n-1,\mathrm{Q})}$
Δ_3	β_A	$\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$

Note that Δ_3 contains only the Boolean sub-formulas of the given $\beta_A(u_0,\dots,u_k)$, which result from substituting closed terms into formulas and, therefore, are sentences.

Let $Proof_A(u_0,\dots,u_k,x,y)$ mean that x is a partial cut free proof of y, where y has form $\Delta_0,\Delta_1,\Delta_2,\Delta_3$ and Δ_3 depends on $\beta_A(u_0,\dots,u_k)$. By formal induction on the length ($lh(z)$) of all putative sub-proofs z of x, we want to show that the universal closure of

$$\bigvee(\Delta_0\cup\Delta_1\cup\Delta_2\cup\Delta_3)$$

is “true” in a specific sense. In Remark 12.16, we will indicate how to derive, in case $n>2$, that $\neg T_n(f(x))$, where $f(x)$ is a formal transformation of axioms in $\mathsf{FR}(n-1,\mathrm{Q})$, as indicated in Remark 12.14. Generally, by the considerations in Remark 12.14, we have, for $i<3$, a formal transformation f_i of the universal closure of $\bigvee\Delta_i$ and a formal representation f_3 of $\bigvee\Delta_3$ such that we may prove the following by formal induction on w, using x as parameter.

$$\begin{aligned}(\forall v < w)(\forall y, z \leq x)\big(Proof_A(u_0, \dots, u_k, z, y) \wedge \\ lh(z) = v + \overline{1} \rightarrow \\ \big(\bigvee\nolimits_{i<3} \neg T_n(f_i(y)) \vee T_A(u_0, \dots, u_k, f_3(y))\big)\end{aligned} \tag{12.7}$$

Notice that (12.7) is $\Sigma_0(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})})$ in (12.5). Therefore, we can rely on Proposition 9.13 and use (12.7) as induction formula.

To conclude the proof, suppose that $\Delta_0 = \Delta_1 = \Delta_2 = \varnothing$ and that $\Delta_3 = \{\beta_A(u_0, \dots, u_k)\}$. Then we get

$$T_A(u_0, \dots, u_k, \beta_A(u_0, \dots, u_k))$$

and, by Lemma 12.13, finally $A(val(u_0), \dots, val(u_k))$. □

Remark 12.16. We add a technical aside. Namely, to prove formulas from $\mathsf{FR}(n, \mathrm{Q})$, where $n > 2$, formally in

$$\mathbf{A}(n, 3) \cup \bigcup_{m=1}^{n} \mathsf{FKF}(m, \mathrm{Q}) \cup \mathsf{IS}(\mathrm{P}\Sigma_1^{(n,\mathrm{Q})}), \tag{12.8}$$

we have to take into consideration that $\mathbf{FR}(n-1, \mathrm{Q})$ comprises axioms from $\mathsf{FR}(n-1, \mathrm{Q})$, that is,

$$\mathsf{URS}\big(\mathbf{FR}(n-2, \mathrm{Q}), \Sigma_0(\mathrm{P}\Sigma_1^{(n-1,\mathrm{Q})})\big). \tag{12.9}$$

We should draw the reader's attention also to how these may be treated in the major Lemma 12.15 above.

Let x be some (possibly non-standard) code of a $\Sigma_0(\mathrm{P}\Sigma_1^{(n-1,\mathrm{Q})})$ formula and let (as formal counterpart of

$$\ulcorner A \urcorner[u_0, \dots, u_k / \ulcorner v_0 \urcorner, \dots, \ulcorner v_k \urcorner]$$

with internalized arity) $x[u]$ be the result of substituting a given sequence u of closed $\mathcal{L}(0, \mathrm{Q})$ terms for the free variables of x. By reasoning very similarly as in the major Lemma 12.15 above, we can derive in (12.8) that

$$Prov_{\mathsf{S}}^{\mathcal{A}}(x[u]) \to \neg T_n(f(x[u]))$$

where S is $\mathbf{FR}(n-2, \mathrm{Q})$ and f formally transforms sentences as indicated in Remark 12.14 before. Finally, by using the axioms from $\mathsf{FKF}(n, \mathrm{Q})$ and by commuting f, we have for arbitrary $x[u]$ that

$$\neg T_n\big(f\big([\ulcorner Prov_{\mathsf{S}}^{\mathcal{A}}(v)\urcorner[num(x[u])/\ulcorner v\urcorner] \wedge [\neg x[u]]]\big)\big)$$

and, by $\mathsf{FKF}(n, \mathrm{Q})$, that the existential closure of every negated formal instance of (12.9) is untrue in the sense of $\neg T_n$.

So, this concludes our estimate of the proof-theoretic upper bounds.

12.5 Conclusion

It only remains to combine our earlier findings and conclude the main theorem of the present chapter: For any $n > 0$, the theory $\mathbf{FR}(n, \mathrm{Q})$ is *the same theory* as $\mathbf{FKF}(n, \mathrm{Q})$ stated (by translation $\mathcal{A}$) without auxiliary function symbols.

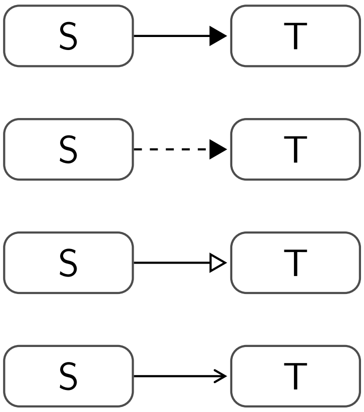
Theorem 12.17. *For $n > 1$, the following are identical.*

- $\mathbf{FR}(n, \mathrm{Q})$
- $\mathbf{B}(\omega, 3) \cup \bigcup_{m=1}^{n} \big(\mathsf{FKF}(m, \mathrm{Q}) \cup \mathsf{TI}(m, \mathrm{Q})\big)^{\mathcal{A}}$

Proof. In the wording of Section 2.4, we need to show that one theory is retractable to the other via identity mapping. But in view of Corollary 3.54 (1) and Remark 11.28, this follows by the Lemmas 12.5, 12.6, and 12.15. □

13 Summary

We conclude our investigations with a graphical overview of the major proof-theoretic relations that we have established. For any pair of theories S and T, we depict:



S is *retractable* to T

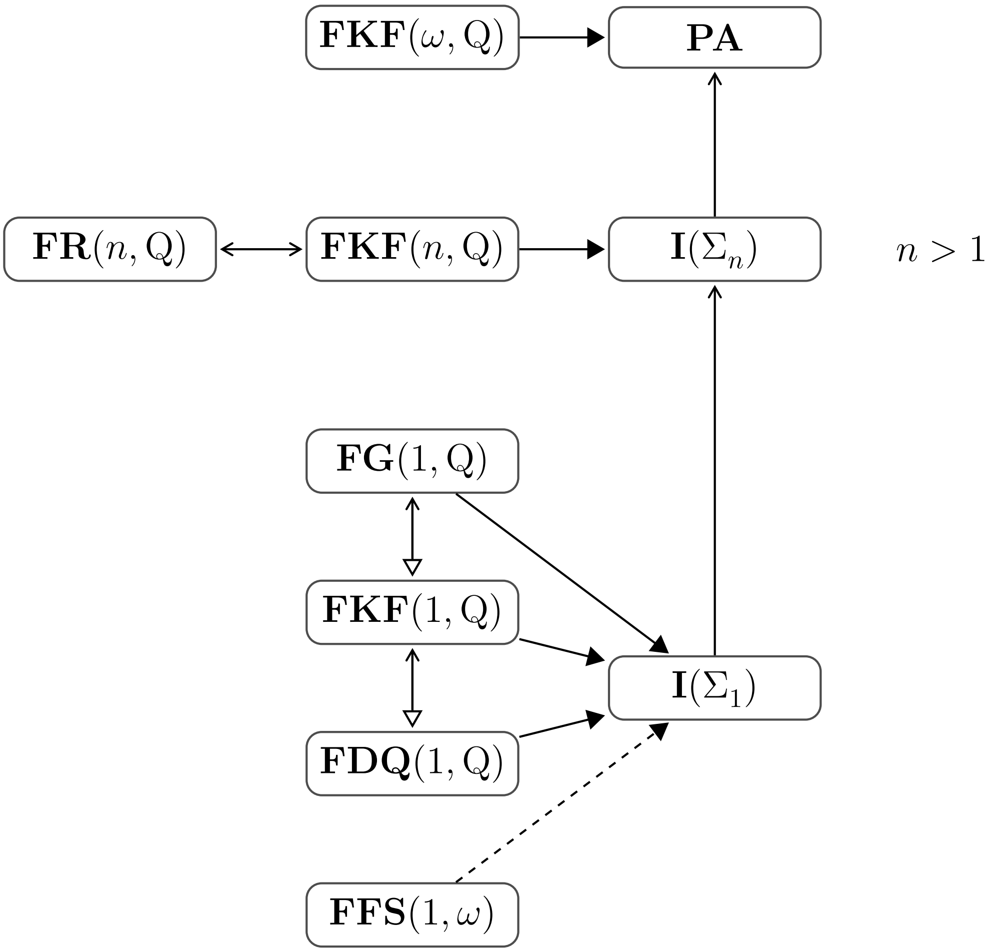
S is *locally retractable* to T

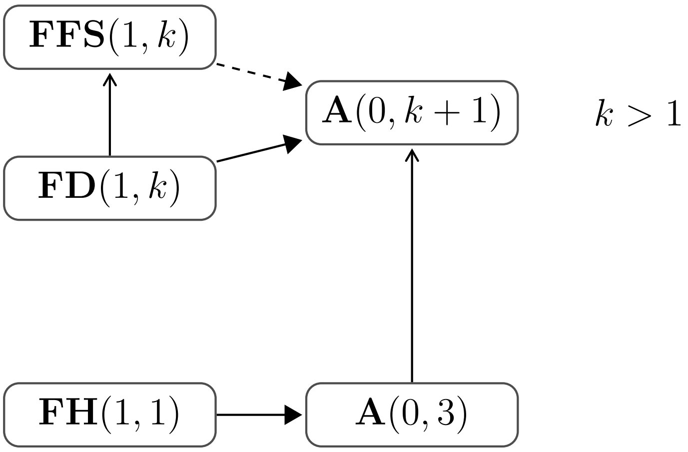
S is *interpretable* in T

S is *sub-theory* of T

Remember that, while retractability entails interpretability (in the sense of 2.27 (1)), the converse may (mainly due to failure of 2.26 (1)) not hold. Let us also recall that, by definition, if S is retractable to T, then T is a sub-theory of S.

Thus, we have the following diagram.


$\mathbf{FKF}(\omega, \mathrm{Q})$
$\mathbf{PA}$
$\mathbf{FR}(n, \mathrm{Q})$
$\mathbf{FKF}(n, \mathrm{Q})$
$\mathbf{I}(\Sigma_n)$
$n > 1$
$\mathbf{FG}(1, \mathrm{Q})$
$\mathbf{FKF}(1, \mathrm{Q})$
$\mathbf{I}(\Sigma_1)$
$\mathbf{FDQ}(1, \mathrm{Q})$
$\mathbf{FFS}(1, \omega)$


$\mathbf{FFS}(1, k)$
$\mathbf{A}(0, k+1)$
$k > 1$
$\mathbf{FD}(1, k)$
$\mathbf{FH}(1, 1)$
$\mathbf{A}(0, 3)$

Notations

$\neg$	negation, 29
$\mathsf{S} \vdash A$	derivability, 31
$\vec{t}$	sequence of terms, 31
$t(\vec{s}/\vec{x})$	substitution (terms), 31
$A(\vec{s}/\vec{x})$	substitution (formulas), 32
$\neq$	not equal, 33
$<$	strictly smaller, 33
$\rightarrow$	conditional, 33
$\leftrightarrow$	biconditional, 33
$(\exists x \leq t)A$	bounded existential quantifier, 33
$(\forall x \leq t)A$	bounded universal quantifier, 33
$(\exists x \in t)A$	existential quantifier (sequence), 62
$(\forall x \in t)A$	universal quantifier (sequence), 62

$\mathsf{KF}(1, \mathrm{Q})$	Kripke-Feferman, 202
$\mathsf{FKF}(1, \mathrm{Q})$	Finitist Kripke-Feferman, 204
$\mathsf{FKF}(m, \mathrm{Q})$	Iterated Finitist Kripke-Feferman, 264
$\mathsf{GS}(\Gamma)$	Groundedness, 236
$\mathsf{TI}(1, 1)$	Truth Induction ($\mathcal{L}(1, 1)$), 150
$\mathsf{TI}(1, \mathrm{Q})$	Truth Induction ($\mathcal{L}(1, \mathrm{Q})$), 207
$\mathsf{TI}(m, \mathrm{Q})$	Truth Induction ($\mathcal{L}(m, \mathrm{Q})$), 266
$\mathsf{TI}(1, k)$	Truth Induction ($\mathcal{L}(1, k)$), 175
$\mathsf{URS}(\mathsf{S}, \Gamma)$	Uniform Reflection Schema, 283
$\mathsf{FR}(n, \mathrm{Q})$	Finitist Reflection, 284
$\mathsf{UTS}(\Gamma)$	Uniform Tarski Schema, 209
$\mathsf{UTS}(m, \Gamma)$	Uniform Tarski Schema (indexed), 268

Complexity

$\Sigma_n^{\mathcal{L}}, \Pi_n^{\mathcal{L}}$	strictly, 34
Σ_n, Π_n	strictly in $\mathcal{L}(0, \mathrm{Q})$, 57
$\Sigma_n^{(-,-)}, \Pi_n^{(-,-)}$	strictly in $\mathcal{L}(-, -)$, 66
$\mathrm{E}\Sigma_n^{\mathcal{L}}, \mathrm{E}\Pi_n^{\mathcal{L}}$	essentially, 34
$\mathrm{E}\Sigma_n, \mathrm{E}\Pi_n$	essentially in $\mathcal{L}(0, \mathrm{Q})$, 57
$\mathrm{E}\Sigma_n^{(-,-)}, \mathrm{E}\Pi_n^{(-,-)}$	essentially in $\mathcal{L}(-, -)$, 66
$\mathrm{PE}\Sigma_1^{(-,-)}$	positive and essentially in $\mathcal{L}(-, -)$, 208

$\mathrm{P}\Sigma_1^{(-,-)}$	positive in $\mathcal{L}(-,-)$, 215
$\Sigma_0(\Gamma)$	Σ_0 in Γ, 238

Formulas

$ClTerm$	closed term, 114
DQ	disquotation, 257
Exp	(base 2) exponentiation, 69
$Form/Form_{-,-}$	formula/$\mathcal{L}(-,-)$ formula, 114
$\mathbb{O}Fip$	operator fixed point, 224
Fun^{k+1}	$(k+1)$-ary function, 241
G_k	k-th Grzegorczyk function, 93
$\mathbb{T}$	truth form, 218
$\mathbb{T}_m$	truth form (indexed), 270
$\mathbb{O}Seq$	operator sequence, 220
$\mathbb{O}Seq^<$	operator sequence (bounded), 221
$\mathbb{O}_mSeq^<$	operator sequence (bounded, indexed), 274
$Pair$	pairing, 57
$\mathbb{C}$	coding form, 243
$Pred$	predecessor, 73
$Proof_{\mathsf{S}}$	proof in S, 282
$Prov_{\mathsf{S}}$	provable in S, 282

$Sent/Sent_{\_,\_}$	sentence/$\mathcal{L}(-,-)$ sentence, 114
Seq	sequence, 61
$\in$	member of sequence, 62
$Subtr$	subtraction, 73
$Term/Term_{\_}$	term/$\mathcal{L}(0,-)$ term, 113
$\llbracket t \rrbracket(x)$	denotation form, 79
Var	variable, 112

Languages

$\mathcal{L}(0,\mathrm{Q})$	Robinson Arithmetic, 54
$\mathcal{L}(m,\mathrm{Q})$	truth index, 76
$\mathcal{L}(m,k)$	truth index and Grzegorczyk index, 93
$\mathcal{L}(1,\omega)$	Grzegorczyk index (unified), 194
$\mathcal{L}(\omega,3)$	truth index (unified), 267
$\mathcal{L}^*$	auxiliary functions (step 1), 78
$\mathcal{L}^+$	auxiliary functions (step 2), 88
$\mathfrak{L}_{\mathrm{PRA}}$	Primitive Recursive Arithmetic, 36

Mappings

$\circledast$	auxiliary functions (step 1), 80
$\oplus$	auxiliary functions (step 2), 89
$\mathcal{A}$	auxiliary functions (steps 1,2), 91

$val(x)$	valuation, 124
$val_k(x)$	valuation (indexed), 134

Theories

$\mathbf{B}(m,k)$	base theory, 96
$\mathbf{A}(m,k)$	base theory (auxiliary), 95
$\mathbf{B}(\omega,3)$	base theory (unified), 284
$\mathbf{A}(1,\omega)$	base theory (auxiliary, unified), 194
$\mathbf{A}(\omega,3)$	base theory (auxiliary, unified), 267
S^*	auxiliary theory (step 1), 78
S^+	auxiliary theory (step 2), 87
$\mathbf{FKF}(1,\mathrm{Q})$	Finitist Kripke-Feferman, 207
$\mathbf{FKF}(n,\mathrm{Q})$	Iterated FKF, 267
$\mathbf{FKF}(\omega,\mathrm{Q})$	Iterated FKF (unified), 279
$\mathbf{FG}(1,\mathrm{Q})$	Finitist Grounded, 239
$\mathbf{FDQ}(1,\mathrm{Q})$	Finitist Disquotation, 260
$\mathbf{FD}(1,k)$	Finitist Decidable, 175
$\mathbf{FH}(1,1)$	Finitist Heck, 150
$\mathbf{FFS}(1,k)$	Finitist Friedman-Sheard, 178
$\mathbf{FFS}(1,\omega)$	FFS (unified), 195
$\mathbf{I}(\Gamma)$	induction, 55

$\mathbf{I}^{<}(\Gamma)$	induction (bounded), 55
PRA	Primitive Recursive Arithmetic, 36
$\mathbf{Q}$	Robinson Arithmetic, 55
$\mathbf{FR}(n, \mathrm{Q})$	Finitist Reflection, 284

Index

References

[1] Belnap, N. and Gupta, A. (1993), *The Revision Theory of Truth*, Cambridge: MIT Press.

[2] Boolos, G. S., Burgess, J. P., and Jeffrey, R. C. (2007), *Computability and Logic*, Cambridge: Cambridge University Press, 5th edn.

[3] Burgess, J. P. (2009), Friedman and the Axiomatization of Kripke's Theory of Truth, conference paper in honor of the 60th birthday of Harvey Friedman at Ohio State University, URL `www.princeton.edu/jburgess/Friedman.pdf`.

[4] Buss, S. R. (1998), First-Order Proof Theory of Arithmetic, *in:* S. R. Buss (Ed.), *Handbook of Proof Theory*, Amsterdam: Elsevier, pp. 79–147.

[5] ——— (1998), An Introduction to Proof Theory, *in:* S. R. Buss (Ed.), *Handbook of Proof Theory*, Amsterdam: Elsevier, pp. 1–78.

[6] Cantini, A. (1989), Notes on Formal Theories of Truth, *Mathematical Logic Quarterly*, **35** (2), 97–130.

[7] ——— (1996), *Logical Frameworks for Truth and Abstraction: An Axiomatic Study*, *Studies in Logic and the Foundations of Mathematics*, vol. 135, Amsterdam: Elsevier.

[8] ——— (1997), Proof-Theoretic Aspects of Self-Referential Truth, *in:* M. L. Dalla Chiara, K. Doets, D. Mundici, and J. Van Benthem (Eds.), *Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence, August 1995*, vol. 1, Amsterdam: Kluwer, pp. 7–27.

[9] Eberhard, S. (2013), *Weak Applicative Theories, Truth, and Computational Complexity*, Ph.D. thesis, University of Bern.

[10] Feferman, S. (1964), Systems of Predicative Analysis, *The Journal of Symbolic Logic*, **29** (1), 1–30.

[11] ——— (1982), Iterated Inductive Fixed Point Theories: Application to Hancock's Conjecture, *in:* G. Metakides (Ed.), *Patras Logic Symposion*, Amsterdam: North-Holland, pp. 171–196.

[12] ——— (1991), Reflecting on Incompleteness, *The Journal of Symbolic Logic*, **56** (1), 1–49.

[13] Feferman, S. and Strahm, T. (2010), The Unfolding of Finitist Arithmetic, *The Review of Symbolic Logic*, **3** (4), 665–689.

[14] Friedman, H. and Sheard, M. (1987), An Axiomatic Approach to Self-Referential Truth, *Annals of Pure and Applied Logic*, **33** (1), 1–21.

[15] Fujimoto, K. (2010), Relative Truth Definability of Axiomatic Truth Theories, *The Bulletin of Symbolic Logic*, **16** (3), 305–344.

[16] ——— (2011), Autonomous Progression and Transfinite Iteration of Self-Applicable Truth, *The Journal of Symbolic Logic*, **76** (3), 914–945.

[17] Gupta, A. (1982), Truth and Paradox, *Journal of Philosophical Logic*, **11** (1), 1–60.

[18] Hàjek, P. and Pudlák, P. (1993), *Metamathematics of First-Order Arithmetic*, Berlin/Heidelberg: Springer.

[19] Halbach, V. (1994), A System of Complete and Consistent Truth, *Notre Dame Journal of Formal Logic*, **35** (3), 311–327.

[20] ——— (2009), Reducing Compositional to Disquotational Truth, *The Review of Symbolic Logic*, **2** (4), 786–798.

[21] ——— (2011), *Axiomatic Theories of Truth*, Cambridge: Cambridge University Press.

[22] Halbach, V. and Horsten, L. (2006), Axiomatizing Kripke's Theory of Truth, *Journal of Symbolic Logic*, **71** (2), 677–712.

[23] Halbach, V. and Leigh, G. E. (2018), Axiomatic Theories of Truth, *in:* E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, spring 2018 edn., URL `https://plato.stanford.edu/archives/spr2018/entries/truth-axiomatic/`.

[24] Heck, R. K. (2007), Self-reference and the Languages of Arithmetic, *Philosophia Mathematica*, **15** (1), 1–29.

[25] Herzberger, H. G. (1982), Notes on Naive Semantics, *Journal of Philosophical Logic*, **11** (1), 61–102.

[26] Hilbert, D. and Bernays, P. (1970), *Grundlagen der Mathematik*, vol. 1, Berlin/Heidelberg/New York: Springer, 2nd edn.

[27] Horsten, L. (2011), *The Tarskian Turn. Deflationism and Axiomatic Truth*, Harvard: MIT Press.

[28] Horwich, P. (1998), *Truth*, Oxford: Oxford University Press.

[29] Jäger, G., Kahle, R., Setzer, A., and Strahm, T. (1999), The Proof-Theoretic Analysis of Transfinitely Iterated Fixed Point Theories, *The Journal of Symbolic Logic*, **64** (1), 53–67.

[30] Kaye, R. (1991), *Models of Peano Arithmetic*, Oxford Logic Guides, Oxford: Oxford University Press.

[31] Kreisel, G. and Lévy, A. (1968), Reflection Principles and Their Use for Establishing the Complexity of Axiomatic Systems, *Mathematical Logic Quarterly*, **14** (7-12), 97–142.

[32] Kripke, S. (1975), Outline of a Theory of Truth, *The Journal of Philosophy*, **72** (19), 690–716.

[33] Leigh, G. E. (2015), Conservativity for Theories of Compositional Truth via Cut Elimination, *The Journal of Symbolic Logic*, **80** (3), 845–865.

[34] Leitgeb, H. (2005), What Truth Depends On, *Journal of Philosophical Logic*, **34** (2), 155–192.

[35] Leivant, D. (1983), The Optimality of Induction as an Axiomatization of Arithmetic, *The Journal of Symbolic Logic*, **48** (1), 182–184.

[36] McGee, V. (1985), How Truthlike Can a Predicate Be? A Negative Result, *Journal of Philosophical Logic*, **14** (4), 399–410.

[37] Moschovakis, Y. N. (2008), *Elementary Induction on Abstract Structures*, Mineola/New York: Dover Publications.

[38] Myhill, J. (1950), A System Which can Define its own Truth, *Fundamenta Mathematicae*, **37**, 190–192.

[39] Nelson, E. (1986), *Predicative Arithmetic. Mathematical Notes*, Princeton: Princeton University Press.

[40] Parsons, C. (1970), On a Number-Theoretic Choice Schema and its Relation to Induction, *in:* A. Kino, J. Myhill, and R. E. Vesley (Eds.), *Intuitionism and Proof Theory: Proceedings of the Summer Conference at Buffalo N.Y. 1968*, Amsterdam: North-Holland, pp. 459–473.

[41] Pohlers, W. (2016), From Subsystems of Analysis to Subsystems of Set Theory, *in:* R. Kahle, T. Strahm, and T. Studer (Eds.), *Advances in Proof Theory*, Progress in Computer Science and Applied Logic, Basel: Birkhäuser, pp. 319–338.

[42] Reinhardt, W. (1986), Some Remarks on Extending and Interpreting Theories, with a Partial Predicate for Truth, *Journal of Philosophical Logic*, **15** (2), 219–251.

[43] Schwichtenberg, H. and Wainer, S. S. (2012), *Proofs and Computations*, Cambridge: Cambridge University Press.

[44] Shoenfield, J. R. (1967), *Mathematical Logic*, Natick, Massachusetts: A K Peters, 2nd edn.

[45] Sieg, W. (1985), Fragments of Arithmetic, *Annals of Pure and Applied Logic*, **28** (1), 33–71.

[46] Skolem, T. (1967), The Foundations of Elementary Arithmetic Established by Means of the Recursive Mode of Thought Without the Use of Apparent Variables Ranging over Infinite Domains, *in:* J. van Heijenoort (Ed.), *From Frege to Gödel: A Source Book in Mathematical Logic (1879–1931)*, Cambridge: Harvard University Press, pp. 302–333.

[47] Tait, W. W. (1968), Normal Derivability in Classical Logic, *in:* J. Barwise (Ed.), *The Syntax and Semantics of Infinitary Languages*, no. 72 in Lecture Notes in Mathematics, Berlin: Springer, pp. 204–236.

[48] ——— (1981), Finitism, *The Journal of Philosophy*, **78** (9), 524–546.

[49] Tarski, A. (1983), The Concept of Truth in Formalized Languages, *in:* J. Corcoran (Ed.), *Logic, Semantics, Metamathematics*, Indianapolis: Hackett Publishing Company, 2nd edn., pp. 152–278.

[50] Yablo, S. (1982), Grounding, Dependence, and Paradox, *Journal of Philosophical Logic*, **11** (1), 117–137.

[51] ——— (1993), Paradox Without Self-Reference, *Analysis*, **53** (4), 251–252.

[52] Zach, R. (2016), Hilbert's Program, *in:* E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, Metaphysics Research Lab, Stanford University, spring 2016 edn., URL `https://plato.stanford.edu/archives/spr2016/entries/hilbert-program/`.