Two proofs of Peano Arithmetic's Consistency

Sevastou Nikoleta
711 514 22 00015

Supervisors:
Koutras Costas,
College of Engineering and Technology,
American University of the Middle East,
Kuwait.
Rigas Nikos,
Associate Lecturer,
Department of Information Technology,
School of Liberal Arts & Sciences,
The American College of Greece

Examination committee:
Koutras Costas (supervisor).
Rigas Nikos (supervisor).
Rondogiannis Panagiotis, Department of Informatics and Telecommunication (Professor), NKUoA.

αὐτὸς γυμνὸς ἐξῆλθον ἐκ κοιλίας μητρός μου, γυμνὸς καὶ ἀπελεύσομαι ἐκεῖ· ὁ Κύριος ἔδωκεν, ὁ Κύριος ἀφείλατο· ὡς τῷ Κυρίῳ ἔδοξεν, οὕτω καὶ ἐγένετο· εἴη τὸ ὄνομα Κυρίου εὐλογημένον εἰς τοὺς αἰῶνας.
(Ιωβ α' 21)

Thou hast made me known to friends whom I knew not. Thou hast given me seats in homes not my own. Thou hast brought the distant near and made a brother of the stranger.
I am uneasy at heart when I have to leave my accustomed shelter; I forget that there abides the old in the new, and that there also thou abidest.
(R. Tagore, Gitangali LXIII)

ABSTRACT

This thesis is about the two most renowned proofs for Peano Arithmetic's consistency. One proof is Gentzen's and the other Gödel's. We also present at the end some objections that have come up since their first publication.

Keywords: Gentzen, Gödel, consistency proof, Peano Arithmetic, first-order arithmetic, PA, objections, proof of consistency of PA, Dialectica interpretation, D-interpretation, Hilbert's program, Double negation, Gödel-Gentzen interpretation

ΣΥΝΟΨΗ

Η διπλωματική αυτή έχει να κάνει με τις δύο πιο γνωστές αποδείξεις της συνέπειας της αριθμητικής Peano. Η πρώτη είναι η απόδειξη του Gentzen και η άλλη του Gödel. Στο τέλος παρουσιάζουμε και κάποιες ενστάσεις που έχουν εμφανιστεί από όταν πρωτοδημοσιεύτηκαν.

Λέξεις-Κλειδιά: Gentzen, Gödel, απόδειξη συνέπειας, Αριθμητική Peano, PA, ενστάσεις, απόδειξη συνέπειας της αριθμητικής, ερμηνεία Dialectica, πρόγραμμα του Hilbert, πρωτοβάθμια αριθμητική, πρωτοτάξια αριθμητική

CONTENTS

CHAPTER 1

INTRODUCTION

The proofs of consistency of Peano Arithmetic constitute an important part within the field of Mathematical Logic. This thesis presents the two most well-known ones from a historical, mathematical and philosophical point of view.

The first one, which is also the oldest of the two we will present, was given by Gerhard Gentzen in 1938 and the most widely known. Although, most of his non-published proofs aren't preserved, the proof that we will see is believed to be his fourth (out of five) [40, p. 6-7, 32–34, 37, sec. 5] and the second published by him [4, p. 63]. The second one we will see is Gödel's and it (partially) appeared in print in 1958 in the Dialectica journal [45]. Other proofs as for example Ackerman's [1] and Khlodovskii's [54], are not presented here.

Before the presentation of the proofs, we discuss their scientific-historical background in order to point out that logic blossomed thanks to them. If one considers also the political, economical and sociological circumstances under which Gentzen and Gödel worked, (s)he respects these scientists and their achievements even more [40, p. 14, 9, 38]. We will not mention though the non-scientific history. The bibliography referenced is sufficient.

The reader can find notes for each proof at the introduction of the corresponding chapter. An index of definitions and symbols, can be found at the end of the document, so that we can ease studying.

Lastly, we consider the proofs from a philosophical point of view. To be more specific, due to the reasons that called for them, a part of the scientific community was -and is even today- doubting their *epistemological* value [27, p. 8]; do they really prove the consistency of Arithmetic? However, in this thesis, we confine ourselves in presenting arguments in defence of them. The last chapter, where the philosophical point of view is discussed, presupposes an elementary acquaintance with the two proofs, so it can be read almost independently of previous chapters.

Unfortunately, some background in logic -both classical and intuitionistic- is needed, so it won't probably be an easy read for undergraduate students. However, we propose chapters and books appropriate for understanding. We also refer to the source that is directly linked to the text.

Of course this thesis is nothing but incomplete in many aspects, especially because many details have been concealed in favour of briefness; particularly in Gentzen's proof. Moreover, although many thorough examinations of the text, it is sure that many mistakes, of all kind, have slipped through the cracks. I apologize in advance, because in such a field as logic is, mistakes can get in the way of understanding.

I must not forget to express my gratitude to my supervisors Pr. C. Koutras and Pr. N. Rigas. Each one of them contributed to the final result in his own way, but especially Mr. Rigas guided me in the most delicate parts of Gödel's proof, which are mostly invisible to the reader.

A special thanks is owed to Pr. C. Dimitracopoulos for reading an early version of my thesis and pointing out many typos, to Pr. G. Koletsos for his course "Lambda calculus and the Curry-Howard isomorphism", to Pr. P. Rondogiannis for taking part in the examination committee on such a short notice, despite his many duties, due to changes to the by-laws of A.L.MA. and for offering the course "Semantics of Programming Languages" and to Pr. C. Poulios for introducing me to the world of logic and set theory as an undergraduate. Finally, I am indebted to Mrs Ioanna for helping me brush up my writing and reading skills in English and for teaching me so many new grammatical and linguistic phenomenons and to the `tex.stackexchange` community for answering many questions of mine concerning typographic issues.

I am also thankful to all my co-students (both undergraduate and postgraduate) and teachers in all four departments that offer courses to A.L.MA. students. Unfortunately, I can't name them one by one as I will probably forget someone. They are too many! Each one of them has taught me many things, has helped me in many ways and has tolerated me for many days! I apologize for sorrowing plenty of them during these past eight years of my studies. I hope that they forgive me.

I wish that the result be effortlessly studied!

Athens, March 2025

CHAPTER 2 HISTORICAL BACKGROUND

To appreciate the two proofs, that we will see in next chapters, we must put them in perspective. In our days they have lost their glory and remain unknown for most scientists. But, back in those days, they were momentous. Let us see why.

2.1 Logic in the 19th century

During the nineteenth century logic was approached in an algebraic way. This can be viewed in Boole's and Schröder's work. However, around 1879, Friedrich Ludwig Gottlob Frege published a rather marginal -during that period- work[1] in which he invents modern quantification logic [6, 37].

After approximately one decade, in 1890, Giuseppe Peano publishes a paper[2] in which he is representing formally proofs in arithmetic [6, 37].

2.2 Paradoxes in set theory

In the late 1890's, set theory, as it was founded by Georg Cantor (1845-1918) by a series of publications, the first of which was in 1874, had started to collapse [53, ch. 0, p. 1, 14]:

> *"it was a widespread idea that pure mathematics is nothing but an elaborate form of arithmetic. Thus, it was usual to talk about the "arithmetisation" of mathematics, and how it had brought about the highest standards of rigor. With Dedekind and Hilbert, this viewpoint led to the idea of grounding all of pure mathematics in set theory. The most difficult steps in bringing forth this viewpoint had been the establishment of a theory of the real numbers, and a set-theoretic reduction of the natural numbers. Both problems had been solved by the work of Cantor and Dedekind. But precisely when mathematicians were celebrating that 'full rigor' had been*

[1] *Begriffsschrift eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*

[2] *Arithmetices principia, nova methodo exposita*

> *finally attained, serious problems emerged for the foundations of set theory. First Cantor, and then Russell, discovered the paradoxes in set theory" [14].*

Cantor and Burali-Forti had already proved that the properties

$$P_1(x) : x \text{ is a cardinal number}$$
$$\text{and}$$
$$P_2(x) : x \text{ is an ordinal number}$$

lead to contradictions. Despite these paradoxes, Russell's paradox (1902) shook the foundations of set theory. It was well believed until then that a built-in error was giving paradoxes in advanced parts of this set theory and, sooner or later, it would be discovered and, thus, overtaken by some "smart handling". After all, analysis had been through the same problems without this affecting its vital parts.

> *"When it isn't a mistake[3], a 'paradox' is simply a fact which runs counter to our intuitions, and set theorists already knew several such 'paradoxes' before Russell announced this one in 1902, in a historic letter to the leading German philosopher and founder of mathematical logic Gottlob Frege. These other paradoxes, however, were technical and affected only some of the most advanced parts of Cantor's theory." [32, sec. 3.5, p. 22]*

> *"They (Cantor and Burali-Forti paradoxes) hadn't troubled significantly the mathematical community, because the numbers that they were related to were new and transfinite numbers, that set theory had constructed to deal with problems that it had posed itself. These numbers weren't used for founding mathematics into Cantor's set theory...Moreover, it was believed that some small changes in the definitions of these transfinite numbers would prevent these contradictions from developing." [4][53, ch. 1, 30, p. 18]*

> *"One could imagine that higher set theory had a systematic error built in, something like allowing a careless 'division by 0' which would soon be discovered and disallowed, and then everything would be fixed. After all, contradictions and paradoxes had plagued the 'infinitesimal calculus' of Newton and Leibnitz and they all went away after the rigorous foundation of the theory which was just being completed in the 1890s, without affecting the vital parts of the subject. Russell's paradox, however, was something else again: simple and brief, it affected directly the fundamental notion of set and the 'obvious' principle of comprehension on which set theory had been built. It is not an exaggeration to say that Russell's paradox brought a foundational crisis of doubt, first to set theory and through it, later, to all of mathematics, which took over thirty years to overcome." [32, sec. 3.5, p. 22]*

It was in 1936 that mathematics would partially recover from this paradox as we will see soon. [32, sec. 3.5, p. 22, 27, p. 16-17]

[3]Here I replaced the first few words of the text with the translation of the corresponding quote in the greek version, because I believe it is more accurate.

[4]Original in greek

2.3 Hilbert's program

David Hilbert had long been involved with geometry and had significantly contributed in its foundations with his book "*Foundations of Geometry*" (1899). Hence, he deeply believed that should any effort be made, it should be in rigorously axiomatizing mathematics and not just avoiding paradoxes, as Cantor thought, by distinguishing between *"consistent multiplicities" or sets, and "inconsistent multiplicities"* [39, 51, 30, p. 18]. Intuition should not affect negatively the axiomatization and, thereby, the theory developed would stay intact.

For him it was of equally great importance to prove the consistency of the axiomatization proposed, i.e., no contradiction should be feasible. Geometry, for instance, could be proved to be consistent via a reduction to analysis, if analysis was proved to be consistent; which was proved to be more difficult than initially expected.

> "*Hilbert also realized that axiomatic investigations required a well worked-out logical formalism. At the time (1900's) he relied on a conception of logic based on the algebraic tradition, in particular, on Schröder's work, which was not particularly suited as a formalism for the axiomatization of mathematics.*" [51]

In 1900 an International Congress of Mathematicians was held in Paris and there he presented a list of twenty three problems that, in his opinion, would much interest mathematicians during the new century. The second in the list was the Consistency of Analysis. From this list, only some are considered to have been solved even in our days.

Between 1910 and 1913 Alfred North Whitehead and Bertrand Russell published *Principia Mathematica* in three volumes. Russell "*took up Frege's logic, but used the notation and formal rules of proof of Peano*" in his previous work *The Principles of Mathematics* (1903), but in Principia, the axiomatic system (which now has become standard due to this book) was changed and it followed Frege's. Mathematics was now "*reduced to logic and its proofs were presented in an axiomatic pattern*" [39, 3].

In this work an axiomatization of arithmetic is proposed in Volume II (Part III section *Cardinal Arithmetic*). The axiomatization is presented, with notation and terminology a bit different than today, as follows (the numbers on the left indicate numeration in Principia) [3]:

1. 0 is a natural number.
   *120 · 12 $\quad 0 \in \mathbb{N}$

2. The successor of any number is a number.
   *120 · 121 $\quad n \in \mathbb{N} \supset n +_c 1 \in \mathbb{N}$

3. No two numbers have the same successor (assuming the axiom of Infinity)
   *120 · 31 Axiom of Infinity$\supset (n +_c 1 = m +_c 1 \supset n = m)$

4. 0 is not the successor of any number.
   *120 · 124 $\quad n +_c 1 \neq 0$

5. Any property ϕ which belongs to 0, and belongs to the successor of m provided that it belongs to m, belongs to all natural numbers n.
   *120 · 13 $\quad \forall n\{[n \in \mathbb{N}\ \&\ \forall m\ (\phi m \supset \phi(m +_c 1))\ \&\ \phi 0] \supset \phi n\}$

"*Frege's and Peano-Russell's approach to logic became the universally accepted one, especially through the influence of Hilbert and his co-workers in the 1920s*" [39]. After the publication of Principia Mathematica, Heinrich Behman and other Hilbert's students studied the system portrayed there and in 1917, and after many years of "silence", Hilbert published "*Axiomatisches Denken*" (Axiomatic Thought); his first contribution to mathematical foundations since 1905. There, he states that the main focus of the mathematical community should be proving the consistency of arithmetic and set theory [51, sec. 1.2].

Between 1917 and 1921, and after Paul Bernays (1888-1977) had joined his "team", with the assistance of Bernays and Behmann, Hilbert contributed significantly in the field of logic [51, sec. 1.2]. One endeavor made by Hilbert and his school was towards founding intuition objectively and safely. So finitism was born.

Although finitism was never rigorously defined, it is believed that finitistic methods are captured, more or less, by the system of Primitive Recursive Arithmetic (PRA) [27, sec. 2, 29, p. 7, 42, sec.2]. In finitism we have "real" (or elementary, finitist) mathematical objects, such as integers, formal proofs (which can be reduced to integers via an encoding) and "real properties" of "real objects". All these have an essence of "elementarity", i.e., we can accept them as real and meaningful (epistemologically speaking) knowledge, properties, proofs etc, they are as intuitive as ordinary finitary number theory and the finitistic (elementary) properties can be formulated by Π^0_1 formulas (of the form $\forall x_1 ... \forall x_n P(x_1, ..., x_n) = 0$ where P is a primitive recursive function). Furthermore, "*a finitist proof must use finitary constructions (i.e., real objects) and elementary properties*" and no unbounded existential quantifiers are permitted, in order to keep ourselves inside finitism[5] [16, p. 34, 29, p. 6].

On the other side, we have "abstract objects" (e.g. ultrafilters). These don't exist outside the realm of our thought. All infinitary objects and other frequently used mathematical objects that are connected to them, are abstract objects. There are also "abstract properties" of either real, or abstract objects and "non-elementary proofs". At this point one might think that finitism fulfills the wish to "not just avoid paradoxes", but in a truly brute way. It ostracizes all infinitary/non-real things. The truth is that:

> *"Hilbert's idea is that although, practically speaking, abstract objects are absolutely necessary (they provide short elegant proofs), they can be, at least in theory, eliminated from proofs of elementary properties....Hilbert's program is an attempt to prove this fact" [16, p. 34].*

2.4 The Consistency Problem

The consistency of set theory or arithmetic is an elementary property. So given a proof of arithmetic's consistency that uses abstract methods, according to Hilbert's program, we can transform it to a proof in the realm of "elementary" reasoning. If it is inconsistent, this is formalized with the formula $\mathbf{0} = \boldsymbol{s}(\mathbf{0})$, which is also an elementary statement and it can be proved by elementary methods [16, p. 35].

In 1930, Kurt Gödel (1906-1978) "*began to pursue Hilbert's program for establishing the consistency of formal axiom systems for mathematics by finitary means....(He) started by working on the consistency problem for analysis, which he sought to reduce to that for arithmetic, but this plan led him to an obstacle related to the well-known*

[5] Hilbert rejected in some of his writings unbounded existential quantifiers as parts of real formulas and he proposed to replace them by bounded [16, p. 35].

paradoxes of truth and definability in ordinary language" [9]. Shortly, he made a new discovery. He presented it in a conference; the first incompleteness theorem which roughly stated that a consistent theory that contains a minimum of arithmetic, enough to encode formulas, can't prove its own consistency. Actually, he proved the existence of an *elementary* formula that this consistent theory can't prove [16, p. 36].

About a year later, the second incompleteness theorem was also published:

> *Arithmetic's consistency can't be proved by means available in its axiomatization (as proposed in Principia Mathematica).* [52, p. 82-83]

In other words, a stronger tool outside the system of arithmetic should be used for the proof, if such exists. Even if we summoned all elementary objects and methods in the theory and added this formula in our assumptions, there exists another formula that ruins our effort [52, pp. 82–83, 16, pp. 64–65]. And so on forever. This also means that we can't produce absolutely "elementary" consistency proofs, but only close to "elementary". As a result we have the fall of Hilbert's program, but not that of the consistency proof search [16, pp. 36, 37].

2.5 Proving the consistency of Peano arithmetic

In early 1932 Gerard Gentzen, at the age of 23, in a letter to his old teacher Hellmuth Kneser, mentioned that he had set as his specific task "*to find a proof of the consistency of logical deduction in arithmetic*", i.e., that from true assumptions we can't derive a contradiction in the logical system of arithmetic. This would be succeeded through the formalization of logical deduction. He had been working on this since 1931 for his PhD thesis, which was ultimately published between 1934 and 1935, but with another topic [40, 37].

He observed that proofs in mathematics "*are not based on axioms expressed in a logical language, as in Hilbert's axiomatic proof theory*", but on assumptions. So he developed and studied the system of Natural Deduction, **NK**. He observed at that time, the existence of what is nowadays known as *cut elimination theorem* and he later proved it. He believed that a corollary of this theorem would be a proof for the consistency of arithmetic, as he would extend deduction to a system of arithmetic by adding a rule for the principle of complete induction [37].

Early in 1933 he quit this idea due to some difficulties and introduced another logical calculus, **LK**. That system would be the main topic of his thesis and would turn to be the first satisfactory formulation of a proof system of classical logic in general [37]. In contrast to the axiomatic system for first order arithmetic proposed in Principia, this system is a "*tool for proof construction*" [43].

*"After his thesis work on **NK** and **LK** he continued his plan of proving the consistency"*. His supervisor that time, Bernays, discussed his second manuscript of the proof with Gödel in 1935. This manuscript was unclear to them, so he wrote the proof over again and in 1936, he published the first ever consistency proof for arithmetic that is compatible with the second incompleteness theorem. The last (fifth to our knowledge) version of the proof was published by Gentzen some years before his early death in 1945 [37, 40].

In a manuscript of Gödel, dating back to 1933, we can see that he "*had speculated about a revised version of Hilbert's program using constructive means that extend the limited finitist ones without being as wide and problematic as the intuitionistic ones*"

[36]. In 1958, he, being cognizant of Gentzen's proof, publishes a *sketch* of a different proof for arithmetic's consistency in the German journal "Dialectica", in honor of P. Bernays' 70th birthday. However, ideas found in this paper date back to 1941 [45]. While the Dialectica interpretation is spelled out in the 1958 paper, "*no details are given of the proof that HA is interpreted in* $\mathcal{T}$" [10, p. 9]. Some of these details were marked in an, unpublished by Gödel, 1972 version of his paper found in his personal archive.

Many great mathematicians contributed in forming Gödel's proof as it is known today, since Gödel himself, being in poor health, didn't write much in the Dialectica paper; Tait, Bernays, Boron, Grzegorczyk, Dragalin to name a few. As mentioned in [43] Grzegorczyk and Dragalin contributed to the modern style of system $\mathcal{T}$ with their papers [21] and [8] respectively. Boron translated from German Gödel's famous Dialectica paper, titled originally "*Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes*" [45]. We should also thank Bernays for his extensive correspondence with Gödel, as the latter presents details of his results there that are not presented elsewhere [11]. Finally, Tait contributed to the proof of the strong normalization theorem with his well-known *Tait method* [24, 43].

CHAPTER 3

GENTZEN'S PROOF

3.1 Introductory notes

We will mostly use [15] for the presentation of Gentzen's proof, but alter the notation to fit that of [43]; minor change in terms of understanding. More specifically, in [15] there is a distinction between free and bounded variables. We won't make such a distinction, as it wouldn't facilitate the presentation of Gödel's proof, which is to come next.

3.2 The Language of the formulas

Definition 3.1. Language L_0 is the language of arithmetic and it consists of [16, 24]:

1. a set of symbols
   - a constant 0 (zero)
   - a unary function symbol s (successor)
   - a denumerable set of variables
   - two binary function symbols $+, \cdot$ (addition, multiplication)
   - one binary predicate symbol $=$ (equality)
   - the connectives $\wedge, \vee, \neg, \rightarrow$
   - the quantifiers $\forall, \exists$
2. the terms (or algebraic terms) that are inductively defined as follows:
   - every variable is a term
   - every constant is a term
   - if $t_1, ..., t_n$ are terms and f is an n-ary function symbol ($n = 1, 2$), then $f(t_1, ..., t_n)$ is a term
   - the only terms are given from the above
     A term in which no variables appear is called a closed term.

3. the formulas that are inductively defined as follows:
   - if t, s are terms $t = s$ is an atomic formula
   - if ϕ, ψ are formulas and x is a variable $\phi \wedge \psi$, $\phi \vee \psi$, $\phi \to \psi$, $\neg\phi$, $\forall x \phi[x]$, $\exists x \phi[x]$ are formulas[1]
   - the only formulas are given from the above

   A formula is called sentence or closed iff it has no free variables and open iff it isn't closed, i.e., it contains at least one free variable not bound by any quantifier.

Remark 3.2. We will also use the symbols "$\neq$" and "$\leftrightarrow$" as abbreviations and not as part of the language. In particular, $\neg(t = s) \equiv (t \neq s)$ and $(\phi \to \psi) \wedge (\psi \to \phi) \equiv \phi \leftrightarrow \psi$

Definition 3.3. For a formula ϕ we define the set of its free variables FV(ϕ) as follows:

- If t is a term in L_0 then FV(t) is the set of all variables (not constants) occurring in t.
- FV$(t_1 = t_2) =$ FV$(t_1) \cup$ FV(t_2) for t_1, t_2 terms over L_0
- FV$(\beta \to \psi) =$ FV$(\beta \wedge \psi) =$ FV$(\beta \vee \psi) =$ FV$(\beta) \cup$ FV(ψ)
- FV$(\forall x \psi) =$ FV$(\exists x \psi) =$ FV$(\psi) \backslash \{x\}$

For a set of formulas $\Gamma = \{\phi_1, ..., \phi_n\}$ we define FV(Γ) as FV$(\phi_1) \cup ... \cup$ FV(ϕ_n) and if Σ is also a set of formulas FV$(\Gamma, \Sigma) =$ FV$(\Gamma) \cup$ FV(Σ).

3.3 Sequent Calculus, Peano Arithmetic

"*Gentzen formulated sequent calculus, denoted **LK**, so that it gave an intuitionistic calculus, denoted **LJ**, as a special case.... He then proved the analogue of the normalization theorem for the classical calculus, the calculus and the proof carefully formulated so that the result for the intuitionistic calculus was a special case of the one for the classical calculus. In **LJ** and **LK**, L stands for "logistic", a term by which Gentzen refers to the axiomatic calculus of logic of Frege, Russell, and Hilbert and Bernays. In such calculi, each line in a derivation is correct in itself, i.e., a logical truth, whence the term. The letters K and J come from the German words klassisch and intuitionistisch. (The latter should thus be upper case 'I', but older German uses upper case 'J' for upper case 'I'.)*" [37, sec. 4]

In our definitions for sequent calculus we will use *sequences* of formulas. Other variations use sets or multi-sets (sets that allow multiple copies of a formula). With sets one has to worry neither for the order of the formulas, nor for multiple copies of a formula. With multi-sets only the order is indifferent. We though follow [15] and [43] in which Sequent Calculus is presented as originally defined by Gentzen.

Definition 3.4. If Γ, Δ are sequences of formulas $A_1, ..., A_n$ and $B_1, ..., B_m$ respectively (possibly empty) then

$$\Gamma \vdash \Delta$$

[1] We will denote by $\phi[x_1, ..., x_n]$ that in formula ϕ the variables $x_1, ..., x_n$ occur free, without that meaning that they are the only free variables occurring. Thus, ϕ might denote a formula for which we don't care or it is unknown if it has a free variable.

is a sequent, Γ is called the antecedent of the sequent and Δ is called the succedent of the sequent.

The sequent where both Γ and Δ are empty (notation $\vdash$) is called empty sequent.

A sequent is intuitionistic if Δ has at most one formula.

Remark 3.5. Intuitively, if we can produce a derivation of $\Gamma \vdash \Delta$, it means that if all formulas in the antecedent are true, at least one formula in the succedent is true. If Γ is empty, at least one in Δ is true. If Δ is empty, then at least one formula in Γ is false. The next table summarises the above for the sequent $A_1, ..., A_n \vdash B_1, ..., B_m$

	Truth value analogous to
$n \neq 0, m \neq 0$	$A_1 \wedge ... \wedge A_n \to B_1 \vee ... \vee B_m$
$n = 0, m \neq 0$	$B_1 \vee ... \vee B_m$
$n \neq 0, m = 0$	$\neg(A_1 \wedge ... \wedge A_n)$
$n = 0, m = 0$	$\perp$

Note that in contrast to **NK** we don't necessarily get that all formulas of Δ are true.

Definition 3.6. The rules of **LK** come in pairs (Left and Right) and they are the following [43, sec. 7.1, sec. 8.3, 15, chap. 5] for $\Gamma, \Theta, \Sigma, \Delta$ sequences of formulas:

1. Axioms

$$\phi \vdash \phi \text{ (Ax)}$$

2. Structural rules

Weakening

$$\frac{\Gamma \vdash \Sigma}{\Gamma, \phi \vdash \Sigma}(\text{LW}) \qquad \frac{\Gamma \vdash \Sigma}{\Gamma \vdash \phi, \Sigma}(\text{RW})$$

Contraction

$$\frac{\Gamma, \phi, \phi \vdash \Sigma}{\Gamma, \phi \vdash \Sigma}(\text{LC}) \qquad \frac{\Gamma \vdash \phi, \phi, \Sigma}{\Gamma \vdash \phi, \Sigma}(\text{RC})$$

Interchange

$$\frac{\Gamma, \phi, \psi, \Delta \vdash \Sigma}{\Gamma, \psi, \phi, \Delta \vdash \Sigma}(\text{LI}) \qquad \frac{\Gamma \vdash \Sigma, \phi, \psi, \Delta}{\Gamma \vdash \Sigma, \psi, \phi, \Delta}(\text{RI})$$

Formulas ϕ, ψ in structural rules are called principal formulas. Sometimes we will omit the indicator L or R for the structural rules.

3. Logical/Operational rules

Conjunction

$$\frac{\Gamma, \phi \vdash \Sigma}{\Gamma, \phi \wedge \psi \vdash \Sigma}(\text{L1}\wedge) \qquad \frac{\Gamma, \psi \vdash \Sigma}{\Gamma, \phi \wedge \psi \vdash \Sigma}(\text{L2}\wedge) \qquad \frac{\Gamma \vdash \psi, \Sigma \quad \Gamma \vdash \phi, \Sigma}{\Gamma \vdash \phi \wedge \psi, \Sigma}(\text{R}\wedge)$$

Disjunction

$$\frac{\Gamma, \psi \vdash \Sigma \quad \Gamma, \phi \vdash \Sigma}{\Gamma \vdash \phi \vee \psi, \Sigma}(\text{L}\vee) \qquad \frac{\Gamma \vdash \psi, \Sigma}{\Gamma \vdash \psi \vee \phi, \Sigma}(\text{R1}\vee) \qquad \frac{\Gamma \vdash \phi, \Sigma}{\Gamma \vdash \psi \vee \phi, \Sigma}(\text{R2}\vee)$$

Conditional

$$\frac{\Gamma \vdash \phi, \Sigma \quad \Delta, \psi \vdash \Lambda}{\Gamma, \Delta, \phi \to \psi \vdash \Delta, \Lambda}(\mathrm{L} \to) \qquad \frac{\Gamma, \phi \vdash \psi, \Sigma}{\Gamma \vdash \phi \to \psi, \Sigma}(\mathrm{R} \to)$$

Negation

$$\frac{\Gamma \vdash \phi, \Sigma}{\Gamma, \neg\phi \vdash \Sigma}(\mathrm{L}\neg) \qquad \frac{\Gamma, \phi \vdash \Sigma}{\Gamma \vdash \neg\phi, \Sigma}(\mathrm{R}\neg)$$

Universal quantifier

$$\frac{\Gamma, \phi[x := t] \vdash \Sigma}{\Gamma, \forall x\phi[x] \vdash \Sigma}(\mathrm{L}\forall) \qquad \frac{\Gamma \vdash \phi[x], \Sigma}{\Gamma \vdash \forall x\phi[x], \Sigma}(x \notin \mathrm{FV}(\Gamma \cup \Sigma))(\mathrm{R}\forall)$$

Existential quantifier

$$\frac{\Gamma, \phi[x] \vdash \Sigma}{\Gamma, \exists x\phi[x] \vdash \Sigma}(x \notin \mathrm{FV}(\Gamma \cup \Sigma))(\mathrm{L}\exists) \qquad \frac{\Gamma \vdash \phi[x := t], \Sigma}{\Gamma \vdash \exists x\phi[x], \Sigma}(\mathrm{R}\exists)$$

The formula newly introduced in every logical rule either on the left, or on the right side is called principal formula.

The variable x in R$\forall$ and L$\exists$ rules is called eigenvariable of the rule.

4. Cut rule

$$\frac{\Gamma \vdash \phi, \Delta \quad \Theta, \phi \vdash \Sigma}{\Gamma, \Theta \vdash \Delta, \Sigma}(\mathrm{CUT})$$

Formula ϕ in the CUT-rule is called CUT-formula or simply CUT.

Remark 3.7. In [43, sec. 7.1] it is mentioned that if we don't take $\perp$ as a primitive (part of our alphabet), we can take negation as primitive and have rules (R$\neg$) and (L$\neg$), which is what we did. Else, i.e., if we take $\perp$ as a primitive, it is considered an atomic formula (see [15, sec. 2.10.1]) and $\neg\phi$ is equivalent to $\phi \to \perp$. In that case if Δ is empty, we can write instead $\Gamma \vdash \perp$.

We now want to define a sequent calculus system for classical arithmetic. This system will extend sequent calculus with some "axioms" and an extra rule for induction.

Definition 3.8. We define the system of **PA** to be **LK** extended with *all* the substitution instances of the following (non-logical) axiom sequents and the induction rule below, i.e., substitutions of a, b with terms and all possible formulas ϕ [15, sec. 7.1, footn.10 sec. 7.5].

P1 $\vdash a = a$

P2 $a = b \vdash b = a$

P3 $a = b, b = c \vdash a = c$

P4 $s(a) = 0 \vdash$

P5 $a = b \vdash s(a) = s(b)$

P6 $s(a) = s(b) \vdash a = b$

P7 $\vdash a + 0 = a$

P8 $\vdash a + s(b) = s(a + b)$

P9 $\vdash a \cdot 0 = 0$

P10 $\vdash a \cdot s(b) = a \cdot b + a$

P11 $a = b \vdash a + c = b + c$

P12 $a = b \vdash c + a = c + b$

P13 $a = b \vdash a \cdot c = b \cdot c$

P14 $a = b \vdash c \cdot a = c \cdot b$

Complete Induction Rule for Γ, Σ sets of formulas and $\phi[x]$ any formula

$$\frac{\phi[a], \Gamma \vdash \Sigma, \phi[a := s(a)]}{\phi[a := 0], \Gamma \vdash \Sigma, \phi[a := t]}\text{CJ} \qquad a \notin \text{FV}(\Gamma, \Sigma, \phi[a := 0], \phi[a := t])$$

$\phi[a := 0], \phi[a := t]$ are called principal formulas of CJ.
The variable a of the CJ rule is called eigenvariable of the rule.

Remark 3.9. Observe that in axiom P3, for example, we have only atomic formulas and not a conjunction. This is of use for the proof of the Cut Elimination theorem, where we will assume that we are given a proof that originates from axiom sequents that contain only atomic formulas.

Definition 3.10. A proof or derivation of $\Gamma \vdash \Delta$ in **PA** is a finite tree with its nodes being sequents, the bottom-most sequent (or end-sequent) being $\Gamma \vdash \Delta$, its leaves (called top-most or initial sequents) being axiom sequents of **PA** and every node and its children match some of the rules of **PA** [43, def. 7.1.1].

Example 3.11. Let's see an example of a derivation. We will prove the Law of Pseudo-Scotus (LPS) $\phi, \neg\phi \vdash \psi$ which is equivalent to ex falso [15, sec. 5.5.1].

$$\cfrac{\cfrac{\cfrac{\phi \vdash \phi}{\neg\phi, \phi \vdash}\ \text{L}\neg}{\phi, \neg\phi \vdash}\ \text{LI}}{\phi, \neg\phi \vdash \psi}\ \text{RW}$$

Definition 3.12. The terms of the form $s(s...(s0)...)$ where s occurs n times will be denoted as $s^n(0)$ and are called numerals.

Remark 3.13. A double line in a proof-tree with many or no labels in the side will denote multiple applications of rules.

Definition 3.14. The degree of a formula ϕ is defined as:

- $d(\phi) = 0$, if ϕ is atomic
- $d(\phi) = d(\psi) + 1$, if ϕ is $\neg\psi,\ \forall x\psi[x]$ or $\exists x\psi[x]$
- $d(\phi) = d(\psi) + d(\beta) + 1$, if ϕ is $\psi \vee \beta, \psi \wedge \beta$ or $\psi \to \beta$

- undefined otherwise

The degree of a CUT or a CJ inference is the degree of the CUT-formula or the induction formula $\phi[a]$ respectively [15, def. 2.5, 2.45, sec. 9.1].

Definition 3.15. The level of a sequent S in a proof δ is the maximum degree of all CUT and CJ inferences below S in δ[2]. If there aren't any the level of S is 0.

We call the level transition for S the topmost inference below S where the premise has the same level as S and the conclusion has a lower level, i.e., the first inference below S where we have a decrease in the level. It can only be a CUT or a CJ inference. [15, sec. 9.1].

Definition 3.16. In any inference I occurring in a proof, a formula A' in the conclusion is the successor of a formula A (and A the predecessor of A') in the premise iff one of the following holds [15, def. 7.22]:

1. In logical inferences A' is the principal formula of the conclusion sequent ($\neg\phi, \phi \wedge \psi, \phi \vee \psi, \phi \rightarrow \psi, \forall x\phi[x], \exists x\phi[x]$) and A is of one of its immediate sub-formula(s) occurring in the premise(s) (i.e., ϕ or ψ).

2. In contraction rules A' is ϕ and A one of the occurrences of ϕ in the premise.

3. In interchange rules A' is ϕ or ψ and A is respectively ϕ or ψ.

4. In induction rule A' is $\phi[0]$ or $\phi[t]$ and A is respectively $\phi[a]$ or $\phi[s(a)]$.

5. In any rule, A' is a formula in Γ, Σ or Δ and A is an occurrence of the same formula in Γ, Σ or Δ in the premise.

Remark 3.17. A formula has no predecessor if it is a weakening formula or occurs in an initial sequent and has no successor if it is a CUT formula or occurs in the end-sequent.

Definition 3.18. A proof π in **PA** is regular if every eigenvariable occurring in π is the eigenvariable of exactly one R$\forall$, L$\exists$ or CJ inference and only occurs above it [15, def. 7.15].

Proposition 3.19. Any proof can be transformed into a regular proof

Proof. See [15, prop. 7.16] $\square$

3.4 Ordinal notations

We assume in this section that the reader is familiar with ordinal numbers and order relations -in particular definitions, properties, addition, multiplication, exponentiation. We suggest for anyone interested [32] or the more concise presentation in [15, ch. 8] as resources. The two books give different but equivalent definitions of the notion of ordinals[3]. Hence, we will omit proofs very often. However, to be clear, the proof *doesn't* rely on set theoretic ordinals (which are transfinite objects), but on strings that resemble

[2]Notice that if S is the premise of a CUT or a CJ inference, this inference is included in counting the maximum degree.

[3]However, I personally had in mind the lecture notes from the undergraduate class of set theory in the Mathematics department of NKUoA during the spring semester of 2019-20 as taught by professor C. Poulios. In the bibliography one can find a link to these notes. For the purposes of the thesis I cited results in other sources.

(or come from) ordinals in many aspects (properties, addition, notation). This means that the fact that they are well-ordered "*follows elementary combinatorial principles about orderings of sequences. In other words, our presentation of ordinal notations is almost finitary [15, p. ix].*"

We will now introduce these very useful *strings*.

Definition 3.20. We define the set $\mathcal{O}$ of ordinal notations $\prec \varepsilon_0$ in stages as follows:

1. In the first stage there is only $\mathbb{0}$, which is the smallest element of $\mathcal{O}$. It has height 0. It is smaller than all other ordinal notations according to the ordering $\preccurlyeq$ of $\mathcal{O}$ ($\prec$ if the two sides are not equal).

2. If we have already defined stages $0, ..., k$ -and ordinal notations of height $\leqslant k$-, $a_1 \succcurlyeq ... \succcurlyeq a_n$ is a non-increasing sequence of ordinal notations of height $\leqslant k$ and a_1 is of height k, then
$$\omega^{a_1} \boxplus ... \boxplus \omega^{a_n}$$
is an ordinal notation of height $k+1$. An ordinal notation of height $k+1$ is greater than any other of smaller height.

   If $\alpha = \omega^{a_1} \boxplus ... \boxplus \omega^{a_n}$ and $\beta = \omega^{\beta_1} \boxplus ... \boxplus \omega^{\beta_m}$ are ordinal notations of height $k+1$, then $\alpha \preccurlyeq \beta$ iff either there exists a j such that $\alpha_i = \beta_i$, for all $i = 1, ..., j$ and $\alpha_{j+1} \preccurlyeq \beta_{j+1}$ or $n < m$ and $\alpha_i = \beta_i$, for all $i = 1, ..., n$.

We also denote $\omega^{\mathbb{0}}$ by $\mathbb{1}$ and $\mathbb{1} \boxplus ... \boxplus \mathbb{1}$ where $\mathbb{1}$ appears n times by $\mathbb{n}$.

Definition 3.21. If α, β are ordinal notations then the natural sum $\alpha \sharp \beta$ is defined as [15, def. 8.27, 8.28]:

- if $\alpha = \mathbb{0}$ then $\alpha \sharp \beta = \beta$
- if $\beta = \mathbb{0}$ then $\alpha \sharp \beta = \alpha$
- Otherwise, there are natural numbers n, m and ordinal notations α_i, β_i such that
$$\alpha = \omega^{\alpha_1} \boxplus ... \boxplus \omega^{\alpha_n}$$
$$\beta = \omega^{\beta_1} \boxplus ... \boxplus \omega^{\beta_m}$$
and also there exists a sequence of ordinal notations γ such that $\gamma_{n+m} \preccurlyeq ... \preccurlyeq \gamma_1$ and
$$\{\gamma_1, ..., \gamma_{n+m}\} = \{\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_m\}$$

We abbreviate $\underbrace{\omega^{\gamma_1} \boxplus ... \boxplus \omega^{\gamma_1}}_{c_1 \text{ copies}} \boxplus ... \boxplus \underbrace{\omega^{\gamma_k} \boxplus ... \boxplus \omega^{\gamma_k}}_{c_k \text{ copies}}$ as
$$\omega^{\gamma_1} \cdot c_1 \boxplus ... \boxplus \omega^{\gamma_k} \cdot c_k$$
if we have at least one $c_i > 0$ and $\mathbb{0}$ otherwise for $\{\gamma_1, ..., \gamma_k\} = \{\alpha_1, ..., \alpha_n, \beta_1, ..., \beta_m\}$ and $\gamma_k \prec ... \prec \gamma_1$.

Remark 3.22. Remember that *ordinal* multiplication is not commutative (as well as addition). So there is a reason for the presence of c_i after ω.

Proposition 3.23. The natural sum of ordinal notations *is* commutative, i.e., $\alpha \sharp \beta = \beta \sharp \alpha$ [15, prop. 8.29]

Proposition 3.24. $(\mathcal{O}, \preccurlyeq)$ is a well ordered set.

Proof. See [15, prop. 8.46] □

Proposition 3.25. If

$$a = \omega^{a_1} \# \ldots \# \omega^{a_n} \qquad \text{and}$$
$$b = \omega^{b_1} \# \ldots \# \omega^{b_m}$$

then $a \prec b$ iff for some j, $a_i = b_i$ when $i \leqslant j$ and either $j = n < m$ (i.e., a is "shorter" than b but the exponents agree) or $j < n$ and $a_{j+1} \prec b_{j+1}$.

Proof. See [15, prop. 8.25] □

Proposition 3.26. There is no infinite (strictly) decreasing sequence of ordinal notations and every non-increasing sequence thereof eventually stabilizes.

Proof. One can see [32, theorem 12.15] or [15, prop. 8.5]. □

3.4.1 Ordinal notation assignment to proofs

Definition 3.27. We define the function $\omega_n(a)$ for an ordinal notation a as:

$$\omega_n(a) = \omega^{\cdot^{\cdot^{\cdot^{\omega^a}}}}$$

where we have a tower of n ω's and a the exponent of the last one for $n = 0, 1, \ldots$. $\omega_0(a) = a$ [15, def. 8.39].

Definition 3.28. The ordinal notation of an inference I $o(I;\pi)$ and ordinal notation of a sequent S $o(S;\pi)$ in a proof π are defined as:

- $o(S;\pi) = \mathbb{1}$ for S an initial sequent of **PA**
- $o(I;\pi) = o(S;\pi)$ for I a structural rule with premise S
- $o(I;\pi) = o(S;\pi) \# \mathbb{1}$ for I a logical rule with one premise S
- $o(I;\pi) = o(S;\pi) \sharp o(S';\pi)$ for I a logical rule with two premises S and S'
- $o(I;\pi) = o(S;\pi) \sharp o(S';\pi)$ for I a CUT rule with premises S and S'
- $o(I;\pi) = \omega^{a_1 \sharp \mathbb{1}}$ for I a CJ rule with premise S and $o(S;\pi) = \omega^{a_1} \# \ldots \# \omega^{a_n}$
- $o(S;\pi) = \omega_{k-l}(o(I;\pi))$ if S is the conclusion of I, the level of the premise(s) of I is k and the level of the conclusion is l (k is strictly greater than l only if I is a CUT or a CJ inference. Otherwise, since k can't be smaller than l, if $k = l$, $o(S;\pi) = o(I;\pi)$)

If S is the end-sequent of a proof π we define the ordinal notation of π as $o(\pi) = o(S;\pi)$ [15, def. 9.3].

Remark 3.29. For structural rules the ordinal notation remains the same from premise to conclusion. For operational rules with one premise it is increased by one, for operational rules with two premises we assign to the sequent the natural sum of the ordinal notations of the two premises. If we have a CUT and there is a change in the level from premises to conclusion or a CJ, we assign a tower of ω's to the conclusion sequent.

3.5 Cut elimination-Hauptsatz for PA-Hilfssatz

We will only delineate the proof of the analogue of Cut Elimination theorem for **PA** in this section, as it is rather long, and state what is needed for the consistency proof. It can be found in detail in [15, ch. 7]. Note that chapter 6 in [15], although titled "*Cut Elimination Theorem*", it refers only to **LK** and can't be straightforwardly applied for the proof we want because of the CJ-rule. Gentzen managed to prove the consistency of PA in a similar, but a bit different, manner, which used again similar, but a bit different, definitions.

For **LK** and **LJ** we can prove that all CUT-free proofs have the sub-formula property. This means that every formula occurring in the CUT-free derivation "*is the sub-formula of either the end-formula or an open assumption*" [15, p. 121]. In **PA** the case is a bit different. If a CUT-free proof of the empty sequent exists, then "*every rule other than CUT, leads from premises containing at least one formula to a conclusion containing at least one formula. So every proof in **PA**, not using the CUT-rule, contains at least one formula in every sequent, including the end-sequent. Since the empty sequent contains no formulas, there cannot be a proof of the empty end-sequent*" [15, p. 275].

Since the CJ-rule is preventing a direct application in **PA** of the Cut Elimination theorem for **LK**, we need a different notion of "simple" proofs, other than CUT-free proofs, and a bit more complicated procedure to prove a similar "Cut Elimination theorem" for **PA** [15, p. 276].

Definition 3.30. An application of the CUT-rule (or simply a CUT[4]) is called atomic if the CUT-formula is atomic and complex otherwise.

Definition 3.31. A proof in **PA** is called simple if it doesn't use logical rules or the CJ-rule and all formulas are atomic and closed. This means that we can only use structural rules and atomic CUTs.

Definition 3.32. An atomic *closed* formula $r = s$ is called "true" if $\mathrm{val}(r) = \mathrm{val}(s)$ and "false" otherwise, where

$$\begin{aligned} \mathrm{val}(0) &= \mathit{0} \\ \mathrm{val}(s(t)) &= \mathrm{val}(t) + \mathit{1} \\ \mathrm{val}(r + t) &= \mathrm{val}(r) + \mathrm{val}(t) \\ \mathrm{val}(r \cdot t) &= \mathrm{val}(r) \cdot \mathrm{val}(t) \end{aligned}$$

The numbers in italics denote the ordinary (not metamathematical) number as well as addition and multiplication in the right side denote the ordinary (not metamathematical) addition and multiplication.

A sequent of only atomic closed formulas is "true" if it contains a "false" formula in the antecedent or a "true" formula in the succedent and "false" otherwise [15, prop. 7.10].

Proposition 3.33. Every sequent in a simple proof is "true".

Proof. For a proof one can see [15, prop. 7.10, 7.12, footnote p. 290] □

> "*Gentzen's consistency proof developed a method for transfroming any proof in **PA** whose end-sequent contains only closed atomic formulas into*

[4]Remember that we denote with CUT the application of CUT-rule, the rule itself and the CUT-formula.

> *a simple proof of the same end-sequent. This yields a consistency proof of* ***PA****. Indeed, if there were a proof of the empty sequent in* ***PA****, the procedure of the consistency proof would produce a simple proof of the empty-sequent".* But *"the end-sequent of any simple proof is 'true'. On the other hand, the empty sequent is not 'true', since it contains neither a 'false'formula in the antecedent, nor a 'true' formula in the succedent".* [15, p. 280]

Definition 3.34. A bundle is a sequence of formula occurrences $A_1, ..., A_n$ in a proof such that A_{i+1} is the successor of A_i, A_1 has no predecessor and A_n has no successor. In other words, a bundle is a branch of the "ancestral tree" of the formula A_n. If $i < j$, then A_i is the ancestor of A_j and A_j is the descendant of A_i [15, def. 7.24].

Definition 3.35. A bundle is implicit if its *last* formula is the CUT-formula in a CUT and explicit otherwise. Furthermore, an inference in a proof is implicit if its principal formula belongs to an implicit bundle and explicit otherwise. Another definition can be this: An inference in a proof is implicit if it precedes a CUT inference and explicit otherwise [15, def. 7.25, 23].

Definition 3.36. A thread is a sequence of occurrences of *sequents* in the proof where each sequent is a premise of an inference the conclusion of which is the following in the thread [15, p. 286].

Definition 3.37. End-part of a proof is the *smallest* part of the proof that satisfies:

1. The end-sequent belongs to the end-part
2. If the conclusion of the inference belongs to the end-part, so do the premises, unless the inference is an implicit logical inference, i.e., the principal formula of the inference is a CUT-formula in a following CUT [15, def. 7.26]

Definition 3.38. An inference is boundary if it is one of the lowermost[5] implicit logical inferences in a proof, i.e., if it is one of the lowermost inferences which introduce a principal formula that is later (not necessarily immediately) cut off by a CUT [15, p. 286].

Remark 3.39. More intuitively, "*the end-part consists of threads which lie between the end-sequent and the conclusions of the lowermost logical inferences (inclusive), the principal formulas of which 'disappear' below them using CUTS -i.e., the bundle terminates in a CUT-rule and the formula doesn't reach all the way down to the end-sequent- and the end-sequent*" or in other words between boundary inferences (if they exist) or between initial sequents and the end-sequent, if that thread doesn't contain an implicit logical inference. [15, p. 286, 309]. Another definition for the end-part could be: all sequents which are not above an implicit logical inference [23].

3.5.1 Steps of Hilfssatz

Assume that we have a regular proof π of only *atomic* initial sequents (which can be achieved for all proofs because of propositions 5.17 and 7.16 in [15]) and end-sequent of *only* atomic closed formulas and in which (i.e., proof π) the only free variables are

[5]Remember that we can compare the height only between sequents in the same thread. So we might have many lowermost implicit inferences that aren't in the same "level" above the end-sequent.

eigenvariables. We want to make a new proof π^* out of π, which is simple (it has no free variables, no logical rules, no CJ rules and it has only atomic formulas). For this we repeat steps 1-3 below.

Since the end-sequent doesn't contain any complex formulas, all operational inferences in π must be implicit and the end-part contains only structural rules, CUTs and CJ inferences. If there was an explicit operational inference, its complex principal formula would finally reach the end-sequent, which can't happen. Hence, the end-part consists of all sequents between either an initial sequent and the end-sequent or a lowermost (implicit as explained) operational inference and the end-sequent.

Before we proceed let's pose a question. Why don't we examine proofs of complex end-sequents? Practically speaking, they are indifferent for proving the consistency since, we only need to verify that the empty sequent can't have a simple proof and the empty sequent can be considered atomic. Moreover, a contradicting complex sequent can derive the empty sequent, so there is no reason to consider the case of complex end-sequents.

In what follows we will assign ordinal notation to the sequents of the proofs that will be explained in section 3.5.2.

STEP 1: Remove induction inferences (applications of CJ) from the end-part.
To explain how this step helps we define the notion of *induction chain*.

Definition 3.40. A sequence of CJ inferences $I_1, ..., I_k$ in the end-part δ is an induction chain iff for each I_j, the inference I_{j+1} occurs below I_j in δ, and there are no CJ inferences above I_1, between I_j and I_{j+1}, or below I_k. The length of the induction chain is k. Further, $m(\delta)$ is the maximum length that an induction chain in δ can achieve and $r(\delta)$ is the number of induction chains in δ of maximum length $m(\delta)$[15, prop. 7.28].

We choose one of the lowermost CJ inferences in the end-part. Which one? One, say I_m, that belongs to an induction chain of maximal (since there might be more than one) length m. Since the end-sequent of the proof π has no free variables (i.e., all formulas are closed), we can suppose (with the use of some propositions omitted here, [15, prop. 7.27]) that the conclusion of I_m is of the form $\phi[0], \Gamma \vdash \Theta, \phi[t]$ where t is closed. Assume that t is the numeral $s^n(0)$. Otherwise we use $s^{val(t)}(0)$ and add a CUT with the sequent $\phi[s^{val(t)}(0)] \vdash \phi[t]$ after we prove that it has a simple proof [15, footnote 10 sec. 7.5].

Let's suppose that the sub-proof ending in this lowermost CJ inference, I_m, that we picked is $\pi(a)$, for a being a free variable.

$$
\begin{array}{c}
\vdots\, \pi(a) \\
\dfrac{a = \omega^{a_1} \# \ldots \# \omega^{a_m}, k \qquad \phi[a], \Gamma \vdash \Theta, \phi[s(a)]}{\omega_{k-l}(\omega^{a_1 \# 1}), l \qquad \phi[0], \Gamma \vdash \Theta, \phi[s^n(0)]}\ \text{CJ}, I_m \\
\vdots\, \pi_0 \\
\Pi \vdash \Xi
\end{array}
$$

We can replace the CJ inference by a larger proof with the same "end"-sequent that uses a series of CUT inferences in the following way:

$$
\text{I,C}\ \dfrac{\dfrac{\begin{matrix}\vdots\,\pi(0)\\ a,k\quad \phi[0],\Gamma\vdash\Theta,\phi[s(0)]\end{matrix}\qquad\begin{matrix}\vdots\,\pi(s(0))\\ a,k\quad \phi[s(0)],\Gamma\vdash\Theta,\phi[s^2(0)]\end{matrix}}{a\sharp a,k\quad \phi[0],\Gamma,\Gamma\vdash\Theta,\Theta,\phi[s^2(0)]}\ \text{CUT } J_1}{a\sharp a,k\quad \phi[0],\Gamma\vdash\Theta,\phi[s^2(0)]}
$$

$$
\text{CUT } J_2\ \dfrac{a\sharp a,k\quad \phi[0],\Gamma\vdash\Theta,\phi[s^2(0)]\qquad\begin{matrix}\vdots\,\pi(s^2(0))\\ a,k\quad \phi[s^2(0)],\Gamma\vdash\Theta,\phi[s^3(0)]\end{matrix}}{\text{I,C}\ \dfrac{a\sharp a\sharp a,k\quad \phi[0],\Gamma,\Gamma\vdash\Theta,\Theta,\phi[s^3(0)]}{a\sharp a\sharp a,k\quad \phi[0],\Gamma\vdash\Theta,\phi[s^3(0)]}}
$$

$$
\vdots
$$

$$
\text{CUT } J_{n-1}\ \dfrac{a\sharp\ldots\sharp a,k\quad \phi[0],\Gamma\vdash\Theta,\phi[s^{n-1}(0)]\qquad\begin{matrix}\vdots\,\pi(s^{n-1}(0))\\ a,k\quad \phi[s^{n-1}(0)],\Gamma\vdash\Theta,\phi[s^n(0)]\end{matrix}}{\text{I,C}\ \dfrac{\omega_{k-l}(a\sharp\ldots\sharp a),l\quad \phi[0],\Gamma,\Gamma\vdash\Theta,\Theta,\phi[s^n(0)]}{\begin{matrix}\omega_{k-l}(a\sharp\ldots\sharp a),l\quad \phi[0],\Gamma\vdash\Theta,\phi[s^n(0)]\\ \vdots\,\pi_0\\ \Pi\vdash\Xi\end{matrix}}}
$$

In the sequent $\phi[0],\Gamma\vdash\Theta,\phi[s^i(0)]$ we assign ordinal notation $a\sharp\ldots\sharp a$, where a appears i times.

Although this new proof has more CJ inferences than the original (since we have added n copies of $\pi(a)$), we have only added sub-proofs containing induction chains of length $\lneq m$, because $\pi(a)$ contains strictly less than m CJ inferences. Moreover, below J_{n-1} in the original proof there are no other CJ inferences by definition. By repeating this for all induction chains of maximal length we get a new proof of maximal length of induction chains $\lneq m$. By induction on $(m(\delta), r(\delta))$ we eliminate CJ inferences from the end-part [15, sec. 7.5].

STEP 2: Remove weakenings from the end-part [15, sec. 7.8].
In step 3 we need "*a complex CUT in which the CUT-formula is descended on the left and the right from principal formulas of boundary inferences....In general, however, not every complex formula need be descended from a principal formula of a corresponding logical inference; it might also be descendant of a formula introduced by weakening*". So we need to remove weakenings (this will become clearer after explaining step 3). This is easily achieved by induction on the last inference of the end-part considering the cases of weakening, contraction, interchange and CUT rules.

By removing weakenings though we may change the end-part and the end-sequent, as, if we have a proof π of an end-sequent $\Gamma\vdash\Delta$ of atomic formulas (atomic end-sequent) and the end-part contains no free variables and no CJ inferences, we will end up with a proof π^* of $\Gamma^*\vdash\Delta^*$ where Γ and Δ have yielded Γ^* and Δ^* respectively by deleting some formulas. Remember that we don't have sets but sequences of formulas, so we might also delete redundant copies of a formula. [15, sec. 7.8].

One may ask "How can it be allowed for us to delete formulas probably needed for the proof? It is a severe change". In fact we are not interested in preserving the original proof at all. Besides, we gradually "chop off" parts of it. The idea is that, if we have ended up in the empty sequent, this must have happened due to "false" assumptions and not because of some inherent weakness of the system of **PA**. Weakenings though (either "true" or "false") are incapable of causing problems. So it is not a problem for us to delete them.

STEP 3: Reduce suitable CUTs in the end-part [15, sec. 9.4].

We assume that we have a proof π with no CJs and no weakenings in the end-part. If π is not its own end-part, it can be proved [15, prop. 7.36] that there is a complex CUT suitable for elimination. This formula must be in an implicit bundle and a descendant of a principal formula of an implicit operational inference, because there are no weakenings capable of introducing it, the initial sequents are by assumption atomic and the formula is missing from the end-sequent.

We will gradually eliminate suitable CUTs by a top-down procedure. Suppose that the topmost suitable CUT formula belongs to CUT inference I and the (common) level transition of the premises of I is an inference J. The procedure alters the sub-proof which ends at J. There are two cases: (1) the CUT inference in which the suitable CUT-formula occurs is itself a level transition (J coincides with I) and (2) there is a level transition below the CUT of interest (J is below I).

<ins>CASE A</ins>: The CUT, I, is a level transition. We will exhibit here only one case; that of the universal quantifier.

Suppose that we have a suitable complex CUT-formula $C \equiv \forall x \phi[x]$ and the proof π is:

$$
\begin{array}{c}
\begin{array}{cc}
\begin{array}{c}
\vdots\ \pi_1(a) \\
\beta_1 \quad \Gamma_1 \vdash \Theta_1, \phi[a] \\
\hdashline
\beta_1 \# \mathbb{1} \quad \Gamma_1 \vdash \Theta_1, \forall x \phi[x] \\
\vdots\ \pi'_1
\end{array}\ \text{R}\forall
&
\begin{array}{c}
\vdots\ \pi_2 \\
\beta_2 \quad \phi[s^n(0)], \Delta_1 \vdash \Lambda_1 \\
\hdashline
\beta_2 \# \mathbb{1} \quad \forall x \phi[x], \Delta_1 \vdash \Lambda_1 \\
\vdots\ \pi'_2
\end{array}\ \text{L}\forall
\end{array}
\\
\text{CUT I}\quad \dfrac{\alpha_1, r \quad \Gamma \vdash \Theta, \forall x \phi[x] \qquad\qquad \alpha_2, r \quad \forall x \phi[x], \Delta \vdash \Lambda}{\omega_{r-s}(a_1 \# a_2), s \quad \Gamma, \Delta \vdash \Theta, \Lambda}\ \ \pi'
\\
\vdots\ \pi_4
\\
\Pi \vdash \Xi
\end{array}
$$

We have indicated boundary inferences with dashed lines. The ordinal notations and the level of each sequent are indicated with blue letters in the side of it and they will be explained further in section 3.5.2.

Since we supposedly have a level transition CUT inference, r is strictly greater than s ($r > s$). For the CUT is suitable, it has ancestors on the left and on the right that are principal formulas of R$\forall$ and L$\forall$ inferences. Moreover, the degree of C is the same as the level of its premises, because I is a level transition.

We will transform π into π^* by replacing π' with this proof:

$$
\begin{array}{c}
\text{CUT I'}\quad \dfrac{
\begin{array}{c} \vdots\ \pi_L \\ \gamma_L, r-1 \quad \Gamma, \Delta \vdash \phi[s^n(0)], \Theta, \Lambda \\ \hline\hline \gamma_L, r-1 \quad \Gamma, \Delta \vdash \Theta, \Lambda, \phi[s^n(0)] \end{array}
\qquad
\begin{array}{c} \vdots\ \pi_R \\ \gamma_R, r-1 \quad \Gamma, \Delta, \phi[s^n(0)] \vdash \Theta, \Lambda \\ \hline\hline \gamma_R, r-1 \quad \phi[s^n(0)], \Gamma, \Delta \vdash \Theta, \Lambda \end{array}
}{\omega_{(r-1)-s}(\gamma_L \# \gamma_R), s \quad \Gamma, \Delta, \Gamma, \Delta \vdash \Theta, \Lambda, \Theta, \Lambda}\ \ \pi''
\\ \hline\hline
\omega_{(r-1)-s}(\gamma_L \# \gamma_R), s \quad \Gamma, \Delta \vdash \Theta, \Lambda
\end{array}
$$

where π_L is:

$$
\text{CUT } I_L\ \dfrac{
\begin{array}{c}
\dfrac{\begin{array}{c}\vdots\ \pi_1(s^n(0)) \\ \beta_1 \quad \Gamma_1 \vdash \Theta_1, \phi[s^n(0)] \\ \hline\hline \beta_1 \quad \Gamma_1 \vdash \phi[s^n(0)], \Theta_1 \end{array}}{\beta_1 \quad \Gamma_1 \vdash \phi[s^n(0)], \Theta_1, \forall x\phi[x]}\ \text{RW} \\ \vdots\ \pi_1'' \\ \alpha_1', r \quad \Gamma \vdash \phi[s^n(0)], \Theta, \forall x\phi[x]
\end{array}
\qquad
\begin{array}{c}
\vdots\ \pi_2 \\
\beta_2 \quad \phi[s^n(0)], \Delta_1 \vdash \Lambda_1 \\ \hdashline \beta_2 \# 1 \quad \forall x\phi[x], \Delta_1 \vdash \Lambda_1 \quad \text{L}\forall \\ \vdots\ \pi_2' \\ \alpha_2, r \quad \forall x\phi[x], \Delta \vdash \Lambda
\end{array}
}{\gamma_L = \omega^{\alpha_1' \sharp \alpha_2}, r-1 \quad \Gamma, \Delta \vdash \phi[s^n(0)], \Theta, \Lambda}
$$

π_1'' is like π_1' but we have added $\phi[s^n(0)]$ to the succedents of the thread that starts with $\Gamma_1 \vdash \phi[s^n(0)], \Theta_1, \forall x\phi[x]$ and ends with $\Gamma \vdash \phi[s^n(0)], \Theta, \forall x\phi[x]$ and π_R is:

$$
\text{CUT } I_R\ \dfrac{
\begin{array}{c}
\vdots\ \pi_1(a) \\
\beta_1 \quad \Gamma_1 \vdash \Theta_1, \phi[a] \\ \hdashline \beta_1 \# 1 \quad \Gamma_1 \vdash \Theta_1, \forall x\phi[x] \quad \text{R}\forall \\ \vdots\ \pi_1' \\ \alpha_1, r \quad \Gamma \vdash \Theta, \forall x\phi[x]
\end{array}
\qquad
\begin{array}{c}
\dfrac{\begin{array}{c}\vdots\ \pi_2 \\ \beta_2 \quad \phi[s^n(0)], \Delta_1 \vdash \Lambda_1 \\ \hline\hline \beta_2 \quad \Delta_1, \phi[s^n(0)] \vdash \Lambda_1 \end{array}}{\beta_2 \quad \forall x\phi[x], \Delta_1, \phi[s^n(0)] \vdash \Lambda_1}\ \text{LW} \\ \vdots\ \pi_2'' \\ \alpha_2', r \quad \forall x\phi[x], \Delta, \phi[s^n(0)] \vdash \Lambda
\end{array}
}{\gamma_L = \omega^{\alpha_1 \sharp \alpha_2'}, r-1 \quad \Gamma, \Delta, \phi[s^n(0)] \vdash \Theta, \Lambda}
$$

where π_2'' is like π_2' but we have added $\phi[s^n(0)]$ to the antecedents of the thread that starts with $\forall x\phi[x], \Delta_1, \phi[s^n(0)] \vdash \Lambda_1$ and ends with $\forall x\phi[x], \Delta, \phi[s^n(0)] \vdash \Lambda$.

CASE B: The topmost CUT inference, I, isn't a level transition. So the premises of the CUT and its conclusion have the same level, say $r > 0$[6]. This means that its conclusion can't be the end-sequent, since its level is zero. So there must be at least one other complex CUT below it that is a level transition. We choose the topmost such CUT, say J.

Suppose that the suitable complex CUT-formula is again $C \equiv \forall x\phi[x]$ (again we

[6]Even if we only have one CUT (remember that we have already eliminated CJs) the level of the CUT inference is the degree of the complex CUT formula which is always > 0 by the definition of the degree of a formula. But in this case it must be a level transition inference, if there are no CJs below it.

won't consider other cases) and the proof π is:

$$
\left.\begin{array}{c}
\text{CUT J}\;\dfrac{
\begin{array}{c}
\text{CUT I}\;\dfrac{
\begin{array}{c}
\vdots\,\pi_1(a) \\
\dfrac{\beta_1 \quad \Gamma_1 \vdash \Theta_1, \phi[a]}{\beta_1 \# \mathbb{1} \quad \Gamma_1 \vdash \Theta_1, \forall x \phi[x]}\,\mathrm{R}\forall \\
\vdots\,\pi_1' \\
\alpha_1, r \quad \Gamma \vdash \Theta, \forall x \phi[x]
\end{array}
\qquad
\begin{array}{c}
\vdots\,\pi_2 \\
\dfrac{\beta_2 \quad \phi[s^n(0)], \Delta_1 \vdash \Lambda_1}{\beta_2 \# \mathbb{1} \quad \forall x \phi[x], \Delta_1 \vdash \Lambda_1}\,\mathrm{L}\forall \\
\vdots\,\pi_2' \\
\alpha_2, r \quad \forall x \phi[x], \Delta \vdash \Lambda
\end{array}
}{a_1 \sharp a_2, r \quad \Gamma, \Delta \vdash \Theta, \Lambda} \\
\vdots\,\pi_3 \\
\lambda_1, r \quad \Gamma_3 \vdash \Theta_3, \psi
\end{array}
\qquad
\begin{array}{c}
\vdots\,\pi_4 \\
\lambda_2, r \quad \psi, \Gamma_4 \vdash \Theta_4
\end{array}
}{\omega_{r-s}(\lambda_1 \sharp \lambda_2), s \quad \Gamma_3, \Gamma_4 \vdash \Theta_3, \Theta_4}
\end{array}\right\}\pi'
\\
\begin{array}{c}
\vdots\,\pi_5 \\
\Pi \vdash \Xi
\end{array}
$$

We have again indicated boundary inferences with dashed lines. We will transform π into π^* which has the form:

$$
\left.\begin{array}{c}
\text{CUT I'}\;\dfrac{
\begin{array}{c}
\vdots\,\pi_L \\
\dfrac{\mu_1, t \quad \Gamma_3, \Gamma_4 \vdash \phi[s^n(0)], \Theta_3, \Theta_4}{\mu_1, t \quad \Gamma_3, \Gamma_4 \vdash \Theta_3, \Theta_4, \phi[s^n(0)]}
\end{array}
\qquad
\begin{array}{c}
\vdots\,\pi_R \\
\dfrac{\mu_2, t \quad \Gamma_3, \Gamma_4, \phi[s^n(0)] \vdash \Theta_3, \Theta_4}{\mu_2, t \quad \phi[s^n(0)], \Gamma_3, \Gamma_4 \vdash \Theta_3, \Theta_4}
\end{array}
}{\dfrac{\omega_{t-s}(\mu_1 \sharp \mu_2), s \quad \Gamma_3, \Gamma_4, \Gamma_3, \Gamma_4 \vdash \Theta_3, \Theta_4, \Theta_3, \Theta_4}{\omega_{t-s}(\mu_1 \sharp \mu_2), s \quad \Gamma_3, \Gamma_4 \vdash \Theta_3, \Theta_4}}
\end{array}\right\}\pi''
\\
\begin{array}{c}
\vdots\,\pi_5 \\
\Pi \vdash \Xi
\end{array}
$$

where π_L is:

$$
\dfrac{
\dfrac{
\dfrac{
\dfrac{
\dfrac{\vdots\,\pi_1(s^n(0))}{\beta_1\quad \Gamma_1 \vdash \Theta_1, \phi[s^n(0)]}
}{\beta_1\quad \Gamma_1 \vdash \phi[s^n(0)], \Theta_1}
}{\beta_1\quad \Gamma_1 \vdash \phi[s^n(0)], \Theta_1, \forall x \phi[x]}\ \mathrm{RW}
}{\vdots\,\pi_1''}
}{\alpha_1', r\quad \Gamma \vdash \phi[s^n(0)], \Theta, \forall x \phi[x]}
\qquad
\dfrac{
\dfrac{
\dfrac{\vdots\,\pi_2}{\beta_2\quad \phi[s^n(0)], \Delta_1 \vdash \Lambda_1}
}{\beta_2 \mathbin{+\!\!\!+} \mathbb{1}\quad \forall x \phi[x], \Delta_1 \vdash \Lambda_1}\ \mathrm{L}\forall
}{
\dfrac{\vdots\,\pi_2'}{\alpha_2, r\quad \forall x \phi[x], \Delta \vdash \Lambda}
}
$$

$$
\text{CUT } I_L\ \dfrac{\alpha_1', r\quad \Gamma \vdash \phi[s^n(0)], \Theta, \forall x \phi[x] \qquad \alpha_2, r\quad \forall x \phi[x], \Delta \vdash \Lambda}{a_1' \sharp a_2, r\quad \Gamma, \Delta \vdash \phi[s^n(0)], \Theta, \Lambda}
$$

$$
\text{CUT } J_L\ \dfrac{\dfrac{\vdots\,\pi_3'}{\lambda_L, r\quad \Gamma_3 \vdash \phi[s^n(0)], \Theta_3, \psi} \qquad \dfrac{\vdots\,\pi_4}{\lambda_2, r\quad \psi, \Gamma_4 \vdash \Theta_4}}{\mu_1 = \mathfrak{w}_{r-t}(\lambda_L \sharp \lambda_2), t\quad \Gamma_3, \Gamma_4 \vdash \phi[s^n(0)], \Theta_3, \Theta_4}
$$

π_1'', π_3' are like π_1' and π_3 but we have added $\phi[s^n(0)]$ on the far left of the succedents and π_R is:

$$
\dfrac{
\dfrac{
\dfrac{\vdots\,\pi_1(a)}{\beta_1\quad \Gamma_1 \vdash \Theta_1, \phi[a]}
}{\beta_1 \mathbin{+\!\!\!+} \mathbb{1}\quad \Gamma_1 \vdash \Theta_1, \forall x \phi[x]}\ \mathrm{R}\forall
}{
\dfrac{\vdots\,\pi_1'}{\alpha_1, r\quad \Gamma \vdash \Theta, \forall x \phi[x]}
}
\qquad
\dfrac{
\dfrac{
\dfrac{
\dfrac{\vdots\,\pi_2}{\beta_2\quad \phi[s^n(0)], \Delta_1 \vdash \Lambda_1}
}{\beta_2\quad \Delta_1, \phi[s^n(0)] \vdash \Lambda_1}
}{\beta_2\quad \forall x \phi[x], \Delta_1, \phi[s^n(0)] \vdash \Lambda_1}\ \mathrm{LW}
}{
\dfrac{\vdots\,\pi_2''}{\alpha_2', r\quad \forall x \phi[x], \Delta, \phi[s^n(0)] \vdash \Lambda}
}
$$

$$
\text{CUT } I_R\ \dfrac{\alpha_1, r\quad \Gamma \vdash \Theta, \forall x \phi[x] \qquad \alpha_2', r\quad \forall x \phi[x], \Delta, \phi[s^n(0)] \vdash \Lambda}{a_1 \sharp a_2', r\quad \Gamma, \Delta, \phi[s^n(0)] \vdash \Theta, \Lambda}
$$

$$
\text{CUT } J_R\ \dfrac{\dfrac{\vdots\,\pi_3''}{\lambda_R, r\quad \Gamma_3 \vdash \phi[s^n(0)], \Theta_3, \psi} \qquad \dfrac{\vdots\,\pi_4}{\lambda_2, r\quad \psi, \Gamma_4 \vdash \Theta_4}}{\mu_2 = \mathfrak{w}_{r-t}(\lambda_R \sharp \lambda_2), t\quad \Gamma_3, \Gamma_4, \phi[s^n(0)] \vdash \Theta_3, \Theta_4}
$$

where π_2'', π_3'' are like π_2' and π_3 but we have added $\phi[s^n(0)]$ on the far right of the antecedents.

But how is that eliminating CUTs? We have even added CUTs and the proof is two times bigger.

> *"First, in the new proof π^*, some copies of π_1 and π_2 now end in weakening inferences instead of operational inferences. Thus, more of these sub-proofs belong to the end-part of π^*. We have 'raised the boundary', and brought more of the proof into the scope of the reduction steps 1-3. The end-part of the new proof may now contain CJ inferences (inside these two copies of π_1 and π_2) which step 1 can remove and replace by CUT inferences. Eventually, the procedure raises the boundary on all threads until all boundary inferences are removed". [15, sec. 7.11]*

We shall skip most of other details of steps 1-3.

> "*In order to prove that applying steps 1-3 repeatedly and in that (particular) order eventually results in a simple proof, we have to find a way of measuring the complexity of proofs in* ***PA*** *in such a way that the complexity of proofs produced (by application of steps 1-3) decreases.... The measures Gentzen developed for this purpose are called ordinal notations*" [15, sec. 7.11]

3.5.2 Termination of the procedure

It can be proved [15, sec. 9.2-9.4] that, by accepting some "conventions" in step 2, when applying steps 1-3 of the CUT-elimination procedure, the ordinal notations of the proofs produced are actually decreased (or not increased if the proof is in the desired form). Since we know that there is no infinite, strictly decreasing sequence of ordinal notations, the proof will eventually become simple.

Proposition 3.41. It holds that a proof is simple if and only if its ordinal notation is $\preccurlyeq \omega^{\mathbb{1}}$.

Proof. See prop. 9.8 in [15]. □

Since we replace sub-proofs by new sub-proofs when applying steps 1-3, we would like to know that the ordinal notation of the entire proof is not increasing, without re-computing the ordinal notation of the end-sequent.

Proposition 3.42. If π_1 and π_2 are the proofs:

$$\begin{array}{cccc} & \vdots\, \pi_1' & & \vdots\, \pi_2' \\ b_1 & S' & b_2 & S' \\ & \vdots\, \pi_3 & & \vdots\, \pi_3 \\ a_1 & S & a_2 & S \end{array}$$

$b_2 \preccurlyeq b_1$ and there are no CJs below S' in π_3, then $a_2 \preccurlyeq a_1$.

Proof. See prop. 9.9 in [15]. □

So it suffices to show that after applying the procedure we have already described, the new sub-proof has smaller ordinal notation.

STEP 1

The CUT formulas $\phi[s^i(0)]$ have all the same degree. This means that the premise and the conclusion of J_{n-2} have the same level k and this is also the case for $J_1, ..., J_{n-3}$, since $\phi[s^i(0)]$ can't increase it.

For the conclusion of the CUT J_{n-1} the level might be $l < k$, if it is a level transition or k if it isn't. So its ordinal notation is $\omega_{k-l}(a\sharp...\sharp a)$ where a appears n times while the conclusion of the original CJ inference was $\omega_{k-l}(\omega^{a_1\sharp\mathbb{1}})$. We need to prove that $\omega_{k-l}(a\sharp...\sharp a) \preccurlyeq \omega_{k-l}(\omega^{a_1\sharp\mathbb{1}})$ [15, sec. 9.2]. It holds that:

Proposition 3.43. If $a \preccurlyeq b$, then $\omega_n(a) \preccurlyeq \omega_n(b)$ for all $n \in \mathbb{N}$.

Proof. See prop. 8.41 in [15]. □

So it suffices to show that $a\sharp ... \sharp a \prec \omega^{a_1 \sharp \mathbb{1}}$.

Proposition 3.44. $\underbrace{a\sharp ... \sharp a}_{n} \prec \omega^{a_1 \sharp \mathbb{1}}$ for every $n \in \mathbb{N}$ [15, sec. 9.2].

Proof. Suppose that $a = \omega^{a_1} \dotplus ... \dotplus \omega^{a_m}$. So

$$\begin{aligned}\underbrace{a\sharp ... \sharp a}_{n} &= \underbrace{\omega^{a_1} \dotplus ... \dotplus \omega^{a_1}}_{n} \dotplus ... \dotplus \underbrace{\omega^{a_m} \dotplus ... \dotplus \omega^{a_m}}_{n} \\ &= \omega^{a_1} \cdot n \dotplus ... \dotplus \omega^{a_m} \cdot n\end{aligned}$$

If a_1 is $\mathbb{0}$ then $a_2, ..., a_m$ are all equal to $\mathbb{0}$, a is $\mathbb{m}$ and $a\sharp ... \sharp a$ is $\mathbb{1} \cdot k$ where $k = n \cdot m$.

But if a_1 is $\mathbb{0}$ then $a_1 \dotplus \mathbb{1}$ is $\mathbb{1}$ and $\omega^{a_1 \sharp \mathbb{1}}$ is $\omega^{\omega^{\mathbb{0}}}$ which is of height 2 while $\mathbb{1} \cdot k$ is of height 1, hence $\underbrace{a\sharp ... \sharp a}_{n} \prec \omega^{a_1 \sharp \mathbb{1}}$ for every $n \in \mathbb{N}$.

On the other hand if a_1 is not $\mathbb{0}$ then $\omega^{a_1 \sharp \mathbb{1}}$ and $\omega^{a_1} \dotplus ... \dotplus \omega^{a_m}$ are of the same height. By prop. 8.37 in [15] $a_1 \prec a_1 \sharp \mathbb{1}$ and by definition of $\preccurlyeq$ we get that

$$\underbrace{a\sharp ... \sharp a}_{n} = \omega^{a_1} \cdot n \dotplus ... \dotplus \omega^{a_m} \cdot n \prec \omega^{a_1 \sharp \mathbb{1}}$$

□

STEP 2

Since the ordinal notations assigned to sequents and inferences on a proof depend on the level of the sequent and this level might change after removing CJs and CUTs, we must somehow guarantee that the ordinal notations actually don't increase. So we define $o_l(S;\pi)$, for every sequent S in π, to be the ordinal notation that is derived in the following way:

1. We have a proof π that has been derived by removing CJ inferences from a regular proof with only atomic axioms.
2. We get a proof π^* by removing weakenings. We label each sequent with its level in π.
3. We compute the ordinal notation $o_l(S^*;\pi^*)$, for every sequent S^* in π, according to the labels assigned in π^* and *not* the levels of π^* (this is the definition).
4. We change the labels in π^* to match the actual levels of π^*.

We can prove that $o_l(S^*;\pi^*) \preccurlyeq o_l(S;\pi)$. Eventually the end-sequent acquires an ordinal notation that is no greater of that of the original [15, prop. 9.12].

STEP 3

CASE A: Since in π_L the ordinal notation of $\Gamma_1 \vdash \phi[s^n(0)], \Theta_1, \forall x\phi[x]$ is β_1, while in π_1' the corresponding sequent had ordinal notation $\beta_1 \dotplus \mathbb{1}$, we can prove that $a_1' \prec a_1$.

For π_R since the ordinal notation of $\forall x\phi[x], \Delta_1, \phi[s^n(0)] \vdash \Lambda_1$ is β_2 while in π_2' the corresponding sequent had ordinal notation $\beta_2 \dotplus \mathbb{1}$, we can prove that $a_2' \prec a_2$.

The end-sequent $\Gamma, \Delta \vdash \Theta, \Lambda$ in π^* has ordinal notation:

$$\omega_{(r-1)-s}(\gamma_L \sharp \gamma_R) = \omega_{(r-1)-s}(\omega^{a_1' \sharp a_2}\sharp \omega^{a_1 \sharp a_2'})$$

and we want to prove that it is smaller than $\omega_{r-s}(a_1 \sharp a_2)$, which are of the same height. So it suffices to prove that

$$\omega^{a'_1 \sharp a_2} \sharp \omega^{a_1 \sharp a'_2} \prec \omega^{a_1 \sharp a_2}$$

which can be achieved with the use of some propositions by considering cases for the natural sum in the left side [15, sec. 9.4].

CASE B:

First we need to consider the levels of the sequents in π^*.

We have assigned level t at the end-sequents of π_L and π_R and level s at the conclusion of CUT I'. We don't know their exact values, but we can prove that.

Proposition 3.45. $s \leqslant t < r$ where r was the level of CUT I in π.

Proof. The levels in a proof can only decrease, so $s \leqslant t$ is easy. For the second inequality we consider two cases. Remember that CUT I in π *wasn't* a level transition, but CUT I' in π^* may be. We assume that $d(\phi[s^n(0)]) = d$ (the degree of the formula).

CASE 1:

$d > s$ Then I' is a level transition and $t = d$ (the degree of the CUT formula is greater than all the CUTs and CJs below I'). Since I wasn't a level transition, the degree of its CUT formula, $d + 1$ is $\leqslant r$. So

$$t = d < d + 1 \leqslant r \Rightarrow t < r$$

CASE 2:

$d \leqslant s$ So I' isn't a level transition and hence $t = s < r$ (s is the level of CUT J which is a level transition and its premises have level r) □

We will now look at the assigned ordinal notations of π^* and validate that they decrease when applying step 3.

For π_L:

Since $r > t$, CUT J_L is a level transition. The degree of the CUT formula ψ is r (the level of the premises of J_L). Since the level s of $\Gamma_3, \Gamma_4 \vdash \Theta_3, \Theta_4$ in π_5 is lower than r by the previous proposition there is no CUT inference with degree greater than s in π_3 and hence the degree of I_L is r (the degree of $\phi[s^n(0)]$). The levels of the premises of J_L are too r as J_L was chosen so it is the topmost level transition below I_L.

The levels of sequents in $\pi(s^n(0)), \pi''_1, \pi_2$ and π'_2 are the same as the corresponding levels of sequents in π (not known exactly, thus not mentioned explicitly). We again have that $a'_1 \prec a_1$. Thence $a'_1 \sharp a_2 \prec a_1 \sharp a_2$ and $\lambda_L \prec \lambda_1$.

For π_R: As in π_L, since $\beta_2 \prec \beta_2 \# \mathbb{1}$, we get that $a'_2 \prec a_2$, $a_1 \sharp a'_2 \prec a_1 \sharp a_2$ and $\lambda_R \prec \lambda_1$.

For the end-sequent: We need to prove that $\omega_{t-s}(\mu_1 \sharp \mu_2) \prec \omega_{r-s}(\lambda_1 \sharp \lambda_2)$. After that, since the steps are same in π and in π^*, we apply proposition 3.42.

Proof. We can see that $r - s = (t - s) + (r - t)$, so

$$\omega_{r-s}(\lambda_1 \sharp \lambda_2) = \omega_{t-s}(\omega_{r-t}(\lambda_1 \sharp \lambda_2))$$

by definition 3.27. So it suffices to show that

$$\omega_{r-t}(\lambda_L \sharp \lambda_2) \sharp \omega_{r-t}(\lambda_R \sharp \lambda_2) = \mu_1 \sharp \mu_2 \prec \omega_{r-t}(\lambda_1 \sharp \lambda_2)$$

We already know that $\lambda_R \prec \lambda_1$ and $\lambda_L \prec \lambda_1$. It can be proved [15, prop. 8.36] that $\lambda_L \sharp \lambda_2 \prec \lambda_1 \sharp \lambda_2$ and $\lambda_R \sharp \lambda_2 \prec \lambda_1 \sharp \lambda_2$.

By proposition 3.43, we get that

$$\omega_{r-t}(\lambda_L \sharp \lambda_2) \prec \omega_{r-t}(\lambda_1 \sharp \lambda_2) \quad \text{and} \quad \omega_{r-t}(\lambda_R \sharp \lambda_2) \prec \omega_{r-t}(\lambda_1 \sharp \lambda_2)$$

.

With the use of proposition 3.25 we have the result we wanted. □

3.6 Consistency of PA: an overview

3.6.1 A quick summary

Gentzen originally used proof by contradiction and assumed that we can derive the empty sequent in **PA** [38, sec. 4.3]. In the proof we have presented we, however, prove that when starting with "true" assumptions we can't end up with a "false" conclusion. Moreover, in order to prove that PA is consistent, Gentzen formalised in a sequent calculus system (originally created by him) Peano axioms. After that he noticed (actually this was noticed even when he was studying the natural deduction system and before formalising PA) that when an introduction rule is succeeded by the corresponding elimination rule we acquire a "hillock" or detour that can be eliminated as redundant [37].

By applying the procedure explained above we eventually create a simple proof (it has no CJs, no complex CUTs, no logical inferences, we only have atomic closed formulas, atomic CUTs and structural rules) the end-sequent of which isn't necessarily the same as the original $\Gamma \vdash \Delta$, but it is a sequent $\Gamma^* \vdash \Delta^*$ where Γ^* and Δ^* are produced by Γ and Δ by deletion of some occurrences of formulas.

We know that we will eventually get a simple proof because of the assignment of ordinal notations in the end-sequents and the fact that the ordinal notations will stop decreasing eventually. By proposition 3.33 every simple proof from "true" sequents must have a "true" end-sequent. So there can't exist a proof of the empty sequent which incarnates absurdity.

3.6.2 How does the proof work?

Regarding Gentzen's first published proof (that probably assumed the existence of a proof of absurdity) we can see that the idea was that:

> *The atomic formulas of arithmetic are decidable equalities between numerical terms. It follows that the whole propositional part of arithmetic is decidable. Gentzen's reduction procedure is carried over from the classical propositional logic of formulas to sequents, as exemplified by the following: If $A \wedge B$ in the antecedent of a sequent $A \wedge B, \Gamma \vdash C$ is false, one of A and B is false, and each can be tried in turn in the place of $A \wedge B$. If $\neg A$ in $\neg A, \Gamma \vdash C$ is false, it is deleted and the sequent changed into $\Gamma \vdash A$.*
>
> *Gentzen's essential idea is to extend the procedure from the finitary domain to quantified formulas, i.e., to apply the "transfinite sense" of $\forall x A(x)$ in a certain way. Gentzen calls it "the in-itself sense" (der an-sich Sinn).*
>
> *A way to think of the reduction procedure is that the correctness of a sequent $\Gamma \vdash C$ is guaranteed if, in whatever way C may have as a consequence a false claim, it can be shown that some assumption in Γ likewise*

> *presupposes a falsity. Then, whenever the assumptions Γ hold, also C holds. Say, to put it in figurative terms, we have a sequent of the form $\Gamma \vdash \forall x A(x) \wedge \forall x B(x)$ and an omniscient opponent who can* ***reason classically*** *by the in-itself sense of things and to whom the infinity of the natural numbers is not an obstacle. Such a creature can decide when $\forall x A(x) \wedge \forall x B(x)$ is false in its eyes, with, say, $\forall x A(x)$ a false conjunct, next to take a falsifying instance $A(n)$ out of the infinitely many possibilities. Our task is to show that, even if we don't have the opponent's classical and transfinite capacities, we can make finitarily choices after the opponent's choices so that some assumption in Γ turns out false. ...*
>
> *The aim of the reduction procedure is to ensure that a false formula in the antecedent part of a sequent can be produced, whenever a false numerical equation has appeared in the succedent. ... Given a sequent $\Gamma \vdash C$, the result of reduction is, provided the process terminates, a sequent to which no reduction step applies. [38, sec. 4.3].*

We, however, have gone the other way round. Any proof that fulfils the conditions we have claimed (it is simple), can't be a proof of absurdity.

3.7 The ordinal ε_0

The ordinal ε_0 is linked to the proof of consistency of PA as it was proved by Gentzen.

> *"Gentzen's consistency proof of PA proceeds by induction along a well-ordering (namely along the well ordering of ordinal notations)" [15, sec.8.3]*

More specifically, we know that the axiomatic system that is needed for the proof is Primitive Recursive Arithmetic together with transfinite induction for quantifier-free formulas up to ε_0 [27, pp. 8–9]. Moreover, "*the transfinite ordinal ε_0 ... characterizes Peano arithmetic*" [38, p. 106].

> "Gentzen showed in his Habilitation (in 1943) that transfinite induction up to any ordinal $< \varepsilon_0$ is provable in first order arithmetic — and made a constructive justification of how to reach any ordinal $< \varepsilon_0$."

but ordinal ε_0 is the least ordinal number for which PA *can't* prove it is well-founded as an ordering [42, p. 51, 37]. It is also a fixed-point [15, sec. 8.8] so it holds that

$$\varepsilon_0 = \omega^{\varepsilon_0}$$

3.8 Applications of parts of the proof

Gentzen's proof is more widely known than that of Gödel and we can't really state all its direct and indirect outcomes. This is mainly because it was from the beginning presented in its wholeness from Gentzen, while Gödel only sketched his, leaving all the rest for the reader (even Troelstra in [45, p. 222] is hesitant when presenting system $\mathcal{T}$ as the one Gödel had in mind).

It is for sure Gentzen the one to be credited for the creation of the field of proof theory and, as a consequence, all proof theorists' results are an offspring of his dissertation.

"Gentzen's celebrated consistency proof—or proofs, to distinguish the different variations he gave—of Peano Arithmetic in terms of transfinite induction up to the ordinal ε_0 can be considered as the birth of modern proof theory. After the blow which Gödel's incompleteness theorems gave the original Hilbert Programme, Gentzen's result did not just provide ***a*** *consistency proof of formalized Arithmetic, it also opened a new way to deal "positively" with incompleteness phenomena. In addition, Gentzen invented, on the way to his result, structural proof theory, understood as the branch of proof theory studying structural (in contrast to mathematical) properties of formal systems. With the introduction of sequent calculus and natural deduction and the corresponding theorems about cut elimination and normalization, respectively, he revolutionized the concept of derivation calculus, fundamental for all further developments of proof theory." [27]*

CHAPTER 4

GÖDEL'S PROOF

4.1 Introductory notes

Before we present Gödel's proof, we should first make some remarks concerning its presentation.

In various books Gödel's proof and system $\mathcal{T}$ are presented differently than in this thesis. We will follow mostly [16, annexes 7.A, 7.B, 46], although not very faithfully. Many proofs and definitions have been grafted from other books after appropriate modification. Other sources are [24] and [43]. In [24] the writer uses combinatory logic instead of lambda calculus.

We shall omit things that would turn this thesis too long, as for example an introduction to lambda calculus, an introduction to intuitionistic logic, the proof of the strong normalization theorem and some other proofs. We thought that combinatory logic is rarely ever taught in Greece, so we chose lambda calculus which is at least easier to understand and still existent in some departments. We assume that the reader is familiar with the omitted parts. If not, one can read Chapters 1 and 2 in [43] for an introduction to lambda calculus and intuitionistic logic. For the rest we will state all that is needed for understanding the proof and give adequate references where the reader can find the omitted parts in detail.

4.2 Sketch of the proof

We thought that it would be quite helpful, if we made a rough sketch of the proof before the main presentation. Moreover, in the next sections, we will give some intuition for better understanding.

Gödel's proof consists of two reductions. To avoid any misconceptions, these reductions have nothing to do with the well-known reductions of problems contained in polynomial or arithmetical hierarchy. The first reduction is from the formulas of PA through the double-negation translation, also called Gödel-Gentzen translation, to HA and the second reduction is from HA through the Dialectica intepretation, to a typed lambda calculus system, called $\mathcal{T}$, constructed especially for the purposes of the proof by Gödel.

Let ϕ^N symbolize the formula that outputs the double-negation translation for PA formula ϕ and $\phi^D = \exists x \forall y \phi_D$ the formula that outputs the Dialectica translation for an HA-formula ϕ (which will become clear in the next sections). These interpretations preserve provability in the following way:

$$\text{PA} \vdash \phi \Rightarrow \text{HA} \vdash \phi^N \Rightarrow \mathcal{T} \vdash (\phi^N)_D$$

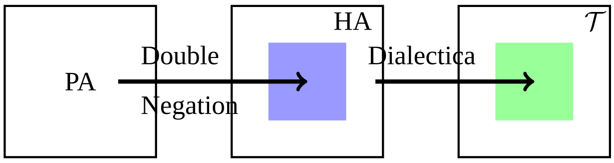

Figure 4.1: The two reductions needed for the proof

We will prove that system $\mathcal{T}$ is consistent, that is we can't prove $(\perp^N)_D$ in it. This property transfers to PA through the two reductions and, thus, PA is consistent. Moreover, its consistency can be proved with the use of PRA and the strong normalization theorem only.

Before the proof of $\mathcal{T}$'s consistency though, we will define all needed notions and explain the underlying ideas of Dialectica's form.

4.3 Peano Arithmetic and Heyting Arithmetic

Definition 4.1. In the language of L_0 (see definition 3.1) we define the theory of Peano Arithmetic (PA) to be the following axiom *schemes* and rules of inference, i.e all substitution instances [25, 46, def. 1.1.3, 1.1.4, 1.3.3]:

1. Equality axioms

   a. $x = x$

   b. $x = y \rightarrow sx = sy$

   c. $x = y \rightarrow (x = z \rightarrow y = z)$

2. Definition axioms

   a. $x + 0 = x$

   b. $x + sy = s(x + y)$

   c. $x \cdot 0 = 0$

   d. $x \cdot sy = x \cdot y + x$

   e. $\neg(sx = 0)$

   f. $sx = sy \rightarrow x = y$

3. Axioms of predicate logic for ϕ, ψ being formulas in L_0 and $\perp \equiv (0 = s0)$ (=absurdity)

   a. $\phi \rightarrow (\phi \wedge \phi)$

   b. $\phi \vee \phi \rightarrow \phi$

   c. $\phi \to (\psi \vee \phi)$
   d. $(\phi \wedge \psi) \to \phi$
   e. $(\phi \wedge \psi) \to (\psi \wedge \phi)$
   f. $\phi \vee \psi \to (\psi \vee \phi)$
   g. $\perp \to \phi$ (ex falso quodlibet)[1]
   h. $\neg\phi \to (\phi \to \perp)$
   i. $(\phi \to \perp) \to \neg\phi$
   j. $\neg\neg\phi \to \phi$ (*)
   k. $\phi[x := Q] \to \exists x\phi[x]$ for Q being a term free for x in $\phi[x]$
   l. $\forall x\phi[x] \to \phi[x := Q]$ for Q being a term free for x in $\phi[x]$

4. We say that $\Gamma \vdash \phi$, for Γ a set of formulas, if there is a sequence of formulas $\phi_1, ..., \phi_n$ such that: ϕ_i is either an axiom of the ones above, or it can be derived by using a rule below and some previous formulas in the sequence. We write $\vdash \phi$ if Γ is empty and call ϕ a theorem of PA.

   The Rules of inference for ϕ, ψ, β being formulas in L_0 are the following:

   a. $\{\phi, \phi \to \psi\} \vdash \psi$ (Modus Ponens)
   b. $\{\phi \to \psi, \psi \to \beta\} \vdash \phi \to \beta$
   c. $\phi \to \beta \vdash \phi \vee \psi \to \beta \vee \psi$
   d. $\phi \to (\psi \to \beta) \vdash (\phi \wedge \psi) \to \beta$
   e. $\phi \to \psi[x] \vdash \phi \to \forall x\psi[x]$ where x is not free in ϕ and if $\Gamma \vdash \phi \to \psi[x]$, then x can't be free in formulas of Γ.
   f. $\psi[x] \to \phi \vdash \exists x\psi[x] \to \phi$ where x is not free in ϕ and if $\Gamma \vdash \psi[x] \to \phi$, then x can't be free in formulas of Γ (Exists Elimination)
   g. $(\phi \wedge \psi) \to \beta \vdash \phi \to (\psi \to \beta)$
   h. $\{\phi[x := 0], \forall x(\phi[x] \to \phi[x := sx])\} \vdash \forall y\phi[x := y]$ (Rule of Induction)[2]

Definition 4.2. The intuitionistic variant of PA is the theory of Heyting Arithmetic (HA), which is obtained from PA by deleting axiom (*), i.e (3j).

Remark 4.3. 1. Because HA has one fewer axiom than PA, it is a weaker theory.

2. Gödel used a system for HA that has the advantage of keeping complexities down to a minimum, i.e., fewer logical symbols appear in the rules and axioms than in other equivalent systems [46, def. 1.1.4]. This is the system that we will also use.

4.4 Double-Negation Interpretation

The Double negation translation embeds classical predicate logic into the "negative" fragment of intuitionistic predicate logic. It was discovered independently by Gentzen and by Gödel. Its name, as we will see, is mostly due to transformation (1) in the definition and its main purpose is to preserve provability between PA and HA.

[1] Latin for "From a false proposition, anything follows".
[2] The assumptions could have been replaced by a single conjunction, but this form is better for the purposes of the soundness theorem.

Definition 4.4. The double-negation translation (or Gödel-Gentzen translation) is defined inductively as follows for ϕ, ψ being formulas of L_0 (in fact we consider them to be formulas of PA) [2]:

1. $\phi^N \equiv \neg\neg\phi$ for ϕ: atomic
2. $(\phi \wedge \psi)^N \equiv \phi^N \wedge \psi^N$
3. $(\phi \vee \psi)^N \equiv \neg(\neg\phi^N \wedge \neg\psi^N)$
4. $(\phi \rightarrow \psi)^N \equiv \phi^N \rightarrow \psi^N$
5. $(\forall x\phi[x])^N \equiv \forall x(\phi[x])^N$
6. $(\exists x\phi[x])^N \equiv \neg\forall x\neg(\phi[x])^N$

Remark 4.5. 1. Via this translation we obtain that if PA $\vdash \phi$ then HA $\vdash \phi^N$. Thus if HA $\nvdash (\bot)^N$, we conclude that PA $\nvdash \bot$ and PA is consistent [2, Collor. 2.1.2].

2. It holds that a formula ϕ and its translation are classically equivalent.

3. It holds that $\exists x\phi[x]$ and $(\exists x\phi[x])^N$ are intuitionistically equivalent as well as $\phi \vee \psi$ and $(\phi \vee \psi)^N$ [7, lem. 6.2.1]. Moreover, we could have chosen their translations as follows: $(\phi \vee \psi)^N \equiv \neg\neg(\phi^N \vee \psi^N)$ and $(\exists x\phi[x])^N \equiv \neg\neg\exists x(\phi[x])^N$ [50].

4. If we consider only the propositional fragment, we obtain that ϕ is a classical tautology iff $\neg\neg\phi$ is an intuitionistic tautology, which is Glivenko's theorem [43, Theorem 2.4.10].

4.5 System $\mathcal{T}$

System $\mathcal{T}$ can be presented in various ways. One for example is free to add or skip adding product types, product terms and projection terms without that affecting its expressive power [43, p. 265]. For types this is because we can interpret $\rho \times \sigma \rightarrow \tau$ as $\rho \rightarrow (\sigma \rightarrow \tau)$ for ρ, σ, τ being types [2]. Moreover, some add extra types as for example **bool** (boolean) [17].

The "original" (we suppose that Gödel had this in mind, but there is no explicit definition in his paper [45, p. 222]) system as presented in [2, 45] has also some constant terms that we won't use, because we follow [16]; the $S_{\rho,\sigma,\tau}$ and $K_{\sigma,\tau}$ typed combinators.

System $\mathcal{T}$ has typed terms, also called programs or algorithms, that always terminate. This is due to the strong normalization theorem that holds for all terms of $\mathcal{T}$. This gives an essence of programming language to $\mathcal{T}$, albeit not all partial recursive functions can be "computed" in it.

It can be proved that functions definable in $\mathcal{T}$ are a strict subset of recursive functions (which are total), called *provably total* [43] and that they are computable by a sequential algorithm, not necessarily though of a desirable time complexity [47].

The idea behind system $\mathcal{T}$ could be summarized in the phrase "*we would like to develop a technique for extracting programs (i.e., terms) from proofs in HA*". In the system, for example, specific axioms of arithmetic correspond to constants of appropriate type as we now explain [43].

Definition 4.6. System $\mathcal{T}$ has types, typed terms (in short terms), formulas, axioms and deduction rules and we can define a relation between its terms. More specifically [43, def. 1.3.1-1.3.3, 10.2.1, 16, def. 7.A.4-7.A.6]:

1. A type in $\mathcal{T}$ is defined inductively as follows:
   - **int** is the only primitive type. It's a type constant for integers.
   - $\sigma \to \tau$ is a type for σ, τ being types. (function type).
   - Nothing else is a type.

   In the system we have no type variables. That is, types can't change once assigned to a term.

2. A (typed) term with type σ in $\mathcal{T}$ is denoted as $M : \sigma, M^\sigma, M \in \sigma$ and is defined inductively as follows:
   - for every variable x from a denumerable set of variables and every type τ, $x : \tau$ is a typed term.
   - $\mathbf{0}$: **int** is a constant term.
   - for $M : \sigma \to \tau$ and $N : \sigma$, $MN : \tau$ is a term (application).
   - for $M : \tau$ and $x : \sigma$, $\lambda x^\sigma.M : \sigma \to \tau$ is a term (abstraction).
   - **s** : **int** $\to$ **int** is a constant term (successor). This gives that for Q : **int**, $s(Q)$ is of type **int**
   - for every type σ, $\boldsymbol{R_\sigma}$: $\sigma \to [(\sigma \to (\textbf{int} \to \sigma)) \to (\textbf{int} \to \sigma)]$ is a constant term (recursor)
   - Nothing else is a typed term.

3. A $\mathcal{T}$-formula is defined inductively as follows:
   - for M, N terms of type **int** in $\mathcal{T}$, $M =_{\mathbf{i}} N$ is an atomic/prime $\mathcal{T}$-formula
   - we define $\bot :\equiv \boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$
   - if ϕ, ψ are $\mathcal{T}$-formulas then $\phi \wedge \psi$, $\phi \vee \psi$, $\phi \to \psi$, $\neg\phi \equiv \phi \to \bot$ [3]
   - Nothing else is a $\mathcal{T}$-formula.

4. We define a relation between terms of $\mathcal{T}$ called reduction and denoted by $\to_{\mathcal{T}}$ as the least binary relation that satisfies the following properties for terms M, N, Z, P, Q of appropriate types and x a (typed) variable:
   - if $M \to_{\mathcal{T}} N$ then $\lambda x.M \to_{\mathcal{T}} \lambda x.N$
   - if $M \to_{\mathcal{T}} N$ then $MZ \to_{\mathcal{T}} NZ$
   - if $M \to_{\mathcal{T}} N$ then $ZM \to_{\mathcal{T}} ZN$
   - $(\lambda x.P)Q \to_{\mathcal{T}} P[x := Q]$[4]
   - $\boldsymbol{R_\sigma} MN\mathbf{0} \to_{\mathcal{T}} M$, where $M \in \sigma$, $N \in \sigma \to (\textbf{int} \to \sigma)$

[3] We will follow similar notational conventions for free variables in terms as in logical formulas. See footnote in definition 3.1

[4] For variable substitution rules see [43, Ch. 1, 31, def.1B.8]. We will make use of simultaneous substitution ([43, def. 1.2.21]) using distinct variables separated by commas instead of vector variables, i.e $\phi[x := M, y := N]$.

- $\boldsymbol{R_\sigma} MN(\boldsymbol{s}(P)) \to_{\mathcal{T}} N(\boldsymbol{R_\sigma} MNP)P$, where $M \in \sigma$, $N \in \sigma \to (\textbf{int} \to \sigma)$, $P \in \textbf{int}$

We will denote the transitive and reflexive closure of $\to_{\mathcal{T}}$ by $\twoheadrightarrow_{\mathcal{T}}$ (multi-step reduction) and the least equivalence relation containing $\to_{\mathcal{T}}$ by $=_{\mathcal{T}}$.

5. The axioms of $\mathcal{T}$ are the following schemes [16, 46, remark 1.5.8, def. 1.6.7, def. 1.6.13]:

   a. for x, y, M of type **int**

   $$(x =_{\mathbf{i}} y) \to M[z := x] =_{\mathbf{i}} M[z := y]$$

   if x, y are free for z in M.

   b. $\neg(\boldsymbol{s}(x) =_{\mathbf{i}} \mathbf{0})$

   c. $(\boldsymbol{s}(x) =_{\mathbf{i}} \boldsymbol{s}(y)) \to x =_{\mathbf{i}} y$

   d. $[x =_{\mathbf{i}} y \wedge y =_{\mathbf{i}} z] \to x =_{\mathbf{i}} z$

6. For M, N, Q, V terms of type **int**, $P \in \textbf{int} \to \textbf{int}$ and $x, w, y \in \textbf{int}$ variables, we have the following rules of inference for prime formulas [16, def. 7.B.1, 44, sec. 4.3.2]:

   a. If $M =_{\mathcal{T}} N$ then $\mathcal{T} \vdash M =_{\mathbf{i}} N$

   b. $M =_{\mathbf{i}} N \vdash N =_{\mathbf{i}} M$ (symmetry)

   c. $M =_{\mathbf{i}} N,\ N =_{\mathbf{i}} Q \vdash M =_{\mathbf{i}} Q$ (transitivity)

   d. $M =_{\mathbf{i}} N \vdash PM =_{\mathbf{i}} PN$[5]

7. For $\mathcal{T}$-formulas ϕ, ψ, ϑ and Γ a set of $\mathcal{T}$-formulas, we have also the following rules of inference for $\mathcal{T}$-formulas [43, def. 2.2.1]:

   a. $\Gamma, \phi \vdash \phi$ (Axiom rule)

   b. If $\Gamma, \phi \vdash \psi$, then $\Gamma \vdash \phi \to \psi$ ($\to$I)

   c. If $\Gamma \vdash \phi \to \psi$ and $\Gamma \vdash \phi$, then $\Gamma \vdash \psi$ ($\to$E)

   d. If $\Gamma \vdash \phi$ and $\Gamma \vdash \psi$, then $\Gamma \vdash \phi \wedge \psi$ ($\wedge$I)

   e. If $\Gamma \vdash \phi \wedge \psi$, then $\Gamma \vdash \phi$ ($\wedge$1E)

   f. If $\Gamma \vdash \phi \wedge \psi$, then $\Gamma \vdash \psi$ ($\wedge$2E)

   g. If $\Gamma \vdash \phi$, then $\Gamma \vdash \phi \vee \psi$ ($\vee$1I)

   h. If $\Gamma \vdash \psi$, then $\Gamma \vdash \phi \vee \psi$ ($\vee$2I)

   i. If $\{\Gamma, \phi\} \vdash \vartheta$, $\{\Gamma, \psi\} \vdash \vartheta$ and $\Gamma \vdash \phi \vee \psi$, then $\Gamma \vdash \vartheta$ ($\vee$E)

   j. If $\Gamma \vdash \bot$, then $\Gamma \vdash \phi$ ($\bot$E)

   k. $\Gamma \cup \{\phi[x := \mathbf{0}], \phi[x := v] \to \phi[x := \boldsymbol{s}(v)]\} \vdash \phi[x := V]$ for $v, V \in \textbf{int}$ a variable and a term respectively (Rule of Induction, Ind)

Definition 4.7. We say that $\Gamma \vdash \phi$ (read "Γ proves ϕ") for Γ a set of $\mathcal{T}$-formulas[6] iff there is a sequence of applications of inference rules that leads to ϕ and every application uses as assumptions $=_{\mathcal{T}}$ equivalences, axioms of $\mathcal{T}$, $\mathcal{T}$-formulas of Γ or is the conclusion of a previous application in the sequence. If Γ is empty then we say that $\mathcal{T} \vdash \phi$ and ϕ is a theorem of $\mathcal{T}$ [43, def. 2.2.1, 52, def. 2.4.2].

[5] Note that we can't have $\lambda x.M =_{\mathbf{i}} \lambda x.N$ if $M =_{\mathbf{i}} N$ since we want both sides of the equation to have type **int**.

[6] Semantical equivalences "$=_{\mathcal{T}}$" are not $\mathcal{T}$-formulas!!!

Remark 4.8. 1. The three constant terms $\boldsymbol{0}, \boldsymbol{s}, \boldsymbol{R_\sigma}$ correspond to the following axioms of PA respectively:
-$0 \in \mathbb{N}$ (1st)
-$s : \mathbb{N} \to \mathbb{N}$ is a function (2nd)
-for every $A \subseteq \mathbb{N}$ for which it holds that if $0 \in A$ and if $\forall n \in A$, then $s(n) \in A$ we have that $A = \mathbb{N}$ (5th)

2. The recursor is presented differently in other books. For all symbols and parentheses conventions we followed mostly [43], but we have changed a bit the definition of its reduction rules.

3. We will omit the types of terms if it isn't necessary.

4. It holds that $\mathcal{T} \vdash M =_{\mathbf{i}} N$ iff $M =_{\mathcal{T}} N$ for M, N being closed terms, as we will see in theorem 4.38. "$=_{\mathbf{i}}$" is connected to denotational semantics and "$=_{\mathcal{T}}$" is connected to operational semantics [25, sec. 6A, def. 6.2].

5. We use the same symbols for connectives in $\mathcal{T}$ and for connectives in PA or HA, so we leave it up to the reader to discern between $\mathcal{T}$-formulas and regular ones. We also use the word "term" both for lambda typed terms and for terms of predicate logic, but the notation is different.

6. Constant terms have no free variables.

Definition 4.9. The terms of the form $\boldsymbol{s}(\boldsymbol{s}...(\boldsymbol{s0})...)$ where $\boldsymbol{s}$ occurs n times will be denoted as $\boldsymbol{n}$ or $\boldsymbol{s}^n(\boldsymbol{0})$ and are called <u>numerals</u>.

Remark 4.10. One can consider the recursor $R_\sigma MNV$ as the following recursive program for term $V \in$ **int** without free variables:

Algorithm 1 $\boldsymbol{R_\sigma}(M, Q, V)$

```
// N = λy^σ.λx^int.Q. Q is a program with at least two inputs x, y
k ← natural s.t. V =_T s^k(0)
if k = 0 then
    // Base case
    w ← M
else if k ≠ 0 then
    // Recursively compute program Q
    w ← Q(x := s^{k-1}(0), y := R_σ(M, Q, s^{k-1}(0)))
end if
return w
```

4.6 Strong Normalization

Definition 4.11. A term M in system $\mathcal{T}$ is in <u>normal form</u> (notation $M \in NF_{\mathcal{T}}$) iff there is no term N such that $M \to_{\mathcal{T}} N$ [43, def. 1.3.2].

Definition 4.12. A term M in system $\mathcal{T}$ is <u>normalizing</u> (notation $M \in WN_{\mathcal{T}}$) iff there *exists* a sequence from M ending in a normal form N. We say that N is the normal form of M [43, def. 1.5.1].

Definition 4.13. A term M in system $\mathcal{T}$ is strongly normalizing (notation $M \in SN_{\mathcal{T}}$ or $M \in SN$) iff *all* reduction sequences starting with M are finite [43, def. 1.5.1].

Theorem 4.14 (Strong normalization theorem). Every term in system $\mathcal{T}$ is strongly normalizing.

Proof. For a detailed proof one can read [43, subsec. 10.3] and for an alternative approach [24, S1G]. Here we will make a quick presentation of the proof in section 4.6.1. □

Theorem 4.15. (Church-Rosser Property) If $M \twoheadrightarrow_{\mathcal{T}} M_1$ and $M \twoheadrightarrow_{\mathcal{T}} M_2$, there exists a term N such that $M_1 \twoheadrightarrow_{\mathcal{T}} N$ and $M_2 \twoheadrightarrow_{\mathcal{T}} N$.

$$\begin{tikzcd}
 & M \arrow[dl, two heads] \arrow[dr, two heads] & \\
M_1 \arrow[dr, two heads] & & M_2 \arrow[dl, two heads] \\
 & N &
\end{tikzcd}$$

Figure 4.2: Church-Rosser Property

Proof. For a proof one can see [43, lem. 3.6.2, theor. 3.6.3] □

Remark 4.16. From theorem 4.15 we can conclude that if M reduces to two terms that are both in normal form, then these terms coincide. So from the strong normalization theorem and the Church-Rosser property, we have the existence and uniqueness of normal forms of terms. This fact is necessary (and sufficient) for proving the consistency of $\mathcal{T}$. While uniqueness, is formalizable in PA, the existence of the normal form isn't, as we will see in 4.6.1.

Definition 4.17. A term (in $\mathcal{T}$) is said to be closed if it has no free variables.

Lemma 4.18. Every closed normal term of type **int** is one of the following: $\mathbf{0}, \boldsymbol{s}(P)$ for some terms $P \in$ **int**.

Proof.
- *Basis*
  For $\mathbf{0}$ the claim holds, as it is closed, normal and of type **int**.
- *Inductive step*
  Suppose that $P \in$ **int** is a closed normal term for which the induction hypothesis holds. Then $\boldsymbol{s}(P) \in$ **int** is closed and normal.
- *Extremal clause*: Nothing else is a closed normal term of type **int**
  Suppose that $M \in$ **int** is normal with at most x free and $N \in$ **int** is closed and normal.

  -$(\lambda x.M)N$ is further reducible, thus not normal
  -$\lambda x.M$ is not of type **int**

  Suppose that $P, M \in$ **int**, $N \in$ **int** $\rightarrow$ (**int** $\rightarrow$ **int**) are closed normal. We can't form closed normal $\boldsymbol{R}_{\mathbf{int}}MNP \in$ **int** because P can't be $\mathbf{0}, \boldsymbol{s}(Q)$ (for Q closed and normal) or anything else, as otherwise $\boldsymbol{R}_{\mathbf{int}}MNP$ would be further reducible.

□

Theorem 4.19. For every closed term M of type **int** there is a unique number $n \in \mathbb{N}$ such that $M =_{\mathcal{T}} \boldsymbol{s}^n(\mathbf{0})$

Proof. We will prove the theorem by induction on terms. The uniqueness is given from the uniqueness of the normal form of M which exists for every term.

- *Basis*
  For $\mathbf{0}$ the theorem holds.

- *Inductive step*
  Suppose that for term Q the induction hypothesis holds. Namely, $Q =_{\mathcal{T}} \boldsymbol{s}^n(\mathbf{0})$. We have that $\boldsymbol{s}(Q)$ is a closed normal term of type **int** and it holds that

$$\boldsymbol{s}(Q) =_{\mathcal{T}} \boldsymbol{s}(\boldsymbol{s}^n(\mathbf{0})) \equiv \boldsymbol{s}^{n+1}(\mathbf{0})$$

From lemma 4.18 and the above induction the theorem holds for all closed normal terms of type **int**. □

Definition 4.20. We define the following closed typed terms for x, y terms of type **int** (the proofs of the properties mentioned are omitted) [16, example 7.B.2]:

1.
   - $\hat{\neg} = \boldsymbol{R}_{\mathbf{int}}\boldsymbol{s}(\mathbf{0})\lambda z^{\mathbf{int}}.\lambda y^{\mathbf{int}}.\mathbf{0}$ of type $\mathbf{int} \to \mathbf{int}$
   - $\hat{\neg}(\mathbf{0}) =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$
   - $\hat{\neg}(\boldsymbol{s}(x)) =_{\mathcal{T}} \mathbf{0}$
2.
   - $\hat{\vee} = \lambda x^{\mathbf{int}}.\boldsymbol{R}_{\mathbf{int}}\mathbf{0}\lambda z^{\mathbf{int}}.\lambda y^{\mathbf{int}}.x$ of type $\mathbf{int} \to (\mathbf{int} \to \mathbf{int})$
   - $\hat{\vee}x\mathbf{0} =_{\mathcal{T}} \mathbf{0}$
   - $\hat{\vee}x\boldsymbol{s}(y) =_{\mathcal{T}} x$
3.
   - $\hat{\wedge} = \lambda x^{\mathbf{int}}.\boldsymbol{R}_{\mathbf{int}}x\lambda z^{\mathbf{int}}.\lambda y^{\mathbf{int}}.\boldsymbol{s}(\mathbf{0})$ of type $\mathbf{int} \to (\mathbf{int} \to \mathbf{int})$
   - $\hat{\wedge}x\mathbf{0} =_{\mathcal{T}} x$
   - $\hat{\wedge}x\boldsymbol{s}(y) =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$
4.
   - $\hat{\to} = \lambda x^{\mathbf{int}}.\boldsymbol{R}_{\mathbf{int}}\mathbf{0}\lambda z^{\mathbf{int}}.\lambda y^{\mathbf{int}}.\hat{\neg}(x)$ of type $\mathbf{int} \to (\mathbf{int} \to \mathbf{int})$
   - $\hat{\to}x\mathbf{0} =_{\mathcal{T}} \mathbf{0}$
   - $\hat{\to}x\boldsymbol{s}(y) =_{\mathcal{T}} \hat{\neg}(x)$
5.
   - $\boldsymbol{E} = \boldsymbol{R}_{\mathbf{int}\to\mathbf{int}}(\boldsymbol{R}_{\mathbf{int}}\mathbf{0}\lambda z^{\mathbf{int}}.\lambda y^{\mathbf{int}}.\boldsymbol{s}(\mathbf{0}))[\lambda z^{\mathbf{int}\to\mathbf{int}}.\lambda y^{\mathbf{int}}.\boldsymbol{R}_{\mathbf{int}}\boldsymbol{s}(\mathbf{0})\lambda w^{\mathbf{int}}.\lambda x^{\mathbf{int}}.z(x)]$ of type $\mathbf{int} \to (\mathbf{int} \to \mathbf{int})$
   - $\boldsymbol{E}\mathbf{0}\mathbf{0} =_{\mathcal{T}} \mathbf{0}$
   - $\boldsymbol{E}\mathbf{0}\boldsymbol{s}(y) =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$
   - $\boldsymbol{E}\boldsymbol{s}(x)\mathbf{0} =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$
   - $\boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathcal{T}} \boldsymbol{E}xy$

Remark 4.21. The first four terms of definition 4.20 represent the propositional connectives in the following way: if M, N are *closed* terms of type **int**, the terms $\hat{\neg}M$, $\hat{\vee}MN$, $\hat{\wedge}MN$ and $\hat{\to}MN$ are of type **int** and they are equivalent to numerals according to theorem 4.19. Namely, it can be proved that [24, lem. S1.24]:

$$\begin{array}{rcl} \hat{\neg} M =_{\mathcal{T}} \mathbf{0} & \Leftrightarrow & M =_{\mathcal{T}} \boldsymbol{s}^n(\mathbf{0}) \text{ for some } n \in \mathbb{N} \setminus \{0\} \\ \hat{\wedge} MN =_{\mathcal{T}} \mathbf{0} & \Leftrightarrow & M =_{\mathcal{T}} \mathbf{0} \text{ and } N =_{\mathcal{T}} \mathbf{0} \\ \hat{\vee} MN =_{\mathcal{T}} \mathbf{0} & \Leftrightarrow & M =_{\mathcal{T}} \mathbf{0} \text{ or } N =_{\mathcal{T}} \mathbf{0} \\ \hat{\rightarrow} MN =_{\mathcal{T}} \mathbf{0} & \Leftrightarrow & N =_{\mathcal{T}} \mathbf{0} \text{ or } M =_{\mathcal{T}} \boldsymbol{s}^n(\mathbf{0}) \text{ for some } n \in \mathbb{N} \setminus \{0\} \end{array}$$

The fifth term represents equality between terms and it can be proved that (see theorem 4.38):

$$\boldsymbol{E} MN =_{\mathcal{T}} \mathbf{0} \Leftrightarrow M =_{\mathcal{T}} N \Leftrightarrow \mathcal{T} \vdash M =_{\mathbf{i}} N$$

if M, N are closed terms of type **int**.

4.6.1 Discussion

If we consider terms of system $\mathcal{T}$ as programs, we can say that the strong normalization theorem states that all programs in the system terminate. None of them falls into an infinite loop. In fact, the functions that can be computed via these programs are, as stated already, a strict subset of the recursive functions, which are total; the set of *provably total* functions [17, sec. 15.1.3].

Another important remark that should be made is one concerning the relation between Gödel's second incompleteness theorem and the strong normalization theorem. Due to Gödel's theorem we know that any consistency proof of the theory of PA must involve a step that can't be formalized in the theory. This step is the proof of the strong normalization theorem. All the other parts of the proof can be formalized in PA. Why is that?

The proof of the strong normalization theorem contains the concept of "computable term" which can't be formalized in PA. It requires quantification over sets. More specifically, the theorem itself could be expressed roughly as "for each (Gödel coding of a) term M in $\mathcal{T}$, there is a number n such that all reduction paths starting with M consist of at most n steps". This statement can be formalized in PA. However, in its proof we bump into the statement that "every computable term is strongly normalizing". One might think that this can be encoded, but a closer look in the definition of computable terms ruins this thought. Firstly, we define, by induction on type complexity, the notion of "computable terms" of certain type. Afterwards, we prove that every typed term and every constant in $\mathcal{T}$ are computable. This means that every term belongs to a class of computable terms $[\![\tau]\!]$ for some type τ. Next, we prove that for a given type τ, the class of computable terms of type τ has only strongly normalizing terms.

Even if we tried to create a formula $computable_\tau(M)$ stating that $M \in [\![\tau]\!]$, we wouldn't have a general solution, as the notion of computability has no *uniform* definition for all types. Hence, it needs an inductive approach, which doesn't yield an explicit formula [43, sec. 10.3].

4.7 Dialectica interpretation

The Dialectica-intepretation is also called D-interpretation and is a functional interpretation of intuitionistic (Heyting) arithmetic in a quantifier-free theory. It took its name by the journal in which Gödel presented his paper for the first time.

Before we move on, and although we kindly disagree, we should mention some criticism concerning Dialectica. In [16, p. 442] it is stated that:

> *it is not the most outstanding of all Gödel's achievements...One of the main reasons is that the interpretation is extremely complicated and usually unmanageable*

This is for the extra reason that there are alternative proofs (see for example [43, sec. 10.4]) that use simpler functions to link provability in HA and in system $\mathcal{T}$. That being said we move to the presentation of Dialectica.

Definition 4.22. We define the following terms of system $\mathcal{T}$ (the proofs for the properties mentioned are omitted) [16, example 7.B.2]:

1. - $\oplus \equiv \lambda x^{\mathbf{int}}.\boldsymbol{R}_{\mathbf{int}} x \lambda z^{\mathbf{int}}.\lambda y^{\mathbf{int}}.\boldsymbol{s}(z)$ of type **int** → (**int** → **int**)
   - $\oplus x\mathbf{0} =_{\mathcal{T}} x$
   - $\oplus x\boldsymbol{s}(y) =_{\mathcal{T}} \boldsymbol{s}(\oplus xy)$
2. - $\bullet \equiv \lambda x^{\mathbf{int}}.\boldsymbol{R}_{\mathbf{int}} \mathbf{0} \lambda z^{\mathbf{int}}.\lambda y^{\mathbf{int}}. \oplus zx$ of type **int** → (**int** → **int**)
   - $\bullet x\mathbf{0} =_{\mathcal{T}} \mathbf{0}$
   - $\bullet x\boldsymbol{s}(y) =_{\mathcal{T}} \oplus(\bullet xy)x$
3. - $\boldsymbol{P} = \boldsymbol{R}_{\mathbf{int}} \mathbf{0} \lambda z^{\mathbf{int}}.\lambda y^{\mathbf{int}}.y$ of type **int** → **int**
   - $\boldsymbol{P}\mathbf{0} =_{\mathcal{T}} \mathbf{0}$
   - $\boldsymbol{P}\boldsymbol{s}(x) =_{\mathcal{T}} x$
4. - $\dot{-} = \lambda x^{\mathbf{int}}.\boldsymbol{R}_{\mathbf{int}} x \lambda z^{\mathbf{int}}.\lambda y^{\mathbf{int}}.\boldsymbol{P}(z)$ of type **int** → (**int** → **int**)
   - $\dot{-} x\mathbf{0} =_{\mathcal{T}} x$
   - $\dot{-} x\boldsymbol{s}(y) =_{\mathcal{T}} \boldsymbol{P}(\dot{-} xy)$

Remark 4.23. 1. The above terms represent addition, multiplication, predecessor function and cut-off subtraction function respectively.

2. We will omit the types of the above terms.

3. In what follows capital letters $X, Y, Z....$ will be sometimes used for variables when we want to stress that they have a function type.

4. We will try to not encumber the reader by using vectors of variables, i.e., $\bar{x} = (x_1, ..., x_n)$. All the results for system $\mathcal{T}$ can be turned to results for vector variables by considering the following notational conventions, with the presupposition that they have appropriate types for application to take place [46, sec. 1.6.5]:
   - $\bar{x}\bar{y} \equiv x_1 y_1 ... y_m, ..., x_n y_1 ... y_m$ (it yields a sequence of application terms)
   - $\lambda\bar{x} \equiv \lambda x_1 .... \lambda x_n \equiv \lambda x_1 .... x_n$
   - $\exists\bar{x} \equiv \exists x_1 .... \exists x_n \equiv \exists x_1 .... x_n$
   - $\forall\bar{x} \equiv \forall x_1 .... \forall x_n \equiv \forall x_1 .... x_n$

Definition 4.24. For every formula $\phi \in L_0$ (in fact we assume it is of HA) we associate a quantified $\mathcal{T}$-formula $\phi^D \equiv \exists x_1, ..., x_n \forall y_1, ..., y_m \phi_D$, where ϕ_D is quantifier free and its free variables consist of the ones free in ϕ together with the sequence of the (typed) variables $x_1, ..., x_n, y_1, ..., y_m$, as follows [2, sec. 2.3, 24, def. S1.31, 46, def. 3.5.2, 16, def. 7.B.10]:

1. for every variable x of L_0, x^D is $x^{\mathbf{int}}$
2. 0^D is $\mathbf{0}$
3. $(sQ)^D$ is $\boldsymbol{s}(Q^D)$[7]
4. $(M = N)^D$ is $(M^D =_{\mathbf{i}} N^D)$ (atomic formulas)[8]
5. $(M + N)^D$ is $\oplus M^D N^D$
6. $(M \cdot N)^D$ is $\bullet M^D N^D$
7. For logically complex formulas $A, B \in L_0$ if (for simplicity) $A^D \equiv \exists x^\sigma \forall y^\tau A_D[x^\sigma, y^\tau]$ and $B^D \equiv \exists z^\alpha \forall w^\beta B_D[z^\alpha, w^\beta]$:
   - $(A \wedge B)^D \equiv \exists x^\sigma \exists z^\alpha \forall y^\tau \forall w^\beta (A_D[x^\sigma, y^\tau] \wedge B_D[z^\alpha, w^\beta])$
   - $(A \vee B)^D \equiv \exists p^{\mathbf{int}} \exists x^\sigma \exists z^\alpha \forall y^\tau \forall w^\beta (A_D[x^\sigma, y^\tau] \wedge p =_{\mathbf{i}} \mathbf{0}) \vee (B_D[z^\alpha, w^\beta] \wedge p =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}))$
   - $(\neg A)^D \equiv \exists Y^{\sigma \to \tau} \forall x^\sigma \neg A_D[x^\sigma, Yx]$
   - $(A \to B)^D \equiv \exists Y^{\sigma \to (\beta \to \tau)} \exists X^{\sigma \to \alpha} \forall x^\sigma \forall w^\beta \{A_D[x^\sigma, Yxw] \to B_D[Xx, w^\beta]\}$
   - $(\forall p A[p])^D \equiv \exists X^{\mathbf{int} \to \sigma} \forall y^\tau \forall p^{\mathbf{int}} A_D[Xp, y^\tau, p^{\mathbf{int}}]$
   - $(\exists p A[p])^D \equiv \exists p^{\mathbf{int}} \exists x^\sigma \forall y^\tau A_D[x^\sigma, y^\tau, p^{\mathbf{int}}]$

Remark 4.25. 1. To equality, addition, successor, multiplication and the variable-formulas we associate the "obvious" terms in $\mathcal{T}$. Similarly, for $\exists p A[p]$ the corresponding $\mathcal{T}$-formula has nothing strange.

2. For $A \wedge B$ the variables x, y, z, w must be distinct. If needed, we rename them [45, sec. 3.2, p225].

3. For formulas that are quantifier-free the interpretation is quantifier-free. This means that in the interpretation of complex formulas we might not have $A^D \equiv \exists x^\sigma \forall y^\tau A_D[x^\sigma, y^\tau]$ and $B^D \equiv \exists z^\alpha \forall w^\beta B_D[z^\alpha, w^\beta]$, but $A^D \equiv A_D$ and $B^D \equiv B_D$

4. The translation of $A \vee B$ has a certificate p of type **int** which "tells" which of the disjuncts to "choose" according to its value. We will see that it is decidable to choose between $p =_{\mathbf{i}} \mathbf{0}$ and $p =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})$.

5. The interpretation of $\forall p A[p]$ is obtained by prefixing A^D with a universal quantifier. After that we apply the axiom of choice

$$(AC) \qquad \forall x^\sigma \exists y^\tau A_D[x^\sigma, y^\tau] \to \exists Y^{\sigma \to \tau} \forall x^\sigma A_D[x^\sigma, Yx]$$

which is accepted in classical logic as it comes from set theory and it is accepted by many (yet not all) constructivists. Hence, the non-constructive parts of the proof are diminished to the minimum. With the application of (AC) we bring the existential quantifier in front ("Skolemization").

[7]Notice that the second "s" is bold. That's because it corresponds to $\mathcal{T}$'s "s" and not that of L_0.

[8]See remark 4.25

6. The interpretation of $A \to B$ needs a more extended explanation. We have the following equivalences:

$$
\begin{aligned}
(A \to B)^D &\equiv (A^D \to B^D) \\
&\leftrightarrow \forall x^\sigma(\forall y^\tau A_D[x^\sigma, y^\tau] \to \exists z^\alpha \forall w^\beta B_D[z^\alpha, w^\beta]) && (4.1)\\
&\leftrightarrow \forall x^\sigma \exists z^\alpha(\forall y^\tau A_D[x^\sigma, y^\tau] \to \forall w^\beta B_D[z^\alpha, w^\beta]) && (4.2)\\
&\leftrightarrow \forall x^\sigma \exists z^\alpha \forall w^\beta(\forall y^\tau A_D[x^\sigma, y^\tau] \to B_D[z^\alpha, w^\beta]) && (4.3)\\
&\leftrightarrow \forall x^\sigma \exists z^\alpha \forall w^\beta \exists y^\tau(A_D[x^\sigma, y^\tau] \to B_D[z^\alpha, w^\beta]) && (4.4)\\
&\leftrightarrow \forall x^\sigma \exists z^\alpha \exists Y_1^{\beta \to \tau} \forall w^\beta(A_D[x^\sigma, Y_1 w] \to B_D[z^\alpha, w^\beta]) && (4.5)\\
&\leftrightarrow \exists Z^{\sigma \to \alpha} \exists Y^{\sigma \to (\beta \to \tau)} \forall x^\sigma \forall w^\beta(A_D[x^\sigma, Yxw] \to B_D[Zx, w^\beta])
\end{aligned}
$$

   Equivalences 4.1 and 4.3 are justified by classicaland intuitionistic logic [7, lem. 6.2.1]. Equivalence 4.5 and the last one are justified by one and two applications of (AC) respectively. Equivalence 4.2 is an instance of the Independence of Premise schema

$$(IP') \qquad (\forall x C_1 \to \exists y C_2) \to \exists y(\forall x C_1 \to C_2)$$

   where C_1, C_2 are quantifier-free, y is not free in C_1 and x is not free in C_2.

   (IP') says that with the assumption $\forall x C_1 \to \exists y C_2$, we can a priori indicate y, independently of the truth of $\forall x C_1$. Intuitionistically, given a proof of $\forall x C_1$, we can find a y (possibly depending on the given proof), so that we construct a proof of C_2.

   Equivalence 4.4 can be justified by a generalization of Markov's principle

$$\neg \forall y \phi \to \exists y \neg \phi$$

   where ϕ is quantifier-free. More specifically, classically we have that the law of excluded middle holds. So if $B_D[z^\alpha, w^\beta]$ is true, then 4.4 is justified. If $B_D[z^\alpha, w^\beta]$ is false, then we apply (MP'). Intuitionistically the things are a bit more complex. There is "*no evident way to choose a witness y for $\neg\phi$ from a given proof that $\forall y \phi$ leads to contradiction. However, if y ranges over the natural numbers one can search for such a y given that one accepts its existence*" [2]. This is so that (MP') can be turned into an alternative form, which is accepted by some constructivists. Again, the purpose is to diminish non-constructive parts of the interpretation. Otherwise, "*any standard recipe for rewriting the formulas in prenex normal form, followed by a number of applications of AC so as to bring '$\exists$' in front, would do the job*" [45, p. 231].

7. Only $(A \to B)^D$ needs more than intuitionistically valid transitions to be an equivalent formula that will cover our needs.

8. Because Dialectica grows in complexity very fast, sometimes, it is preferable to interpret simpler equivalent formulas instead of the ones intended at first [16, p. 474].

9. For investigating the relationship between $\mathcal{T}$ and HA one can also use modified realizability as introduced by Kreisel [43, sec. 10.4]

10. The interpretation of $\neg A$ comes from the equivalence $\neg A \equiv A \to \bot$ and the interpretation of $A \to B$.

11. With classical logic and (AC) a formula ϕ and its D-interpretation are proved to be equivalent. For a more detailed explanation one can see [24, S1.32] and [46, def. 3.5.2, lem. 3.5.7].

4.8 Provability in system $\mathcal{T}$

In this (rather long) section we present some results that have to do with provability of formulas in $\mathcal{T}$. Most of them are linked to theorem 4.40. Our main goal is to demonstrate that for theorem 4.40 it is enough to have atomic formulas, i.e., have an equation calculus [44, ch. 4 paragr. 1.3, sec. 3]. The first three lemmas, although necessary, are quite technical and could be skipped at first. They only assist in proving lemma 4.34(a).

Remark 4.26. We will make use of the reverse of (→I) which holds because:

if $\Gamma \vdash \phi \to \psi$, we have that $\Gamma, \phi \vdash \phi \to \psi$ (weakening). Because $\Gamma, \phi \vdash \phi$ from axiom rule, we can apply (→E) and obtain $\Gamma, \phi \vdash \psi$.

Lemma 4.27. The following are theorems of $\mathcal{T}$ [16, lemma 7.B.3]:

(1) $\boldsymbol{E}xx =_{\mathsf{i}} \mathbf{0}$

(2) $\boldsymbol{E}x\mathbf{0} =_{\mathsf{i}} \mathbf{0} \vee \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}(x)) =_{\mathsf{i}} \mathbf{0}$

(3) $(\boldsymbol{E}x\mathbf{0} =_{\mathsf{i}} \mathbf{0} \wedge \phi[x]) \to \phi[x := \mathbf{0}]$ for ϕ a quantifier-free $\mathcal{T}$-formula

(4) $(\boldsymbol{E}x\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0} \wedge \phi[x]) \to \phi[x := \boldsymbol{s}(\mathbf{0})]$ for ϕ a quantifier-free $\mathcal{T}$-formula

(5) $(\boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}(x)) =_{\mathsf{i}} \mathbf{0} \wedge \phi[x]) \to \phi[x := \boldsymbol{s}(\boldsymbol{P}(x))]$ for ϕ a quantifier-free $\mathcal{T}$-formula

(6) $\boldsymbol{E}xy =_{\mathsf{i}} \mathbf{0} \to \boldsymbol{E}\boldsymbol{P}(x)\boldsymbol{P}(y) =_{\mathsf{i}} \mathbf{0}$

(7) $\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(y))(\dot{-}xy) =_{\mathsf{i}} \mathbf{0}$

(8) $\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0}$

(9) $\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x))\mathbf{0} =_{\mathsf{i}} \mathbf{0}$

(10) $\boldsymbol{E}(\dot{-}x\boldsymbol{P}(y))\mathbf{0} =_{\mathsf{i}} \mathbf{0} \to \boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathsf{i}} \mathbf{0}$

Proof. (1) We assume that $\phi[x] :\equiv \boldsymbol{E}xx =_{\mathsf{i}} \mathbf{0}$. We know that $\boldsymbol{E}\mathbf{0}\mathbf{0} =_{\mathcal{T}} \mathbf{0}$, so by rule 6a $\mathcal{T} \vdash \phi[x := \mathbf{0}]$. We also know that $\boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(x) =_{\mathcal{T}} \boldsymbol{E}xx$ and by axiom rule $\boldsymbol{E}xx =_{\mathsf{i}} \mathbf{0} \vdash \boldsymbol{E}xx =_{\mathsf{i}} \mathbf{0}$. So $\mathcal{T} \vdash \boldsymbol{E}xx =_{\mathsf{i}} \mathbf{0} \to \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(x) =_{\mathsf{i}} \mathbf{0}$. By Ind rule we have $\mathcal{T} \vdash \phi[x]$.

(2) $\phi[x] :\equiv \boldsymbol{E}x\mathbf{0} =_{\mathsf{i}} \mathbf{0} \vee \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}(x)) =_{\mathsf{i}} \mathbf{0}$. $\phi[x := \mathbf{0}] \equiv \boldsymbol{E}\mathbf{0}\mathbf{0} =_{\mathsf{i}} \mathbf{0} \vee \boldsymbol{E}\mathbf{0}\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0}$ which is a theorem due to $\mathcal{T} \vdash \boldsymbol{E}\mathbf{0}\mathbf{0} =_{\mathsf{i}} \mathbf{0}$ and rule (∨1I).

$\phi[x := \boldsymbol{s}(x)] \equiv \boldsymbol{E}\boldsymbol{s}(x)\mathbf{0} =_{\mathsf{i}} \mathbf{0} \vee \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(\boldsymbol{P}(\boldsymbol{s}(x))) =_{\mathsf{i}} \mathbf{0}$

$\boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(\boldsymbol{P}(\boldsymbol{s}(x))) =_{\mathcal{T}} \boldsymbol{E}x\boldsymbol{P}(\boldsymbol{s}(x)) =_{\mathcal{T}} \boldsymbol{E}xx.$

By (1) we have that

$$\mathcal{T} \vdash \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(\boldsymbol{P}(\boldsymbol{s}(x))) =_{\mathsf{i}} \mathbf{0}$$

We then apply rule (∨2I) and Ind rule.

(3) $\psi[x] :\equiv (\boldsymbol{E}x\mathbf{0} =_{\mathsf{i}} \mathbf{0} \wedge \phi[x]) \to \phi[x := \mathbf{0}]$

For $\psi[x := \mathbf{0}] \equiv (\boldsymbol{E}\mathbf{0}\mathbf{0} =_{\mathsf{i}} \mathbf{0} \wedge \phi[x := \mathbf{0}]) \to \phi[x := \mathbf{0}]$ to be a theorem, by reverse →I, it suffices to show that $\boldsymbol{E}\mathbf{0}\mathbf{0} =_{\mathsf{i}} \mathbf{0} \wedge \phi[x := \mathbf{0}] \vdash \phi[x := \mathbf{0}]$, which is true from ∧2E rule.

For $\psi[x := \boldsymbol{s}(x)] \equiv \boldsymbol{E}\boldsymbol{s}(x)\mathbf{0} =_{\mathsf{i}} \mathbf{0} \wedge \phi[x := \boldsymbol{s}(x)] \to \phi[x := \mathbf{0}]$ we have that:

1. $\boldsymbol{Es}(x)\mathbf{0} =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$ so $\mathcal{T} \vdash \boldsymbol{Es}(x)\mathbf{0} =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})$ (6a)
2. $\boldsymbol{Es}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \boldsymbol{s}(x)] \vdash \boldsymbol{Es}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0}$ ($\wedge$1E)
3. $\boldsymbol{Es}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0}, \boldsymbol{Es}(x)\mathbf{0} =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \vdash \boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \equiv \bot$ (1, 2, 6c)
4. $\boldsymbol{Es}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \boldsymbol{s}(x)] \vdash \boldsymbol{Es}(x)\mathbf{0} =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \rightarrow \bot$ (2, 3, $\rightarrow$I)
5. $\boldsymbol{Es}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \boldsymbol{s}(x)] \vdash \bot$ (2, 4, $\rightarrow$E)
6. $\boldsymbol{Es}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \boldsymbol{s}(x)] \vdash \phi[x := \mathbf{0}]$ (5, $\bot$E)

By weakening $\mathcal{T} \vdash \psi[x] \rightarrow \psi[x := \boldsymbol{s}(x)]$. We then apply Ind rule.

(4) $\psi[x] :\equiv (\boldsymbol{E}x\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x]) \rightarrow \phi[x := \boldsymbol{s}(\mathbf{0})]$
$\psi[x := \mathbf{0}] \equiv \boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \mathbf{0}]) \rightarrow \phi[x := \boldsymbol{s}(\mathbf{0})]$

1. $\boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$ so $\mathcal{T} \vdash \boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})$ (6a)
2. $\boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}), \boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \equiv \bot$ (6c)
3. $\boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := 0] \vdash \boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$ ($\wedge$1E)
4. $\boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := 0], \boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \vdash \bot$ (2, 3, $\rightarrow$E)
5. $\boldsymbol{E}0\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := 0] \vdash \phi[x := \boldsymbol{s}(\mathbf{0})]$ ($\bot$E, 4, weakening 1, $\rightarrow$E)

For $\psi[x := \boldsymbol{s}(x)] \equiv \boldsymbol{Es}(x)\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \boldsymbol{s}(x)] \rightarrow \phi[x := \boldsymbol{s}(\mathbf{0})]$
It holds that $\boldsymbol{Es}(x)\boldsymbol{s}(\mathbf{0}) =_{\mathcal{T}} \boldsymbol{E}x\mathbf{0}$. By using 4.27(3) replacing $\phi[x]$ by $\phi[x := \boldsymbol{s}(x)]$, we get that $\psi[x := \boldsymbol{s}(x)]$ is a theorem. We then apply Ind rule.

(5) $\psi[x] :\equiv (\boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}(x)) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x]) \rightarrow \phi[x := \boldsymbol{s}(\boldsymbol{P}(x))]$
For $\psi[x := \mathbf{0}] \equiv \boldsymbol{E}0\boldsymbol{s}(\boldsymbol{P}(\mathbf{0})) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \mathbf{0}]) \rightarrow \phi[x := \boldsymbol{s}(\boldsymbol{P}(\mathbf{0}))]$:
we know that $\boldsymbol{E}0\boldsymbol{s}(\boldsymbol{P}(\mathbf{0})) =_{\mathcal{T}} \boldsymbol{E}0\boldsymbol{s}(\mathbf{0})$ and $\phi[x := \boldsymbol{s}(\boldsymbol{P}(\mathbf{0}))] =_{\mathcal{T}} \phi[x := \boldsymbol{s}(\mathbf{0})]$, so we get a theorem from (4).
For $\psi[x := \boldsymbol{s}(x)] \equiv \boldsymbol{Es}(x)\boldsymbol{s}(\boldsymbol{P}(\boldsymbol{s}(x))) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \boldsymbol{s}(x)] \rightarrow \phi[x := \boldsymbol{s}(\boldsymbol{P}(\boldsymbol{s}(x)))]$:
-$\boldsymbol{Es}(x)\boldsymbol{s}(\boldsymbol{P}(\boldsymbol{s}(x))) =_{\mathcal{T}} \boldsymbol{E}x\boldsymbol{P}(\boldsymbol{s}(x)) =_{\mathcal{T}} \boldsymbol{E}xx$
-$\boldsymbol{s}(\boldsymbol{P}(\boldsymbol{s}(x))) =_{\mathcal{T}} \boldsymbol{s}(x)$
so $\psi[x := \boldsymbol{s}(x)]$ is a theorem by $\wedge$2E. We then apply Ind rule.

(6) To avoid presenting a very long and cumbersome proof, we will omit some steps explained now. Whenever we omit steps there will be three dots.

- If we can prove that $M =_{\mathbf{i}} \mathbf{0}$ and $M =_{\mathbf{i}} \boldsymbol{s}(x)$, then, by 5b, 6c, $\rightarrow$E we can prove $\bot$. The technique has already appeared above.
- We will make use of the following theorems of **NJ**, for ϕ, ψ, ϑ $\mathcal{T}$-formulas, without proving them (and for the first two without even mentioning them), since our definition of $\mathcal{T}$ is in fact an extension of **NJ** [7, lem. 6.2.1]. We will mention only the corresponding number on the left and omit the deductions. We will also omit some of the applications of $\rightarrow$E.
  - ◊ (6.2.1(1)) $\mathcal{T} \vdash \phi \vee \psi \leftrightarrow \psi \vee \phi$
  - ◊ (6.2.1(2)) $\mathcal{T} \vdash \phi \wedge \psi \leftrightarrow \psi \wedge \phi$
  - ◊ (6.2.1(5)) $\mathcal{T} \vdash \phi \vee (\psi \wedge \vartheta) \leftrightarrow (\phi \vee \psi) \wedge (\phi \vee \vartheta)$

◊ (6.2.1(6)) $\mathcal{T} \vdash \phi \wedge (\psi \vee \vartheta) \leftrightarrow (\phi \wedge \psi) \vee (\phi \wedge \vartheta)$

◊ (6.2.1(8)) $\mathcal{T} \vdash [\phi \to (\psi \to \vartheta)] \leftrightarrow [(\phi \wedge \psi) \to \vartheta]$

We will also use that

$$\mathcal{T} \vdash \boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \mathbf{0}] \to \phi[x]$$

which holds because

1. $\mathcal{T} \vdash \boldsymbol{E}\mathbf{0}\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \mathbf{0}] \to \phi[x := \mathbf{0}]$ (∧2E, →I)

   ...

2. $\underbrace{\boldsymbol{E}\boldsymbol{s}(x)\mathbf{0}}_{=_{\mathcal{T}}\boldsymbol{s}(\mathbf{0})} =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \mathbf{0}] \vdash \bot$ (5b, ∧1E)
3. $\mathcal{T} \vdash \boldsymbol{E}\boldsymbol{s}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := \mathbf{0}] \to \phi[x := \boldsymbol{s}(x)]$ (2, ⊥E)

From Ind rule we get the result, to which we will refer as theorem (*). (Reminder:↔ is an abbreviation as in PA)

1. $\mathcal{T} \vdash \boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}x) =_{\mathbf{i}} \mathbf{0}$ (4.27(2))
2. $\mathcal{T} \vdash \boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}$ (4.27(2))
3. $\mathcal{T} \vdash (\boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}x) =_{\mathbf{i}} \mathbf{0}) \wedge (\boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0})$ (1, 2 , ∧I)

   ...

4. $\mathcal{T} \vdash (\boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}x) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}x) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}) \equiv \phi_1 \vee \phi_2 \vee \phi_3 \vee \phi_4$ (several applications of 6.2.1(1, 2, 5, 6))

---

5. $\mathcal{T} \vdash \boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\mathbf{0}y =_{\mathbf{i}} \mathbf{0}$ (4.27(3))
6. $\mathcal{T} \vdash \boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\mathbf{0}y =_{\mathbf{i}} \mathbf{0} \to \underbrace{\boldsymbol{E}\mathbf{0}\mathbf{0}}_{=_{\mathcal{T}}\boldsymbol{E}\boldsymbol{P}(\mathbf{0})\boldsymbol{P}(\mathbf{0})} =_{\mathbf{i}} \mathbf{0}$ (4.27(3))
7. $\mathcal{T} \vdash \boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}\boldsymbol{P}(\mathbf{0})\boldsymbol{P}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$ (5, 6, →E)

   ...

8. $\mathcal{T} \vdash \boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}\boldsymbol{P}(x)\boldsymbol{P}(y) =_{\mathbf{i}} \mathbf{0}$ (multiple times theorem (*), 7, ∧1E, ∧2E, ∧I, →E)

---

9. $\mathcal{T} \vdash \boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\mathbf{0}y =_{\mathbf{i}} \mathbf{0}$ (4.27(3))
10. $\mathcal{T} \vdash \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\mathbf{0}y =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{E}\mathbf{0}\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}$ (4.27(5))
11. $\mathcal{T} \vdash \boldsymbol{E}\mathbf{0}\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})$ (6a)

    ...

12. $\mathcal{T} \vdash \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\mathbf{0}y =_{\mathbf{i}} \mathbf{0} \to \bot$ (10, 11, →E)
13. $\mathcal{T} \vdash \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\mathbf{0}y =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{E}\boldsymbol{P}(x)\boldsymbol{P}(y) =_{\mathbf{i}} \mathbf{0}$ (⊥E, 12)
14. $\mathcal{T} \vdash \boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}\boldsymbol{P}(x)\boldsymbol{P}(y) =_{\mathbf{i}} \mathbf{0}$ (9, 13, →E)

---

15. $\mathcal{T} \vdash \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}x) =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\boldsymbol{s}(\boldsymbol{P}x)y =_{\mathbf{i}} \mathbf{0}$ (4.27(5))
16. $\mathcal{T} \vdash \boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\boldsymbol{s}(\boldsymbol{P}x)y =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{E}\boldsymbol{s}(\boldsymbol{P}x)\mathbf{0} =_{\mathbf{i}} \mathbf{0}$ (4.27(3))
17. $\mathcal{T} \vdash \boldsymbol{E}\boldsymbol{s}(\boldsymbol{P}x)\mathbf{0} =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})$ (6a)

    ...
18. $\mathcal{T} \vdash \boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\boldsymbol{s}(\boldsymbol{P}x)y =_{\mathbf{i}} \mathbf{0} \to \perp$ ($\to$E, 16, 17, 5b)
19. $\mathcal{T} \vdash \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}x) =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \perp$ (15, 18, $\to$E)
20. $\mathcal{T} \vdash \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}x) =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}\boldsymbol{P}(x)\boldsymbol{P}(y) =_{\mathbf{i}} \mathbf{0}$ (19, $\perp$E)

    ---
21. $\mathcal{T} \vdash \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}x) =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\boldsymbol{s}(\boldsymbol{P}x)y =_{\mathbf{i}} \mathbf{0}$ (4.27(5))
22. $\mathcal{T} \vdash \boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}\boldsymbol{s}(\boldsymbol{P}x)y =_{\mathbf{i}} \mathbf{0} \to \underbrace{\boldsymbol{E}\boldsymbol{s}(\boldsymbol{P}x)\boldsymbol{s}(\boldsymbol{P}y)}_{=_{\mathcal{T}}\boldsymbol{E}(\boldsymbol{P}x)(\boldsymbol{P}y)} =_{\mathbf{i}} \mathbf{0}$ (4.27(5))
23. $\mathcal{T} \vdash \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}x) =_{\mathbf{i}} \mathbf{0} \wedge (\boldsymbol{E}y\boldsymbol{s}(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}(\boldsymbol{P}x)(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}$ (21, 22, $\to$E)

    ...
24. $\phi_1, \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{E}(\boldsymbol{P}x)(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}$ (6.2.1(8), 8, reverse $\to$I)
25. $\phi_2, \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{E}(\boldsymbol{P}x)(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}$ (6.2.1(8), 14, reverse $\to$I)
26. $\phi_3, \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{E}(\boldsymbol{P}x)(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}$ (6.2.1(8), 20 reverse $\to$I)
27. $\phi_4, \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{E}(\boldsymbol{P}x)(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}$ (6.2.1(8), 23, reverse $\to$I)
28. $\boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{E}(\boldsymbol{P}x)(\boldsymbol{P}y) =_{\mathbf{i}} \mathbf{0}$ (4, 24-27, $\vee$E)

(7) We will use induction on y. For $\psi[y] :\equiv \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(y))(\dot{-}xy) =_{\mathbf{i}} \mathbf{0}$ and

$$\begin{aligned}\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(\mathbf{0}))(\dot{-}x\mathbf{0}) &=_{\mathcal{T}} \boldsymbol{E}(\boldsymbol{P}(\dot{-}\boldsymbol{s}(x)\mathbf{0}))(x)\\ &=_{\mathcal{T}} \boldsymbol{E}(\boldsymbol{P}\boldsymbol{s}(x))(x)\\ &=_{\mathcal{T}} \boldsymbol{E}xx\end{aligned}$$

we get that $\psi[y := \mathbf{0}]$ is a theorem of $\mathcal{T}$ by 4.27(1).

We take as assumption (using reverse $\to$ I) $\psi[y]$. It holds that

$$\psi[y := \boldsymbol{s}(y)] \equiv \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(\boldsymbol{s}(y)))(\dot{-}x\boldsymbol{s}(y)) =_{\mathcal{T}} \boldsymbol{E}(\boldsymbol{P}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(y)))(\boldsymbol{P}(\dot{-}xy))$$

By setting $x := \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(y)$ and $y := \dot{-}xy$ in 4.27(6) we get that

$$\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(y))(\dot{-}xy) \to \boldsymbol{E}(\boldsymbol{P}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(y)))(\boldsymbol{P}(\dot{-}xy)) \equiv \psi[y] \to \psi[y := \boldsymbol{s}(y)]$$

is provable due to the above equivalence. Then we apply Ind rule.

(8) We will use induction on x. For $\psi[x] :\equiv \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$

-$\psi[x := \mathbf{0}] :\equiv \boldsymbol{E}(\dot{-}\boldsymbol{s}(\mathbf{0})\mathbf{0})\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$

Since $\dot{-}\boldsymbol{s}(\mathbf{0})\mathbf{0} =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$ $\psi[x := \mathbf{0}]$ is a theorem.

-We take as assumption $\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$. By 4.27(4) for $\phi[y] :\equiv \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0}$, $x := \dot{-}\boldsymbol{s}(\boldsymbol{s}(x))\boldsymbol{s}(x)$ and $y := \dot{-}\boldsymbol{s}(x)x$ we get that

$$(\boldsymbol{E}(\dot{-}\boldsymbol{s}(\boldsymbol{s}(x))\boldsymbol{s}(x))\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}(\dot{-}\boldsymbol{s}(\boldsymbol{s}(x))\boldsymbol{s}(x))(\dot{-}\boldsymbol{s}(x)x) =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{E}(\dot{-}\boldsymbol{s}(\boldsymbol{s}(x))\boldsymbol{s}(x))\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$$

is a theorem.

By 4.27(7) for $x := \boldsymbol{s}(x)$ and $y := x$ we get that $\boldsymbol{E}(\dot{-}\boldsymbol{s}(\boldsymbol{s}(x))\boldsymbol{s}(x))(\dot{-}\boldsymbol{s}(x)x) =_{\mathbf{i}} \mathbf{0}$ is a theorem. We can then eliminate

$$\boldsymbol{E}(\dot{-}\boldsymbol{s}(\boldsymbol{s}(x))\boldsymbol{s}(x))\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}(\dot{-}\boldsymbol{s}(\boldsymbol{s}(x))\boldsymbol{s}(x))(\dot{-}\boldsymbol{s}(x)x) =_{\mathbf{i}} \mathbf{0}$$

with the use of ($\wedge$I), the assumption and ($\to$E) and get the result. We then apply Ind rule.

(9) By setting $x := \dot{-}\boldsymbol{s}(x)x$ and $y := \boldsymbol{s}(\mathbf{0})$ in 4.27(6), using reverse ($\to$I) and 4.27(8), we can deduce

$$\mathcal{T} \vdash \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)(\boldsymbol{s}(\mathbf{0})) =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{E}(\boldsymbol{P} \dot{-} \boldsymbol{s}(x)x)\boldsymbol{P}\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$$

We get the result, because $\boldsymbol{P}(\dot{-}\boldsymbol{s}(x)x) =_{\mathcal{T}} \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x)$ and $\boldsymbol{P}(\boldsymbol{s}(\mathbf{0})) =_{\mathcal{T}} \mathbf{0}$.

(10) We will use induction on y for $\psi[y] :\equiv \boldsymbol{E}(\dot{-}x\boldsymbol{P}(y))\mathbf{0} =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathbf{i}} \mathbf{0}$

-$\psi[y := \mathbf{0}] \equiv \boldsymbol{E}(\dot{-}x\boldsymbol{P}(\mathbf{0}))\mathbf{0} =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{E}(\dot{-}x\mathbf{0})\mathbf{0} =_{\mathbf{i}} \mathbf{0}$

Because $\dot{-}x\boldsymbol{P}(\mathbf{0}) =_{\mathcal{T}} \dot{-}x\mathbf{0} =_{\mathcal{T}} x$, $\psi[y := \mathbf{0}]$ is a theorem of $\mathcal{T}$ by axiom rule and ($\to$I).

-$\psi[y := \boldsymbol{s}(y)] \equiv \boldsymbol{E}(\dot{-}x\boldsymbol{P}(\boldsymbol{s}(y)))\mathbf{0} =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{E}(\dot{-}x\boldsymbol{s}(y))\mathbf{0} =_{\mathbf{i}} \mathbf{0}$

Because $\dot{-}x\boldsymbol{P}(\boldsymbol{s}(y)) =_{\mathcal{T}} \dot{-}xy$ and $\boldsymbol{E}(\dot{-}x\boldsymbol{s}(y))\mathbf{0} =_{\mathcal{T}} \boldsymbol{E}(\boldsymbol{P}(\dot{-}xy))\boldsymbol{P}(\mathbf{0})$ from 4.27(6) by setting $x := \dot{-}xy$ and $y := \mathbf{0}$ we get that $\psi[y := \boldsymbol{s}(y)]$ is a theorem and thus $\psi[y] \to \psi[y := \boldsymbol{s}(y)]$ is a theorem. We can then apply Ind rule.

□

Remark 4.28. Our system has no axiom or rule that can be used in order to substitute terms that are equal with respect to "$=_{\mathbf{i}}$" in $\mathcal{T}$-formulas (supposing that this substitution is feasible according to our conventions). This is solved by the above lemma, for the cases needed for the following lemmas.

Lemma 4.29. If $\phi[x, y]$ is a quantifier-free $\mathcal{T}$-formula $x, y \in$ **int** and $\phi[x := \mathbf{0}, y]$, $\phi[x, y := \boldsymbol{s}(y)] \to \phi[x := \boldsymbol{s}(x), y]$ are theorems of $\mathcal{T}$, then $\phi[x, y := \mathbf{0}]$ is a theorem of $\mathcal{T}$ [16, lemma 7.B.4].

Proof. For the proof below we will use as in (4.27(6)) some helpful theorems of $\mathcal{T}$ to which we add

◊ (6.2.1(16)) $\mathcal{T} \vdash \bot \leftrightarrow \phi \wedge \neg\phi$

and the following, to which we will refer as helpful lemma (*):

1. $\Gamma, \phi \vdash \vartheta$ (hypothesis)

2. $\Gamma, \psi \vdash \vartheta$ (hypothesis)

3. $\Gamma, \phi \vee \psi, \phi \vdash \vartheta$ (1, weakening)

4. $\Gamma, \phi \vee \psi, \psi \vdash \vartheta$ (2, weakening)

5. $\Gamma, \phi \vee \psi \vdash \phi \vee \psi$ (axiom rule)

6. $\Gamma, \phi \vee \psi \vdash \vartheta$ (3, 4, 5, $\vee$E)

We set $\psi[x, y] :\equiv \boldsymbol{E}(\dot{-}x\boldsymbol{P}(y))\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := y, y := \dot{-}xy]$.

$\psi[x, y := \mathbf{0}] \equiv \boldsymbol{E}(\dot{-}x\boldsymbol{P}(\mathbf{0}))\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := \mathbf{0}, y := \dot{-}x\mathbf{0}] \equiv \boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := \mathbf{0}, y := x]$

1. $\mathcal{T} \vdash \phi[x := \mathbf{0}, y := x]$ (hypothesis)

2. $\phi[x := \mathbf{0}, y := x] \vdash \boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := \mathbf{0}, y := x] \equiv \psi[x, y := \mathbf{0}]$ (1,$\vee$2I)

3. $\mathcal{T} \vdash \psi[x, y := \mathbf{0}]$ (1, 2, $\to$E)

$$\begin{aligned}\psi[x, y] \to \psi[x, y := \boldsymbol{s}(y)] \equiv\; & (\boldsymbol{E}(\dot{-}x\boldsymbol{P}(y))\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := y, y := \dot{-}xy]) \to \\ & (\boldsymbol{E}(\dot{-}x\underbrace{\boldsymbol{P}(\boldsymbol{s}(y))}_{=_{\mathcal{T}} y})\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := \boldsymbol{s}(y), y := \dot{-}x\boldsymbol{s}(y)]) \\ \equiv\; & (\boldsymbol{E}(\dot{-}x\boldsymbol{P}(y))\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := y, y := \dot{-}xy]) \to \\ & (\boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := \boldsymbol{s}(y), y := \dot{-}x\boldsymbol{s}(y)])\end{aligned}$$

1. $\boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vdash \psi[x, y := \boldsymbol{s}(y)]$ ($\vee$1I)

2. $\boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathbf{i}} \mathbf{0}, \psi[x, y] \vdash \psi[x, y := \boldsymbol{s}(y)]$ (1, weakening)

---

3. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}xy)\boldsymbol{s}(\underbrace{\boldsymbol{P}(\dot{-}xy)}_{=_{\mathcal{T}} \dot{-}x\boldsymbol{s}(y)}) =_{\mathbf{i}} \mathbf{0} \wedge \phi[x := y, y := \dot{-}xy] \to \phi[x := y, y := \boldsymbol{s}(\dot{-}x\boldsymbol{s}(y))]$ (4.27(5))
   ...

4. $\boldsymbol{E}(\dot{-}xy)\boldsymbol{s}(\boldsymbol{P}(\dot{-}xy)) =_{\mathbf{i}} \mathbf{0} \vdash \phi[x := y, y := \dot{-}xy] \to \phi[x := y, y := \boldsymbol{s}(\dot{-}x\boldsymbol{s}(y))]$ (6.2.1(8), reverse $\to$I)

5. $\mathcal{T} \vdash \phi[x := y, y := \boldsymbol{s}(\dot{-}x\boldsymbol{s}(y))] \to \phi[x := \boldsymbol{s}(y), y := \dot{-}x\boldsymbol{s}(y)]$ (hypothesis)
   ...

6. $\boldsymbol{E}(\dot{-}xy)\boldsymbol{s}(\boldsymbol{P}(\dot{-}xy)) =_{\mathbf{i}} \mathbf{0} \vdash \phi[x := y, y := \dot{-}xy] \to \phi[x := \boldsymbol{s}(y), y := \dot{-}x\boldsymbol{s}(y)]$ (4, 5, $\to$E, reverse $\to$I)

7. $\boldsymbol{E}(\dot{-}xy)\boldsymbol{s}(\boldsymbol{P}(\dot{-}xy)) =_{\mathbf{i}} \mathbf{0}, \phi[x := y, y := \dot{-}xy] \vdash \boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := \boldsymbol{s}(y), y := \dot{-}x\boldsymbol{s}(y)] \equiv \psi[x, y := \boldsymbol{s}(y)]$ (6, $\vee$2I, reverse $\to$I)

---

8. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}x\boldsymbol{P}(y))\mathbf{0} =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathbf{i}} \mathbf{0}$ (4.27(10))

9. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}x\boldsymbol{P}(y))\mathbf{0} =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \phi[x := \boldsymbol{s}(y), y := \dot{-}x\boldsymbol{s}(y)] \equiv \psi[x, y := \boldsymbol{s}(y)]$ (8, reverse $\to$I, $\vee$1I)

10. $\boldsymbol{E}(\dot{-}xy)\boldsymbol{s}(\boldsymbol{P}(\dot{-}xy)) \vdash \boldsymbol{E}(\dot{-}x\boldsymbol{P}(y))\mathbf{0} =_{\mathsf{i}} \mathbf{0} \to \boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathsf{i}} \mathbf{0} \vee \phi[x := \boldsymbol{s}(y), y := \dot{-}x\boldsymbol{s}(y)] \equiv \psi[x, y := \boldsymbol{s}(y)]$ (9, weakening, $\to$I)
...

11. $\boldsymbol{E}(\dot{-}xy)\boldsymbol{s}(\boldsymbol{P}(\dot{-}xy)) \vdash \underbrace{\boldsymbol{E}(\dot{-}x\boldsymbol{P}(y))\mathbf{0} =_{\mathsf{i}} \mathbf{0} \vee \phi[y, \dot{-}xy]}_{\psi[x,y]} \to \psi[x, y := \boldsymbol{s}(y)]$ (7, 10, helpful lemma (*))

12. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}xy)\mathbf{0} =_{\mathsf{i}} \mathbf{0} \vee \boldsymbol{E}(\dot{-}xy)\boldsymbol{s}(\boldsymbol{P}(\dot{-}xy)) =_{\mathsf{i}} \mathbf{0}$ (4.27(2))

13. $\mathcal{T} \vdash \psi[x, y] \to \psi[x, y := \boldsymbol{s}(y)]$ (2, 11, 12, $\vee$E)

From Ind Rule $\psi[x, y]$ is a theorem.

1. $\mathcal{T} \vdash \psi[x := \boldsymbol{s}(x), y := \boldsymbol{s}(x)] \equiv \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)\underbrace{\boldsymbol{P}(\boldsymbol{s}(x))}_{=_{\mathcal{T}} x})\mathbf{0} =_{\mathsf{i}} \mathbf{0} \vee \phi[x := \boldsymbol{s}(x), y := \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x)]$

2. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0} \wedge \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0} \to \underbrace{\boldsymbol{E}\mathbf{0}\boldsymbol{s}(\mathbf{0})}_{=_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})} =_{\mathsf{i}} \mathbf{0}$ (4.27(3))
...

3. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0} \wedge \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0} \to \perp$ (2, 5b, $\to$E)
...

4. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0} \to (\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0} \to \perp)$ (6.2.1(1, 8))
...

5. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0}$ (4.27(8))

6. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0} \to \perp$ (4, 5, $\to$E)

7. $\mathcal{T} \vdash \Big[\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0} \vee \phi[x := \boldsymbol{s}(x), y := \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x)]\Big] \wedge \neg(\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0})$ (1, 6, $\wedge$I)
...

8. $\mathcal{T} \vdash \Big[(\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0}) \wedge \neg(\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0})\Big] \vee \Big[\phi[x := \boldsymbol{s}(x), y := \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x)] \wedge \neg(\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0})\Big]$ (6.2,1(1, 6))
...

9. $(\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0}) \wedge \neg(\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0}) \vdash \perp$ (6.2.1(16))

10. $\perp \vdash \phi[x := \boldsymbol{s}(x), y := \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x)]$ ($\perp$E)

11. $\phi[x := \boldsymbol{s}(x), y := \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x)] \wedge \neg(\boldsymbol{E}(\dot{-}\boldsymbol{s}(x)x)\mathbf{0} =_{\mathsf{i}} \mathbf{0}) \vdash \phi[x := \boldsymbol{s}(x), y := \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x)]$ ($\wedge$1E)

12. $\mathcal{T} \vdash \phi[x := \boldsymbol{s}(x), y := \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x)]$ (8, 10, 11, $\vee$E)

13. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x))\mathbf{0} =_{\mathsf{i}} \mathbf{0}$ (4.27(9))

14. $\mathcal{T} \vdash \boldsymbol{E}(\dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x))\mathbf{0} =_{\mathsf{i}} \mathbf{0} \wedge \phi[x := \boldsymbol{s}(x), y := \dot{-}\boldsymbol{s}(x)\boldsymbol{s}(x)] \to \phi[x := \boldsymbol{s}(x), y := \mathbf{0}]$ (4.27(3))
...

15. $\mathcal{T} \vdash \phi[x := \boldsymbol{s}(x), y := \mathbf{0}]$ ($\to$E, 12, 13, 14, 6.2.1(8), reverse $\to$I multiple times)

16. $\mathcal{T} \vdash \phi[x := \mathbf{0}, y := \mathbf{0}]$ (hypothesis)

17. $\mathcal{T} \vdash \phi[x, y := \mathbf{0}]$ (15, 16, Ind Rule)

$\square$

Lemma 4.30. If $\phi[x, z]$ is a quantifier-free $\mathcal{T}$-formula $x \in \mathbf{int}$, $z \in \sigma$ such that

$$\begin{aligned} &\mathcal{T} \vdash \phi[x := \mathbf{0}, z] \\ &\mathcal{T} \vdash \phi[x, z := Q[x, z]] \to \phi[x := \boldsymbol{s}(x), z], \qquad Q \in \sigma \end{aligned}$$

then $\mathcal{T} \vdash \phi[x, z]$ [16, lemma 7.B.5].

Proof. We set
$$\psi[x, y] :\equiv \phi[x, z := Fxzy]$$
for $x, y \in \mathbf{int}$, $F :\equiv \lambda a^{\mathbf{int}}.\lambda z^{\sigma}.\boldsymbol{R_\sigma} z(\lambda w^{\tau}.\lambda b^{\mathbf{int}}.Q[x, w])$, $Q \in \sigma$.

$\psi[x := \mathbf{0}, y] \equiv \phi[x := \mathbf{0}, z := F\mathbf{0}zy]$ is a theorem by hypothesis.

$\psi[x, y := \boldsymbol{s}(y)] \to \psi[x := \boldsymbol{s}(x), y] \equiv \phi[x, z := Fxz(\boldsymbol{s}(y))] \to \phi[x := \boldsymbol{s}(x), z := F(\boldsymbol{s}(x))zy]$

We have that

$$Fxz(\boldsymbol{s}(y)) =_{\mathcal{T}} \ldots =_{\mathcal{T}} Q[x, z := \boldsymbol{R_\sigma} z(\lambda w^{\tau}.\lambda b^{\mathbf{int}}.Q[x, w])y]$$

and
$$F(\boldsymbol{s}(x))zy =_{\mathcal{T}} \ldots =_{\mathcal{T}} \boldsymbol{R_\sigma} z(\lambda w^{\tau}.\lambda b^{\mathbf{int}}.Q[x, w])y$$

So $\psi[x, y := \boldsymbol{s}(y)] \to \psi[x := \boldsymbol{s}(x), y]$ is a theorem by hypothesis for $z := \boldsymbol{R_\sigma} z(\lambda w^{\tau}.\lambda b^{\mathbf{int}}.Q[x, w])y$.

By lemma 4.29 $\mathcal{T} \vdash \psi[x, y := \mathbf{0}] \equiv \phi[x, z := Fxz\mathbf{0}] \equiv \phi[x, z]$, where the last equality comes from $\boldsymbol{R_\sigma} z(\lambda w^{\tau}.\lambda b^{\mathbf{int}}.Q[x, w])\mathbf{0} =_{\mathcal{T}} z$

$\square$

Remark 4.31. Remember that when dealing with properties of naturals of the form $P(n, m)$ for $n, m \in \mathbb{N}$, we use double induction. That is, we prove that

- $P(0, 0)$ holds
- supposing that $P(k, l)$ holds for (a) $k < n$ (and all l) (b)$k = n$ but $l < m$

in order to obtain that $P(n, m)$ holds for all $n, m \in \mathbb{N}$ [15, sec. 4.2]. Since our induction rule is only for one variable, the previous two lemmas make double induction possible in a unnoticeable way so as to prove 4.34(a).

Definition 4.32. For $\mathcal{T}$-formulas ϕ, ψ, β we define inductively t_ϕ as follows:

- for ϕ being $M =_{\mathbf{i}} N$ (M, N of type **int**), t_ϕ is $\boldsymbol{E}MN$
- for ϕ being $\neg\psi$, t_ϕ is $\hat{\neg} t_\psi$
- for ϕ being $\beta \vee \psi$, t_ϕ is $\hat{\vee} t_\beta t_\psi$
- for ϕ being $\beta \wedge \psi$, t_ϕ is $\hat{\wedge} t_\beta t_\psi$
- for ϕ being $\beta \to \psi$, t_ϕ is $\hat{\to} t_\beta t_\psi$

Remark 4.33. The free variables of t_ϕ are exactly those of ϕ and if ϕ is a sentence, t_ϕ is a closed term.

Lemma 4.34. For M, N being *any* terms of type **int**

(a) $\mathcal{T} \vdash M =_{\mathbf{i}} N \leftrightarrow \boldsymbol{E}MN =_{\mathbf{i}} \mathbf{0}$

(b) $\mathcal{T} \vdash M \neq_{\mathbf{i}} \mathbf{0} \leftrightarrow (\hat{\neg} M) =_{\mathbf{i}} \mathbf{0}$

(c) $\mathcal{T} \vdash (M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}) \leftrightarrow \hat{\vee} MN =_{\mathbf{i}} \mathbf{0}$

(d) $\mathcal{T} \vdash (M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0}) \leftrightarrow \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0}$

(e) $\mathcal{T} \vdash (M =_{\mathbf{i}} \mathbf{0} \rightarrow N =_{\mathbf{i}} \mathbf{0}) \leftrightarrow \hat{\rightarrow} MN =_{\mathbf{i}} \mathbf{0}$

[16, lem. 7.B.6, 7.B.7]

Proof. (a) It suffices to prove that (1) $M =_{\mathbf{i}} N \vdash \boldsymbol{E}MN =_{\mathbf{i}} \mathbf{0}$ and (2) $\boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0} \vdash x =_{\mathbf{i}} y$. For (1) we have the following proof:

1. $M =_{\mathbf{i}} N \vdash \boldsymbol{E}MM =_{\mathbf{i}} \boldsymbol{E}MN$ (rule (6d) and $\boldsymbol{E}M \in$ **int** $\rightarrow$ **int**)
2. $\boldsymbol{E}MM =_{\mathcal{T}} \mathbf{0} \vdash \boldsymbol{E}MM =_{\mathbf{i}} \mathbf{0}$ (rule (6a))
3. $\boldsymbol{E}MM =_{\mathbf{i}} \boldsymbol{E}MN, \boldsymbol{E}MM =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{E}MN =_{\mathbf{i}} \mathbf{0}$ (rule (6c), 1, 2)

For (2) firstly we will use induction on y for $\psi[y] :\equiv \boldsymbol{E}\mathbf{0}y =_{\mathbf{i}} \mathbf{0} \rightarrow \mathbf{0} =_{\mathbf{i}} y$:
$\psi[y := \mathbf{0}] \equiv \boldsymbol{E}\mathbf{0}\mathbf{0} =_{\mathbf{i}} \mathbf{0} \rightarrow \mathbf{0} =_{\mathbf{i}} \mathbf{0}$ which is a theorem by axiom rule.

1. $\boldsymbol{E}\mathbf{0}\boldsymbol{s}(y) =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$ so $\mathcal{T} \vdash \boldsymbol{E}\mathbf{0}\boldsymbol{s}(y) =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})$ (rule 6a)
2. $\boldsymbol{E}\mathbf{0}\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})$, $\boldsymbol{E}\mathbf{0}\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$ (rule 6c)
   ...
3. $\boldsymbol{E}\mathbf{0}\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \vdash \perp$ (1, 2, 5b, $\rightarrow$E)
4. $\boldsymbol{E}\mathbf{0}\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \vdash \mathbf{0} =_{\mathbf{i}} \boldsymbol{s}(y)$ (3, $\perp$E)

By Ind rule $\psi[y]$ is a theorem.
We set $\phi[x, y] :\equiv \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0} \rightarrow x =_{\mathbf{i}} y$. We want to prove that

$$\mathcal{T} \vdash \vartheta[x, y] :\equiv \phi[x, y := \boldsymbol{P}(y)] \rightarrow \phi[x := \boldsymbol{s}(x), y]$$

i.e.,

$$\mathcal{T} \vdash [\boldsymbol{E}x\boldsymbol{P}(y) =_{\mathbf{i}} \mathbf{0} \rightarrow x =_{\mathbf{i}} \boldsymbol{P}(y)] \rightarrow [\boldsymbol{E}\boldsymbol{s}(x)y =_{\mathbf{i}} \mathbf{0} \rightarrow \boldsymbol{s}(x) =_{\mathbf{i}} y]$$

so that we can use 4.30 for $Q[x, y] :\equiv \boldsymbol{P}(y)$.
By induction on y for $\vartheta[x, y]$ we get:

$$\vartheta[x, y := \mathbf{0}] \equiv [\boldsymbol{E}x\boldsymbol{P}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \rightarrow x =_{\mathbf{i}} \boldsymbol{P}(\mathbf{0})] \rightarrow [\boldsymbol{E}\boldsymbol{s}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \rightarrow \boldsymbol{s}(x) =_{\mathbf{i}} \mathbf{0}]$$

and since $\boldsymbol{P}(\mathbf{0}) =_{\mathcal{T}} \mathbf{0}$

$$\mathcal{T} \vdash [\boldsymbol{E}x\mathbf{0} =_{\mathbf{i}} \mathbf{0} \rightarrow x =_{\mathbf{i}} \mathbf{0}] \rightarrow [\boldsymbol{E}\boldsymbol{s}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \rightarrow \boldsymbol{s}(x) =_{\mathbf{i}} \mathbf{0}]$$

Furthermore, we can also prove that $\mathcal{T} \vdash \neg(\boldsymbol{E}\boldsymbol{s}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0})$ as we have done before and obtain $\mathcal{T} \vdash \boldsymbol{E}\boldsymbol{s}(x)\mathbf{0} =_{\mathbf{i}} \mathbf{0} \rightarrow \boldsymbol{s}(x) =_{\mathbf{i}} \mathbf{0}$.

Hence, $\mathcal{T} \vdash \vartheta[x, y := \mathbf{0}]$.

We want to prove that

$$\vartheta[x, y := \boldsymbol{s}(y)] \equiv [\boldsymbol{E}x\boldsymbol{P}(\boldsymbol{s}(y)) =_{\mathbf{i}} \mathbf{0} \to x =_{\mathbf{i}} \boldsymbol{P}(\boldsymbol{s}(y))] \to [\boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{s}(x) =_{\mathbf{i}} \boldsymbol{s}(y)]$$

is a theorem but since $\boldsymbol{P}(\boldsymbol{s}(y)) =_{\mathcal{T}} y$ and

$$\boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathcal{T}} \boldsymbol{E}xy$$

it suffices to show that

$$\mathcal{T} \vdash \boldsymbol{E}xy =_{\mathbf{i}} \mathbf{0} \to x =_{\mathbf{i}} y, \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{s}(x) =_{\mathbf{i}} \boldsymbol{s}(y)$$

or

$$\mathcal{T} \vdash \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \to x =_{\mathbf{i}} y, \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{s}(x) =_{\mathbf{i}} \boldsymbol{s}(y)$$

1. $\mathcal{T} \vdash x =_{\mathbf{i}} y \to \boldsymbol{s}(z)[z := x] =_{\mathbf{i}} \boldsymbol{s}(z)[z := y]$ (rule 6a)
2. $\boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \to x =_{\mathbf{i}} y, \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \to x =_{\mathbf{i}} y$ (axiom rule)
3. $\boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \to x =_{\mathbf{i}} y, \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0}$ (axiom rule)
4. $\boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \to x =_{\mathbf{i}} y, \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \vdash x =_{\mathbf{i}} y$ (2, 3, $\to$ E)
5. $\boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \to x =_{\mathbf{i}} y, \boldsymbol{E}\boldsymbol{s}(x)\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{s}(x) =_{\mathbf{i}} \boldsymbol{s}(y)$ (1, 4, $\to$ E)

By lemma 4.30(2) we have that $\mathcal{T} \vdash \phi[x, y]$.

(b) For ($\to$):

1. $\mathcal{T} \vdash \boldsymbol{E}M\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vee \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}(M)) =_{\mathbf{i}} \mathbf{0}$ (lemma 4.27(2))
2. $\boldsymbol{E}M\mathbf{0} \vdash M =_{\mathbf{i}} \mathbf{0}$ (lemma 4.34(a))
3. $\boldsymbol{E}M\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vdash M =_{\mathbf{i}} \mathbf{0} \vee M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (2, $\vee$1I)
4. $\boldsymbol{E}M\boldsymbol{s}(\boldsymbol{P}(M)) =_{\mathbf{i}} \mathbf{0} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (lemma 4.34(a))
5. $\boldsymbol{E}M\boldsymbol{s}(\boldsymbol{P}(M)) =_{\mathbf{i}} \mathbf{0} \vdash M =_{\mathbf{i}} \mathbf{0} \vee M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (4, $\vee$2I)
6. $\mathcal{T} \vdash M =_{\mathbf{i}} \mathbf{0} \vee M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (1, 3, 5 $\vee$E)
7. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M)) \vdash \hat{-}M =_{\mathbf{i}} \hat{-}\boldsymbol{s}(\boldsymbol{P}(M))(=_{\mathcal{T}} \mathbf{0})$ (6a)
8. $\neg(M =_{\mathbf{i}} \mathbf{0}) \vdash M =_{\mathbf{i}} \mathbf{0} \vee M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (6, weakening)
9. $\neg(M =_{\mathbf{i}} \mathbf{0}), M =_{\mathbf{i}} \mathbf{0} \vdash \perp$ (axiom, rule, reverse $\to$I)
10. $\neg(M =_{\mathbf{i}} \mathbf{0}), M =_{\mathbf{i}} \mathbf{0} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (9, $\perp$E)
11. $\neg(M =_{\mathbf{i}} \mathbf{0}), M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M)) \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (axiom rule)
12. $\neg(M =_{\mathbf{i}} \mathbf{0}) \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (8, 10, 11, $\vee$E)
13. $\neg(M =_{\mathbf{i}} \mathbf{0}) \vdash \hat{-}M =_{\mathbf{i}} \mathbf{0}$ (12, 7, $\to$E)

For ($\leftarrow$):

1. $M =_{\mathbf{i}} \mathbf{0} \vdash \hat{-}M =_{\mathbf{i}} \hat{-}\mathbf{0}(=_{\mathcal{T}} \boldsymbol{s}(\mathbf{0}))$ (6a)

2. $\hat{\neg} M =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}), \hat{\neg} M =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}$ (6c)
   ...
3. $\hat{\neg} M =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}), \hat{\neg} M =_{\mathbf{i}} \mathbf{0} \vdash \bot$ (5b, reverse →I, →E, 2)
4. $\hat{\neg} M =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \vdash \neg(\hat{\neg} M =_{\mathbf{i}} \mathbf{0})$ (→I, 3)
5. $M =_{\mathbf{i}} \mathbf{0} \vdash \neg(\hat{\neg} M =_{\mathbf{i}} \mathbf{0})$ (1, 4, →E)
6. $M =_{\mathbf{i}} \mathbf{0}, \hat{\neg} M =_{\mathbf{i}} \mathbf{0} \vdash \bot$ (5, reverse →I)
7. $\hat{\neg} M =_{\mathbf{i}} \mathbf{0} \vdash \neg(M =_{\mathbf{i}} \mathbf{0})$ (6, →I)

(c) For (→):

1. $M =_{\mathbf{i}} \mathbf{0} \vdash \hat{\vee} MN =_{\mathbf{i}} \underbrace{\hat{\vee}\mathbf{0}N}_{=_{\mathcal{T}}\mathbf{0}}$ (6a, weakening, reverse →I, $\hat{\vee}\mathbf{0}\mathbf{0} =_{\mathcal{T}} \mathbf{0}$, $\hat{\vee}\mathbf{0}\boldsymbol{s}(y) =_{\mathcal{T}} \mathbf{0}$, Ind)
2. $N =_{\mathbf{i}} \mathbf{0} \vdash \hat{\vee} MN =_{\mathbf{i}} \underbrace{\hat{\vee}M\mathbf{0}}_{=_{\mathcal{T}}\mathbf{0}}$ (6a, reverse →I)
3. $M =_{\mathbf{i}} \mathbf{0}, M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0} \vdash \hat{\vee} MN =_{\mathbf{i}} \mathbf{0}$ (weakening, 1)
4. $N =_{\mathbf{i}} \mathbf{0}, M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0} \vdash \hat{\vee} MN =_{\mathbf{i}} \mathbf{0}$ (weakening, 2)
5. $M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0} \vdash M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ (axiom rule)
6. $M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0} \vdash \hat{\vee} MN =_{\mathbf{i}} \mathbf{0}$ (∨E, 3, 4, 5)

For (←): As in 4.27(6) we will make use of some theorems in [7, lem. 6.2.1] without proof, such as

- ◊ (6.2.1(1)) $\mathcal{T} \vdash \phi \vee \psi \leftrightarrow \psi \vee \phi$
- ◊ (6.2.1(2)) $\mathcal{T} \vdash \phi \wedge \psi \leftrightarrow \psi \wedge \phi$
- ◊ (6.2.1(5)) $\mathcal{T} \vdash \phi \vee (\psi \wedge \vartheta) \leftrightarrow (\phi \vee \psi) \wedge (\phi \vee \vartheta)$
- ◊ (6.2.1(6)) $\mathcal{T} \vdash \phi \wedge (\psi \vee \vartheta) \leftrightarrow (\phi \wedge \psi) \vee (\phi \wedge \vartheta)$
- ◊ (6.2.1(8)) $\mathcal{T} \vdash [\phi \to (\psi \to \vartheta)] \leftrightarrow [(\phi \wedge \psi) \to \vartheta]$
- ◊ (6.2.1(9)) $\mathcal{T} \vdash \phi \to (\psi \to \phi)$

Moreover, due to 4.34(a) we can use $M =_{\mathbf{i}} N$ and $\boldsymbol{E}MN =_{\mathbf{i}} \mathbf{0}$ interchangeably. This will happen below implicitly.

1. $M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0} \to [\hat{\vee} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}]$ (6.2.1(9))
   ...
2. $\mathcal{T} \vdash (\boldsymbol{E}M\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\mathbf{0} =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}M\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\boldsymbol{s}(\boldsymbol{P}N) =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}M\boldsymbol{s}(\boldsymbol{P}M) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\mathbf{0} =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}M\boldsymbol{s}(\boldsymbol{P}M) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\boldsymbol{s}(\boldsymbol{P}N) =_{\mathbf{i}} \mathbf{0}) \equiv \phi_1 \vee \phi_2 \vee \phi_3 \vee \phi_4$ (As in 4.27(6))
   ...
3. $\phi_1 \vdash M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ (∧1E or ∧2E and then ∨1I or ∨2I)
4. $\phi_2 \vdash M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ (∧1E and then ∨1I)
5. $\phi_3 \vdash M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ (∧2E and then ∨2I)
6. $\phi_1 \vdash \hat{\vee} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ (1, 3, →E)
7. $\phi_2 \vdash \hat{\vee} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ (1, 4, →E)
8. $\phi_3 \vdash \hat{\vee} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ (1, 5, →E)

9. $\mathcal{T} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \vdash N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N)$ ($\wedge$2E)
10. $\mathcal{T} \vdash \boldsymbol{E}N\boldsymbol{s}(\boldsymbol{P}N) =_{\mathbf{i}} \mathbf{0} \wedge (M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge \hat{\vee} MN =_{\mathbf{i}} \mathbf{0}) \to (M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge \underbrace{\hat{\vee} M\boldsymbol{s}(\boldsymbol{P}N)}_{=_{\mathcal{T}} M} =_{\mathbf{i}} \mathbf{0})$ (4.27(5))
...
11. $\mathcal{T} \vdash N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \wedge (M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge \hat{\vee} MN =_{\mathbf{i}} \mathbf{0}) \to \boldsymbol{s}(\boldsymbol{P}M) =_{\mathbf{i}} \mathbf{0}$ (10, $\wedge$2E, $\wedge$1E, 6c)
...
12. $\mathcal{T} \vdash N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \wedge (M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge \hat{\vee} MN =_{\mathbf{i}} \mathbf{0}) \to \bot$ (5b, 11, $\to$E)
13. $\mathcal{T} \vdash N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \wedge (M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge \hat{\vee} MN =_{\mathbf{i}} \mathbf{0}) \to M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ ($\bot$E, 12)

---

14. $\mathcal{T} \vdash \hat{\vee} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ (6, 7, 8, 13, $\vee$E, 6.2.1(8))

(d) For ($\to$):

1. $N =_{\mathbf{i}} \mathbf{0} \vdash \hat{\wedge} MN =_{\mathbf{i}} \underbrace{\hat{\wedge} M\mathbf{0}}_{=_{\mathcal{T}} M}$ (axiom 6a, reverse $\to$I)
2. $M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0} \vdash N =_{\mathbf{i}} \mathbf{0}$ ($\wedge$2E)
3. $M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0} \vdash \hat{\wedge} MN =_{\mathbf{i}} M$ ($\to$E, 1, 2)
4. $M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0} \vdash M =_{\mathbf{i}} \mathbf{0}$ ($\wedge$1E)
5. $M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0} \vdash \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0}$ (3, 4, 6c)

For ($\leftarrow$): As in 4.34(c) some steps will be omitted.

1. $\mathcal{T} \vdash M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0} \to [\hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0}]$ (6.2.1(9))

---

2. $\mathcal{T} \vdash M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}\boldsymbol{N}) \wedge \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}\boldsymbol{N}) \wedge \hat{\wedge} \mathbf{0}N =_{\mathbf{i}} \mathbf{0}$ (4.27(3))
3. $\mathcal{T} \vdash N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \wedge \hat{\wedge} \mathbf{0}N =_{\mathbf{i}} \mathbf{0} \to \underbrace{\hat{\wedge} \mathbf{0}\boldsymbol{s}(\boldsymbol{P}\boldsymbol{N})}_{=_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})} =_{\mathbf{i}} \mathbf{0}$ (4.27(5))
...
4. $\mathcal{T} \vdash N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \wedge \hat{\wedge} \mathbf{0}N =_{\mathbf{i}} \mathbf{0} \to \bot$ (3, 5b)
5. $\mathcal{T} \vdash M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}\boldsymbol{N}) \wedge \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to \bot$ (2, 4, $\to$E)
6. $\mathcal{T} \vdash M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}\boldsymbol{N}) \wedge \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0}$ (5, $\bot$E)

---

7. $\mathcal{T} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}\boldsymbol{M}) \wedge N =_{\mathbf{i}} \mathbf{0} \wedge \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to N =_{\mathbf{i}} \mathbf{0} \wedge \hat{\wedge} \boldsymbol{s}(\boldsymbol{P}\boldsymbol{M})N =_{\mathbf{i}} \mathbf{0}$ (4.27(5))
8. $\mathcal{T} \vdash N =_{\mathbf{i}} \mathbf{0} \wedge \hat{\wedge} \boldsymbol{s}(\boldsymbol{P}M)N =_{\mathbf{i}} \mathbf{0} \to \underbrace{\hat{\wedge} \boldsymbol{s}(\boldsymbol{P}\boldsymbol{M})\mathbf{0}}_{=_{\mathcal{T}} \boldsymbol{s}(\boldsymbol{P}\boldsymbol{M})} =_{\mathbf{i}} \mathbf{0}$ (4.27(3))
...
9. $\mathcal{T} \vdash N =_{\mathbf{i}} \mathbf{0} \wedge \hat{\wedge} \boldsymbol{s}(\boldsymbol{P}M)N =_{\mathbf{i}} \mathbf{0} \to \bot$ (8, 5b)
...
10. $\mathcal{T} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}\boldsymbol{M}) \wedge N =_{\mathbf{i}} \mathbf{0} \wedge \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0}$ (7, 8, 9, 5b, $\to$E, $\bot$E)

---

11. $\mathcal{T} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PM}) \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PN}) \wedge \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PN}) \wedge \hat{\wedge}\boldsymbol{s}(\boldsymbol{PM})N =_{\mathbf{i}} \mathbf{0}$ (4.27(5))
12. $\mathcal{T} \vdash N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PN}) \wedge \hat{\wedge}\boldsymbol{s}(\boldsymbol{PM})N =_{\mathbf{i}} \mathbf{0} \to \underbrace{\hat{\wedge}\boldsymbol{s}(\boldsymbol{PM})\boldsymbol{s}(\boldsymbol{PN})}_{=_{\mathcal{T}}\boldsymbol{s}(\mathbf{0})} =_{\mathbf{i}} \mathbf{0}$ (4.27(5))

   ...
13. $\mathcal{T} \vdash N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PN}) \wedge \hat{\wedge}\boldsymbol{s}(\boldsymbol{PM})N =_{\mathbf{i}} \mathbf{0} \to \perp$ (12, 5b)

   ...
14. $\mathcal{T} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PM}) \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PN}) \wedge \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0}$ (11, 13, $\to$E, $\perp$E)

   ...
15. $\mathcal{T} \vdash (\boldsymbol{E}M\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\mathbf{0} =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}M\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\boldsymbol{s}(\boldsymbol{PN}) =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}M\boldsymbol{s}(\boldsymbol{PM}) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\mathbf{0} =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}M\boldsymbol{s}(\boldsymbol{PM}) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\boldsymbol{s}(\boldsymbol{PN}) =_{\mathbf{i}} \mathbf{0}) \equiv \phi_1 \vee \phi_2 \vee \phi_3 \vee \phi_4$ (As in 4.27(6))
16. $\mathcal{T} \vdash \phi_2 \to \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0}$ (6, 6.2.1(8))
17. $\mathcal{T} \vdash \phi_3 \to \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0}$ (10, 6.2.1(8))
18. $\mathcal{T} \vdash \phi_4 \to \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0}$ (14, 6.2.1(8))
19. $\mathcal{T} \vdash \hat{\wedge} MN =_{\mathbf{i}} \mathbf{0} \to M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0}$ (1, 16, 17, 18, $\vee$E)

(e) For ($\to$):

We set $\phi[y] :\equiv (M =_{\mathbf{i}} \mathbf{0} \to y =_{\mathbf{i}} \mathbf{0}) \to \hat{\to} My =_{\mathbf{i}} \mathbf{0}$

1. $\mathbf{0} =_{\mathcal{T}} \mathbf{0}$ so $\mathcal{T} \vdash \mathbf{0} =_{\mathbf{i}} \mathbf{0}$ (6a)
2. $M =_{\mathbf{i}} \mathbf{0} \to \mathbf{0} =_{\mathbf{i}} \mathbf{0} \vdash \underbrace{\hat{\to} M\mathbf{0}}_{=_{\mathcal{T}}\mathbf{0}} =_{\mathbf{i}} \mathbf{0}$ (1, weakening)

Thus $\phi[y := \mathbf{0}]$ is a theorem.

1. $(M =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0}), M =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0}$ ($\to$E)
2. $(M =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0}), M =_{\mathbf{i}} \mathbf{0} \vdash \neg(\boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0})$ (5b, weakening)
3. $(M =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0}), M =_{\mathbf{i}} \mathbf{0} \vdash \perp$ (1, 2, $\to$E)
4. $(M =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0}) \vdash \neg(M =_{\mathbf{i}} \mathbf{0})$ (3, $\to$I)

   ...
5. $(M =_{\mathbf{i}} \mathbf{0} \to \boldsymbol{s}(y) =_{\mathbf{i}} \mathbf{0}) \vdash \underbrace{\hat{\neg} M}_{=_{\mathcal{T}}\hat{\to} M\boldsymbol{s}(y)} =_{\mathbf{i}} \mathbf{0}$ (4.34(b))

So $\phi[y := \boldsymbol{s}(y)]$ is a theorem and we have the desired result from Ind Rule.

For ($\leftarrow$): As in 4.34(c) we will omit some steps. We will further use [7, lem. 6.2.1(13)] and a proof from a succeeding lemma (4.41).

◊ (6.2.1(13)) $\mathcal{T} \vdash (\neg\phi \vee \psi) \to (\phi \to \psi)$

1. $\mathcal{T} \vdash M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PN}) \wedge \hat{\to} MN =_{\mathbf{i}} \mathbf{0} \to N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PN}) \wedge \hat{\to}\mathbf{0}N =_{\mathbf{i}} \mathbf{0}$ (4.27(3))
2. $\mathcal{T} \vdash N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{PN}) \wedge \hat{\to}\mathbf{0}N =_{\mathbf{i}} \mathbf{0} \to \underbrace{\hat{\to}\mathbf{0}\boldsymbol{s}(\boldsymbol{PN})}_{=_{\mathcal{T}}\hat{\neg}\mathbf{0}=_{\mathcal{T}}\boldsymbol{s}(\mathbf{0})} =_{\mathbf{i}} \mathbf{0}$ (4.27(5))

   ...

3. $\mathcal{T} \vdash M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \wedge \hat{\rightarrow} MN =_{\mathbf{i}} \mathbf{0} \rightarrow \perp$ (1, 2, →E, 5b)
4. $\mathcal{T} \vdash M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \wedge \hat{\rightarrow} MN =_{\mathbf{i}} \mathbf{0} \rightarrow (M =_{\mathbf{i}} \mathbf{0} \rightarrow N =_{\mathbf{i}} \mathbf{0})$ (3, ⊥E)

5. $\mathcal{T} \vdash (M =_{\mathbf{i}} \mathbf{0} \rightarrow N =_{\mathbf{i}} \mathbf{0}) \rightarrow [\hat{\rightarrow} MN =_{\mathbf{i}} \mathbf{0} \rightarrow (M =_{\mathbf{i}} \mathbf{0} \rightarrow N =_{\mathbf{i}} \mathbf{0})]$ (6.2.1(9))
6. $M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0} \vdash N =_{\mathbf{i}} \mathbf{0}$ (∧2E)
7. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge N =_{\mathbf{i}} \mathbf{0} \vdash N =_{\mathbf{i}} \mathbf{0}$ (∧2E)
8. $N =_{\mathbf{i}} \mathbf{0} \vdash \neg(M =_{\mathbf{i}} \mathbf{0}) \vee N =_{\mathbf{i}} \mathbf{0}$ (∨2E)
   ...
9. $\mathcal{T} \vdash \neg(M =_{\mathbf{i}} \mathbf{0}) \vee N =_{\mathbf{i}} \mathbf{0} \rightarrow (M =_{\mathbf{i}} \mathbf{0} \rightarrow N =_{\mathbf{i}} \mathbf{0})$ (6.2.1(13))

   ...
10. $\mathcal{T} \vdash M =_{\mathbf{i}} \mathbf{0} \wedge N =_{\mathbf{i}} \mathbf{0} \rightarrow [\hat{\rightarrow} MN =_{\mathbf{i}} \mathbf{0} \rightarrow (M =_{\mathbf{i}} \mathbf{0} \rightarrow N =_{\mathbf{i}} \mathbf{0})]$ (6, 8, 9, 5, →E multiple times)

    ...
11. $\mathcal{T} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge N =_{\mathbf{i}} \mathbf{0} \rightarrow [\hat{\rightarrow} MN =_{\mathbf{i}} \mathbf{0} \rightarrow (M =_{\mathbf{i}} \mathbf{0} \rightarrow N =_{\mathbf{i}} \mathbf{0})]$ (7, 8, 9, 5, →E multiple times)

    ...
12. $\mathcal{T} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \leftrightarrow \neg M =_{\mathbf{i}} \mathbf{0}$ (an in 4.41)
13. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M)$ (∧1E)
14. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \vdash \neg M =_{\mathbf{i}} \mathbf{0}$ (12, 13, →E)
15. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \vdash \neg M =_{\mathbf{i}} \mathbf{0} \vee N =_{\mathbf{i}} \mathbf{0}$ (14, ∨1I)
16. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}M) \wedge N =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}N) \vdash [\hat{\rightarrow} MN =_{\mathbf{i}} \mathbf{0} \rightarrow (M =_{\mathbf{i}} \mathbf{0} \rightarrow N =_{\mathbf{i}} \mathbf{0})]$ (15, 9, 5, →E multiple times)

    ...
17. $\mathcal{T} \vdash (\boldsymbol{E}M\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\mathbf{0} =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}M\mathbf{0} =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\boldsymbol{s}(\boldsymbol{P}N) =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}M\boldsymbol{s}(\boldsymbol{P}M) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\mathbf{0} =_{\mathbf{i}} \mathbf{0}) \vee (\boldsymbol{E}M\boldsymbol{s}(\boldsymbol{P}M) =_{\mathbf{i}} \mathbf{0} \wedge \boldsymbol{E}N\boldsymbol{s}(\boldsymbol{P}N) =_{\mathbf{i}} \mathbf{0}) \equiv \phi_1 \vee \phi_2 \vee \phi_3 \vee \phi_4$ (As in 4.27(6))

18. $\mathcal{T} \vdash \hat{\rightarrow} MN =_{\mathbf{i}} \mathbf{0} \rightarrow (M =_{\mathbf{i}} \mathbf{0} \rightarrow N =_{\mathbf{i}} \mathbf{0})$ (4, 10, 11, 16, 17, ∨E, 6.2.1)

□

Theorem 4.35. For every quantifier-free $\mathcal{T}$-formula ϕ it holds that $\mathcal{T} \vdash \phi \leftrightarrow t_\phi =_{\mathbf{i}} \mathbf{0}$. ($\phi$ and $t_\phi =_{\mathbf{i}} \mathbf{0}$ are called provably equivalent in $\mathcal{T}$, i.e., $\phi \vdash t_\phi =_{\mathbf{i}} \mathbf{0}$ and $t_\phi =_{\mathbf{i}} \mathbf{0} \vdash \phi$ in $\mathcal{T}$) [16, lem. 7.B.7]

Proof. By induction on ϕ.

If ϕ is $M =_{\mathbf{i}} N$ for M, N terms of type **int**, by lemma 4.34 it holds.

Suppose that we have ψ and β such that $\mathcal{T} \vdash \psi \leftrightarrow t_\psi =_{\mathbf{i}} \mathbf{0}$ and $\mathcal{T} \vdash \beta \leftrightarrow t_\beta =_{\mathbf{i}} \mathbf{0}$. Let ϕ be:

- $\psi \vee \beta$. We have that $\mathcal{T} \vdash t_\phi =_{\mathbf{i}} \mathbf{0} \leftrightarrow t_\psi =_{\mathbf{i}} \mathbf{0} \vee t_\beta =_{\mathbf{i}} \mathbf{0}$ from lemma 4.34:

1. $t_\phi =_\mathbf{i} \mathbf{0} \vdash t_\psi =_\mathbf{i} \mathbf{0} \lor t_\beta =_\mathbf{i} \mathbf{0}$ (reverse of →I)
2. $t_\psi =_\mathbf{i} \mathbf{0} \vdash \psi$ (inductive hypothesis)
3. $t_\beta =_\mathbf{i} \mathbf{0} \vdash \beta$ (inductive hypothesis)
4. $\psi \vdash \psi \lor \beta$ ($\lor$1I)
5. $\beta \vdash \psi \lor \beta$ ($\lor$2I)
6. $t_\beta =_\mathbf{i} \mathbf{0} \vdash \psi \lor \beta$ (3, 5, →E)
7. $t_\psi =_\mathbf{i} \mathbf{0} \vdash \psi \lor \beta$ (4, 2, →E)
8. $t_\phi =_\mathbf{i} \mathbf{0}, t_\beta =_\mathbf{i} \mathbf{0} \vdash \psi \lor \beta$ (weakening 6)
9. $t_\phi =_\mathbf{i} \mathbf{0}, t_\psi =_\mathbf{i} \mathbf{0} \vdash \psi \lor \beta$ (weakening 7)
10. $t_\phi =_\mathbf{i} \mathbf{0} \vdash \psi \lor \beta (\equiv \phi)$ (1, 8, 9, $\lor$E) ✓

For the reverse:

1. $t_\psi =_\mathbf{i} \mathbf{0} \lor t_\beta =_\mathbf{i} \mathbf{0} \vdash t_\phi =_\mathbf{i} \mathbf{0}$
2. $\psi \vdash t_\psi =_\mathbf{i} \mathbf{0}$ (inductive hypothesis)
3. $\beta \vdash t_\beta =_\mathbf{i} \mathbf{0}$ (inductive hypothesis)
4. $t_\psi =_\mathbf{i} \mathbf{0} \vdash t_\psi =_\mathbf{i} \mathbf{0} \lor t_\beta =_\mathbf{i} \mathbf{0}$ ($\lor$1I)
5. $t_\beta =_\mathbf{i} \mathbf{0} \vdash t_\psi =_\mathbf{i} \mathbf{0} \lor t_\beta =_\mathbf{i} \mathbf{0}$ ($\lor$2I)
6. $\psi \vdash t_\psi =_\mathbf{i} \mathbf{0} \lor t_\beta =_\mathbf{i} \mathbf{0}$ (2, 4, →E)
7. $\beta \vdash t_\psi =_\mathbf{i} \mathbf{0} \lor t_\beta =_\mathbf{i} \mathbf{0}$ (3, 5, →E)
8. $\phi, \psi \vdash t_\psi =_\mathbf{i} \mathbf{0} \lor t_\beta =_\mathbf{i} \mathbf{0}$ (weakening, 6)
9. $\phi, \beta \vdash t_\psi =_\mathbf{i} \mathbf{0} \lor t_\beta =_\mathbf{i} \mathbf{0}$ (weakening, 7)
10. $\phi \vdash \psi \lor \beta$ (Axiom rule)
11. $\phi \vdash t_\psi =_\mathbf{i} \mathbf{0} \lor t_\beta =_\mathbf{i} \mathbf{0}$ (8, 9, 10, $\lor$E)
12. $\phi \vdash t_\phi =_\mathbf{i} \mathbf{0}$ (1, 11, →E) ✓

- $\psi \land \beta$. We have that $\mathcal{T} \vdash t_\phi =_\mathbf{i} \mathbf{0} \leftrightarrow (t_\psi =_\mathbf{i} \mathbf{0} \land t_\beta =_\mathbf{i} \mathbf{0})$ from lemma 4.34.

1. $t_\phi =_\mathbf{i} \mathbf{0} \vdash t_\psi =_\mathbf{i} \mathbf{0} \land t_\beta =_\mathbf{i} \mathbf{0}$
2. $t_\psi =_\mathbf{i} \mathbf{0} \vdash \psi$ (inductive hypothesis)
3. $t_\beta =_\mathbf{i} \mathbf{0} \vdash \beta$ (inductive hypothesis)
4. $t_\psi =_\mathbf{i} \mathbf{0} \land t_\beta =_\mathbf{i} \mathbf{0} \vdash t_\psi =_\mathbf{i} \mathbf{0}$ ($\land$1E)
5. $t_\psi =_\mathbf{i} \mathbf{0} \land t_\beta =_\mathbf{i} \mathbf{0} \vdash t_\beta =_\mathbf{i} \mathbf{0}$ ($\land$2E)
6. $t_\psi =_\mathbf{i} \mathbf{0} \land t_\beta =_\mathbf{i} \mathbf{0} \vdash \psi$ (2, 4, →E)
7. $t_\psi =_\mathbf{i} \mathbf{0} \land t_\beta =_\mathbf{i} \mathbf{0} \vdash \beta$ (3, 5, →E)
8. $t_\psi =_\mathbf{i} \mathbf{0} \land t_\beta =_\mathbf{i} \mathbf{0} \vdash \psi \land \beta$ (6, 7, $\land$I)
9. $t_\phi =_\mathbf{i} \mathbf{0} \vdash \phi$ (1, 8, →E) ✓

For the reverse:

1. $t_\psi =_\mathbf{i} \mathbf{0} \land t_\beta =_\mathbf{i} \mathbf{0} \vdash t_\phi =_\mathbf{i} \mathbf{0}$
2. $\psi \vdash t_\psi =_\mathbf{i} \mathbf{0}$ (inductive hypothesis)

3. $\beta \vdash t_\beta =_{\mathbf{i}} \mathbf{0}$ (inductive hypothesis)
4. $\phi \vdash \psi$ ($\wedge$1 E)
5. $\phi \vdash \beta$ ($\wedge$2 E)
6. $\phi \vdash t_\psi =_{\mathbf{i}} \mathbf{0}$ (2, 4, $\to$E)
7. $\phi \vdash t_\beta =_{\mathbf{i}} \mathbf{0}$ (3, 5, $\to$E)
8. $\phi \vdash t_\psi =_{\mathbf{i}} \mathbf{0} \wedge t_\beta =_{\mathbf{i}} \mathbf{0}$ (6, 7, $\wedge$I)
9. $\phi \vdash t_\phi =_{\mathbf{i}} \mathbf{0}$ (1, 8, $\to$E) ✓

- $\psi \to \beta$. We have that $\mathcal{T} \vdash t_\phi =_{\mathbf{i}} \mathbf{0} \leftrightarrow (t_\psi =_{\mathbf{i}} \mathbf{0} \to t_\beta =_{\mathbf{i}} \mathbf{0})$ from lemma 4.34.

  1. $t_\phi =_{\mathbf{i}} \mathbf{0} \vdash t_\psi =_{\mathbf{i}} \mathbf{0} \to t_\beta =_{\mathbf{i}} \mathbf{0}$
  2. $t_\psi =_{\mathbf{i}} \mathbf{0} \vdash \psi$ (inductive hypothesis)
  3. $t_\beta =_{\mathbf{i}} \mathbf{0} \vdash \beta$ (inductive hypothesis)
  4. $t_\phi =_{\mathbf{i}} \mathbf{0}, t_\psi =_{\mathbf{i}} \mathbf{0} \vdash t_\beta =_{\mathbf{i}} \mathbf{0}$ (1, reverse of $\to$I)
  5. $t_\phi =_{\mathbf{i}} \mathbf{0}, t_\psi =_{\mathbf{i}} \mathbf{0} \vdash \beta$ (3, 4, $\to$E)
  6. $t_\phi =_{\mathbf{i}} \mathbf{0}, \psi \vdash t_\psi =_{\mathbf{i}} \mathbf{0}$ (inductive hypothesis $\psi \vdash t_\psi =_{\mathbf{i}} \mathbf{0}$, weakening)
  7. $\psi, t_\phi =_{\mathbf{i}} \mathbf{0} \vdash t_\psi =_{\mathbf{i}} \mathbf{0} \to \beta$ (5, $\to$I, weakening)
  8. $\psi, t_\phi =_{\mathbf{i}} \mathbf{0} \vdash \beta$ (6, 7, $\to$E)
  9. $t_\phi =_{\mathbf{i}} \mathbf{0} \vdash \psi \to \beta (\equiv \phi)$ (8, $\to$I) ✓

  For the reverse:

  1. $t_\psi =_{\mathbf{i}} \mathbf{0} \to t_\beta =_{\mathbf{i}} \mathbf{0} \vdash t_\phi =_{\mathbf{i}} \mathbf{0}$
  2. $\phi \vdash \psi \to \beta$ (axiom rule)
  3. $\phi, \psi \vdash \beta$ (2, reverse of $\to$I)
  4. $\beta \vdash t_\beta =_{\mathbf{i}} \mathbf{0}$ (inductive hypothesis)
  5. $\phi, \psi \vdash t_\beta =_{\mathbf{i}} \mathbf{0}$ (3, 4, $\to$E)
  6. $t_\psi =_{\mathbf{i}} \mathbf{0} \vdash \psi$ (inductive hypothesis)
  7. $\psi \vdash \phi \to t_\beta =_{\mathbf{i}} \mathbf{0}$ (5, $\to$I)
  8. $t_\psi =_{\mathbf{i}} \mathbf{0} \vdash \phi \to t_\beta =_{\mathbf{i}} \mathbf{0}$ (6, 7, $\to$E)
  9. $t_\psi =_{\mathbf{i}} \mathbf{0}, \phi \vdash t_\beta =_{\mathbf{i}} \mathbf{0}$ (8, reverse of $\to$I)
  10. $\phi \vdash t_\psi =_{\mathbf{i}} \mathbf{0} \to t_\beta =_{\mathbf{i}} \mathbf{0}$ (9, $\to$I)
  11. $\phi \vdash t_\phi =_{\mathbf{i}} \mathbf{0}$ (1, 10, $\to$E) ✓

- $\neg\psi \equiv \psi \to \bot$ We have that it holds, because:
  - $\mathcal{T} \vdash t_\psi \neq_{\mathbf{i}} \mathbf{0} \leftrightarrow (\hat{\neg} t_\psi) =_{\mathbf{i}} \mathbf{0}$ (def. 4.32, lem. 4.34)
  - $\neg(t_\psi =_{\mathbf{i}} \mathbf{0}) \equiv (t_\psi =_{\mathbf{i}} \mathbf{0}) \to \bot$
  - From the previous proof $\mathcal{T} \vdash t_\phi =_{\mathbf{i}} \mathbf{0} \leftrightarrow (t_\psi =_{\mathbf{i}} \mathbf{0} \to t_\bot =_{\mathbf{i}} \mathbf{0})$ and $t_\bot \equiv \boldsymbol{E0s}(\mathbf{0}) =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$ ✓.

□

Remark 4.36. One can find a different way of proving that all terms in $\mathcal{T}$ have a provably equivalent atomic formula in [44, ch. 4], where the term $\boldsymbol{E}$ is defined using the almost minus term $\dot{-}$ and the proofs are justified because of [46, sec. 1.6.14, 1.6.9] in which the hatted terms $\hat{\neg}$, $\hat{\rightarrow}$, $\hat{\vee}$, $\hat{\wedge}$ we introduced above are linked to primitive recursive functions / truth functions. However, we provided syntactic proofs.

> *"This equation calculus is (with trivial modifications) an extension of primitive recursive arithmetic in the sense of [20]. ... Using the induction rule (IR), we can prove the usual properties of those[9] functions as in recursive arithmetic. .... It was Gödel's intention that formulas of $\mathcal{T}$ should be built up by means of propositional truth-functions. ... In this way, all tautologies of classical propositional calculus become provable in $\mathcal{T}'$[10]" [44, pp. 116–117]*

In [24] the above method is (almost) explicitly presented; all is smartly avoided by staying inside an equation calculus and modifying the definition of the interpretation accordingly. This is possible because the author replaces connectives by appropriate (hatted) combinator terms, which work similarly to regular connectives without extending the calculus, and links every primitive recursive function with a closed combinator term. The proof is admittedly more elegant this way, although not much shorter.

Remark 4.37. Ex falso rule ($\perp$E) is actually omissible in our definition of $\mathcal{T}$ as we now prove. This fact is important for the proof of the next theorem and is a consequence of the definition of $\perp$ which isn't an inherent distinct symbol, but a prime formula (see also [15, pp. 37–38, 77–78]).

Proof. 1. $\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0} \vdash \mathbf{0} =_{\mathsf{i}} \mathbf{0}$ (axiom rule, weakening)

2. $\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0}, x =_{\mathsf{i}} \mathbf{0} \vdash \boldsymbol{s}(x) =_{\mathsf{i}} \boldsymbol{s}(\mathbf{0})$ (5a)

3. $\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0}, x =_{\mathsf{i}} \mathbf{0} \vdash \boldsymbol{s}(x) =_{\mathsf{i}} \mathbf{0}$ (1, 2, 6c)

By Ind rule $\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0} \vdash M =_{\mathsf{i}} \mathbf{0}$ for every $M \in$ **int**. From that and 6c we also get that $\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0} \vdash M =_{\mathsf{i}} N$ for all $M, N \in$ **int**.

If $\Gamma = \{\boldsymbol{s}(\mathbf{0}) =_{\mathsf{i}} \mathbf{0}\}$ and ψ, ϑ are $\mathcal{T}$-formulas such that $\Gamma \vdash \psi$, $\Gamma \vdash \vartheta$, then if ϕ is:

- $\psi \rightarrow \vartheta$ by weakening the hypothesis for ϑ and ($\rightarrow$I) $\Gamma \vdash \psi \rightarrow \vartheta$
- $\psi \wedge \vartheta$ by ($\wedge$I), $\Gamma \vdash \psi \wedge \vartheta$
- $\psi \vee \vartheta$ by ($\vee$1I) or ($\vee$1I), $\Gamma \vdash \psi \vee \vartheta$

If $\Gamma \vdash \phi[x := \mathbf{0}]$, $\Gamma \vdash \phi[x := w] \rightarrow \phi[x := \boldsymbol{s}(w)]$ then by Ind rule $\Gamma \vdash \phi[x := W]$. ∎

Theorem 4.38. For any $\mathcal{T}$-sentence ϕ, $\mathcal{T} \vdash t_\phi =_{\mathsf{i}} \mathbf{0}$ iff $t_\phi =_{\mathcal{T}} \mathbf{0}$.

Proof. ($\Leftarrow$) is given by 6a rule. For ($\Rightarrow$) suppose that $M, N, Q \in$ **int** are closed terms. By 4.19 there exist unique numbers $m, n, k \in \mathbb{N}$ such that $M =_{\mathcal{T}} \boldsymbol{s}^m(\mathbf{0})$, $N =_{\mathcal{T}} \boldsymbol{s}^n(\mathbf{0})$ and $Q =_{\mathcal{T}} \boldsymbol{s}^k(\mathbf{0})$. We will use induction on the proof of $\mathcal{T} \vdash t_\phi =_{\mathsf{i}} \mathbf{0}$.

Basis

[9] Functions that can be defined with the use of closed terms $\hat{+}, \hat{\cdot}, \hat{\dot{-}}$ similar to the ones we defined.

[10] The equation calculus variant of our $\mathcal{T}$.

- If ϕ is axiom 5a $t_\phi \equiv \hat{\rightarrow}(\boldsymbol{E}MN)(\boldsymbol{E}Q[z := M]Q[z := N]))$.
(Case a) $Q[z := M]$ and $Q[z := N]$ coincide. Then $\boldsymbol{E}Q[z := M]Q[z := N] =_{\mathcal{T}} \mathbf{0}$ and $t_\phi =_{\mathcal{T}} \mathbf{0}$. (Case b) $Q[z := M]$ and $Q[z := N]$ don't coincide, so $\boldsymbol{E}MN =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$ because M, N are different and $t_\phi =_{\mathcal{T}} \mathbf{0}$.

- If ϕ is axiom 5b $t_\phi \equiv \hat{\neg}\boldsymbol{s}(M) =_{\mathcal{T}} \mathbf{0}$.

- If ϕ is axiom 5c $t_\phi \equiv \hat{\rightarrow}(\underbrace{\boldsymbol{E}\boldsymbol{s}(M)\boldsymbol{s}(N)}_{=_{\mathcal{T}}\boldsymbol{E}MN})(\boldsymbol{E}MN))$.
(Case a) $m \neq n$. Then $\boldsymbol{E}MN =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$ and $t_\phi =_{\mathcal{T}} \mathbf{0}$. (Case b) $m = n$ so $\boldsymbol{E}MN =_{\mathcal{T}} \mathbf{0}$ and $t_\phi =_{\mathcal{T}} \mathbf{0}$.

- If ϕ is axiom 5d $t_\phi \equiv \hat{\rightarrow}(\hat{\wedge}(\boldsymbol{E}MN)(\boldsymbol{E}MQ))(\boldsymbol{E}MQ)$.
(Case a) $m = k$. Then $\boldsymbol{E}MQ =_{\mathcal{T}} \mathbf{0}$ and $t_\phi =_{\mathcal{T}} \mathbf{0}$. (Case b1) $m \neq k$ and $m = n$ so $\boldsymbol{E}MQ =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$ and $t_\phi =_{\mathcal{T}} \mathbf{0}$. (Case b2) $m \neq k$ and $m \neq n$ so $\hat{\wedge}(\boldsymbol{E}MN)(\boldsymbol{E}MQ) =_{\mathcal{T}} \boldsymbol{s}^\lambda(\mathbf{0})$ for some $\lambda \geqslant 1$ and $t_\phi =_{\mathcal{T}} \mathbf{0}$.

Inductive step
For rules 6a-6d the result comes from the fact that $=_{\mathcal{T}}$ is an equivalence relation that respects these rules.

Assume that ϕ, ψ, ϑ are $\mathcal{T}$-sentences such that the last rule of the proof is:

- $\Gamma, \phi \vdash \phi$ and the hypothesis holds for ϕ ✓.

- $\Gamma \vdash \phi \rightarrow \psi$ coming from $\Gamma, \phi \vdash \psi$, where the hypothesis holds for ϕ, ψ. But $\hat{\rightarrow} t_\phi t_\psi =_{\mathcal{T}} \hat{\rightarrow}\mathbf{0}\mathbf{0} =_{\mathcal{T}} \mathbf{0}$ ✓.

- ($\rightarrow$E) If $\Gamma \vdash \phi \rightarrow \psi$ and $\Gamma \vdash \phi$, then $\Gamma \vdash \psi$ and hypothesis holds for $\phi \rightarrow \psi, \phi$. If $t_\psi =_{\mathcal{T}} \boldsymbol{s}^k(\mathbf{0})$, for $k > 0$ then $\hat{\rightarrow} t_\phi t_\psi =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$ which contradicts the hypothesis. So $t_\psi =_{\mathcal{T}} \mathbf{0}$ ✓.

- ($\wedge$I) If $\Gamma \vdash \phi$ and $\Gamma \vdash \psi$, then $\Gamma \vdash \phi \wedge \psi$ and the hypothesis holds for ϕ, ψ. $\hat{\wedge} t_\phi t_\psi =_{\mathcal{T}} \hat{\wedge}\mathbf{0}\mathbf{0} =_{\mathcal{T}} \mathbf{0}$ ✓.

- ($\wedge$1E) or ($\wedge$2E) If $\Gamma \vdash \phi \wedge \psi$, then $\Gamma \vdash \phi$ or $\Gamma \vdash \psi$ where the hypothesis holds for $\phi \wedge \psi$. But if $t_\phi =_{\mathcal{T}} \boldsymbol{s}^{k+1}(\mathbf{0})$ or $t_\psi =_{\mathcal{T}} \boldsymbol{s}^{k+1}(\mathbf{0})$ for some $k \in \mathbb{N}$ then $\hat{\wedge} t_\phi t_\psi \neq_{\mathcal{T}} \mathbf{0}$. So $t_\phi =_{\mathcal{T}} t_\psi =_{\mathcal{T}} \mathbf{0}$ ✓.

- ($\vee$1I) or ($\vee$2I) If $\Gamma \vdash \phi$ or $\Gamma \vdash \psi$, then $\Gamma \vdash \phi \vee \psi$ where the hypothesis holds for ϕ in the first case and for ψ in the second. $\hat{\vee} t_\phi t_\psi =_{\mathcal{T}} \mathbf{0}$ if at least one of t_ϕ, t_ψ is equivalent to $\mathbf{0}$ ✓.

- If $\{\Gamma, \phi\} \vdash \vartheta$, $\{\Gamma, \psi\} \vdash \vartheta$ and $\Gamma \vdash \phi \vee \psi$, then $\Gamma \vdash \vartheta$ ($\vee$E) where the hypothesis holds for $\phi \vee \psi, \phi \rightarrow \vartheta, \psi \rightarrow \vartheta$. So at least one of t_ϕ, t_ψ is equivalent to $\mathbf{0}$. But if $t_\vartheta \neq_{\mathcal{T}} \mathbf{0}$ we have a contradiction in at least one case ✓.

- ($\perp$E) is omissible in the definition of $\mathcal{T}$.

- (Ind) $\{\phi[x := \mathbf{0}], \phi[x := V] \rightarrow \phi[x := \boldsymbol{s}(V)]\} \vdash \phi[x := W]$ for $V, W \in$ **int** closed. By 4.19 there exist unique numbers $n, k \in \mathbb{N}$ such that $V =_{\mathcal{T}} \boldsymbol{s}^n(\mathbf{0})$ and $W =_{\mathcal{T}} \boldsymbol{s}^k(\mathbf{0})$. The hypothesis holds for $\phi[x := \mathbf{0}], \phi[x := V] \rightarrow \phi[x := \boldsymbol{s}(V)]$. If $k = 0$ it is immediate. If $k > 0$ then we notice that by substituting gradually $m = 0, 1, ..., k-1$ in $\hat{\rightarrow} t_\phi[x := \boldsymbol{s}^m(\mathbf{0})] t_\phi[x := \boldsymbol{s}^{m+1}(\mathbf{0})] =_{\mathcal{T}} \mathbf{0}$, we obtain that $t_\phi[x := \boldsymbol{s}^k(\mathbf{0})] =_{\mathcal{T}} \mathbf{0}$ ✓.

$\square$

Remark 4.39.
- If ex falso wasn't omissible, since we haven't yet proved that $\mathcal{T}$ is consistent, we are neither cognizant that $\perp$ isn't provable; which would make the previous proof not true.
- For a random ϕ that is not closed we don't generally have the above result.

4.8.1 Consistency of system $\mathcal{T}$

In [24] and in [16] we have a definition of "truth" for closed quantifier-free formulas, which "*amounts to define truth of closed prime formulas $M =_{\mathbf{i}} N$*". The essence of this definition is that a closed prime formula is "true" if the normal forms of the two sides (which are reduced to numerals due to theorem 4.19 since they have type **int**) coincide.

We use this definition in order to instantiate every provable formula by all its possible closed instantiations; that is, we substitute all free variables with all possible closed terms. Here is where the strong normalization theorem enters. Yet, these steps are not that obvious in our proof.

We now present the final theorem which completes the proof of consistency for $\mathcal{T}$.

Theorem 4.40. System $\mathcal{T}$ is consistent. More specifically, under the assumption that all terms in $\mathcal{T}$ have a unique normal form, the consistency is provable in PRA.

Proof. Lets assume for the sake of contradiction that $\mathbf{0} =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})$ is provable in $\mathcal{T}$. By theorems 4.38 and 4.35 since both $\mathbf{0}$ and $\boldsymbol{s}(\mathbf{0})$ are in closed normal form, we have that $\mathbf{0} =_{\mathcal{T}} \boldsymbol{s}(\mathbf{0})$, which doesn't hold because they are distinct normal terms. Hence, $\mathbf{0} =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})$ isn't provable in $\mathcal{T}$. $\square$

Lastly, the following lemma is useful for some cases of the soundness theorem, as it implies decidability of prime formulas. Our proof requires the consistency of the system.

Lemma 4.41. For any $\mathcal{T}$-formula ϕ we can prove that $\mathcal{T} \vdash t_\phi =_{\mathbf{i}} \mathbf{0} \lor t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi))$ [46, sec. 1.6.11] and $\mathcal{T} \vdash \neg(t_\phi =_{\mathbf{i}} \mathbf{0}) \leftrightarrow t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi))$.

Proof. 1. $\mathcal{T} \vdash \boldsymbol{E}t_\phi\mathbf{0} =_{\mathbf{i}} \mathbf{0} \lor \boldsymbol{E}x\boldsymbol{s}(\boldsymbol{P}(t_\phi)) =_{\mathbf{i}} \mathbf{0}$ (lemma 4.27(2))

2. $\boldsymbol{E}t_\phi\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vdash t_\phi =_{\mathbf{i}} \mathbf{0}$ (lemma 4.34(a))

3. $\boldsymbol{E}t_\phi\mathbf{0} =_{\mathbf{i}} \mathbf{0} \vdash t_\phi =_{\mathbf{i}} \mathbf{0} \lor t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi))$ (2, $\lor$1I)

4. $\boldsymbol{E}t_\phi\boldsymbol{s}(\boldsymbol{P}(t_\phi)) =_{\mathbf{i}} \mathbf{0} \vdash t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi))$ (lemma 4.34(a))

5. $\boldsymbol{E}t_\phi\boldsymbol{s}(\boldsymbol{P}(t_\phi)) =_{\mathbf{i}} \mathbf{0} \vdash t_\phi =_{\mathbf{i}} \mathbf{0} \lor t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi))$ (4, $\lor$2I)

6. $\mathcal{T} \vdash t_\phi =_{\mathbf{i}} \mathbf{0} \lor t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi))$ (1, 3, 5 $\lor$E)

We can also prove that:

(1) we can't prove both $t_\phi =_{\mathbf{i}} \mathbf{0}$ and $t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi))$ for a formula ϕ, i.e $\mathcal{T} \vdash \neg(t_\phi =_{\mathbf{i}} \mathbf{0} \land t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi)))$.

(2) $\mathcal{T} \vdash \neg(t_\phi =_{\mathbf{i}} \mathbf{0}) \leftrightarrow t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi))$, i.e., we can substitute $t_\phi =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(t_\phi))$ for $\neg(t_\phi =_{\mathbf{i}} \mathbf{0})$ and vice versa.

Suppose that $M :\equiv t_\phi$.

For (1):

1. $M =_{\mathbf{i}} \mathbf{0} \wedge M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M)) \vdash M =_{\mathbf{i}} \mathbf{0}$ ($\wedge$1E)

2. $M =_{\mathbf{i}} \mathbf{0} \wedge M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M)) \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ ($\wedge$2E)

3. $M =_{\mathbf{i}} \mathbf{0} \wedge M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M)) \vdash \boldsymbol{s}(\boldsymbol{P}(M)) =_{\mathbf{i}} \mathbf{0}$ (1, 2, 6c)
   ...

4. $\mathcal{T} \vdash \neg(M =_{\mathbf{i}} \mathbf{0} \wedge M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M)))$ (3, 5b)

For (2):

1. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))), M =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{s}(\boldsymbol{P}(M))) =_{\mathbf{i}} \mathbf{0}$ (6c)

2. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))), M =_{\mathbf{i}} \mathbf{0} \vdash \boldsymbol{s}(\boldsymbol{P}(M))) =_{\mathbf{i}} \mathbf{0} \rightarrow \perp$ (5b, weakening)

3. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))), M =_{\mathbf{i}} \mathbf{0} \vdash \perp$ (1,2 $\rightarrow$E)

4. $M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))) \vdash M =_{\mathbf{i}} \mathbf{0} \rightarrow \perp \equiv \neg(M =_{\mathbf{i}} \mathbf{0})$ (3,$\rightarrow$I)

and

1. $\neg(M =_{\mathbf{i}} \mathbf{0}), M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M)) \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M)))$ (Axiom rule)

2. $\neg(M =_{\mathbf{i}} \mathbf{0}), M =_{\mathbf{i}} \mathbf{0} \vdash \perp$ (Axiom rule, reverse $\rightarrow$I)

3. $\neg(M =_{\mathbf{i}} \mathbf{0}), M =_{\mathbf{i}} \mathbf{0} \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M)))$ (2, $\perp$E)

4. $\neg(M =_{\mathbf{i}} \mathbf{0}) \vdash M =_{\mathbf{i}} \mathbf{0} \vee M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (as above, weakening)

5. $\neg(M =_{\mathbf{i}} \mathbf{0}) \vdash M =_{\mathbf{i}} \boldsymbol{s}(\boldsymbol{P}(M))$ (1, 3, 4, $\vee$E)

$\square$

4.9 Soundness theorem

We need a result that would guarantee that provability in HA yields provability in $\mathcal{T}$. If this is achieved, then the consistency of system $\mathcal{T}$ is transferred to HA.

Theorem 4.42. Assume that HA $\vdash \phi$ and let $\phi^D = \exists x^\sigma \forall y^\tau \phi_D[x^\sigma, y^\tau]$ be the functional interpretation of ϕ; then it is possible to find a term U of type σ not containing y free such that:

$$\mathcal{T} \vdash \phi_D[U, y^\tau]$$

[16, theorem 7.B.12]

Proof. We apply induction on the length of the proof of HA $\vdash \phi$. [46, theor. 3.5.4, sec. 1.7.5]

1. *Equality axioms*
   It suffices to show that the D-interpretations of the equality axioms are provable in $\mathcal{T}$. Afterwards, we have what is asked for any term U of type **int** not containing y (and z for some) free. The checkmark means that $\mathcal{T} \vdash \phi_D[U, y^\tau]$

   (a) $(x = x)^D : x^{\textbf{int}} =_{\mathbf{i}} x^{\textbf{int}}$. We have that $x^{\textbf{int}} =_{\mathcal{T}} x^{\textbf{int}}$. We then apply rule 6a. ✓

(b) $(x = y \to sx = sy)^D : x^{\mathbf{int}} =_{\mathbf{i}} y^{\mathbf{int}} \to \boldsymbol{s}(x) =_{\mathbf{i}} \boldsymbol{s}(y)$, which is provable by axiom 5a. ✓

(c) $(x = y \to (x = z \to y = z))^D : x^{\mathbf{int}} =_{\mathbf{i}} y^{\mathbf{int}} \to (x^{\mathbf{int}} =_{\mathbf{i}} z^{\mathbf{int}} \to y^{\mathbf{int}} =_{\mathbf{i}} z^{\mathbf{int}})$. But $x^{\mathbf{int}} =_{\mathbf{i}} y^{\mathbf{int}}, x^{\mathbf{int}} =_{\mathbf{i}} z^{\mathbf{int}} \vdash y^{\mathbf{int}} =_{\mathbf{i}} z^{\mathbf{int}}$ by rules 6c, 6b. ✓

2. *Definition axioms* (We will suppress variable types in D-interpretation. Besides they are all of type **int**)

(a) $(x + 0 = x)^D : \oplus x\mathbf{0} =_{\mathbf{i}} x$. We have that $\oplus x\mathbf{0} =_{\mathcal{T}} x$ by definition 4.22 ✓.

(b) $(x + sy = s(x + y))^D : \oplus x\boldsymbol{s}(y) =_{\mathbf{i}} \boldsymbol{s}(\oplus xy)$ We have that $\oplus x\boldsymbol{s}(y) =_{\mathcal{T}} \boldsymbol{s}(\oplus xy)$ ✓.

(c) $(x \cdot 0 = 0)^D : \bullet x\mathbf{0} =_{\mathbf{i}} \mathbf{0}$, but $\bullet x\mathbf{0} =_{\mathcal{T}} \mathbf{0}$ ✓.

(d) $(x \cdot sy = x \cdot y + x)^D : \bullet x\boldsymbol{s}(y) =_{\mathbf{i}} \oplus(\bullet xy)x$, but $\bullet x\boldsymbol{s}(y) =_{\mathcal{T}} \oplus(\bullet xy)x$. ✓

(e) $(\neg(sx = 0))^D : \neg(\boldsymbol{s}(x) =_{\mathbf{i}} \mathbf{0})$. $\hat{\neg}(\boldsymbol{E}\boldsymbol{s}(x)\mathbf{0}) =_{\mathcal{T}} \hat{\neg}\boldsymbol{s}(\mathbf{0}) =_{\mathcal{T}} \mathbf{0}$ by definition 4.20. By rule 6a and theorem 4.35 we get what we want. ✓

(f) $(sx = sy \to x = y)^D : \boldsymbol{s}(x) =_{\mathbf{i}} \boldsymbol{s}(y) \to x =_{\mathbf{i}} y$. We know that $\hat{\to}(\boldsymbol{E}(\boldsymbol{s}(x))(\boldsymbol{s}(y)))(\boldsymbol{E}xy) =_{\mathcal{T}} \hat{\to}(\boldsymbol{E}xy)(\boldsymbol{E}xy)$. But $\hat{\to}xx =_{\mathbf{i}} \mathbf{0}$ is provable in $\mathcal{T}$ because we can set $\vartheta[x] \equiv \hat{\to}xx =_{\mathbf{i}} \mathbf{0}$.

-$\vartheta[x := \mathbf{0}] \equiv \underbrace{\hat{\to}\mathbf{0}\mathbf{0}}_{=_{\mathcal{T}}\mathbf{0}} =_{\mathbf{i}} \mathbf{0}$ which is provable in $\mathcal{T}$ by rule 6a.

-$\vartheta[x := \boldsymbol{s}(x)] \equiv \underbrace{\hat{\to}\boldsymbol{s}(x)\boldsymbol{s}(x)}_{=_{\mathcal{T}}\hat{\neg}\boldsymbol{s}(x) =_{\mathcal{T}}\mathbf{0}} =_{\mathbf{i}} \mathbf{0}$ which is provable in $\mathcal{T}$ by rule 6a.

and we then apply Ind rule for $\vartheta[x]$. ✓

3. *Axioms of predicate logic*

(We avoid referring to more sequences of variables -as Troelstra does in [46]- for convenience. "$\leftrightarrow$" below means "provably equivalent formulas". We just omit $\mathcal{T} \vdash$ at the beginning)

Assume that the inductive hypothesis holds of formulas ϕ, ψ, i.e., HA $\vdash \phi$, HA $\vdash \psi$ and for $\phi^D = \exists x^\sigma \forall y^\tau \phi_D[x^\sigma, y^\tau]$ and $\psi^D = \exists z^\alpha \forall w^\beta \psi_D[z^\alpha, w^\beta]$ we have that there exist terms U, U' such that (i) $\mathcal{T} \vdash \phi_D[U^\sigma, y^\tau]$, $\mathcal{T} \vdash \psi_D[U'^\alpha, w^\beta]$ and (ii) they don't contain y, w free respectively.

(a)

$$\begin{aligned}[]
[\phi \to (\phi \wedge \phi)]^D &\leftrightarrow \{\exists x\forall y\phi_D[x,y] \to \exists x'\exists x''\forall y'\forall y''\phi_D[x',y'] \wedge \phi_D[x'',y'']\}^D \\
\ldots\ldots &\leftrightarrow \text{(omitted "skolemization" as explained in section 4.7)} \\
&\leftrightarrow \exists Y\exists X'\exists X''\forall x\forall y'\forall y''\{\phi_D[x, Yxy'y''] \to \phi_D[X'x,y'] \wedge \\
&\qquad \phi_D[X''x,y'']\}
\end{aligned}$$

$X', X'' \in \sigma \to \sigma$ and $Y \in \sigma \to (\tau \to (\tau \to \tau))$. For X' and X'' take $T' \equiv \lambda x.x$ and for Y take a term $T'' \equiv \lambda xy'y''.(\boldsymbol{R}_\tau y'(\lambda z^\tau b^{\mathbf{int}}.y'')t_{\phi_D})$ where t_{ϕ_D} is as in definition 4.32 for $\phi_D[x,y']$.

(b)

$$\begin{aligned}
(\phi \vee \phi \to \phi)^D &\leftrightarrow \Big[\exists p^{\mathbf{int}}\exists x\exists x'\forall y\forall y'\{(p =_{\mathbf{i}} \mathbf{0} \wedge \phi_D[x,y]) \vee (p =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \\
&\qquad \wedge\phi_D[x',y'])\} \to \exists x''\forall y''\phi_D[x'',y'']\Big]^D \\
\ldots\ldots &\leftrightarrow \exists Y\exists Y'\exists X''\forall p^{\mathbf{int}}\forall x\forall x'\forall y''\{[(p =_{\mathbf{i}} \mathbf{0} \wedge \phi_D[x, Ypxx'y'']) \\
&\qquad \vee(p =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \wedge \phi_D[x', Y'pxx'y''])] \to \phi_D[X''pxx',y'']\}
\end{aligned}$$

For Y, Y' take $\lambda pxx'y''.y''$ and for X'' take $T'' \equiv \lambda pxx'.\big(\boldsymbol{R}_\sigma x(\lambda z^\sigma b^{\textbf{int}}.x')p\big)$.

(c)

$$\begin{aligned}(\phi \to (\psi \vee \phi))^D \leftrightarrow [\exists x \forall y \phi_D[x,y] \to \exists p^{\textbf{int}} \exists x' \exists z \forall y' \forall w \big((p =_{\mathbf{i}} \mathbf{0} \wedge \phi_D[x',y'])\\ \vee (p =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \wedge \psi_D[z,w])\big)]^D\\ \ldots.. \leftrightarrow \exists Y \exists P \exists X' \exists Z \forall x \forall y' \forall w \{\phi_D[x, Yxy'w] \to [\big(Px =_{\mathbf{i}} \mathbf{0}\\ \wedge \phi_D[X'x,y']\big) \vee \big((Px =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \wedge \psi_D[Zx,w])\big)]\}\end{aligned}$$

For Y take $\lambda xy'w.y'$, for P take $\lambda x.\mathbf{0}$, for Z take $\lambda x.x$ and for X' take $\lambda x.z$.

(d)

$$\begin{aligned}[(\phi \wedge \psi) \to \phi]^D \leftrightarrow [\exists x \exists z \forall y \forall w (\phi_D[x,y] \wedge \psi_D[z,w]) \to \exists x' \forall y' \phi_D[x',y']]^D\\ \ldots.. \leftrightarrow \exists X' \exists Y \exists W \forall x \forall z \forall y' (\phi_D[x, Yxzy'] \wedge \psi_D[z, Wxzy'])\\ \to \phi_D[X'xz, y']\end{aligned}$$

For Y take $\lambda xzy'.y'$, for W take $\lambda xzy'.w$ and for X' take $\lambda xz.x$.

(e)

$$\begin{aligned}[(\phi \wedge \psi) \to (\psi \wedge \phi)]^D \leftrightarrow \quad & [\exists x \exists z \forall y \forall w (\phi_D[x,y] \wedge \psi_D[z,w]) \to\\ & \exists z' \exists x' \forall w' \forall y' (\psi_D[z',w'] \wedge \phi_D[x',y'])]^D\\ \ldots.. \leftrightarrow \quad & \exists Z' \exists X' \exists Y \exists W \forall x \forall z \forall w' \forall y' [(\phi_D[x, Yxzw'y']\\ & \wedge \psi_D[z, Wxw'y']) \to (\psi_D[Z'xz, w'] \wedge \phi_D[X'xz, y'])]\end{aligned}$$

For Z' take $\lambda xz^z.$, for X' take $\lambda xz^x.$, for Y take $\lambda xzw'y'.y'$ and for W take $\lambda xw'y'.w'$.

(f)

$$\begin{aligned}[(\phi \vee \psi) \to (\psi \vee \phi)]^D \leftrightarrow \quad & [\exists p^{\textbf{int}} \exists x \exists z \forall y \forall w \big((p =_{\mathbf{i}} \mathbf{0} \wedge \phi_D[x,y]) \vee (p =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})\\ & \wedge \psi_D[z,w])\big) \to \exists p' \exists z' \exists x' \forall w' \forall y' \big((p' =_{\mathbf{i}} \mathbf{0} \wedge \psi_D[z',w'])\\ & \vee (p' =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \wedge \phi_D[x',y'])\big)]^D\\ \ldots.. \leftrightarrow \quad & \exists P' \exists Z' \exists X' \exists Y \exists W \forall p \forall x \forall z \forall w' \forall y' \{\big((p =_{\mathbf{i}} \mathbf{0}\\ & \wedge \phi_D[x, Ypxzw'y']) \vee (p =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \wedge \psi_D[z, Wpxzw'y'])\big)\\ & \to \big((P'pxz =_{\mathbf{i}} \mathbf{0} \wedge \psi_D[Z'pxz, w']) \vee (P'pxz =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0})\\ & \wedge \phi_D[X'pxz, y'])\big)\}\end{aligned}$$

For Z' take $\lambda pxz.z$, for X' take $\lambda pxz.x$, for Y take $\lambda pxzw'y'.y'$, for P' $\lambda pxz.(\boldsymbol{R}_{\textbf{int}} \boldsymbol{s}(\mathbf{0})(\lambda a^{\textbf{int}} b^{\textbf{int}}.\mathbf{0})p)$. and for W take $\lambda pxzw'y'.w'$.

(g)

$$\begin{aligned}(\bot \to \phi)^D &\leftrightarrow [\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \to \exists x \forall y \phi_D[x,y]]^D\\ &\leftrightarrow \exists x \forall y (\boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0} \to \phi_D[x,y])\end{aligned}$$

For x one can take any term of type σ that doesn't contain y free.
The provability is immediate by ($\bot$E) and axiom 5b.

(h) We make use of [7, lem. 6.2.1(25, 26)] for the first equivalence

$$\begin{aligned}[\neg\phi \to (\phi \to \bot)]^D \leftrightarrow & \Big[\exists Y \forall x \neg\phi_D(x, Yx) \to \forall x' \exists y' \big(\phi_D[x', y'] \\ & \to \boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}\big)\Big]^D \\ \leftrightarrow & \Big[\exists Y \forall x \neg\phi_D(x, Yx) \to \exists Y' \forall x' \big(\phi_D[x', Y'x'] \\ & \to \boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}\big)\Big]^D \\ \ldots.. \leftrightarrow & \exists Y' \exists X \forall Y \forall x' \{\neg\phi_D[XYx', Y(XYx')] \to \\ & \big(\phi_D[x', Y'Yx'] \to \boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}\big)\}\end{aligned}$$

For X take $\lambda Yx'.x'$ and for Y' take $\lambda Yx'.Yx'$.

(i)

$$\begin{aligned}((\phi \to \bot) \to \neg\phi)^D \leftrightarrow & [(\exists x \forall y \phi_D[x, y] \to \bot) \to \exists Y' \forall x' \neg\phi_D[x', Y'x']]^D \\ \ldots.. \leftrightarrow & \exists Y' \exists X \forall Y \forall x' [\big(\phi_D[XYx', Y(XYx')] \to \\ & \boldsymbol{s}(\mathbf{0}) =_{\mathbf{i}} \mathbf{0}\big) \to \neg\phi_D[x', Y'Yx]]\end{aligned}$$

For X take $\lambda Yx'.Yx'$ and for Y' take $\lambda y^\tau.y$.

(j) for Q being a term free for x in $\psi[x]$

$$\begin{aligned}(\psi[x := Q] \to \exists x \psi[x])^D \leftrightarrow & \Big[(\psi[x' := Q])^D \to \exists x^{\textbf{int}} \exists z^\alpha \forall w^\beta \psi_D[z^\alpha, w^\beta, x^{\textbf{int}}]\Big]^D \\ \leftrightarrow & \Big[\big(\exists z' \forall w' \psi_D[z', w', x'^{\textbf{int}} := Q^D]\big) \to \exists x^{\textbf{int}} \exists z^\alpha \forall w^\beta \\ & \psi_D[z^\alpha, w^\beta, x^{\textbf{int}}]\Big]^D \\ \ldots \leftrightarrow & \exists X \exists Z \exists W' \forall z' \forall w [\psi_D[z', W'z'w, x'^{\textbf{int}} := Q^D] \\ & \to \psi_D[Zz', w, Xz']\big)\end{aligned}$$

For Z take $\lambda z'.z'$, for W' take $\lambda zw.w$ and for X take $\lambda z'.Q^D$.

(k) for Q being a term free for x in $\phi[x]$

$$\begin{aligned}(\forall x \phi[x] \to \phi[x := Q])^D \leftrightarrow & \exists X \forall y \forall p^{\textbf{int}} (\phi_D[Xp, y, p] \to \exists x' \forall y' \phi_D[x', y', p' := Q^D] \\ \ldots \leftrightarrow & \exists X' \exists Y \exists P \forall X \forall y' \Big[\phi_D[X(PXy'), YXy', PXy'] \\ & \to \phi_D[X'X, y', p' := Q^D]\Big]\end{aligned}$$

For X', Y, P we take respectively $\lambda X.XQ^D$, $\lambda Xy'.y'$ and $\lambda Xy'.Q^D$.

4. Rules of inference for ϕ, ψ being formulas in L_0

(a) $\{\phi, \phi \to \psi\} \vdash \psi$
Assume that

$$\mathcal{T} \vdash \phi_D[U, y] \tag{4.6}$$

$$\mathcal{T} \vdash \phi_D[x, Yxw] \to \psi_D[Xx, w] \tag{4.7}$$

In 4.6 we take $y \equiv Yxw$ and in 4.7 we take $x \equiv U$. So

$$\mathcal{T} \vdash \phi_D[U, Yxw] \tag{4.8}$$
$$\mathcal{T} \vdash \phi_D[U, Yxw] \to \psi_D[XU, w] \tag{4.9}$$

By 4.8, 4.9 and ($\to$E) we get $\mathcal{T} \vdash \psi_D[XU, w]$ and we take for X the term $\lambda x.x$ (we suppose that w isn't free in ϕ, otherwise we rename it from the beginning. So U doesn't contain w free).

(b) $\{\phi \to \psi, \psi \to \beta\} \vdash \phi \to \beta$
Assume that

$$\mathcal{T} \vdash \phi_D[x, Yxw] \to \psi_D[Xx, w]$$
$$\mathcal{T} \vdash \psi_D[z, Y'zd] \to \beta_D[X'z, d]$$

We take $w \equiv Y'zd$ and $z \equiv Xx$. So we have

$$\mathcal{T} \vdash \phi_D[x, Yx(Y'(Xx)d)] \to \beta_D[X'(Xx), d]$$

Since
$$(\phi \to \beta)^D \equiv \exists Y \exists C \forall x \forall d\ \phi_D[x, Yxd] \to \beta_D[Cx, d]$$

we want appropriate terms Y, C such that

$$\mathcal{T} \vdash \phi_D[x, Yxd] \to \beta_D[Cx, d] \tag{4.10}$$

So in 4.10 we take for Y the term $\lambda xd.Yx(Y'(Xx)d)$ and for C $\lambda x.X'(Xx)$.

(c) $\phi \to \beta \vdash \phi \vee \psi \to \beta \vee \psi$
Assume that

$$\mathcal{T} \vdash \phi_D[x, Yxd] \to \beta_D[Cx, d] \tag{4.11}$$

$$\begin{aligned}
[(\phi \vee \psi) \to (\beta \vee \psi)]^D \leftrightarrow \quad & [\exists p^{\textbf{int}} \exists x \exists z \forall y \forall w \big((p =_{\mathbf{i}} \mathbf{0} \wedge \phi_D[x, y]) \vee (p =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \\
& \wedge \psi_D[z, w])\big) \to \exists p'^{\textbf{int}} \exists c \exists z' \forall d \forall w' \big((p' =_{\mathbf{i}} \mathbf{0} \wedge \beta_D[c, d]) \\
& \vee (p' =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \wedge \psi_D[z', w'])\big)]^D \\
..... \leftrightarrow \quad & \exists P' \exists C \exists Z' \exists Y \exists W \forall p \forall x \forall z \forall d \forall w' \{\big((p =_{\mathbf{i}} \mathbf{0} \\
& \wedge \phi_D[x, Ypxzdw']) \vee (p =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \wedge \psi_D[z, Wpxzdw'])\big) \\
& \to \big((P'pxz =_{\mathbf{i}} \mathbf{0} \wedge \beta_D[Cpxz, d]) \vee (P'pxz =_{\mathbf{i}} \boldsymbol{s}(\mathbf{0}) \\
& \wedge \psi_D[Zpxz, w'])\big)\}
\end{aligned}$$

Take $P' \equiv \lambda pxz.p$, $Y \equiv \lambda pxzdw'.Yxd$, $C \equiv \lambda pxz.Cx$, $W \equiv \lambda pxzdw'.w'$ and $Z' \equiv \lambda pxz.z$.

(d) $\phi \to (\psi \to \beta) \vdash (\phi \wedge \psi) \to \beta$

$$
\begin{aligned}
[\phi \to (\psi \to \beta)]^D \leftrightarrow \exists x \exists y\, \phi_D[x,y] \to \big(\exists W \exists C \forall z \forall d\, \psi_D[z, Wzd] \\
\to \beta_D[Cz,d]\big) \\
\ldots.. \leftrightarrow \exists W \exists C \exists Y \forall x \forall z \forall d\, \phi_D[x, Yxzd] \to \\
\big(\psi_D[z, Wxzd] \to \beta_D[Cxz,d]\big)
\end{aligned}
\tag{4.12}
$$

$$
\begin{aligned}
[(\phi \wedge \psi) \to \beta]^D \leftrightarrow \exists x \exists z \forall y \forall w \big(\phi_D[x,y] \wedge \psi_D[z,w]\big) \to \\
\exists c \forall d \beta_D[c,d] \\
\ldots.. \leftrightarrow \exists C \exists Y \exists W \forall x \forall z \forall d \big(\phi_D[x, Yxzd] \wedge \\
\psi_D[z, Wxzd]\big) \to \beta_D[Cxz,d]
\end{aligned}
\tag{4.13}
$$

By [7, lem. 6.2.1(8)]] we get the result.

(e) $\phi \to \psi[x] \vdash \phi \to \forall x \psi[x]$ where x not free in ϕ and if $\Gamma \vdash \phi \to \psi[x]$, then x can't be free in formulas of Γ.

$$
\begin{aligned}
[\phi \to \psi[p]]^D \leftrightarrow \exists x \forall y \phi_D[x,y] \to \exists z \forall w\, \psi_D[z,w,p] \\
\ldots.. \leftrightarrow \exists Z \exists Y \forall x \forall w \big(\phi_D[x, Yxw] \to \psi_D[Zx, w, p]\big)
\end{aligned}
$$

We assume that $\mathcal{T} \vdash \phi_D[x, Yxw] \to \psi_D[Zx, w, p]$.

$$
\begin{aligned}
[\phi \to \forall p \psi[p]]^D \leftrightarrow \exists x \forall y \phi_D[x,y] \to \exists Z \forall w \forall p \psi_D[Zp, w, p] \\
\ldots.. \leftrightarrow \exists Z' \exists Y' \forall x \forall w \forall p \big(\phi_D[x, Y'xwp] \to \psi_D[Z'xp, w, p]\big)
\end{aligned}
$$

Take $Z' \equiv \lambda xwp.Yxw$ and $Y' \equiv \lambda xp.Zx$.

(f) $\psi[x] \to \phi \vdash \exists x \psi[x] \to \phi$ where x not free in ϕ and if $\Gamma \vdash \psi[x] \to \phi$, then x can't be free in formulas of Γ.

$$
\begin{aligned}
[\psi[p] \to \phi]^D \leftrightarrow \exists z \forall w \psi_D[z,w,p] \to \exists x \forall y \phi_D[x,y] \\
\ldots.. \leftrightarrow \exists X \exists W \forall z \forall y \big(\psi_D[z, Wzy, p] \to \phi_D[Xz, y]\big)
\end{aligned}
$$

$$
\begin{aligned}
[\exists p \psi[p] \to \phi]^D \leftrightarrow \exists p \exists z \forall w \psi_D[z,w,p] \to \exists x \forall y \phi_D[x,y] \\
\ldots.. \leftrightarrow \exists X' \exists W' \forall p \forall z \forall y [\psi_D[z, Wpzy, p] \to \phi_D[Xpz, y]]
\end{aligned}
$$

For W', X' we take respectively $\lambda pzy.Wzy$ and $\lambda pz.Xz$.

(g) $(\phi \wedge \psi) \to \beta \vdash \phi \to (\psi \to \beta)$
We use 4.12, 4.13 and we assume that $\mathcal{T} \vdash \big(\phi_D[x, Yxzd] \wedge \psi_D[z, Wxzd]\big) \to \beta_D[Cxz, d]$. By [7, lem. 6.2.1(8)]] we get the result.

(h) $\{\phi[x := 0], \forall x(\phi[x] \to \phi[x := sx])\} \vdash \forall y \phi[x := y]$ (Rule of Induction)
We have that

$$\begin{aligned}
\big(\phi[z := 0]\big)^D &\leftrightarrow \exists x \forall y \phi_D[x, y, z := \mathbf{0}] \\
\big[\forall z \big(\phi[z] \to \phi[z := sz]\big)\big]^D &\leftrightarrow \Big[\forall z^{\text{int}} \big(\exists x \exists y \phi_D[x, y, z] \to \exists x' \forall y' \\
&\qquad \phi_D[x', y', z := \boldsymbol{s}(z)]\big)\Big]^D \\
\ldots.. &\leftrightarrow \exists X' \exists Y \forall z \forall x \forall y' \phi_D[x, Yzxy', z] \to \\
&\qquad \phi_D[X'zx, y', z := \boldsymbol{s}(z)]
\end{aligned} \tag{4.14}$$

(Note that the replacement of z with $\boldsymbol{s}(z)$ in 4.14 is taking place before the substitution of $X'zx$ for x) We assume that:

$$\mathcal{T} \vdash \phi_D[x, y, z := \mathbf{0}] \tag{4.15}$$

$$\mathcal{T} \vdash \phi_D[x, Yzxy', z] \to \phi_D[X'zx, y', z := \boldsymbol{s}(z)] \tag{4.16}$$

If $M :\equiv \boldsymbol{R}_{\text{int}} x(\lambda xz.X'zx)$ we have that

$$M\mathbf{0} =_{\mathcal{T}} x \qquad \text{and} \qquad M(\boldsymbol{s}(z)) =_{\mathcal{T}} X'z(Mz)$$

If we set $x := M\mathbf{0}$, $y := y'$ in 4.15, $x := Mz$ and $Y \equiv \lambda zxy'.y'$ in 4.16, we get that $\psi[z] :\equiv \phi_D[Mz, y, z]$ is provable because:

$$\mathcal{T} \vdash \phi_D[M\mathbf{0}, y, z := \mathbf{0}]$$
$$\mathcal{T} \vdash \phi_D[Mz, y', z] \to \phi_D[M\boldsymbol{s}(z), y', z := \boldsymbol{s}(z)]$$

and we can apply Ind rule.

□

4.10 Applications of parts of the proof

Gödel's interpretation has been adapted and extended both to stronger and weaker theories. A result of such an extension is analysis' consistency proof thanks to C. Spector.

Moreover, the D-interpretation yields interesting constructive information about proofs in classical arithmetic. Specifically,

> "*any recursive function whose totality can be proved either in PA or in HA, is represented by a term in* $\mathcal{T}$" [2].

4.11 Recent results

It's worth mentioning some recent results that have to do with Dialectica. Firstly, we have [34] which "*presents an analysis of Gödel's Dialectica interpretation via a refinement of intuitionistic logic known as linear logic*" and then [35] which "*surveys several computational interpretations of classical linear logic based on two-player one-move games*" including Dialectica. We also have [13] which "*presents a family of functional interpretations of **intuitionistic** linear logic*", instead of classical, which yields simpler results in some cases.

CHAPTER 5

THE TWO PROOFS FROM A PHILOSOPHICAL POINT OF VIEW

In this section we will discuss some objections raised concerning the two proofs presented earlier so as to illuminate some hidden parts of their history.

First and foremost, we should clarify that despite many “false red alarms” that every now and then ring in the mathematical community, the two proofs have been well examined since their appearance and they have no errors concerning their correctness. Any objection raised is mostly around their acceptability from a philosophical point of view.

It is normal to a certain extent that some are skeptical when talking about the foundations of mathematics and what axioms or argumentation we are allowed to use in a proof. Besides, all these began in the first place, as we explained in chapter 2, due to some “careless” mathematical handling that took for granted things that we could not assume. So a greater level of rigour is expected from a proof that will assure mathematicians of the reliability of mathematics as founded. So what are the objections?

One objection is that of the use of stronger methods than the ones available in the system of PA itself [5]. This, if true, makes the proof useless. Moreover, according to Gödel’s second incompleteness theorem, PA is inept at proving its own consistency. So one might assert that the system used for the proofs, i.e., PRA (Primitive Recursive Arithmetic), a subset of PA, and one axiom out of PA, makes the assumptions stronger than the conclusion.

That could be deemed correct if the system (PRA+axiom out of PA) yielded a superset (as a theory) of PA. This however isn’t true. The system is neither a superset, nor a subset (in accordance to Gödel’s second incompleteness theorem). Though, if we really want to use a weaker system than PA, we are led to a deadend. All proofs possibly proposed for PA’s consistency must include something not formalised in PA. So, in that case, we will be left with no proof at all, as no proof will be acceptable. Even for weaker systems that contain a minimum of arithmetic, as ZF (Zermelo-Fraenkel axiomatic system) [52] or Robinson Arithmetic (**Q**) which doesn’t include the induction schema, we are left with no proof.

Similarly, some might accuse Gentzen for involving in his proof “*a viciously circular pattern of epistemic dependence between its premises and its conclusion*” [49, 27] due to the fact that he uses “*induction up to ε_0 to prove the consistency of arithmetic,* *i.e., induction up to ω*" [16, p. 37]. The accusation (Circularity Argumentation) could be presented as follows:

CA1 Justification for the premises used in Gentzen's proof depends on an understanding of the ordinals up to ε_0.

CA2 An understanding of the ordinals up to ε_0 depends on an understanding of the ordinals up to ω, i.e., the natural numbers.

CA3 An understanding of the natural numbers depends on justification for the consistency of PA.

CA4 Therefore Genzen's proof fails to transmit justification for the consistency of PA. [49, p. 15]

If we have a circular pattern, at least one of the premises **CA1-CA3** must depend on justification for PA's consistency. However **CA3** is rather misleading. It is more likely that we are inclined to believe that PA is consistent due to our understanding of natural numbers, than the reverse [49].

Nevertheless, even if this isn't true, i.e., even if the dependence is true reversly, **CA2** isn't completely true. Before we move to the explanation of this assertion, we should first explain two notions related to induction; *height* and *width*.

The height of the induction is "*how far into the ordinals induction can be carried out*", while the width is "*the range of properties or conditions over which the induction can be carried out*". Whereas the height of the induction in Genzen's proof is "greater" than the induction in PA, the width of induction in Genzen's proof is "narrower" than that of PA, as it can be formalised in the system (PRA+ axiom out of PA) which has only quantifier-free formulas, a strict subset of PA's formulas. There is therefore a trade-off between the height and the width, which ruins argumentation about the circularity [49, 27].

One other thing that one could ascribe to these proofs is that they give a result dependent on the consistency of another system, i.e., PA is consistent if the system (PRA+axiom out of PA) is consistent. This seems problematic if we consider that the latter isn't known to be consistent. However, as stated in [5] this objection:

> "*assumes that it is somehow possible to 'pull yourself up by your own bootstraps' by setting up some system whose consistency is guaranteed because it has been proven—presumably in some absolute, unconditional sense. But any consistency proof has to assume s o m e t h i n g, and you can always cast doubt on that 'something' and demand that it be given a consistency proof, and so on ad infinitum. Even if somehow you found a plausible system that proved its own consistency, any doubts you had about its consistency would hardly be allayed just because it vouched for itself! At some point, you simply have to take something for granted without demanding that it be proved from something more basic. This much is obvious, even without Gödel's theorem.*"

The next objection we will refer to is one coming from the constructivists' world. In both proofs we use a stabilization point. In Gödel's proof this is the point where we reach for the first time the unique normal form of a term in system $\mathcal{T}$ in the procedure of reductions, while in Gentzen's it is the point where the descending sequence of ordinals stabilizes. Both these points are not computable in general. Thereby, they aren't considered as constructive enough arguments.

The problem with this objection is that it leads to a dead-end as much as the second one does. If we restrict ourselves to computable mathematics, many well-known theorems will have to fall into oblivion, as they aren't provable in this part. Two cases in point are the Bolzano-Weirstrass theorem and Brouwer's fixed-point theorem [5], as is the consistency of PA too.

From the constructivist's point of view, in Gödel's proof and more precisely in the proof that $(A \to B)^D$ yields an equivalent formula for $A \to B$, we have also Markov's principle causing "problems". This principle, as we said earlier, is not generally intuitionistically valid (one of the schools of constructivism), albeit some of its instances are provable in a constructive context (see also [42]).

In order to prove that a formula ϕ and its D-interpretation are equivalent, we must add to our axiomatic system IP'(variant of Independence of Premise), MP' (variant of Markov's Principle) and AC (axiom of Choice) [45, sec. 3]. This means that many constructivists won't accept Dialectica as a constructive enough tool for the needs of the proof, due to MP'. Maybe that's why Gödel in his 1972 version of the paper, in note **h** [18, 22], was trying to argue that D-interpretation is more constructive than Heyting's interpretation (probably the BHK interpretation as it is mostly known today, which is the standard interpretation of intuitionistic logic) claiming that this comes from (MP'). This argument is questioned by Troelstra [45, p. 232], but "*while the reasoning leading to the form of the D-interpretation is not fully constructive, it can still be used as a tool in constructive mathematics and to derive information from non-constructive proofs... (the not fully constructive reasoning) allows one to use the D-interpretation to extract constructive information from non-constructive proofs*" [2, p. 11].

In closing, we shall say that more objections can be found in the bibliography. Howbeit, it is our firm belief that as many as they might be, it is very hard -though not impossible since consistency proofs are relative proofs, based on the consistency of other systems- for them to demolish the building constructed over these proofs, if they don't "attack" the correctness, but the epistemological value of the proofs.

Additionally we should not forget that since their appearance, mathematics (and all sciences nourished by mathematics) have grown irrespective of the suspicion of inconsistency, because most mathematicians don't think that arithmetic is inconsistent [cf. 27, sec. 5].

So what is a conclusion? In my opinion we can conclude only one sure thing: these objections are indicative of the limits of mathematical knowledge; it is limited; we can't prove everything assuming an axiomatic system, as Gödel's second incompleteness theorem suggests too and this happens with any possible extension of the system. As for the proofs of consistency: there will always be someone that rejects them and there will always be someone that objects to their epistemological value, but this doesn't mean that Arithmetic is necessarily inconsistent [cf. 12, 48, 33].

BIBLIOGRAPHY

[1] W. Ackermann. “Zur Widerspruchsfreiheit der Zahlentheorie.” In: *Math. Ann.* 117 (1940), pp. 162–194. DOI: https://doi.org/10.1007/BF01450016.

[2] J. Avigad and S. Feferman. “Gödel’s functional (”Dialectica”) Interpretation.” In: *Handbook of proof theory*. Elsevier, 1999.

[3] L. Bernard and A. D. Irvine. “Principia Mathematica.” In: *The Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta and U. Nodelman. 2024. URL: https://plato.stanford.edu/entries/principia-mathematica/.

[4] W. Buchholz. “On Gentzen’s First Consistency Proof for Arithmetic.” In: *Gentzen’s Centenary. The Quest for Consistency*. Ed. by R. Kahle and M. Rathjen. Springer, 2015.

[5] T.Y Chow. *The Consistency of Arithmetic*. 2018. URL: http://arxiv.org/abs/1807.05641v1.

[6] R. Cook. “Frege’s Logic.” In: *The Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta and U. Nodelman. 2024. URL: https://plato.stanford.edu/entries/frege-logic/#LogiBegr.

[7] D. van Dalen. *Logic and Structure*. Springer, 2013.

[8] A. G. Dragalin. “The computability of primitive recursive terms of finite type and the primitive recursive realization.” In: *Zapiski Nauchnykh Seminarov POMI* 8 (1968), pp. 32–45.

[9] S. Feferman. “Gödel’s life and work.” In: *Kurt Gödel Collected works, publications 1938-1974*. Ed. by S. Feferman and al. Vol. 1. Oxford University Press, 1986.

[10] S. Feferman. “Gödel’s Dialectica interpretation and its two-way stretch.” In: *Computational Logic and Proof Theory (Proc. 3d Kurt Gödel Colloquium, Brno Aug.)* Vol. 713. Lecture Notes in Computer Science. G. Gottlob et at., 1993, pp. 23–40.

[11] S. Feferman. “Lieber Herr Bernays!, Lieber Herr Gödel! Gödel on finitism, constructivity and Hilbert’s program. Special Issue: Gödel’s dialectica Interpretation.” In: *Dialectica* 62.2 (2008), pp. 179–203. URL: https://math.stanford.edu/~feferman/papers/bernays.pdf.

[12] F. Ferreira and G. Ferreira. "Interpretability in Robinson's Q." In: *The Bulletin of Symbolic Logic* 19.3 (2013), pp. 289–317. ISSN: 10798986. URL: http://www.jstor.org/stable/41955413 (visited on 12/18/2024).

[13] G. Ferreira and P. Oliva. "Functional Interpretations of Intuitionistic Linear Logic." In: *Log. Methods Comput. Sci.* 7.1 (2011). DOI: 10.2168/LMCS-7(1:9)2011. URL: https://doi.org/10.2168/LMCS-7(1:9)2011.

[14] J. Ferreirós. "The Early Development of Set Theory." In: *The Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta and U. Nodelman. 2023. URL: https://plato.stanford.edu/archives/sum2023/entries/settheory-early/.

[15] S. Galvan, P. Mancosu, and R. Zach. *An introduction to proof theory*. Oxford University Press, 2021.

[16] J.-Y. Girard. *Proof theory and logical complexity*. Vol. 1. Bibliopolis, 1987.

[17] J.-Y. Girard, Y. Lafont, and P. Taylor. *Proofs and types*. Vol. 7. Cambridge Tracts in Theoretical Computer Science, 1989.

[18] K. Gödel. "On an extension of finitary mathematics which has not yet been used." In: *Kurt Gödel Collected works, publications 1938-1974*. Ed. by S. Feferman and al. Vol. 2. Oxford University Press.

[19] K. Gödel. "Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes." In: *Kurt Gödel Collected works, publications 1938-1974*. Ed. by S. Feferman and al. Vol. 2. Oxford University Press.

[20] R.L. Goodstein. *Recursive number theory. A development of recursive arithmetic in a logic-free equation calculus*. Studies in Logic and Foundations of mathematics. North-Holland Publishing Company, 1957.

[21] A. Grzegorczyk. "Recursive objects in all finite types." In: *Fundamenta Mathematicae* 54.1 (1964), pp. 73–93. URL: http://eudml.org/doc/213775.

[22] M. Hämeen-Anttila. *Sources and Studies in the History of Mathematics and Physical Sciences. Kurt Gödel: The Princeton Lectures on Intuitionism*. Ed. by M. Hämeen-Anttila and J. von Plato. Springer, 2021.

[23] S. Hetzl. *Lecture Slides: Proof Theory of Induction. Summer School for Proof Theory in First-Order Logic Funchal, Madeira*. Institute of Discrete Mathematics and Geometry, Vienna University of Technology, 2017. URL: https://www.dmg.tuwien.ac.at/hetzl/teaching/slides_induction.pdf.

[24] J.R. Hindley. *S1 Gödel's Consistency-Proof for Arithmetic*. Published in author's personal website (inactive). 2011.

[25] J.R. Hindley and J. P. Seldin. *Lambda-Calculus and Combinators: An introduction*. Cambridge University Press, 2008.

[26] H. R. Jervell. "Climbing Mount ε_0." In: *Gentzen's Centenary. The Quest for Consistency*. Ed. by R. Kahle and M. Rathjen. Springer, 2015.

[27] R. Kahle. "Gentzen's Consistency Proof in Context." In: *Gentzen's Centenary. The Quest for Consistency*. Ed. by R. Kahle and M. Rathjen. Springer, 2015.

[28] J. Kennedy. "Kurt Gödel." In: *The Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta and U. Nodelman. 2020. URL: https://plato.stanford.edu/archives/win2020/entries/goedel/.

[29] K. Kunen. *Set Theory*. Vol. 34. Studies in Logic: Mathematical Logic and Foundations. College Publications, 2013.

[30] K. Kunen. *The foundations of Mathematics*. Vol. 19. Studies in Logic: Mathematical Logic and Foundations. College Publications, 2009.

[31] Y. Moschovakis. *Lecture notes in logic*. 2014.

[32] Y. N. Moschovakis. *Notes on Set Theory*. Springer, 1994.

[33] E. Nelson. "Warning signs of a possible collapse of contemporary mathematics." In: *Infinity: New Frontiers of Research*. Ed. by M. Heller and W.H. Woodin. Cambridge University Press, 2011. DOI: https://doi.org/10.1017/CBO9780511976889.

[34] P. Oliva. "An Analysis of Gödel's Dialectica Interpretation Via Linear Logic." In: *Dialectica* 62.2 (2008), pp. 269–290. DOI: 10.1111/j.1746-8361.2008.01135.x.

[35] P. Oliva. "Functional interpretations of linear and intuitionistic logic." In: *Information and Computation* 208.5 (2010). Special Issue: 14th Workshop on Logic, Language, Information and Computation (WoLLIC 2007), pp. 565–577. ISSN: 0890-5401. DOI: https://doi.org/10.1016/j.ic.2008.11.008. URL: https://www.sciencedirect.com/science/article/pii/S0890540109002417.

[36] C. Parsons and W. Sieg. "Introductory note to *1938a*." In: *Kurt Gödel Collected works, Unpublished Essays and Lectures*. Ed. by S. Feferman and al. Vol. 3. Oxford University Press, 1995.

[37] J. von Plato. ""The Development of Proof Theory." In: *The Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta and U. Nodelman. 2018. URL: https://plato.stanford.edu/entries/proof-theory-development/.

[38] J. von Plato. "From Hauptsatz to Hilfssatz." In: *Gentzen's Centenary. The Quest for Consistency*. Ed. by R. Kahle and M. Rathjen. Springer, 2015.

[39] J. von Plato. "Proof Theory." In: *The Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta and U. Nodelman. 2024. URL: https://plato.stanford.edu/entries/proof-theory-development/.

[40] J. von Plato. *Saved from the Cellar. Gerhard Gentzen's Shorthand Notes on Logic and Foundations of Mathematics*. Springer, 2017.

[41] C. Poulios. *Undergraduate lecture notes in set theory*. 2020. URL: https://sites.google.com/site/poulios314/notes.

[42] A. Setzer. "The Use of Trustworthy Principles in a Revised Hilbert's Program." In: *Gentzen's Centenary. The Quest for Consistency*. Ed. by R. Kahle and M. Rathjen. Springer, 2015.

[43] M.H. Sørensen and P. Urzyczyn. *Lectures on the Curry-Howard Isomorphism*. Vol. 149. Studies in Logic and Foundations of mathematics. Elsevier, 2006.

[44] S. Stenlund. *Combinators, λ—Terms and Proof Theory*. Synthese Library. D. Reidel Publishing Company/Dordrecht-Holland, 1972.

[45] A.S. Troelstra. "Introductory note to *1958* and *1972*." In: *Kurt Gödel Collected works, publications 1938-1974*. Ed. by S. et al. Feferman. Vol. 2. Oxford University Press, 1990.

[46] A.S. Troelstra. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*. Vol. 344. Lecture Notes in Mathematics. Springer-Verlag, 1973.

[47] P. Valarcher. “Intensional semantics of system T of Gödel.” In: *Electronic notes in Theoretical Computer Science. Workshop on Domains IV 2-4 October 1998 • Rolandseck, Germany* 35 (2000). URL: https://www.sciencedirect.com/journal/electronic-notes-in-theoretical-computer-science/vol/35/suppl/C.

[48] V. Voevodsky. *What If Current Foundations of Mathematics Are Inconsistent? Celebration of the eightieth anniversary of the Institute*. Institute for Advanced Study (Princeton). 2010. URL: https://youtu.be/O45LaFsaqMA.

[49] D. Waxman. “Did Gentzen prove the Consistency of Arithmetic?” 2017. URL: https://www.danielwaxman.com/files/gentzen.pdf.

[50] Wikipedia. *Double-negation translation*. July 29, 2024. URL: https://en.wikipedia.org/wiki/Double-negation_translation.

[51] R. Zach. “Hilbert’s Program.” In: *The Stanford Encyclopedia of Philosophy*. Ed. by E.N. Zalta and U. Nodelman. 2023. URL: https://plato.stanford.edu/entries/hilbert-program/.

[52] Κ. Δημητρακόπουλος. *Σημειώσεις στη μαθηματική λογική*. 2002.

[53] Κ. Κάλφα. *Αξιωματική Θεωρία Συνόλων*. Ζήτη, 1990.

[54] И.Н. Хлодовский. «Новое доказательство непротиворечивости арифметики». В: *УМН* 14.6(90) (1959), с. 105—140. URL: http://mi.mathnet.ru/rm7377.

INDEX