

Axiomatic Theories of Partial Ground II

Johannes Korbmacher

Journal of Philosophical Logic
Published under the Auspices of the
Association for Symbolic Logic

ISSN 0022-3611

J Philos Logic
DOI 10.1007/s10992-017-9444-z

Volume 43 • Number 1 • February 2014

**ONLINE
FIRST**

JOURNAL OF
PHILOSOPHICAL LOGIC

EDITORS
JOHN HORTY
FRANK VELTMAN

 Springer

 Springer

Your article is published under the Creative Commons Attribution license which allows users to read, copy, distribute and make derivative works, as long as the author of the original work is cited. You may self-archive this article on your own website, an institutional repository or funder's repository and make it publicly available immediately.

Axiomatic Theories of Partial Ground II

Partial Ground and Hierarchies of Typed Truth

Johannes Korbmacher¹ 

Received: 24 June 2016 / Accepted: 12 July 2017

© The Author(s) 2017. This article is an open access publication

Abstract This is part two of a two-part paper in which we develop an axiomatic theory of the relation of partial ground. The main novelty of the paper is the use of a binary ground predicate rather than an operator to formalize ground. In this part of the paper, we extend the base theory of the first part of the paper with hierarchically typed truth-predicates and principles about the interaction of partial ground and truth. We show that our theory is a proof-theoretically conservative extension of the ramified theory of positive truth up to ϵ_0 and thus is consistent. We argue that this theory provides a natural solution to Fine’s “puzzle of ground” about the interaction of truth and ground. Finally, we show that if we apply the truth-predicate to sentences involving our ground-predicate, we run into paradoxes similar to the semantic paradoxes: we get ground-theoretical paradoxes of self-reference.

Keywords Metaphysical grounding · Axiomatic theories of truth · Predicational theories of ground · Positive truth

1 Introduction

This is part two of a two-part paper in which we develop axiomatic theories of partial ground.¹ Partial ground, remember, is the relation of one truth holding (either

¹For part one, see [12] in this journal.

✉ Johannes Korbmacher
j.korbmacher@uu.nl

¹ Department of Philosophy and Religious Studies, Utrecht University, Janskerkhof 13,
3512 BL Utrecht, The Netherlands

wholly or partially) in virtue of another, cf. [5, 7].² Partial ground in this sense is standardly taken to be *irreflexive*, meaning that no truth partially grounds itself, and *transitive*, meaning that partial grounds are inherited through partial grounding. In other words, partial ground is a *strict partial order* on the truths.³ Consequently, the relation of partial ground induces a *hierarchy of grounds* on the truths, in which the partial grounds of a truth rank “strictly below the truth itself. In the first part of the paper, we axiomatized this hierarchy over the truths of arithmetic using a (binary) ground *predicate* instead of an operator to formalize the relation of partial ground.⁴

When axiomatizing the grounding hierarchy, we explicitly restricted ourselves to applications of the ground predicate to *arithmetical truths* (i.e. true sentences in the language of arithmetic), leaving aside applications of the ground predicate to sentences involving the truth predicate and applications of the truth predicate to sentences involving the ground predicate. In the present paper, we lift these restrictions.

First, we turn our attention to applications of the ground predicate to sentences involving the truth predicate. A plausible principle about the interaction of truth and partial ground states that the truth of any true sentence grounds the truth of the sentence which states that the sentence in question is true. This principle traces back to Aristotle and may be put a bit more explicitly as follows:

(Aristotelian Principle): If φ is a true sentence, then the truth of $Tr(\ulcorner\varphi\urcorner)$ holds either wholly or partially in virtue of the truth of φ .⁵

Unfortunately, as Fine effectively shows in [6], the Aristotelian Principle is inconsistent with standardly accepted principles about the interaction of partial ground and the logical operators: it allows us to derive that the truths of some true sentences partially ground themselves—in direct contradiction to the irreflexivity of partial ground. To give a quick example, take the sentence $\exists x Tr(x)$, which says that there is at least one true sentence. This sentence is itself (provably) true, and so we have that $\exists x Tr(x)$ partially grounds $Tr(\ulcorner\exists x Tr(x)\urcorner)$ by the Aristotelian Principle. By the standardly accepted principle of partial ground that an existential truth is partially grounded in all of its true instances, we get furthermore that $Tr(\ulcorner\exists x Tr(x)\urcorner)$ partially grounds $\exists x Tr(x)$. But by the transitivity of partial ground this immediately gives us that $\exists x Tr(x)$ partially grounds itself. This is (an instance of) what in [6] Fine calls *the puzzle of ground*.

²For a more comprehensive introduction to the concept of (partial) ground we rely on in this paper and further literature references, see Sections 1 and 2 of the first part of this paper.

³This is, in any case, the standard view of partial ground. Some authors have challenged the view, though: [11] challenges the claim that partial ground is irreflexive and [19] challenges the claim that partial ground is transitive. See [16] and [18] for a defense of the standard view against these challenges.

⁴For a precis of the first part of the paper, see Section 1.1 below.

⁵Here we use Tr as our truth-predicate and the Gödel-corners $\ulcorner \urcorner$ as our quotation device. For more on these syntactic matters, see Section 1.1 below.

In this paper, we develop, in quite some formal detail, what Fine [6, p. 108–10] calls a “predicativist” solution to the puzzle of ground. What makes our solution predicativist is our appeal to Tarski’s hierarchy of object-language, meta-language, meta-meta-language, and so on to rule out the problematic cases of apparently self-grounding sentences. We show that by observing Tarski’s distinction between object and meta-language, we can formulate a *consistent* axiomatic theory of partial ground that proves (a predicativist version of) the Aristotelian principle while retaining the irreflexivity of ground. Formally, we obtain this theory using the method *typing*, familiar from theories of truth, where applications of a truth-predicate to sentences involving the same truth-predicate are ruled out.⁶ To the best of our knowledge, this is the first fully worked out proposal for a predicativist solution to the puzzle of ground in the literature.⁷

Then, we turn our attention to applications of the truth predicate to sentences containing the ground predicate. We show that we cannot consistently add axioms to our axiomatic theory of ground that would allow us to prove all the instances of the T-scheme $Tr(\ulcorner \varphi \urcorner) \leftrightarrow \varphi$ over formulas φ involving the ground-predicate; for if we were to do so, we’d get what we call *ground-theoretic paradoxes of self-reference*. We will show that these paradoxes are closely related to the well-known semantic paradoxes of self-reference, such as the infamous *liar paradox*, and furthermore that the paradoxes are genuinely different from Fine’s puzzle of ground. Fine’s puzzle arises from the fact that intuitively plausible principles about partial ground and truth entail that the truths of some sentences partially ground themselves. The ground-theoretic paradoxes of self-reference, in contrast, concern an entirely different aspect of partial ground: the problem is, as we’ll argue, that the ground predicate behaves too much like a truth-predicate.⁸

Here is the plan for the paper: Section 1.1 contains a précis of the first part of the paper. Then, in Section 2, we’ll discuss the interaction of truth and partial ground on a general level and introduce the Aristotelian Principle into our predicational setting for axiomatic theories of partial ground. In Section 3, we’ll develop an axiomatic predicational theory that proves (a predicative/typed version of) the Aristotelian principle. We then prove that the theory is a conservative extension of the *ramified theory of truth* (and is thus, in particular, consistent) and we construct a model. Then, in Section 4, we will show that if we’re not careful, partial ground, just like truth, can give rise to paradoxes of self-reference. In Section 5, we conclude with some general observations and a map of possible responses to the ground-theoretic paradoxes of self-reference.

⁶For a detailed exposition of typed theories of truth, see [8, p. 49–286].

⁷Stephan Krämer [14] and Thomas Donaldson [3] tentatively sketch predicativist solutions without developing the formal details. Fine [6], Correia [2], and Litland [15] propose non-predicativist solutions to the puzzle of ground. While Correia gives up the irreflexivity of ground, Litland and Fine do not.

⁸This generalizes some previously results obtained by the author in [13].

1.1 Précis of “Axiomatic Theories of Partial Ground I. The Base Theory”

In the following sections, we assume that the reader is familiar with the techniques and results obtained in the first part of the paper. Those who have these techniques and results sufficiently present in their mind can safely skip this section. But for the rest, let’s briefly refresh our memory.

In the first part of the paper, we develop an axiomatic theory of partial ground using a binary ground *predicate* rather than an *operator* to express partial ground. We call such a theory a *predicational* theory of partial ground, in contrast to *operational* theories. At the outset of the paper, we provide technical and philosophical reasons for using a predicational approach (Section 2).

In the following sections (Sections 3 and 4), we develop our theory in formal detail and investigated its properties. On the syntax side, we start from the language \mathcal{L} of Peano arithmetic. The language \mathcal{L}_{Tr} is defined as $\mathcal{L} \cup \{Tr\}$, where Tr is the unary truth-predicate, and the language $\mathcal{L}_{Tr}^\triangleleft$ is $\mathcal{L}_{Tr} \cup \{\triangleleft\}$, where \triangleleft is the binary (partial) ground-predicate. We formulate our theory in $\mathcal{L}_{Tr}^\triangleleft$.

On the theory side, we start from the theory PAG of Peano arithmetic (PA) with all the instances of the induction scheme

$$\varphi(0) \wedge \forall x(\varphi(x) \rightarrow \varphi(Sx)) \rightarrow \forall x\varphi(x)$$

over the language $\mathcal{L}_{Tr}^\triangleleft$. To formulate our predicational theory of ground, we additionally rely on standard syntax theory developed in PA . In particular, we use a standard coding function $\#$ to map every expression (variable, term, formula, etc.) σ to a unique code $\#\sigma$ from the natural numbers. If σ is an expression, then we also write $\ulcorner \sigma \urcorner$ for the numeral intended to denote $\#\sigma$. We use function terms $\neg, \wedge, \vee, \exists, \forall, =$ and \triangleleft to represent the corresponding syntactic operations.⁹

Now our predicational theory of ground PG has all the axioms of PAG , plus the following axioms:¹⁰

Basic Ground Axioms:

- G₁ $\forall x\neg(x \triangleleft x)$
- G₂ $\forall x\forall y\forall z(x \triangleleft y \wedge y \triangleleft z \rightarrow x \triangleleft z)$
- G₃ $\forall x\forall y(x \triangleleft y \rightarrow Tr(x) \wedge Tr(y))$

Basic Truth Axioms:

- T₁ $\forall s\forall t(Tr(s \doteq t) \leftrightarrow s^\circ = t^\circ)$
- T₂ $\forall s\forall t(Tr(s \neq t) \leftrightarrow s^\circ \neq t^\circ)$
- T₃ $\forall x(Tr(x) \rightarrow Sent(x))$

Upward Directed Axioms:

- U₁ $\forall x(Tr(x) \rightarrow x \triangleleft \neg\neg x)$
- U₂ $\forall x\forall y((Tr(x) \rightarrow x \triangleleft x \vee y) \wedge (Tr(y) \rightarrow y \triangleleft x \vee y))$
- U₃ $\forall x\forall y(Tr(x) \wedge Tr(y) \rightarrow (x \triangleleft x \wedge y) \wedge (y \triangleleft x \wedge y))$
- U₄ $\forall x\forall y(Tr(\neg x) \wedge Tr(\neg y) \rightarrow (\neg x \triangleleft \neg(x \vee y)) \wedge (\neg y \triangleleft \neg(x \vee y)))$
- U₅ $\forall x\forall y((Tr(\neg x) \rightarrow \neg x \triangleleft \neg(x \wedge y)) \wedge (Tr(\neg y) \rightarrow \neg y \triangleleft \neg(x \wedge y)))$

⁹For a more detailed account of the syntax theory employed throughout the paper, see Section 3 of the first part of this paper or [8, p. 29–35].

¹⁰In formulating these axioms, we make use of the function symbol $^\circ$ to represent the valuation function *val* in the standard model, i.e. $s^\circ = t$ means that the denotation of s is t . Officially, however, this is merely an abbreviation for the more complicated defining formula of the valuation function (which exists by standard representation results). Officially, we cannot have symbol representing the valuation function in our language, since then we’d run the risk of inconsistency. For more on this, compare [8, p. 32].

- U₆ $\forall x \forall t \forall v (Tr(x(t/v)) \rightarrow x(t/v) \triangleleft \exists vx)$
- U₇ $\forall x \forall v (\forall t Tr(\neg x(t/v)) \rightarrow \forall t (\neg x(t/v) \triangleleft \neg \exists vx))$
- U₈ $\forall x \forall v (\forall t (Tr(x(t/v)) \rightarrow \forall t (x(t/v) \triangleleft \forall vx))$
- U₉ $\forall x \forall t \forall v (Tr(\neg x(t/v)) \rightarrow \neg x(t/v) \triangleleft \neg \forall vx)$

Downward Directed Axioms:

- D₁ $\forall x (Tr(\neg \neg x) \rightarrow x \triangleleft \neg \neg x)$
- D₂ $\forall x \forall y (Tr(x \vee y) \rightarrow (Tr(x) \rightarrow x \triangleleft x \vee y) \wedge (Tr(y) \rightarrow y \triangleleft x \vee y))$
- D₃ $\forall x \forall y (Tr(x \wedge y) \rightarrow (x \triangleleft x \wedge y) \wedge (y \triangleleft x \wedge y))$
- D₄ $\forall x \forall y (Tr(\neg(x \wedge y)) \rightarrow (Tr(\neg x) \rightarrow \neg x \triangleleft \neg(x \vee y)) \wedge (Tr(\neg y) \rightarrow \neg y \triangleleft \neg(x \vee y)))$
- D₅ $\forall x \forall y (Tr(\neg(x \vee y)) \rightarrow (\neg x \triangleleft \neg(x \vee y)) \wedge (\neg y \triangleleft \neg(x \vee y)))$
- D₆ $\forall x (Tr(\exists vx(v)) \rightarrow \exists t (x(t/v) \triangleleft \exists vx))$
- D₇ $\forall x \forall v (Tr(\neg \exists vx) \rightarrow \forall t (\neg x(t/v) \triangleleft \neg \exists vx))$
- D₈ $\forall x \forall v (Tr(\forall vx) \rightarrow \forall t (x(t/v) \triangleleft \forall vx))$
- D₉ $\forall x \forall v (Tr(\neg \forall vx) \rightarrow \exists t (\neg x(t/v) \triangleleft \neg \forall vx))$

In the remaining sections of part one of the paper, we show that *PG* is a proof-theoretically conservative extension of the theory *PT* of positive truth (Theorem 1) and provide a model construction (Theorem 2) by extending the standard model of *PT* to a canonical model of *PG*. As an immediate corollary of the conservativity result, we infer that *PG* is consistent (Corollary 3) and that *PG* has the same arithmetic statements as the theory *ACA* of arithmetic comprehension (Corollary 4).

2 The Aristotelian Principle and Typed Truth

Aristotle has the following to say about the interaction of truth and explanation:

It is not because we think truly that you are pale, that you are pale; but because you are pale we who say this have the truth.

(Metaphysics 1051b6–9)

In the context of partial ground, this motivates the following principle: If φ is true, then it is natural to say that φ is true *in virtue of* what it says being the case, and if φ is false, then φ is false *in virtue of* what it says *not* being the case. We will thus call the corresponding informal ground-theoretic principles the *Aristotelian principles* about truth and falsehood respectively.

If we wish to formalize the Aristotelian principles in the context of our predicational theory of ground, we have to make a couple of adjustments. First, we have to assume that we have a Gödel-numbering for the language \mathcal{L}_{Tr} . In particular, we now assume that we have a name $\ulcorner \varphi \urcorner$ for every sentence $\varphi \in \mathcal{L}_{Tr}$. Second, we have to adjust our basic truth axioms. The axioms T₃ and G₃ of *PG* together ensure that both the truth predicate and the ground predicate only apply to sentences of \mathcal{L} . To formalize the Aristotelian principles, however, we need to relax this requirement: we need

for the truth predicate and the ground predicate to apply to sentences of \mathcal{L}_{Tr} that are not already in \mathcal{L} . To allow for this, we adjust the axiom T_3 to the axiom:

$$\forall x(Tr(x) \rightarrow Sent_{Tr}(x)),$$

which we'll label T_3^* . Together with G_3 , this then entails:

$$\forall x\forall y(x \triangleleft y \rightarrow Sent_{Tr}(x) \wedge Sent_{Tr}(y)),$$

as well. We'll refer to the theory that results from replacing T_3 in PG with T_3^* as PGT . Thus, in PGT , we are not only talking about the truths of arithmetic, but also about the truth of the truths of arithmetic.

With these adjustments in place, we can schematically express the Aristotelian principle about truth by saying that for all sentences φ that:

$$\varphi \rightarrow \ulcorner \varphi \urcorner \triangleleft \ulcorner Tr(\ulcorner \varphi \urcorner) \urcorner.$$

For the Aristotelian principle about falsehood, we get:

$$\neg\varphi \rightarrow \ulcorner \neg\varphi \urcorner \triangleleft \ulcorner \neg Tr(\ulcorner \varphi \urcorner) \urcorner,$$

for all sentences φ . Using the same strategy as in the previous paper, we can translate these schemata into the quantified axioms:

- $(AP_T) \quad \forall x(Tr(x) \rightarrow x \triangleleft Tr(\dot{x}))$, and
- $(AP_F) \quad \forall x(Tr(\neg x) \rightarrow \neg x \triangleleft \neg Tr(\dot{x}))$,

where Tr represents the function that maps the code $\#t$ of a term t to the code $\#Tr(t)$ of the application $Tr(t)$ of the truth predicate to the term t . Thus, we have arrived at a quantified axiomatization of the Aristotelian principles.¹¹

Unfortunately, as [6] shows, the Aristotelian principles are *ground-theoretically inconsistent*:¹²

Lemma 1 (Puzzle of Ground) AP_T and AP_F are inconsistent over PGT .

Proof To show that $AP_T \vdash_{PGT} \perp$ consider the following derivation:

1. $\bar{0} = \bar{0}$ (Arithmetic)
2. $Tr(\ulcorner \bar{0} = \bar{0} \urcorner)$ (T-scheme over \mathcal{L})
3. $\ulcorner \bar{0} = \bar{0} \urcorner \triangleleft \ulcorner Tr(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner$ (2., AP_T)
4. $Tr(\ulcorner Tr(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner)$ (3., G_3)
5. $\ulcorner Tr(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner \triangleleft \exists x Tr(x)$ (4., U_6)
6. $Tr(\ulcorner \exists x Tr(x) \urcorner)$ (5., G_3)
7. $\ulcorner \exists x Tr(x) \urcorner \triangleleft \ulcorner Tr(\ulcorner \exists x Tr(x) \urcorner) \urcorner$ (6., AP_T)
8. $Tr(\ulcorner Tr(\ulcorner \exists x Tr(x) \urcorner) \urcorner)$ (7., G_3)

¹¹To illustrate how the quantified axioms work, consider a formula φ such that $Tr(\ulcorner \varphi \urcorner)$. Then the principle AP_T says that $\ulcorner \varphi \urcorner \triangleleft Tr(\ulcorner \varphi \urcorner)$. But the latter is $\ulcorner \varphi \urcorner \triangleleft \ulcorner Tr(\ulcorner \varphi \urcorner) \urcorner$, since $\ulcorner \varphi \urcorner = \ulcorner \ulcorner \varphi \urcorner \urcorner$ and $Tr(\ulcorner \ulcorner \varphi \urcorner \urcorner) = \ulcorner Tr(\ulcorner \varphi \urcorner) \urcorner$. So AP_T allows us to prove $Tr(\ulcorner \varphi \urcorner) \rightarrow \ulcorner \varphi \urcorner \triangleleft \ulcorner Tr(\ulcorner \varphi \urcorner) \urcorner$, for every sentence φ . The quantified principle AP_F works analogously.

¹²Fine works in an operational framework, but his argument can easily be adapted to the present framework. Fine's argument is discussed and refined by [2, 14].

- 9. $\lceil Tr(\lceil \exists x Tr(x) \rceil) \rceil \triangleleft \lceil \exists x Tr(x) \rceil$ (8., U₆)
- 10. $\lceil \exists x Tr(x) \rceil \triangleleft \lceil \exists x Tr(x) \rceil$ (7.,9., G₂)
- 11. $\neg(\lceil \exists x Tr(x) \rceil \triangleleft \lceil \exists x Tr(x) \rceil)$ (G₁)
- 12. \perp (10.,11., \perp -Intro)

To show that $AP_F \vdash_{PGT} \perp$, we can perform an analogous derivation, which is left to the interested reader. □

We are left with a ground-theoretic puzzle about truth.¹³ All the principles involved in the proof of Lemma 1 are intuitively plausible: the basic ground axioms G₂ and G₃ directly arise from the definition of partial ground, the upward directed axiom U₆ about the existential quantifier is plausible in light of the usual semantics for first-order logic, and the Aristotelian principle for truth is plausible from considerations about truth. (The principles required to show that AP_F is ground-theoretically inconsistent are equally plausible.) So what are we to do? In this section, we will propose a solution to the puzzle of ground that preserves the intuition behind all of these principles. We will achieve this by *typing* our truth predicate—a move familiar from *typed theories of truth*.

In typed theories of truth, no applications of the truth predicate to sentences containing the same truth predicate are provable. Thus, typed theories of truth respect Tarski’s distinction between *object-language and meta-language* [20]. Tarski motivates this distinction from the semantic paradoxes, such as the infamous *liar paradox*. This paradox results when we apply the T-scheme to the *liar sentence* which intuitively “says of itself” that it is not true. More specifically, Tarski observed that, in a sufficiently strong background theory, such as *PA*, if we allow applications of the truth predicate to sentences with the same truth predicate in them, we get a sentence λ that is provably equivalent to its own falsehood:

$$\vdash_{PA} \lambda \leftrightarrow \neg Tr(\lceil \lambda \rceil).$$

This follows from the so-called *diagonal lemma*, which is provable in *PA* in the context of an appropriate Gödel-numbering for \mathcal{L}_{Tr} :¹⁴

Lemma 2 *For every sentence $\varphi(x) \in \mathcal{L}_{Tr}$ with exactly one free variable x , there is a sentence $\delta \in \mathcal{L}_{Tr}$ such that*

$$\vdash_{PA} \delta \leftrightarrow \varphi(\lceil \delta \rceil).$$

The existence of the liar sentence λ , then, follows by a simple application of the diagonal lemma to the formula $\neg Tr(x) \in \mathcal{L}_{Tr}$. It is well-known that the existence of a liar sentence is inconsistent with the T-scheme over the language \mathcal{L}_{Tr} .

¹³Fine, in his original paper [6], discusses a range of ground-theoretic puzzles that arise in a similar way from principles similar to the Aristotelian principle. Here we’ll focus on the puzzle about truth and partial ground, because the problem arises most naturally in the present context.

¹⁴For a proof, see for example [1, p. 220–224].

A common intuitive response to the liar paradox is that it somehow arises from the self-reference involved.¹⁵ On this informal view, the problem is that the liar sentence “says something of itself,” namely that it is not true.¹⁶ Thus, so the intuitive response, we should put restrictions on our language that prevent self-reference. Tarski makes this response precise by introducing the distinction between object-language and meta-language. To illustrate the distinction, consider the truths of arithmetic. According to Tarski, if we wish to talk about numbers and their properties, we can do so in the language \mathcal{L} of PA —our object-language for arithmetic. But if we wish to talk about the truth of sentences in \mathcal{L} , we have to do so in the language \mathcal{L}_{Tr} —our meta-language for the truths of arithmetic.¹⁷ Moreover, if we wish to talk about the truth of the truths of arithmetic, i.e. the truths of sentences in \mathcal{L}_{Tr} containing the truth predicate, we need to do so in yet another *meta-meta-language*, which has a distinct truth-predicate for the sentences of \mathcal{L}_{Tr} . And so on. In contrast, Tarski calls a language that can talk about the truths of its own sentences, i.e. a language that has both names for all of its sentences and a truth predicate that applies to these names, *semantically closed*. Thus, a semantically closed language is its own meta-language, as it were, and thus we get self-referential paradoxes. Tarski shows that if we obey the distinction between object-language and meta-language, we can formulate a consistent theory of truth: In an appropriate meta-language, which is not semantically closed, we can consistently affirm the T-scheme for the sentences of the object-language and we never can prove problematic sentences, such as the liar. The liar paradox, on the other hand, shows that if we work in a semantically closed language, disaster ensues: If we have a Gödel-numbering for the terms of \mathcal{L}_{Tr} within \mathcal{L}_{Tr} and at the same time affirm the T-scheme over the sentences of \mathcal{L}_{Tr} , i.e. if we use \mathcal{L}_{Tr} as its own meta-language, we get semantic paradoxes, like the liar paradox. Thus, so Tarski argues, when we wish to talk about truth, we should not never do so in a semantically closed language, but always in an appropriate meta-language. Intuitively, the picture is that semantic truths, such as truths about the truths of arithmetic, are on a “higher level” than non-semantic truths, such as the ordinary truths of arithmetic. Moreover, this can be iterated: the truths about truths about the truths of arithmetic are on yet a “higher level” than the truths about the truths of arithmetic and so on. What emerges is Tarski’s *hierarchy of truths*. Following Tarski, if we work in a semantically closed

¹⁵Here and in the following, we will use a weak concept of self-reference, where a sentence that is (probably) equivalent to a sentence that involves its own Gödel numeral is said to be self-referential. In this sense, the diagonal lemma as stated above gives us a sentence that is self-referential. There is a stronger sense of self-reference, in which a sentence is only self-referential if it is *identical* to a sentence that involves its own Gödel numeral. There are stronger versions of the diagonal lemma that would give us self-reference in this sense as well, but for simplicity’s sake, we’ll stick to the weak concept of self-reference throughout the paper. For a more comprehensive discussion of these issues, see, e.g., [9].

¹⁶However, as [21] shows there are paradoxes without self-reference. Moreover, there are self-referential sentences that are not paradoxical. But still, the intuitive view is that in the case of the liar and similar paradoxes, self-reference plays an essential role.

¹⁷This language is then, of course, an *extension* of the language of PA : it extends the purely arithmetic vocabulary with names for the sentences of \mathcal{L} and a truth predicate for those sentences.

language, we mix the levels of the hierarchy of truths—and ultimately this is the source of the semantic paradoxes.¹⁸

Our original, unmodified predicational theory of ground PG respects Tarski’s distinction between object- and meta-language: We have formulated PG in the language $\mathcal{L}_{Tr}^{\triangleleft}$ in the context of a coding for the language \mathcal{L} of PA . In particular, we have assumed that we have a name $\ulcorner \varphi \urcorner$ for every sentence $\varphi \in \mathcal{L}$, but *not* that we have names $\ulcorner Tr(t) \urcorner$ for sentences of the form $Tr(t) \in \mathcal{L}_{Tr}$ and so on. Moreover, as we have said before, by the axioms T_3 and G_3 , we have ensured that both the truth predicate and the ground predicate only apply to the sentences of \mathcal{L} . Thus, we have used the language $\mathcal{L}_{Tr}^{\triangleleft}$ as an appropriate meta-language for our object-language \mathcal{L} —in compliance with Tarski’s distinction. When we move to the modified theory PGT , however, we no longer conform with Tarski’s distinction. Since PGT is formulated in $\mathcal{L}_{Tr}^{\triangleleft}$ in the context of a Gödel numbering for \mathcal{L}_{Tr} , PGT is formulated in a semantically closed language. Now, the truth predicate may apply to sentences with the same truth predicate in them: Since in PGT we work in the context of a coding for \mathcal{L}_{Tr} , we have names for all the sentences of $\mathcal{L}_{Tr}^{\triangleleft}$ within $\mathcal{L}_{Tr}^{\triangleleft}$ itself. Moreover, by the axiom T_3^* we have allowed for these terms to occur truly in the context of the truth predicate and the ground predicate. In other words, when we formulated PGT , we have used $\mathcal{L}_{Tr}^{\triangleleft}$ as its own meta-language—we talked about the truths of $\mathcal{L}_{Tr}^{\triangleleft}$ within $\mathcal{L}_{Tr}^{\triangleleft}$ itself.

Based on this observation, we argue that the semantic closure of $\mathcal{L}_{Tr}^{\triangleleft}$ is (at least part of) the reason for why the puzzle of ground arises. Note that the semantic closure of $\mathcal{L}_{Tr}^{\triangleleft}$ is required for the proof of Lemma 1. In the third step of the derivation, we applied the ground predicate to the truth predicate in $\ulcorner \bar{0} = \bar{0} \urcorner \triangleleft \ulcorner Tr(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner$. Moreover, in the fourth step, we inferred $Tr(\ulcorner Tr(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner)$ from this and thus applied the truth predicate to a sentence containing the same truth predicate. The main difference between the liar paradox and the puzzle of ground is that, in the case of the liar, we get a *truth-theoretic* inconsistency, i.e. an inconsistency with plausible principles for truth, while in the case of the puzzle, we get a *ground-theoretic* inconsistency, i.e. an inconsistency with plausible principles for partial ground. Still, the problematic sentences in both cases are quite similar. In both cases some intuitive form of self-reference is involved: while the liar sentence λ intuitively *says* something of itself, the principles of partial ground entail that truth of $\exists x Tr(x)$ partially *grounds* itself. Thus, we can say that the self-reference in the case of the liar is *semantic*, while the self-reference in the case of the puzzle of ground is *ground-theoretic*.¹⁹

In analogy to typed theories of truth, we propose a *typed solution* to the puzzle of ground. To formulate this solution, we will move to a slightly different framework, where instead of a single truth predicate Tr , we have a family Tr_1, Tr_2, \dots of *typed* truth predicates. These truth predicates intuitively express truth on the first, second,

¹⁸This also applies to paradoxes without self-reference, such as Yablo’s paradox: Yablo similarly formulates his paradox in a semantically closed language.

¹⁹There is also a kind of semantic self-reference involved in the case of puzzle. The existential quantifier in $\exists x Tr(x)$ semantically ranges over all sentences of \mathcal{L}_{Tr} , including $\exists x Tr(x)$ itself. Thus the truth of $\exists x Tr(x)$ is partially witnessed by the truth of $\exists x Tr(x)$. But here we do not wish to push this point any further.

... level of Tarski's hierarchy. In the remainder of this section, we will develop a *consistent* theory of partial ground and typed truth using these typed truth predicates. This theory will contain typed versions of the axioms of *PG* plus typed versions of the Aristotelian principles. Much like in the case of typed theories of truth, this will mean that no sentence is provable in which the truth predicate is applied to a sentence containing the same truth predicate. We will show that this restriction is sufficient to obtain a *consistent* theory of partial ground and typed truth.

3 Axiomatic Theories of Partial Ground and Typed Truth

Typed theories of truth aim to axiomatize Tarski's hierarchy of truths.²⁰ For this purpose, in typed theories of truth, we have different truth predicates for the different levels of the hierarchy. Correspondingly, we get a hierarchy of languages with a different language for every level of the hierarchy. To illustrate, we start with $\mathcal{L}_0 =_{def} \mathcal{L}$ —the language of *PA*. The truth predicate for sentences of arithmetic is, then, Tr_1 and the language \mathcal{L}_1 extends \mathcal{L}_0 with Tr_1 . The truth predicate for sentences of \mathcal{L}_1 , in turn, is Tr_2 and the language \mathcal{L}_2 extends \mathcal{L}_1 with Tr_2 . And so on. Thus, typed theories of truth are formulated using a hierarchical family of truth predicates Tr_1, Tr_2, \dots that intuitively correspond to truth on the different levels of Tarski's hierarchy.

3.1 Language and Background Theory

We will now formally define a hierarchy of languages, such that on every level we can talk about the truth of sentences on the lower levels. For reasons of generality, we will define this hierarchy in such a way that it includes even infinitary levels. Specifically, we assume that for every ordinal $0 < \alpha < \epsilon_0$, we have a different truth predicate T_α that intuitively expresses truth at the level α : we have $Tr_1, \dots, Tr_\omega, \dots, Tr_{\omega^\omega}, \dots, Tr_{\omega^{\omega^\omega}}, \dots$, where for $\alpha \neq \beta < \epsilon_0$, we have $Tr_\alpha \neq Tr_\beta$.²¹ For all ordinals $1 \leq \alpha \leq \epsilon_0$, we define the language $\mathcal{L}_{<\alpha}$ as the language \mathcal{L} of *PA* extended with all the truth predicates T_β for $0 < \beta < \alpha$:

$$\mathcal{L}_{<\alpha} =_{def} \mathcal{L} \cup \{Tr_\beta \mid 0 < \beta < \alpha\}.$$

Then we set:

$$\mathcal{L}_\alpha =_{def} \mathcal{L}_{<\alpha+1},$$

²⁰For more on axiomatizations of Tarski's hierarchy, see [8, p. 125–29].

²¹We assume that the reader is familiar with the basic theory of ordinals. For the relevant definitions, see [10, p. 17–26], for example. The ordinal ϵ_0 is the limit of the sequence $1, \omega, \omega^\omega, \omega^{\omega^\omega}, \dots$; in other words, ϵ_0 is the first ordinal that satisfies the equation $\omega^x = x$. This ordinal ϵ_0 is still *countable*, i.e. it has the same cardinality as the set of the natural numbers. But it provides a natural stopping point for our infinitary hierarchy, since (i) we can code the ordinals below ϵ_0 and (ii) *PA* represents the well-ordering of the ordinals below ϵ_0 . We will not go into the details here, as the infinitary nature of our hierarchy is not particularly important to our *philosophical* point. Nevertheless, for reasons of generality, we will extend our hierarchy to this level, since ϵ_0 is the limit up to where we can apply the methods of this paper.

for all ordinals $0 \leq \alpha < \epsilon_0$. Thus, the language \mathcal{L}_0 is \mathcal{L} , the language \mathcal{L}_1 is $\mathcal{L} \cup \{Tr_1\}$, and so on. Intuitively, for an ordinal $1 \leq \alpha \leq \epsilon_0$, the language $\mathcal{L}_{<\alpha}$ talks about the truths at the levels strictly below α and \mathcal{L}_α talks about the truths at all levels up to and including α . When we are operating on the ordinal level α , the language $\mathcal{L}_{<\alpha}$ is our intended object-language, i.e. we wish to talk about grounding between the truths of sentences in $\mathcal{L}_{<\alpha}$. For most informal purposes, however, we already stop at the level of $\mathcal{L}_{<2} = \mathcal{L} \cup \{Tr_1\}$. The reason for this is that $\mathcal{L}_{<2}$ is the first language in which grounding between arithmetic truths and truths involving a truth predicate occurs. For all ordinals $1 \leq \alpha \leq \epsilon_0$, the language $\mathcal{L}_{<\alpha}^\triangleleft$ is $\mathcal{L}_{<\alpha}$ extended with our binary ground predicate \triangleleft :

$$\mathcal{L}_{<\alpha}^\triangleleft =_{def} \mathcal{L}_{<\alpha} \cup \{\triangleleft\}.$$

And we set:

$$\mathcal{L}_\alpha^\triangleleft =_{def} \mathcal{L}_\alpha \cup \{\triangleleft\},$$

for all ordinals $0 \leq \alpha < \epsilon_0$. When we are operating on the ordinal level α , we'll use $\mathcal{L}_\alpha^\triangleleft$ as our meta-language for the object-language language $\mathcal{L}_{<\alpha}$. Again, for expository purposes, we'll usually stop at $\mathcal{L}_2^\triangleleft = \mathcal{L}_2 \cup \{\triangleleft\}$, which is the first language in which we can talk about grounding in $\mathcal{L}_{<2} = \mathcal{L}_1$.

For an ordinal $0 < \alpha < \epsilon_0$, the theory $PAT_{<\alpha}$ is the result of extending PA with all the instances of the induction scheme over the language $\mathcal{L}_{<\alpha}$ and the theory $PAG_{<\alpha}$ is the result of extending $PAT_{<\alpha}$ with all the missing instances of the induction scheme over $\mathcal{L}_\alpha^\triangleleft$. For $0 \leq \alpha < \epsilon_0$, the theory PAT_α , then, is $PAT_{<\alpha+1}$ and similarly PAG_α is $PAG_{<\alpha+1}$. Thus, PAT_0 is PAT and PAG_0 is PAG . In PAT_α , we can develop a syntax theory for the languages $\mathcal{L}_{<\alpha}$ analogously to the way we developed the syntax theory in the first part of this paper. When we work on an ordinal level α , we assume that in $\mathcal{L}_\alpha^\triangleleft$, via some appropriate Gödel coding, we have names $\ulcorner \varphi \urcorner$ for all formulas $\varphi \in \mathcal{L}_{<\alpha}$.²² Moreover, we assume that for every $0 < \beta < \alpha$, we have a function symbol Tr_β that represents the function which maps the code $\#t$ of a term t to the code $\#Tr_\beta(t)$ of the formula $Tr_\beta(t) \in \mathcal{L}_{<\alpha}$. And we abbreviate the formula that allows us to (strongly) represent the (set of codes of) sentences in $\mathcal{L}_{<\alpha}$ by $Sent_{<\alpha}$.

3.2 Axioms for Partial Ground and Typed Truth

With the syntax in place, we extend our theory PG to account for partial ground between truths on the same level of Tarski's hierarchy. We'll define this extension from the perspective of some ordinal level $0 < \alpha < \epsilon_0$. Thus, we wish to talk about ground between truths on all the ordinal levels $0 < \beta < \alpha$. To achieve this, we have to modify the basic ground axiom G_3 and the basic truth axioms T_1 , T_2 , and T_3 . The axiom G_3 splits up in the following pair, for all ordinals $0 < \beta < \alpha$:

$$- \quad (G_{3\alpha}^\beta) \quad \forall x \forall y (x \triangleleft y \rightarrow (Sent_{<\beta}(x) \rightarrow Tr_\beta(x)))$$

²²Here it is important that we're restricting ourselves to countable ordinals, because otherwise we would "run out of codes" at some point.

$$- (G_{3b}^\beta) \quad \forall x \forall y (x \triangleleft y \rightarrow (Sent_{<\beta}(y) \rightarrow Tr_\beta(y)))$$

These axioms formalize the factivity of ground in a typed context. In particular, the axiom G_{3a}^β says that if the truth of some sentence grounds the truth of another, and if the sentence is below the level β in the hierarchy, then it is true at level β of the hierarchy. The axiom G_{3b}^β , on the other hand, says the same thing the other way around: if the truth of some sentence is grounded in the truth of another, and if the former sentence is below level β , then it is true at level β . To illustrate, if we let $\alpha = 2$, we get the following axiom pair:

$$- (G_{3a}^1) \quad \forall x \forall y (x \triangleleft y \rightarrow (Sent_{<1}(x) \rightarrow Tr_1(x)))$$

$$- (G_{3b}^1) \quad \forall x \forall y (x \triangleleft y \rightarrow (Sent_{<1}(y) \rightarrow Tr_1(y)))$$

Thus, for sentences $\varphi, \psi \in \mathcal{L}_{<1}$, the axioms G_{3a}^1 and G_{3b}^1 together say that the truth of φ can only ground the truth of ψ , if φ and ψ are both true at level one—i.e. if they are truths of arithmetic.

Next, in the typed versions of T_1 and T_2 , we wish to make sure that true equations are true at every level below α . Thus, for all ordinals $0 < \beta < \alpha$, we postulate:

$$- (T_1^\beta) \quad \forall s \forall t (Tr_\beta(s \doteq t) \leftrightarrow s^\circ = t^\circ)$$

$$- (T_2^\beta) \quad \forall s \forall t (Tr_\beta(s \not\doteq t) \leftrightarrow s^\circ \neq t^\circ)$$

Thus, we get for example $Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner)$, $Tr_2(\ulcorner \bar{0} = \bar{0} \urcorner)$, . . . and so on, for all ordinal levels below α .

In the case of T_3 , we wish to make sure that a truth predicate Tr_β , for an ordinal $0 < \beta < \alpha$, only applies to sentences on levels below β —in compliance with Tarski’s distinction. Thus, we postulate for all $0 < \beta < \alpha$:

$$- (T_3^\beta) \quad \forall x (Tr_\beta(x) \rightarrow Sent_{<\beta}(x))$$

Thus, for example, we get $\forall x (Tr_1(x) \rightarrow Sent_{<1}(x))$, which intuitively says that the predicate Tr_1 only applies to sentences of arithmetic.

Another adjustment is needed: Note that now $G_{3a/b}^\beta$ and T_3^β do *not* entail anymore that the ground predicate applies only to sentences. To ensure this, we postulate the following final basic ground axiom:

$$- (G_4^\alpha) \quad \forall x \forall y (x \triangleleft y \rightarrow Sent_{<\alpha}(x) \wedge Sent_{<\alpha}(y))$$

Thus, on the level $\alpha = 2$, we get that $\forall x \forall y (x \triangleleft y \rightarrow Sent_{<2}(x) \wedge Sent_{<2}(y))$. In words: if $x \triangleleft y$, then both x and y are sentences of $\mathcal{L}_{<2}$, which is just the language of arithmetic \mathcal{L} extended with the truth predicate Tr_1 . Taken together, all of these modified axioms entail that our new theory respects Tarski’s distinction between object- and meta-language. Finally, we have to modify our upwards and downwards directed grounding axioms to apply on all levels below α . We achieve this by postulating that the axioms apply on all of these levels. Take the axioms U_1 and D_1 for example. They become the new typed set of axioms for all ordinals $0 < \beta < \alpha$:

$$- (U_1^\beta) \quad \forall x (Tr_\beta(x) \rightarrow x \triangleleft \ulcorner \neg \neg x \urcorner)$$

$$- (D_1^\beta) \quad \forall x (Tr_\beta(\ulcorner \neg \neg x \urcorner) \rightarrow x \triangleleft \ulcorner \neg \neg x \urcorner)$$

Thus, on every level $0 < \beta < \alpha$, if a sentence is true on that level, then the sentence grounds its double-negation and if a double negation is true on the level, it is grounded by the sentence it is a double negation of.

Putting all of this together, we get:

Definition 1 For all ordinals $0 \leq \alpha \leq \epsilon_0$, the *predicational theory* $PG_{<\alpha}$ of ground up to α , consists of the axioms of $PAG_{<\alpha}$ plus the following axioms for all $0 < \beta < \alpha$:

Typed Ground Axioms:

- $G_1 \quad \forall x \neg(x \triangleleft x)$
- $G_2 \quad \forall x \forall y \forall z (x \triangleleft y \wedge y \triangleleft z \rightarrow x \triangleleft z)$
- $G_{3a}^\beta \quad \forall x \forall y (x \triangleleft y \rightarrow (Sent_{<\beta}(x) \rightarrow Tr_\beta(x)))$
- $G_{3b}^\beta \quad \forall x \forall y (x \triangleleft y \rightarrow (Sent_{<\beta}(y) \rightarrow Tr_\beta(y)))$
- $G_4^\alpha \quad \forall x \forall y (x \triangleleft y \rightarrow Sent_{<\alpha}(x) \wedge Sent_{<\alpha}(y))$

Typed Truth Axioms:

- $T_1^\beta \quad \forall s \forall t (Tr_\beta(s \doteq t) \leftrightarrow s^\circ = t^\circ)$
- $T_2^\beta \quad \forall s \forall t (Tr_\beta(s \neq t) \leftrightarrow s^\circ \neq t^\circ)$
- $T_3^\beta \quad \forall x (Tr_\beta(x) \rightarrow Sent_{<\beta}(x))$

Typed Upward Directed Axioms:

- $U_1^\beta \quad \forall x (Tr_\beta(x) \rightarrow x \triangleleft \neg\neg x)$
- $U_2^\beta \quad \forall x \forall y (Tr_\beta(x) \rightarrow x \triangleleft x \vee y \wedge Tr_\beta(y) \rightarrow y \triangleleft x \vee y)$
- $U_3^\beta \quad \forall x \forall y (Tr_\beta(x) \wedge Tr_\beta(y) \rightarrow (x \triangleleft x \wedge y) \wedge (y \triangleleft x \wedge y))$
- $U_4^\beta \quad \forall x \forall y (Tr_\beta(\neg x) \wedge Tr_\beta(\neg y) \rightarrow (\neg x \triangleleft \neg(x \vee y)) \wedge (\neg y \triangleleft \neg(x \vee y)))$
- $U_5^\beta \quad \forall x \forall y (Tr_\beta(\neg x) \rightarrow \neg x \triangleleft \neg(x \wedge y) \wedge Tr_\beta(\neg y) \rightarrow \neg y \triangleleft \neg(x \wedge y))$
- $U_6^\beta \quad \forall x \forall t \forall v (Tr_\beta(x(t/v)) \rightarrow x(t/v) \triangleleft \exists v x)$
- $U_7^\beta \quad \forall x \forall v (\forall t Tr_\beta(\neg x(t/v)) \rightarrow \forall t (\neg x(t/v) \triangleleft \neg \exists v x))$
- $U_8^\beta \quad \forall x \forall v (\forall t (Tr_\beta(x(t/v)) \rightarrow \forall t (x(t/v) \triangleleft \forall v x))$
- $U_9^\beta \quad \forall x \forall t \forall v (Tr_\beta(\neg x(t/v)) \rightarrow \neg x(t/v) \triangleleft \neg \forall v x)$

Typed Downward Directed Axioms:

- $D_1 \quad \forall x (Tr_\beta(\neg\neg x) \rightarrow x \triangleleft \neg\neg x)$
- $D_2 \quad \forall x \forall y (Tr_\beta(x \vee y) \rightarrow (Tr_\beta(x) \rightarrow x \triangleleft x \vee y) \wedge (Tr_\beta(y) \rightarrow y \triangleleft x \vee y))$
- $D_3 \quad \forall x \forall y (Tr_\beta(x \wedge y) \rightarrow (x \triangleleft x \wedge y) \wedge (y \triangleleft x \wedge y))$
- $D_4 \quad \forall x \forall y (Tr_\beta(\neg(x \wedge y)) \rightarrow (Tr_\beta(\neg x) \rightarrow \neg x \triangleleft \neg(x \vee y)) \wedge (Tr_\beta(\neg y) \rightarrow \neg y \triangleleft \neg(x \vee y)))$
- $D_5 \quad \forall x \forall y (Tr_\beta(\neg(x \vee y)) \rightarrow (\neg x \triangleleft \neg(x \vee y)) \wedge (\neg y \triangleleft \neg(x \vee y)))$
- $D_6 \quad \forall x (Tr_\beta(\exists v x(v)) \rightarrow \exists t (x(t/v) \triangleleft \exists v x))$
- $D_7 \quad \forall x \forall v (Tr_\beta(\neg \exists v x) \rightarrow \forall t (\neg x(t/v) \triangleleft \neg \exists v x))$
- $D_8 \quad \forall x \forall v (Tr_\beta(\forall v x) \rightarrow \forall t (x(t/v) \triangleleft \forall v x))$
- $D_9 \quad \forall x \forall v (Tr_\beta(\neg \forall v x) \rightarrow \exists t (\neg x(t/v) \triangleleft \neg \forall v x))$

For $0 \leq \alpha < \epsilon_0$, we define PG_α as $PG_{<\alpha+1}$.

To illustrate what PG_α looks like for different α 's, let's consider at a few examples. First, note that PG_0 is PAG . Next, note PG_1 is a functional analog of our original theory PG , where the truth-predicate has been "renamed" Tr_1 . In particular, we get that PG_1 proves the theory PT of positive truth.

Since for all $1 < \alpha < \epsilon_0$, PG_α contains PG_1 , we can say that PG_α essentially is (in the precise sense sketched above) an extension of PG . For α bigger than one, PG_α essentially consists of α -many copies of PG , one for every \mathcal{L}_β and truth predicate Tr_β , where $1 < \beta < \alpha$. What is new in those theories is that now (names of) sentences involving the truth predicate may occur in the context of the ground predicate and other truth predicates—as long as we respect the typing restriction that for all $0 < \beta \leq \alpha$, if $Tr_\beta(\ulcorner \varphi \urcorner)$, then $Sent_{<\beta}(\ulcorner \varphi \urcorner)$. For example, in PG_2 , we get the following instance of U_1^2 :

$$Tr_2(\ulcorner Tr_1(\overline{0} = \overline{0}) \urcorner) \rightarrow \ulcorner Tr_1(\overline{0} = \overline{0}) \urcorner \triangleleft \ulcorner \neg\neg Tr_1(\overline{0} = \overline{0}) \urcorner.$$

Indeed, using G_{3b}^2 we can infer from this that:

$$\ulcorner Tr_1(\overline{0} = \overline{0}) \urcorner \triangleleft \ulcorner \neg\neg Tr_1(\overline{0} = \overline{0}) \urcorner \rightarrow Tr_2(\ulcorner \neg\neg Tr_1(\overline{0} = \overline{0}) \urcorner),$$

which together with the previous formula gives us:

$$Tr_2(\ulcorner Tr_1(\overline{0} = \overline{0}) \urcorner) \rightarrow Tr_2(\ulcorner \neg\neg Tr_1(\overline{0} = \overline{0}) \urcorner).$$

The other direction:

$$Tr_2(\ulcorner \neg\neg Tr_1(\overline{0} = \overline{0}) \urcorner) \rightarrow Tr_2(\ulcorner Tr_1(\overline{0} = \overline{0}) \urcorner)$$

can be shown analogously using D_1^2 and G_{3a}^2 . Generalizing this idea, we get more substantial truth-theoretic theorems in PG_2 , such as:

$$\forall x(Tr_2(\ulcorner Tr_1(\dot{x}) \urcorner) \leftrightarrow Tr_2(\ulcorner \neg\neg Tr_1(\dot{x}) \urcorner)),$$

for example. But so far, PG_2 does not allow us to prove any theorems of the form $Tr_2(\ulcorner Tr_1(\ulcorner \varphi \urcorner) \urcorner)$, where $\varphi \in \mathcal{L}$. In other words, we can't prove the truth of any sentence involving a truth predicate—even if they respect the typing restrictions. Thus, PG_2 is not really a theory of truth at level 2 yet—it can't even show that $Tr_2(\ulcorner Tr_1(\overline{0} = \overline{0}) \urcorner)$. Moreover, in PG_2 , we can't prove any theorems of the form $\ulcorner \varphi \urcorner \triangleleft \ulcorner Tr_1(\ulcorner \varphi \urcorner) \urcorner$ or the like, where the ground predicate applies to sentence involving a truth predicate. This doesn't change on any level $\alpha > 2$. To get a more substantial theory of ground and partial truth, we need to say something about the grounds of truths involving the truth predicate: we need typed versions of the Aristotelian principles.

Typing the Aristotelian principles for use on an ordinal level α is pretty straightforward. We get the following axioms for every $\gamma < \alpha$:

- $(APU_T^\gamma) \quad \forall x(Tr_\gamma(x) \rightarrow x \triangleleft Tr_\gamma(\dot{x}))$
- $(APU_F^\gamma) \quad \forall x(Tr_\gamma(\neg x) \rightarrow \neg x \triangleleft \neg Tr_\gamma(\dot{x}))$

The axiom APU_T^1 , for example, allows us to derive that $\ulcorner \overline{0} = \overline{0} \urcorner \triangleleft \ulcorner Tr_1(\overline{0} = \overline{0}) \urcorner$ using the fact that by axiom T_1^1 we have $Tr_1(\overline{0} = \overline{0})$. The axioms $APU_{T/F}^\beta$ are upwards directed axioms. For analogous reasons as in the case of the other ground axioms, we also need downward directed axioms for the Aristotelian principles. Again, straight-forwardly, we get for all $\gamma < \beta \leq \alpha$:

- $(APD_T^{\beta,\gamma}) \quad \forall x(Tr_\beta(Tr_\gamma(\dot{x})) \rightarrow x \triangleleft Tr_\gamma(\dot{x}))$

$$- \text{ (APD}_F^{\beta,\gamma} \text{)} \quad \forall x (Tr_\beta(\neg Tr_\gamma(\dot{x})) \rightarrow \neg x \triangleleft \neg Tr_\gamma(\dot{x}))$$

If we add the upward and downward directed versions of the Aristotelian principles to the previous theory, we arrive at our theory of ground and typed truth:

Definition 2 For every ordinal $0 \leq \alpha < \epsilon_0$, the theory PGA_α of *partial ground with the Aristotelian principles up to α* consists of the axioms of PG_α plus the following axioms for all $\gamma < \beta \leq \alpha$:

Upward Directed Aristotelian Principles:

$$- \text{ (APU}_T^\gamma \text{)} \quad \forall x (Tr_\gamma(x) \rightarrow x \triangleleft Tr_\gamma(\dot{x}))$$

$$- \text{ (APU}_F^\gamma \text{)} \quad \forall x (Tr_\gamma(\neg x) \rightarrow \neg x \triangleleft \neg Tr_\gamma(\dot{x}))$$

Downward Directed Aristotelian Principles:

$$- \text{ (APD}_T^{\beta,\gamma} \text{)} \quad \forall x (Tr_\beta(Tr_\gamma(\dot{x})) \rightarrow x \triangleleft Tr_\gamma(\dot{x}))$$

$$- \text{ (APD}_F^{\beta,\gamma} \text{)} \quad \forall x (Tr_\beta(\neg Tr_\gamma(\dot{x})) \rightarrow \neg x \triangleleft \neg Tr_\gamma(\dot{x}))$$

The theory $PGA_{<\alpha}$ is defined as $\bigcup_{\beta < \alpha} PGA_\beta$, for all $0 < \alpha \leq \epsilon_0$.

To see how PGA_α works for different $0 \leq \alpha < \epsilon_0$, let's consider again a few examples. First note that PGA_0 is PG_0 which is just PAG . Similarly, PGA_1 is PG_1 . Things get interesting at the level PGA_2 . Here we get:

$$Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \rightarrow \ulcorner \bar{0} = \bar{0} \urcorner \triangleleft \ulcorner Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner,$$

by instantiating the axiom APU_T^1 with the term $\ulcorner \bar{0} = \bar{0} \urcorner$. Moreover, by instantiating the axiom T_1^1 with the same term, we have:

$$Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner).$$

So putting the two together, we get:

$$\ulcorner \bar{0} = \bar{0} \urcorner \triangleleft \ulcorner Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner.$$

Now, using the instance:

$$\ulcorner \bar{0} = \bar{0} \urcorner \triangleleft \ulcorner Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner \rightarrow (Sent_{<2}(\ulcorner Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner) \rightarrow Tr_2(\ulcorner Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner))$$

of the axiom G_{3b}^2 , and since:

$$Sent_{<2}(\ulcorner Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner)$$

is derivable in PA , we can infer:

$$Tr_2(\ulcorner Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner).$$

So, in PGA_2 , we can indeed derive applications of the truth predicate to sentences with a truth predicate in them. Moreover, by putting APU_T^1 :

$$\forall x (Tr_1(x) \rightarrow x \triangleleft Tr_1(\dot{x}))$$

and $APD_T^{2,1}$:

$$\forall x (Tr_2(Tr_1(\dot{x})) \rightarrow x \triangleleft Tr_1(\dot{x}))$$

together, we can actually prove:

$$\forall x(Tr_2(Tr_1(\dot{x})) \leftrightarrow Tr_1(x))$$

using the axioms $G_{3a/b}^1$ and $T_{1/2}^1$.²³ Thus, PGA_2 proves intuitive truths at level two, such as $Tr_2(\ulcorner Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner)$, as well as quite substantial truth-theoretic principles, such as $\forall x(Tr_2(Tr_1(\dot{x})) \leftrightarrow Tr_1(x))$. In other words, PG_2 proves something that looks like a substantial theory of truth at level two of Tarski’s hierarchy. In the next section, we will show that for $0 < \alpha < \epsilon_0$, PGA_α proves the theory PRT_α of *positive ramified truth up to α* . Indeed, we can show that PGA_α is a *proof-theoretically conservative extension* of PRT_α .

3.3 Conservativity and Models

The theory $PT_{<\alpha}$ of positive ramified truth up to an ordinal level $1 \leq \alpha \leq \epsilon_0$ is formulated in the language $\mathcal{L}_{<\alpha}$ and it is the result of modifying the theory of typed truth with the typed versions of its axioms in a similar way as we developed $PG_{<\alpha}$:

Definition 3 (‘Positive Ramified Truth’) For all ordinals $1 \leq \alpha \leq \epsilon_0$, the theory $PRT_{<\alpha}$ of *positive ramified truth up to α* consists of the axioms of $PAT_{<\alpha}$ plus the following axioms for all $\gamma < \beta < \alpha$:

Typed Truth Axioms:

$$T_1^\beta \quad \forall s \forall t (Tr_\beta(s \doteq t) \leftrightarrow s^\circ = t^\circ)$$

$$T_2^\beta \quad \forall s \forall t (Tr_\beta(s \not\doteq t) \leftrightarrow s^\circ \neq t^\circ)$$

$$T_3^\beta \quad \forall x (Tr_\beta(x) \rightarrow Sent_{<\beta}(x))$$

Positive Ramified Truth Axioms:

$$RP_1^\beta \quad \forall x (Tr_\beta(x) \leftrightarrow Tr_\beta(\ulcorner \neg \neg x \urcorner))$$

$$RP_2^\beta \quad \forall x \forall y (Tr_\beta(x \wedge y) \leftrightarrow Tr_\beta(x) \wedge Tr_\beta(y))$$

$$RP_3^\beta \quad \forall x \forall y (Tr_\beta(\ulcorner \neg(x \wedge y) \urcorner) \leftrightarrow Tr_\beta(\ulcorner \neg x \vee \neg y \urcorner))$$

$$RP_4^\beta \quad \forall x \forall y (Tr_\beta(x \vee y) \leftrightarrow Tr_\beta(x) \vee Tr_\beta(y))$$

$$RP_5^\beta \quad \forall x \forall y (Tr_\beta(\ulcorner \neg(x \vee y) \urcorner) \leftrightarrow Tr_\beta(\ulcorner \neg x \wedge \neg y \urcorner))$$

$$RP_6^\beta \quad \forall x \forall v (Tr_\beta(\forall v x) \leftrightarrow \forall t Tr_\beta(x(t/v)))$$

$$RP_7^\beta \quad \forall x \forall v (Tr_\beta(\ulcorner \neg \forall v x \urcorner) \leftrightarrow \exists t Tr_\beta(\ulcorner \neg x(t/v) \urcorner))$$

$$RP_8^\beta \quad \forall x \forall v (Tr_\beta(\exists v x) \leftrightarrow \exists t Tr_\beta(x(t/v)))$$

$$RP_9^\beta \quad \forall x \forall v (Tr_\beta(\ulcorner \neg \exists v x \urcorner) \leftrightarrow \forall t Tr_\beta(\ulcorner \neg x(t/v) \urcorner))$$

$$RP_{10}^\beta \quad \forall x (Tr_\beta(Tr_\gamma \dot{x}) \leftrightarrow Tr_\gamma(x))$$

$$RP_{11}^\beta \quad \forall x (Tr_\beta(\ulcorner \neg Tr_\gamma \dot{x} \urcorner) \leftrightarrow Tr_\gamma(\ulcorner \neg x \urcorner))$$

$$RP_{12}^{\gamma, \beta} \quad \forall x (Sent_{<\gamma}(x) \rightarrow (Tr_\beta(Tr_\gamma \dot{x}) \leftrightarrow Tr_\beta(x)))$$

²³For the proof note that if we assume that $Tr_1(t)$ for a term t , it follows by axiom T_1^1 that $Sent_{<1}(t)$ and thus we can prove in PA that $Sent_{<2}(Tr_1(t))$. Similarly, if we assume $Tr_2(Tr_1 \dot{i})$, we can prove that $Sent_{<1}(t)$ by T_3^2 and PA .

$$RP_{13}^{\gamma, \beta} \quad \forall x (Sent_{<\gamma}(x) \rightarrow (Tr_{\beta}(\neg Tr_{\gamma} \dot{x}) \leftrightarrow Tr_{\beta}(\neg x)))$$

The theory PRT_{α} , for $0 \leq \alpha < \epsilon_0$, is defined as $PRT_{<\alpha+1}$.

Note that the theory PRT_1 is a functional analog of PT in the same way that PG_1 is a functional analog of PG . The theory $PRT_{<\alpha}$, for $1 < \alpha \leq \epsilon_0$, however, is a much stronger theory of truth than PT —it formalizes the Tarskian hierarchy up to the level α .²⁴ For example, PGT_2 contains the axioms:

$$Tr_2(\ulcorner Tr_1(\ulcorner \bar{0} = \bar{0} \urcorner) \urcorner),$$

and

$$\forall x (Tr_2(Tr_1(\dot{x})) \leftrightarrow Tr_1(x)),$$

just like PGA_2 . Indeed, we get:

Proposition 1 *For all ordinals $1 \leq \alpha \leq \epsilon_0$, the theory $PGA_{<\alpha}$ proves the theory $PRT_{<\alpha}$: $PGA_{<\alpha} \vdash PRT_{<\alpha}$.*

Proof In large parts, the proof is analogous to the proof of Proposition 4.6 of the first part of this paper: we carry out the same argument for all $\beta < \alpha$. The interesting cases are the new axioms RP_{10-13}^{β} , for $\beta < \alpha$, which can be shown from the typed Aristotelian principles $APU_{T/F}^{\gamma}$ and $APD_{T/F}^{\beta, \gamma}$, for $\gamma < \beta \leq \alpha$, the typed truth axiom T_3^{β} , and the typed ground axioms $G_{3a/b}^{\beta}$, for $\beta < \alpha$, and G_4^{α} . Here we only show how to derive RP_{10}^{β} and RP_{11}^{β} , as the other axioms are analogous:

- $\vdash_{PGA_{<\alpha}} \forall x (Tr_{\beta}(Tr_{\gamma} \dot{x}) \leftrightarrow Tr_{\gamma}(x))$ for $\gamma < \beta < \alpha$

Let x be a fresh variable for \forall -Intro. We now prove both directions of the biconditional $Tr_{\beta}(Tr_{\gamma} \dot{x}) \leftrightarrow Tr_{\gamma}(x)$.

(\Rightarrow): Assume (*) $Tr_{\beta}(Tr_{\gamma}(x))$ for a \rightarrow -Intro. By T_3^{β} , we can derive $Sent_{<\beta}(Tr_{\gamma}(x))$. Using PA and T_3^{γ} , we can derive (**) $Sent_{<\gamma}(x)$ from this.

Moreover, using (*) and $APD_T^{\beta, \gamma}$:

$$\forall x (Tr_{\beta}(Tr_{\gamma}(\dot{x})) \rightarrow x \triangleleft Tr_{\gamma}(\dot{x})),$$

we can derive $x \triangleleft Tr_{\gamma}(\dot{x})$. Using (**) and G_4^{γ} :

$$\forall x \forall y (x \triangleleft y \rightarrow (Sent_{<\gamma}(x) \rightarrow Tr_{\gamma}(x))),$$

²⁴ We can formulate a slightly stronger version of $PRT_{<\alpha}$ by replacing the schematic axioms RP_{12}^{β} and RP_{13}^{β} with the quantified axioms: $\forall x \forall \gamma \ulcorner \beta \urcorner (Sent_{<\gamma}(x) \rightarrow (Tr_{\beta}(Tr_{\gamma} \dot{x}) \leftrightarrow Tr_{\beta}(x)))$ and $\forall x \forall \gamma \ulcorner \beta \urcorner (Sent_{<\gamma}(x) \rightarrow (Tr_{\beta}(\neg Tr_{\gamma} \dot{x}) \leftrightarrow Tr_{\beta}(\neg x)))$, where $\forall \gamma$ quantifies over codes of ordinals, $\ulcorner \beta \urcorner$ is a term for a code of the ordinal β , and $<$ represents the well-ordering on ordinals. To properly formulate these axioms, we require a coding of the ordinals up to ϵ_0 , a representation of the natural well-ordering of these ordinals, and a justification for quantifying into ordinal indexes in $Sent_{<\gamma}$ and Tr_{γ} . We will discuss such a coding below, but for reasons of perspicuity, we will stick with the slightly weaker schematic version of $PRT_{<\alpha}$. Also, the (schematic) theory $PRT_{<\alpha}$ is equivalent to the (schematic) theory $RT_{<\alpha}$ of ramified truth up to α , which is the typed version of CT . As in the case of PT and CT , we will take the meta-theorems of $RT_{<\alpha}$ and apply them immediately to $PRT_{<\alpha}$.

we can in turn derive: $Tr_\gamma(x)$. Thus, we get $Tr_\beta(Tr_\gamma(x)) \rightarrow Tr_\gamma(x)$ by \rightarrow -Intro.

(\Leftarrow): Assume (\dagger) $Tr_\gamma(x)$ for another \rightarrow -Intro. Using PA and G_4^γ , we get $Sent_{<\gamma}(x)$. From this and PA , we can derive for all $\gamma < \beta < \alpha$ that ($\dagger\dagger$) $Sent_{<\beta}(Tr_\gamma(\dot{x}))$. Moreover, using (\dagger) and APU_T^γ :

$$\forall x(Tr_\gamma(x) \rightarrow x \triangleleft Tr_\gamma(\dot{x})),$$

we get $x \triangleleft Tr_\gamma(\dot{x})$. From this, using ($\dagger\dagger$) and G_{3b}^β , we get $Tr_\beta(Tr_\gamma(x))$ and thus $Tr_\gamma(x) \rightarrow Tr_\beta(Tr_\gamma(x))$ by \rightarrow -Intro.

Putting both " \Rightarrow " and " \Leftarrow " together, we get $Tr_\beta(Tr_\gamma(\dot{x})) \leftrightarrow Tr_\gamma(x)$ by \leftrightarrow -Intro. And, since x was a fresh variable, we can derive:

$$\forall x(Tr_\beta(Tr_\gamma(\dot{x})) \leftrightarrow Tr_\gamma(x))$$

by \forall -Intro as desired.

- $\vdash_{PGA_{<\alpha}} \forall x(Sent_{<\gamma}(x) \rightarrow (Tr_\beta(Tr_\gamma(\dot{x})) \leftrightarrow Tr_\beta(x)))$ for $\gamma < \beta < \alpha$.

Let x be a fresh variable for \forall -Intro. Assume $Sent_{<\gamma}(x)$ for a \rightarrow -Intro. We now prove both directions of the biconditional $Tr_\beta(Tr_\gamma(\dot{x})) \leftrightarrow Tr_\beta(x)$.

(\Rightarrow): Assume (\ddagger) $Tr_\beta(Tr_\gamma(\dot{x}))$ for \rightarrow -Intro. From this, T_3^β , and PA , we can derive $Sent_{<\beta}(Tr_\gamma(\dot{x}))$ and thus also ($\ddagger\ddagger$) $Sent_{<\beta}(x)$. As before, we get $x \triangleleft Tr_\gamma(\dot{x})$ using $APD_T^{\beta,\gamma}$. Using ($\ddagger\ddagger$) and G_{3a}^β , we can derive $Tr_\beta(x)$. Thus, we have $Tr_\beta(Tr_\gamma(\dot{x})) \rightarrow Tr_\beta(x)$ by \rightarrow -Intro.

(\Leftarrow): Assume $Tr_\beta(x)$ for yet another \rightarrow -Intro. From this and APU_T^β , we get $x \triangleleft Tr_\beta(x)$. Since we have assumed $Sent_{<\gamma}(x)$, we get $Tr_\gamma(x)$ from this and G_{3a}^γ . From this and APU_T^γ , we get (\S) $x \triangleleft Tr_\gamma(\dot{x})$. But now since $\gamma < \beta$, we can show in PA that $Sent_{<\beta}(Tr_\gamma(\dot{x}))$. But from this and (\S), we get $Tr_\beta(Tr_\gamma(\dot{x}))$. So, we have $Tr_\beta(x) \rightarrow Tr_\beta(Tr_\gamma(\dot{x}))$ by \rightarrow -Intro.

Now putting both " \Rightarrow " and " \Leftarrow " together, we get $Tr_\beta(Tr_\gamma(\dot{x})) \leftrightarrow Tr_\beta(x)$ by \leftrightarrow -Intro and so $Sent_{<\gamma}(x) \rightarrow (Tr_\beta(Tr_\gamma(\dot{x})) \leftrightarrow Tr_\beta(x))$ by \rightarrow -Intro. Finally, since x was a fresh variable, we have:

$$\forall x(Sent_{<\gamma}(x) \rightarrow (Tr_\beta(Tr_\gamma(\dot{x})) \leftrightarrow Tr_\beta(x))),$$

as desired. □

This has the immediate consequence that for all ordinals $\beta < \alpha < \epsilon_0$, the theory $PGA_{<\alpha}$ proves the following typed version of the T-scheme for all languages $\mathcal{L}_{<\beta}$:

Lemma 3 For all ordinals $0 < \gamma \leq \beta < \alpha < \epsilon_0$ and for all sentences $\varphi \in \mathcal{L}_{<\gamma}$:

$$\vdash_{PGA_{<\alpha}} \forall t_1, \dots, \forall t_n(Tr_\beta(\ulcorner \varphi(t_1, \dots, t_n) \urcorner) \leftrightarrow \varphi(t_1, \dots, t_n)).$$

Next, we will now show that for all ordinals $1 \leq \alpha < \epsilon_0$, the theory PGA_α is a proof-theoretically conservative extension of the theory PRT_α . But first, we need to introduce some more technical preliminaries: It is well-known that we can extend

the technique of Gödel numbering to get terms for all ordinals below ϵ_0 [17, p. 17–42]. Let’s denote the set of all ordinals below ϵ_0 by $On_{<\epsilon_0}$. We can adjust our coding function $\# : \mathcal{L} \rightarrow \mathbb{N}$ such that we injectively assign every ordinal $\alpha \in On_{<\epsilon_0}$ a unique code $\#\alpha \in \mathbb{N}$ that is different from all the codes $\#\sigma$ of the other expressions σ of \mathcal{L} . For all $\alpha \in On_{<\epsilon_0}$, we define the term $\ulcorner \alpha \urcorner$ to be $\#\alpha$, i.e. our term for α is the numeral of the code $\#\alpha$ of α . Moreover, we extend the axioms of ordinary arithmetic to cover ordinal arithmetic up to ϵ_0 . For simplicity, we’ll use the same terminology for ordinal arithmetic and ordinary arithmetic. Thus, for example, we can now write $\ulcorner \alpha \urcorner \times \ulcorner \beta \urcorner$ in \mathcal{L} to denote the product of ordinals $\alpha, \beta \in On_{<\epsilon_0}$. Moreover, we get: $\vdash_{PA} \ulcorner \alpha \urcorner \times \ulcorner \beta \urcorner = \ulcorner \gamma \urcorner$ iff $\alpha \times \beta = \gamma$, for all ordinals $\alpha, \beta, \gamma \in On_{<\epsilon_0}$. PA can represent the set of codes of ordinals below ϵ_0 and we’ll use $On_{<\epsilon_0}$ as a predicate for this. In particular, we get for all natural numbers $n \in \mathbb{N}$: $\vdash_{PA} On_{<\epsilon_0}(\bar{n})$ iff $n \in \#On_{<\epsilon_0} = \{\#\alpha \mid \alpha \in On_{<\epsilon_0}\}$. Finally, PA can represent the standard well-ordering $<$ of the ordinals below ϵ_0 and we’ll use the relation symbol $<$ to represent this ordering. So we get that for all ordinals $\alpha, \beta \in On_{<\epsilon_0}$: $\vdash_{PA} \ulcorner \alpha \urcorner < \ulcorner \beta \urcorner$ iff $\alpha < \beta$. With these preliminaries in place,²⁵ we’ll define a slightly non-standard notion of complexity for the formulas in $\mathcal{L}_{<\epsilon_0}$:

Definition 4 (‘ ω -complexity’) For all ordinals $1 \leq \alpha \leq \epsilon_0$, we define the function $|\cdot|_\omega : \mathcal{L}_{<\alpha} \rightarrow On_{<\epsilon_0}$ that assigns to every formula $\varphi \in \mathcal{L}_\alpha$ its ω -complexity $|\varphi|_\omega$ recursively by saying that:

- (i) $|\varphi|_\omega = \begin{cases} \omega \times \alpha & \text{if } \varphi = Tr_\alpha(t) \\ 0 & \text{if } \varphi \text{ is another atomic formula} \end{cases}$
- (ii) $|\neg\varphi|_\omega = |\varphi|_\omega + 1$;
- (iii) $|\varphi \circ \psi|_\omega = lub(|\varphi|_\omega, |\psi|_\omega) + 1$, for $\circ = \wedge, \vee$;²⁶ and
- (iv) $|Qx\varphi|_\omega = |\varphi|_\omega + 1$, for $Q = \forall, \exists$.

Note that ω -complexity agrees with ordinary complexity on the formulas of $\mathcal{L}_{<1}$. Note furthermore that the function $x \mapsto \omega \times x$ is strictly monotonically increasing on the ordinals below ϵ_0 :

Lemma 4 For all $\alpha, \beta \in On_{<\epsilon_0}$, if $\alpha < \beta$, then $\omega \times \alpha < \omega \times \beta$.

Note that as a consequence, we get that for all ordinals $0 < \alpha < \epsilon_0$, if $\varphi \in \mathcal{L}_{<\alpha}$, then $|\varphi|_\omega < |Tr_\alpha(\ulcorner \varphi \urcorner)|_\omega$. In other words, ω -complexity has a sort of “tracking property:” it can “track” the levels of Tarski’s hierarchy. Moreover, we can represent ω -complexity in PA . More specifically, the function $c_\omega : \#\mathcal{L} \rightarrow \mathbb{N}$ that maps the code $\#\varphi$ of a formula $\varphi \in \mathcal{L}$ to its ω -complexity $|\varphi|_\omega$ is recursive and thus representable in PA . We represent c_ω by the unary function symbol c_ω . Thus, we get for

²⁵Now we could define quantification over ordinals by saying that $\forall x\varphi$ means $\forall x(On_{<\epsilon_0}(x) \rightarrow \varphi)$ and work with the more general versions of the axioms mentioned before. But for reasons of perspicuity, we refrain from doing so.

²⁶Here lub is the operation of taking the *least upper bound* with respect to the standard partial ordering \leq on the ordinals.

all $\varphi \in \mathcal{L}_{<\epsilon_0}$ and all $\alpha \in On_{<\epsilon_0}$: $\vdash_{PA} c_\omega(\ulcorner\varphi\urcorner) = \ulcorner\alpha\urcorner$ iff $|\varphi|_\omega = \alpha$. Using this representation, we can show that Peano arithmetic proves that ω -complexity has the “tracking-property” in the following sense:

Lemma 5 *For all ordinals $0 < \beta \leq \alpha < \epsilon_0$:*

$$\vdash_{PGA_\alpha} \forall x (Sent_{<\beta}(x) \rightarrow c_\omega(x) \dot{<} c_\omega(Tr_\beta(\dot{x}))).^{27}$$

Using ω -complexity, we’ll obtain the main result of this section:

Theorem 1 *For all $0 \leq \alpha < \epsilon_0$, the theory PGA_α is a proof-theoretically conservative extension of the theory PRT_α .*

Proof First, we define the translation function $\tau : \mathcal{L}_\alpha^{\triangleleft} \rightarrow \mathcal{L}_\alpha$ by saying that:

- $\tau(\varphi) = \begin{cases} Tr_\alpha(s) \wedge Tr_\alpha(t) \wedge c_\omega(s) \dot{<} c_\omega(t) & \text{if } \varphi = s \triangleleft t \\ \varphi & \text{if } \varphi \text{ is another atomic formula} \end{cases}$
- $\tau(\neg\varphi) = \neg\tau(\varphi)$;
- $\tau(\varphi \circ \psi) = \tau(\varphi) \circ \tau(\psi)$, for $\circ = \wedge, \vee$; and
- $\tau(Qx\varphi) = Qx(\tau(\varphi))$, for $Q = \forall, \exists$.

Then, we note that (a) for all $\varphi \in \mathcal{L}_\alpha$, $\tau(\varphi) = \varphi$. Next, we check that (b) for all $\varphi \in \mathcal{L}_\alpha^{\triangleleft}$, if $\vdash_{PGA_\alpha} \varphi$, then $\vdash_{PRT_\alpha} \tau(\varphi)$. The typed truth axioms of PGA_α are also axioms of PRT_α , so we only need to check the typed ground axioms and the typed upward and downward directed axioms. Here we only go through a few cases to illustrate the idea:

- In the case of the axiom G_4^α , we get:

$$\tau(G_4^\alpha) = \forall x \forall y ((Tr_\alpha(x) \wedge Tr_\alpha(y) \wedge c_\omega(x) \dot{<} c_\omega(y)) \rightarrow Sent_{<\alpha}(x) \wedge Sent_{<\alpha}(y))$$
 This is provable (almost) immediately from the typed truth axiom T_3^α of PRT_α :

$$\forall x (Tr_\alpha(x) \rightarrow Sent_{<\alpha}(x)).$$
- Finally, consider the axioms (APU_T^β) :

$$\forall x (Tr_\beta(x) \rightarrow x \triangleleft Tr_\beta(\dot{x})),$$

where $\beta < \alpha$. We get:

$$\tau(APU_T^\beta) = \forall x (Tr_\beta(x) \rightarrow Tr_\alpha(x) \wedge Tr_\alpha(Tr_\beta(\dot{x})) \wedge c_\omega(x) \dot{<} c_\omega(Tr_\beta(\dot{x}))).$$

Now let x be a fresh variable for \forall -Intro and assume $Tr_\beta(x)$ for a \rightarrow -Intro. Using the axiom T_3^β of PRT_α , we can infer that $Sent_{<\beta}(x)$. Moreover, since $\beta < \alpha$ by assumption, we can infer that $Tr_\alpha(Tr_\beta(\dot{x}))$ and $Tr_\beta(x)$ using the axiom RP_{12}^α of PRT_α . Finally, by Lemma 5, we get $Sent_{<\beta}(x) \rightarrow c_\omega(x) \dot{<} c_\omega(Tr_\beta(\dot{x}))$. Since we

²⁷We don’t give the detailed proof here, but it essentially proceeds by using induction on ordinals below ϵ_0 in PA .

know already that $Sent_{<\beta}(x)$, we get the final piece $c_\omega(x) \prec c_\omega Tr_\beta(\dot{x})$. Putting all of this together, by \rightarrow -Intro, we have

$$Tr_\beta(x) \rightarrow Tr_\alpha(x) \wedge Tr_\alpha(Tr_\beta(\dot{x})) \wedge c_\omega(x) \prec c_\omega(Tr_\beta(\dot{x})),$$

and since x was a fresh variable, by \forall -Intro, we get the desired theorem.

Putting (a) and (b) together, the claim follows. □

The theorem has the following immediate consequence:²⁸

Corollary 1 *For all $0 \leq \alpha < \epsilon_0$, the theory PGA_α is consistent.*

The proof of Theorem 1 essentially works because of the “tracking property” of ω -complexity. The idea of the proof is the same as in the proof of the corresponding result in the first part of this paper, but the translation we used there would not work here. Sentences of the form $Tr_\beta(\ulcorner \varphi \urcorner)$ involving the truth predicate all have a classical complexity of zero, while the sentence φ may have arbitrary complexity. Thus, we would not be able to derive the translations of the (typed versions of the) Aristotelian principles under the translation from the previous paper. The trick is to use ω -complexity in the translation—this is what allowed us to prove the result. The technique of the proof works for all ordinals $\alpha < \epsilon_0$, since PA can represent the well-ordering of these ordinals, which is required for the proof. The theory $PGA_{<\epsilon_0}$ is the first theory where our proof doesn’t work anymore, because in this theory we don’t have a “highest” truth predicate as required for the definition of τ . But we can extend our result to this theory using a simple *compactness argument*:

Corollary 2 *The theory $PGA_{<\epsilon_0}$ is a proof-theoretically conservative extension of the theory $PRT_{<\epsilon_0}$.*

Proof Assume that there is a sentence $\varphi \in \mathcal{L}_{<\epsilon_0}$ such that $\vdash_{PGA_{<\epsilon_0}} \varphi$, but $\not\vdash_{PRT_{<\epsilon_0}} \varphi$. Since proofs are finite objects, there can only be finitely many occurrences of different truth predicates $Tr_{\beta_1}, \dots, Tr_{\beta_n}$, for $0 < \beta_1 < \dots < \beta_n < \epsilon_0$, in the proof. But then the proof of φ , is also a proof in PGA_{β_n} and $\varphi \in \mathcal{L}_{\beta_n}$. Now by Theorem 1, PGA_{β_n} is conservative over PRT_{β_n} . This means that $\vdash_{PRT_{\beta_n}} \varphi$ and thus also $\vdash_{PRT_{<\epsilon_0}} \varphi$. Contradiction! Thus, there is no such φ and the claim holds. □

We get immediately:

Corollary 3 *The theory $PGA_{<\epsilon_0}$ is consistent.*

²⁸We could also use the theorem to determine the proof theoretic strength of PGA_α , but there is a small “hiccup:” the version of PRT_α that we discussed here is not exactly the one that is usually discussed in the literature. As mentioned in Footnote 24, PRT_α is usually formulated using quantification over ordinals, which we avoided here for reasons of perspicuity. The version of PRT_α with axioms quantifying over ordinals proves the same arithmetical theorems as the theory RA_α of *ramified analysis up to (and including) α* . For a proof of this result, see [4]. We suspect that the proof theoretic strength of our version of PRT_α is very close to this, although we’re not going to prove anything to this effect.

The theory $PGA_{<\epsilon_0}$ is a natural stopping point for the methods we've developed in this paper.²⁹

We have shown the consistency of our theories $PGA_{<\alpha}$, where $1 \leq \alpha \leq \epsilon_0$, by proof theoretic means. But for reasons of perspicuity, it would also be good to have an idea what models for these theories look like. In the rest of this section, we will show how to extend the construction from the previous paper to obtain models for $PGA_{<\alpha}$, where $1 \leq \alpha \leq \epsilon_0$.

As in the case of PT , there is a standard model of $PRT_{<\alpha}$, for $1 \leq \alpha \leq \epsilon_0$. A model for the language $\mathcal{L}_{<\alpha}$ is a structure of the form $(\mathbf{N}, (S_\beta)_{\beta < \alpha})$, where for $\beta < \alpha$, the set S_β interprets the truth predicate $Tr_\beta \in \mathcal{L}_{<\alpha}$. For $1 \leq \alpha \leq \epsilon_0$, we define the sets $(S_\beta)_{\beta < \alpha}$ by the following (transfinite) recursion:

- $S_1 = \{\#\varphi \mid \varphi \in \mathcal{L}_{<1}, \mathbf{N} \models \varphi\}$;
- $S_{\alpha+1} = S_\alpha \cup \{\#\varphi \mid \varphi \in \mathcal{L}_{<\alpha}, (\mathbf{N}, (S_\beta)_{\beta < \alpha}) \models \varphi\}$
- $S_\alpha = \bigcup_{\beta < \alpha} S_\beta$, if α is a limit ordinal.

Then we get, for all $1 \leq \alpha \leq \epsilon_0$, that $(\mathbf{N}, (S_\beta)_{\beta < \alpha}) \models PRT_{<\alpha}$. For $1 \leq \alpha \leq \epsilon_0$, the model $(\mathbf{N}, (S_\beta)_{\beta < \alpha})$ is the *standard model* of $PRT_{<\alpha}$ —it models Tarski's hierarchy of truths.

We now extend our definition of grounding-trees from the previous paper to *grounding-trees over the standard model of $PRT_{<\alpha}$* :

Definition 5 Let $1 \leq \alpha \leq \epsilon_0$ and let $(\mathbf{N}, (S_\beta)_{\beta < \alpha})$ be the standard model of $PRT_{<\alpha}$. We define the *grounding-trees over $(\mathbf{N}, (S_\beta)_{\beta < \alpha})$* by the following clauses for all formulas $\varphi \in \mathcal{L}_{<\alpha}$:

- (i) $\#\varphi \in \bigcup_{\beta < \alpha} S_\beta$, then $\#\varphi$ is a grounding-tree over $(\mathbf{N}, (S_\beta)_{\beta < \alpha})$ with $\#\varphi$ as its root;
- (ii) if $\widehat{\mathcal{T}}_{\#\varphi}$ is a grounding-tree \mathcal{T} over $(\mathbf{N}, (S_\beta)_{\beta < \alpha})$ with $\#\varphi$ as its root, then $\widehat{\mathcal{T}}_{\#\neg\neg\varphi}$ is a grounding-tree over $(\mathbf{N}, (S_\beta)_{\beta < \alpha})$ with $\#\neg\neg\varphi$ as its root;
- (iii) if $\widehat{\mathcal{T}}_{\#\varphi}$ is a grounding-tree \mathcal{T} over $(\mathbf{N}, (S_\beta)_{\beta < \alpha})$ with $\#\varphi$ as its root, then $\widehat{\mathcal{T}}_{\#(\varphi \vee \psi)}$ is a grounding-tree over $(\mathbf{N}, (S_\beta)_{\beta < \alpha})$ with $\#(\varphi \vee \psi)$ as its root;

²⁹The situation is quite similar to the corresponding theories of truth. For a discussion of the natural stopping point see [8, p. 322-29].

(iv) if $\widehat{\mathcal{T}}$ is a grounding-tree \mathcal{T} over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\psi$ as its root, then $\widehat{\mathcal{T}}$ is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#(\varphi \vee \psi)$ as its root;

(v) if $\widehat{\mathcal{T}}_1, \widehat{\mathcal{T}}_2$ are grounding-trees $\mathcal{T}_1, \mathcal{T}_2$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\varphi, \#\psi$ as their roots respectively, then

$\widehat{\mathcal{T}}$ is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#(\varphi \wedge \psi)$ as its root;

(vi) if $\widehat{\mathcal{T}}$ is a grounding-tree \mathcal{T} over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\varphi(t)$ as its root, then $\widehat{\mathcal{T}}$ is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\exists x\varphi(x)$ as its root;

(vii) if $\widehat{\mathcal{T}}_1, \widehat{\mathcal{T}}_2, \dots$ are grounding-trees $\mathcal{T}_1, \mathcal{T}_2, \dots$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\varphi(t_1), \#\varphi(t_2), \dots$ as their roots respectively, where t_1, t_2, \dots are all and only the terms of \mathcal{L}_{PA} , then $\widehat{\mathcal{T}}$ is a grounding-tree over

$(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\forall x\varphi(x)$ as its root;

(viii) if $\widehat{\mathcal{T}}$ is a grounding-tree \mathcal{T} over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\neg\varphi$ as its root, then $\widehat{\mathcal{T}}$ is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\neg(\varphi \wedge \psi)$ as its root;

(ix) if $\widehat{\mathcal{T}}$ is a grounding-tree \mathcal{T} over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\neg\psi$ as its root, then

- $$\begin{array}{c} \# \neg(\varphi \wedge \psi) \\ | \\ \# \neg\psi \\ \widehat{\mathcal{T}} \end{array}$$
 is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\# \neg(\varphi \wedge \psi)$ as its root;
- (x) if $\widehat{\mathcal{T}}_1, \widehat{\mathcal{T}}_2$ are grounding-trees $\mathcal{T}_1, \mathcal{T}_2$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\# \neg\varphi, \# \neg\psi$ as their roots respectively, then

$$\begin{array}{c} \# \neg(\varphi \vee \psi) \\ / \quad \backslash \\ \# \neg\varphi \quad \# \neg\psi \\ \widehat{\mathcal{T}}_1 \quad \widehat{\mathcal{T}}_2 \end{array}$$
 is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\# \neg(\varphi \vee \psi)$ as its root;
- (xi) if $\widehat{\mathcal{T}}$ is a grounding-tree \mathcal{T} over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\# \neg\varphi(t)$ as its root, then

$$\begin{array}{c} \# \neg\forall x\varphi(x) \\ | \\ \# \neg\varphi(t) \\ \widehat{\mathcal{T}} \end{array}$$
 is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\# \neg\forall x\varphi(x)$ as its root;
- (xii) if $\widehat{\mathcal{T}}_1, \widehat{\mathcal{T}}_2, \dots$ are grounding-trees $\mathcal{T}_1, \mathcal{T}_2, \dots$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\# \neg\varphi(t_1), \# \neg\varphi(t_2), \dots$ as their roots respectively, where t_1, t_2, \dots are all and only the terms of \mathcal{L}_{PA} , then

$$\begin{array}{c} \# \neg\exists x\varphi(x) \\ / \quad | \quad \backslash \\ \# \neg\varphi(t_1) \quad \# \neg\varphi(t_2) \quad \dots \\ \widehat{\mathcal{T}}_1 \quad \widehat{\mathcal{T}}_2 \end{array}$$
 is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\# \forall x\varphi(x)$ as its root;
- (xiii) if $\widehat{\mathcal{T}}$ is a grounding-tree \mathcal{T} over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\varphi$ as its root and $\#\varphi \in \mathbf{S}_\beta$, for $\beta < \alpha$, then

$$\begin{array}{c} \#Tr_\beta(\ulcorner \varphi \urcorner) \\ | \\ \#\varphi \\ \widehat{\mathcal{T}} \end{array}$$
 is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#Tr_\beta(\ulcorner \varphi \urcorner)$ as its root;
- (xiv) if $\widehat{\mathcal{T}}$ is a grounding-tree \mathcal{T} over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\# \neg\varphi$ as its root and

is at least ω and thus the height of this tree is at least $\omega + 1$.³⁰

Now, an important consequence of this observation is that we can't use ordinary induction on the height of trees to prove claims about all grounding-trees. We need to use *transfinite* induction on the height of the grounding-trees. This doesn't add any further complications, but to be explicit let's state the form of the principle that we're going to use. Consider a property of grounding-trees. Then, if we can show that any degenerate grounding-tree has the property and we can show that if we can show that assuming that all trees of a height smaller than a given tree have the property, then the tree itself has the property, it follows that all grounding-trees have the property. Note that in this form of the principle, the induction step also includes *limit cases*, where we consider a tree of the height of a limit ordinal and need to show that the tree has the property in question, given that all trees of a lower height have the property.

Analogously to the case of grounding-trees over (\mathbf{N}, \mathbf{S}) , we can now show that grounding-trees over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ are: (i) *rooted graphs* over $\bigcup_{\beta < \alpha} \mathbf{S}_\beta$; (ii) indeed *rooted trees* over $\bigcup_{\beta < \alpha} \mathbf{S}_\beta$, i.e. they don't contain any cycles; and finally, (iii) transitive.

Lemma 6 *Let $1 \leq \alpha \leq \epsilon_0$, $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ be the standard model of $PRT_{<\alpha}$, and let \mathcal{T} be a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$. Then for all formulas $\varphi \in \mathcal{L}_{<\alpha}$, if $\#\varphi$ is a vertex in \mathcal{T} , then $\#\varphi \in \bigcup_{\beta < \alpha} \mathbf{S}_\beta$.*

Proof The new cases for clauses (xiii) and (xiv) follow by the fact that $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ is a model of $PRT_{<\alpha}$. □

Remember the notion of a code of a formula occurring below another code in a grounding-tree over (\mathbf{N}, \mathbf{S}) . We now adapt this notion to grounding-trees over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ by recursively saying that, for all $1 \leq \alpha \leq \epsilon_0$, no code of any formula occurs below the code of any other formula in a degenerate grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$, and if \mathcal{T} is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ that was constructed from grounding-trees $\mathcal{T}_1, \mathcal{T}_2, \dots$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ according to the rules (ii–xvi) of Definition 5, then all occurrences of all formulas in $\mathcal{T}_1, \mathcal{T}_2, \dots$ occur below the root of \mathcal{T} in \mathcal{T} .

Then we can show:

Lemma 7 *Let $1 \leq \alpha \leq \epsilon_0$ and let $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ be the standard model of $PRT_{<\alpha}$. If \mathcal{T} is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\varphi$ as its root, for some formula $\varphi \in \mathcal{L}_{<\alpha}$. Then, all formulas $\psi \in \mathcal{L}_{<\alpha}$ whose code $\#\psi$ occurs below $\#\varphi$ in \mathcal{T} have a lower ω -complexity than φ .*

Lemma 8 *Let $1 \leq \alpha \leq \epsilon_0$, $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ be the standard model of $PRT_{<\alpha}$, and let \mathcal{T} be a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$. Then between any two nodes $\#\varphi$ and $\#\psi$ in \mathcal{T} , for formulas $\varphi, \psi \in \mathcal{L}_{<\alpha}$, there is exactly one path.*

³⁰In fact, by a slightly more complicated argument we can show that the height of this tree is *exactly* $\omega + 1$.

Lemma 9 Let $1 \leq \alpha \leq \epsilon_0$ and let $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ be the standard model of $PRT_{<\alpha}$. If there is a grounding-tree \mathcal{T}_1 over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\psi$ as its root and $\#\varphi_1, \#\varphi_2, \dots$ as its leaves and there is grounding-tree \mathcal{T}_2 over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\psi, \#\psi_1, \#\psi_2, \dots$ as its leaves and $\#\theta$ as its root, then there is a grounding-tree \mathcal{T}_3 over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with $\#\varphi_1, \#\varphi_2, \dots, \#\psi_1, \#\psi_2, \dots$ as its leaves and $\#\theta$ as its root.

Finally, we define the standard model of $PGA_{<\alpha}$ by saying that:

Definition 7 Let $1 \leq \alpha \leq \epsilon_0$ and let $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ be the standard model of $PRT_{<\alpha}$. We define the relation $\mathbf{R} \subseteq \mathbb{N}^2$ by saying that for all $n, m \in \mathbb{N}$, $\mathbf{R}(m, n)$ iff there is a non-degenerate grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ with n as a leaf and m as its root.

Putting Lemmas 6, 8, and 9 together, we obtain:

Theorem 2 $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha}, \mathbf{R})$ is a model of $PGA_{<\alpha}$, for $1 \leq \alpha \leq \epsilon_0$, i.e. $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha}, \mathbf{R}) \models PGA_{<\alpha}$.

Proof By Lemmas 6, 8, and 9, grounding-trees over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ behave appropriately and satisfy the basic ground axioms. Since $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$ is a model of $PRT_{<\alpha}$, the typed truth axioms are satisfied. The only new cases are the axioms for the typed Aristotelian principles. Here we only show that APU_γ^γ , for $\gamma < \alpha$ holds:

– $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha}, \mathbf{R}) \models \forall x (Tr_\gamma(x) \rightarrow x \triangleleft Tr_\gamma \dot{x})$ for all $\gamma < \alpha$.

Let σ be a variable assignment over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha}, \mathbf{R})$ and σ' some x -variant of σ . Assume that $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha}, \mathbf{R}) \models_{\sigma'} Tr_\gamma(x)$. This means that $\sigma'(x) \in \mathbf{S}_\gamma$. Since $\mathbf{S}_\gamma = \{\#\varphi \mid \varphi \in \mathcal{L}_{<\gamma}, (\mathbf{N}, (\mathbf{S}_\delta)_{\delta < \gamma}) \models \varphi\}$, we know that $\sigma'(x) = \#\varphi$, for some formula $\varphi \in \mathcal{L}_{<\gamma}$. Now, $\#\varphi$ is a degenerate grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$.

But then, by clause (xiii) of Definition 5,
$$\begin{array}{c} \#Tr_\gamma(\ulcorner \varphi \urcorner) \\ | \\ \#\varphi \end{array}$$

is a non-degenerate grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha})$. Moreover, the root of this tree is $\#Tr_\gamma(\ulcorner \varphi \urcorner)$ and its only leaf is $\#\varphi$. Now consider $\sigma'(Tr_\gamma \dot{x})$. Since we know that $\sigma'(x) = \#\varphi$ and Tr_γ expresses the function that maps codes of formulas to the code of Tr_γ applied to the formula, we know that $\sigma'(Tr_\gamma \ulcorner \varphi \urcorner) = \#Tr_\gamma(\ulcorner \varphi \urcorner)$. Thus, $\mathbf{R}(\sigma'(x), \sigma'(Tr_\gamma \dot{x}))$ meaning $\models_{\sigma'} x \triangleleft Tr_\gamma \dot{x}$. And since σ was arbitrary, we get $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta < \alpha}, \mathbf{R}) \models \forall x (Tr_\gamma(x) \rightarrow x \triangleleft Tr_\gamma \dot{x})$, as desired.

We can show analogously that the other axioms hold. □

4 Paradoxes of Self-Referential Ground

In the setting of PG (as well as $PGA_{<\alpha}$ for $0 < \alpha < \epsilon_0$), we can accommodate new n -ary predicates R by stipulating:

$$\begin{aligned} &\forall t_1, \dots, \forall t_n (Tr(R(t_1, \dots, t_n)) \leftrightarrow R(t_1^\circ, \dots, t_n^\circ)) \text{ and} \\ &\forall t_1, \dots, \forall t_n (Tr(\neg R(t_1, \dots, t_n)) \leftrightarrow \neg R(t_1^\circ, \dots, t_n^\circ)). \end{aligned}$$

If we furthermore add a theory for the new predicate R to our background theory, the resulting theory will be a theory of partial ground over the hierarchy of truths of that extended background theory. For example, we could formulate theories of partial ground over the truths of analysis, of mereology, or even of set-theory.³¹ The resulting theory will then, of course, also prove all the instances of the T-scheme $Tr(\ulcorner \varphi \urcorner) \leftrightarrow \varphi$ over sentences of the new language $\mathcal{L} \cup \{R\}$. But this approach has limits. Of course, we can't let R be our unary truth predicate Tr itself. It is well-known that if we affirm:

$$\forall t (Tr(Tr(t)) \leftrightarrow Tr(t^\circ)) \text{ and}$$

$$\forall t (Tr(\neg Tr(t)) \leftrightarrow \neg Tr(t^\circ)),$$

our theory will fall prey to the liar paradox and its ilk. It might be somewhat surprising to learn, however, that we also can't affirm:

$$\forall s \forall t (Tr(s \triangleleft t) \leftrightarrow s^\circ \triangleleft t^\circ) \text{ and}$$

$$\forall s \forall t (Tr(s \not\triangleleft t) \leftrightarrow s^\circ \not\triangleleft t^\circ)$$

in the context of predicational theories of ground. This will be the main result of this section.³² We'll call the resulting problem the *new puzzle of ground*.

Let's move to a setting where we can affirm applications of the ground predicate to sentences involving the ground predicate. For this purpose, we have to make a couple of adjustments to our theory setup. In the following, let $\mathcal{L}_{\triangleleft}$ be the language $\mathcal{L} \cup \{\triangleleft\}$. Now first, we have to assume that we work in the context of a proper coding for $\mathcal{L}_{\triangleleft}$. In particular, we now assume that we have a name $\ulcorner \varphi \urcorner$ for every sentence $\varphi \in \mathcal{L}_{\triangleleft}$. We assume that we have a function symbol \triangleleft such that:

$$\vdash_{PA} s \triangleleft t = \ulcorner s \triangleleft t \urcorner,$$

for all terms s and t . And we let $Sent_{\triangleleft}$ abbreviate the formula that allows us to represent the set (of codes of) sentences of $\mathcal{L}_{\triangleleft}$. In particular, we assume that for all $n \in \mathbb{N}$:

$$\vdash_{PA} Sent_{\triangleleft}(\bar{n}) \text{ iff } n \in \#\mathcal{L}_{\triangleleft} \text{ and}$$

$$\vdash_{PA} \neg Sent_{\triangleleft}(\bar{n}) \text{ iff } n \notin \#\mathcal{L}_{\triangleleft}.$$

Second, we need to adjust the axiom T_3 to:

$$(T_3^{\triangleleft}) \forall x (Tr(x) \rightarrow Sent_{\triangleleft}(x)),$$

which allows sentences involving the ground predicate to occur in the context of the truth predicate (and thus in the context of the ground predicate). We arrive at a modified predicational theory of ground:

³¹The truths of new atomic sentences will, of course, not have any provable grounds in the theory.

³²Here $s \not\triangleleft t$, for terms s and t , is an abbreviation of $\neg(s \triangleleft t)$, analogous to the case of $s \neq t$. Similarly, the notation $s \not\triangleleft t$ is an abbreviation for the complex function term $\neg(s \triangleleft t)$ for terms s and t .

Definition 8 The predicational theory *PUG* of *untyped ground* consists of the axioms of *PG* without the axiom T_3 , plus the axiom T_3^{\triangleleft} and the axioms:

$$(T_{\triangleleft}^1) \quad \forall s \forall t (Tr(s \triangleleft t) \leftrightarrow s^\circ \triangleleft t^\circ) \text{ and}$$

$$(T_{\triangleleft}^2) \quad \forall s \forall t (Tr(s \not\triangleleft t) \leftrightarrow s^\circ \not\triangleleft t^\circ).$$

We will now show that *PUG* is inconsistent. To see this, first note that we can show in the same way as in the case of *PG* that *PUG* proves the uniform T-scheme for sentences involving the ground predicate:

Lemma 10 For all sentences $\varphi \in \mathcal{L}_{\triangleleft}$,

$$\vdash_{PUG} \forall t_1, \dots, \forall t_n (Tr(\varphi(t_1, \dots, t_n)) \leftrightarrow \varphi(t_1^\circ, \dots, t_n^\circ)).$$

Proof By induction on the positive complexity of formulas. The new axioms $T_{\triangleleft}^{1/2}$ take care of the new base-case. □

Note that this lemma doesn't entail yet that *PUG* is inconsistent: the truth predicate *Tr* is not in the language $\mathcal{L}_{\triangleleft}$ and thus Lemma 10 doesn't entail that we're applying the truth predicate to sentences involving the same truth predicate. To see that *PUG* is inconsistent, we need to do some more work. Since we work in the context of a coding for $\mathcal{L}_{\triangleleft}$, we can get the diagonal lemma for the language:

Lemma 11 For all formulas $\varphi \in \mathcal{L}_{\triangleleft}$ with exactly one free variable, there is a formula $\delta \in \mathcal{L}_{\triangleleft}$ such that

$$\vdash_{PA} \delta \leftrightarrow \varphi(\ulcorner \delta \urcorner).$$

With these two lemmas in place, we can show our main result:

Theorem 3 *PUG* is inconsistent.

Proof Let $\varphi(x)$ be the formula $\neg(x \triangleleft \neg \neg x)$. By the diagonal lemma for $\mathcal{L}_{\triangleleft}$, we know that there is a formula $\delta \in \mathcal{L}_{\triangleleft}$ such that:

$$\vdash_{PA} \delta \leftrightarrow \neg(\ulcorner \delta \urcorner \triangleleft \ulcorner \neg \neg \delta \urcorner).$$

It is easily checked that we can now both prove δ and $\neg \delta$. □

We are left with yet another puzzle of ground: by letting the truth predicate (and thus the ground predicate) apply to truths involving the ground predicate, we made our intuitively plausible theory of ground inconsistent. But intuitively, we want to be able to talk about the truth of sentences involving the ground predicate. So what went wrong?

First, note that the new puzzle is different from Fine's puzzle of ground. Fine's puzzle consists in the fact that different intuitively plausible principles for partial ground and truth entail that the truth of some sentences partially ground themselves—in contradiction to the irreflexivity of partial ground. Our puzzle, in contrast, consists

in the fact that letting the truth predicate apply to the ground predicate makes our previously consistent principles of ground inconsistent.

Moreover, note that the use of double negation (and of the corresponding upward directed ground axiom) in the proof of Theorem 3 is dispensable. We could equally well have applied the diagonal lemma to the formula $\neg(x \triangleleft x \vee x)$ or $\neg(x \triangleleft x \wedge x)$ or ... and we would still have gotten the same inconsistency result (using the corresponding upward directed ground axioms for these connectives). The point is that our paradox is not a paradox of double negation, or disjunction, or conjunction, or the like—it has another source.

We argue that the paradox is a paradox of *self-reference* in the context of partial ground. In support of this claim, first note that all the different sentences that we could use for the proof of Theorem 3 have in common that they (provably) “say of themselves” that they violate certain principles of partial ground, and as a consequence they are inconsistent over our theory of untyped ground. Thus, the new puzzle of ground bears a strong resemblance to the classic semantic paradoxes. Indeed, we can make this analogy even more explicit in the following proposition. Let’s define the predicate $Tr_0^=$ by the following explicit definition:

$$\forall x(Tr_0^=(x) \leftrightarrow_{def} \exists s, t(x = (s \equiv t) \wedge s^\circ = t^\circ) \vee \exists s, t(x = (s \not\equiv t) \wedge s^\circ \neq t^\circ))$$

And let’s define the predicate Tr_0^\triangleleft by the following explicit definition:

$$\forall x(Tr_0^\triangleleft(x) \leftrightarrow_{def} \exists s, t(x = (s \triangleleft t) \wedge s^\circ \triangleleft t^\circ) \vee \exists s, t(x = (s \not\triangleleft t) \wedge s^\circ \not\triangleleft t^\circ))$$

So, intuitively $Tr_0^=$ and Tr_0^\triangleleft are truth predicates for the literals of $\mathcal{L}_\triangleleft$. We can then show the following proposition:

Proposition 2 *PUG proves that the formula $Tr_0^=(x) \vee Tr_0^\triangleleft(x) \vee \exists y(y \triangleleft x)$ satisfies the T-scheme for the formulas of $\mathcal{L}_\triangleleft$, i.e. for all $\varphi \in \mathcal{L}_\triangleleft$:*

$$\vdash_{PUG} Tr_0^=(\ulcorner \varphi \urcorner) \vee Tr_0^\triangleleft(\ulcorner \varphi \urcorner) \vee \exists y(y \triangleleft \ulcorner \varphi \urcorner) \leftrightarrow \varphi.$$

Proof By induction on the positive complexity of φ for both directions of the biconditional. The base cases for φ being $s = t$ or $s \neq t$, for terms s and t , are covered by the basic truth axioms T_1 and T_2 and the definition of $Tr_0^=$. Similarly, the base cases for φ being $s \triangleleft t$ or $s \not\triangleleft t$, for terms s and t , are covered by the axioms T_\triangleleft^1 and T_\triangleleft^2 and the definition of Tr_0^\triangleleft . The remaining cases are can be dealt with using Lemma 10. □

In other words, in *PUG* we can define a truth predicate for $\mathcal{L}_\triangleleft$ that satisfies the T-scheme for $\mathcal{L}_\triangleleft$. It follows by Tarski’s *theorem of the undefinability of truth* that *PUG* is inconsistent [20]. Our Theorem 3 is only a special case of this more general fact, as it were. This is the precise sense in which the ground predicate “behaves too much like a truth predicate”—in other words: the new puzzle of ground is at heart a paradox of self-reference.

5 Conclusion

In this paper, we have provided a typed solution to the puzzle of ground. But there are a couple of worries that should be addressed: First, we might be worried that we used “a sledgehammer to crack a walnut:” our use of typing up to ϵ_0 could be seen as a slight overkill to solve such a simple problem as the puzzle of ground. We would like to say two things in response to this worry: First, the use of typing up to ϵ_0 is not really necessary to formulate the solution. We could already have stopped at level two, where in the typed setting grounding between arithmetic truths and truths involving a truth predicate occurs for the first time: it follows from our results that the theory PGA_2 is consistent and proves the following simple version of the typed Aristotelian principle for truth:

$$\forall x(Tr_1(x) \rightarrow x \triangleleft Tr_1(\dot{x})).$$

But there is no real reason to stop at level two. We have shown that the methods of the paper work up to the level ϵ_0 . Moreover, the result was an intuitive and mathematically non-trivial theory of partial ground and typed truth. Thus, the results of this paper support our plea for collaboration between ground-theorists and truth-theorists.

Second, there is the worry that by “going typed” we have “thrown out the baby with the bathwater:” There are several unproblematic self-referential sentences that fall out of the scope of our typed theory. For example, by a simple application of the diagonal lemma to the formula $Tr(x)$, we get the *truth-teller* sentence ρ such that:

$$\vdash_{PA} \rho \leftrightarrow Tr(\ulcorner \rho \urcorner).$$

This sentence can consistently be added to a theory that proves the instance of the T-scheme for it. It might be argued that this shows that we should include it in our best theory of partial ground and truth. PA and the T-scheme don’t decide this sentence: we can both consistently affirm it and its negation (although, of course, not at the same time). Thus, it might be argued, a good theory of partial ground and truth should at least prove the following principle:

$$\rho \rightarrow \ulcorner \rho \urcorner \triangleleft \ulcorner Tr(\ulcorner \rho \urcorner) \urcorner.$$

This sentence, however, is in none of the languages $\mathcal{L}_{<\alpha}^{\triangleleft}$, for $0 < \alpha < \epsilon_0$, and so in these languages we cannot say anything about its partial grounds. To make things worse, it could be argued that the principles of partial ground and truth should even apply to the liar sentence λ such that $\vdash_{PA} \lambda \leftrightarrow \neg Tr(\ulcorner \lambda \urcorner)$. In particular, it might be argued that both

$$\lambda \rightarrow \ulcorner \lambda \urcorner \triangleleft \ulcorner \neg Tr(\ulcorner \lambda \urcorner) \urcorner \text{ and } \neg \lambda \rightarrow \ulcorner \neg \lambda \urcorner \triangleleft \ulcorner Tr(\ulcorner \lambda \urcorner) \urcorner$$

should be the provable. The argument would be that assuming that if λ is true, then since λ “says of itself” that it is not true, the truth of λ grounds the truth of $\neg Tr(\ulcorner \lambda \urcorner)$ and analogously for $\neg \lambda$. In the context of the factivity of ground, this would lead to a non-classical theory of truth. If we have:

$$\forall x \forall y (x \triangleleft y \rightarrow Tr(x) \wedge Tr(y)),$$

then we would immediately get $Tr(\ulcorner \lambda \urcorner)$ and $Tr(\ulcorner \neg \lambda \urcorner)$ from the above principles. But there are non-classical theories of truth on the market, where this would indeed be plausible—so what’s the problem?

So far, we have axiomatized a *classical* theory of partial ground and truth, and we have achieved this by typing our truth predicate. We could also try to axiomatize *non-classical* theories of partial ground and truth, for example in the spirit of the *Kripke-Feferman* theory of truth *KF* [8, p. 195–227]. In such a setting we could consistently affirm the existence of the liar sentence and the truth-teller sentence, and even apply the principles of partial ground to them—which is, of course, ruled out on our approach. We would get a theory of partial ground and *untyped* truth. But such a project would require the development of new methods and this goes beyond the scope of this paper. Most importantly, in such a setting, we would have to find a new solution to the puzzle of ground, since we relied on the typing of the truth predicate for our solution. In the terminology of [6, p. 110], our solution is a *predicativist solution* to the puzzle of ground: On our approach, the grounds of a truth of the form $\exists x Tr_\alpha(x)$ never includes truths of involving a truth predicate Tr_α , where α is some ordinal with $0 < \alpha < \epsilon_0$ —in particular, the grounds of the truth of $\exists x Tr_\alpha(x)$ don't include the truth of $\exists x Tr_\alpha(x)$ itself. We have shown that by respecting the typing restrictions of Tarski's hierarchy, we can obtain a consistent theory of partial ground that proves (typed versions of) the Aristotelian principles. There is, of course, room for impredicativist solutions in the terminology of [6, p. 110], but developing such solutions is beyond the scope of this paper. We leave this for future research.

How should we respond to the paradox of self-referential ground? Three natural ways in which we could try to block the inconsistency theorem suggest themselves: First, we could try to rule out self-referential sentences of ground like the one used in the proof of the inconsistency theorem. Second, we could try to restrict the principles of partial ground used in the proof of the inconsistency theorem. And third, we could try to formulate a non-standard logic of ground that does not sanction the logical principles used in the proof of the inconsistency theorem. The analogy between our inconsistency theorem (Theorem 3) and the Tarski's theorem of the undefinability of truth (Proposition 2) suggests a neat terminology for these approaches. Analogously to theories of truth [8], we get: *typed theories of partial ground*, which avoid paradox by putting type-restrictions on the relation of partial ground, effectively ruling out self-referential sentences like the one in the proof; *untyped theories of partial ground*, which avoid paradox by restricting the principles of partial ground; and finally *non-classical theories of partial ground*, which avoid paradox (or: triviality) by abandoning classical logic in favor of alternative logics.

The results of this paper all point in the direction of a typed theory of partial ground. Indeed our theories $PGA_{<\alpha}$ are typed theories of partial ground: the axiom G_4^α is effectively a typing axiom. Moreover, we have shown that the theories $PGA_{<\alpha}$, for $1 \leq \alpha \leq \epsilon_0$, are consistent (Corollary 1). Thus, there is good evidence that a typed solution to the new puzzle of ground works. A natural direction to take from here would be to extend Tarski's truth-theoretic hierarchy to a truth- and *ground-theoretic* hierarchy. We would type the ground predicate to get a family of predicates: $\triangleleft_1, \triangleleft_2, \dots$. We would end up with a doubly-typed theory: typed with respect to truth and typed with respect to ground. The results of this paper suggest that this theory

will turn out consistent and this provides further support for the claim that we should use typing in the context of theories of ground. But carrying out the details of this proposal is beyond the scope of this paper. Also this we leave for future research.³³

Acknowledgements I'd like to thank Hannes Leitgeb, Benjamin Schnieder, Albert Anglberger, Thomas Schindler, Lavinia Picollo, Johannes Stern, Martin Fischer, Tim Button, the participants of the workshop "Logical and Metaphysical Perspectives on Grounding" at the GAP.9 in Osnabrück, O. Foisch, and two anonymous referees for their helpful comments and suggestions. Part of this research was supported by the Alexander von Humboldt Foundation.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Boolos, G. et al. (2007). *Computability and logic*. Cambridge University Press.
2. Correia, F. (2013). Logical grounds. *Review of Symbolic Logic* 7.1 (pp. 31–59).
3. Donaldson, T. (2016). The (metaphysical) foundations of arithmetic?. *Noûs*. doi:[10.1111/nous.12147](https://doi.org/10.1111/nous.12147).
4. Feferman, S. (1991). Reflecting on incompleteness. *Journal of Symbolic Logic*, 56.1, 1–49.
5. Fine, K. (2012). Guide to Ground. In Correia, F., & Schnieder, B. (Eds.) *Metaphysical Grounding: Understanding the Structure of Reality* (pp. 37–80): Cambridge University Press.
6. Fine, K. (2010). Some puzzles of ground. *Notre Dame Journal of Formal Logic* 51.1 (pp. 97–118).
7. Fine, K. (2012). The pure logic of ground. *Review of Symbolic Logic* 5.1 (pp. 1–25).
8. Halbach, V. (2011). *Axiomatic theories of truth*. Cambridge University Press.
9. Halbach, V., & Visser, A. (2014). Self-reference in arithmetic I. *Review of Symbolic Logic* 7.4 (pp. 671–691).
10. Jech, T. (2002). *Set theory*. Springer.
11. Jenkins, C.S. (2011). Is metaphysical dependence irreflexive?. *The Monist* 94.2 (pp. 267–276).
12. Korbmacher, J. ((forthcoming)). Axiomatic Theories of Partial Ground I. *Journal of Philosophical Logic* (pp. 1–31).
13. Korbmacher, J. (2015). Yet another puzzle of ground. *Kriterion* 29.2 (pp. 1–10).
14. Krämer, S. (2013). A simpler puzzle of ground. *Thought: A Journal of Philosophy* 2.2 (pp. 85–89).
15. Litland, J.E. (2015). Grounding, Explanation, and the limit of internality. *Philosophical review* 124.4.
16. Litland, J.E. (2013). On some counterexamples to the transitivity of grounding. *Essays in Philosophy* 14.1 (pp. 19–32).

³³A particularly interesting point that I'll leave for future research has been brought to my attention by an anonymous referee: next to the kind of self-reference we obtain, for example, by means of the diagonal lemma, there is also what's sometimes called *empirical* self-reference: self-reference that depends on contingent facts about the world. Consider the sentence "the last example sentence in the paper *Axiomatic Theories of Partial Ground II. Partial Ground and Hierarchies of Typed Truth* does not partially ground its own double negation," for example. Properly formalized, using, for example, a description operator, this sentence will create a ground-theoretic paradox of self-reference in the same way as the sentence δ in the proof of Theorem 3—but only because it is in fact the last such example sentence, otherwise it would simply be false. To deal with such forms of self-reference in a typed theory, we would presumably need to introduce some kind of *context dependent* typing. But alas, working out the details goes beyond the scope of this paper.

17. Pohlers, W. (2009). *Proof Theory. The first step into impredicativity*. Springer.
18. Raven, M.J. (2013). Is ground a strict partial order?. *American Philosophical Quarterly* 50.2 (pp. 191–99).
19. Schaffer, J. (2012). Grounding, transitivity, and contrastivity. In Correia, F., & Schnieder, B. (Eds.) *Metaphysical Grounding: Understanding the Structure of Reality* (pp. 122–138): Cambridge University Press.
20. Tarski, A. (1944). The semantic conception of truth: and the foundations of semantics. *Philosophy and Phenomenological Research* 4.3 (pp. 341–76).
21. Yablo, S. (1993). Paradox Without Self-Reference. *Analysis* 53.4 (pp. 251–52).