A defence of Isaacson's thesis
or how to make sense of the boundaries of finite mathematics

Abstract

Daniel Isaacson has advanced an epistemic notion of arithmetical truth according to which the latter is the set of truths that we grasp on the basis of our understanding of the structure of natural numbers alone. Isaacson's thesis is then the claim that Peano Arithmetic (**PA**) is *the* theory of finite mathematics, in the sense that it proves all and only arithmetical truths thus understood. In this paper, we raise a challenge for the thesis and show how it can be overcome. We introduce the concept of purity for theories of arithmetic: a theory of arithmetic is pure when it only proves arithmetical truths. Then, we argue that, under Isaacson's thesis, some **PA**-provable truths—including transfinite induction claims for infinite ordinals and some consistency statements—are seemingly not arithmetical in Isaacson's sense, and hence that Isaacson's thesis might entail the impurity of **PA**. Nonetheless, we conjecture that the advocate of Isaacson's thesis can avoid this undesirable consequence: the arithmetical nature, as understood by Isaacson, of all contentious **PA**-provable statements can be justified. As a case study, we explore how this is done for transfinite induction claims with infinite ordinals below ε_0. To this end, we show that the PA-proof of such claims exclusively employs resources from finite mathematics, and that ordinals below ε_0 are finitary objects despite being infinite.

1 Introduction

Against what has generally been believed to follow from Gödel's work, Daniel Isaacson (1987/1996; 1992), defended the view that **PA** is complete with respect to arithmetical truth. He proposed to conceptualise arithmetical truth as the set of truths that we grasp on the basis of our understanding of the structure of natural numbers alone, and argued that the first-order axiom system for arithmetic **PA** coincides with that set. It is thus that we must understand Isaacson's claim that **PA** is sound and complete with respect to arithmetical truth.

In this paper we identify a reading of Isaacson's work in which the status of certain **PA**-provable sentences as arithmetical, at least in the sense of the word Isaacson proposes, can be called into question. We first introduce the notion of *purity* for theories of arithmetic: a theory of arithmetic is said to be pure if and only if it only proves arithmetical truths. We note that purity thus understood is an important component of Isaacson's thesis. We then argue that, under the aforementioned reading of Isaacson's thesis, **PA** seems to be impure with respect to arithmetical truth—that is, some of the truths proven by **PA** might not be arithmetical truths in the sense of Isaacson, for their proof in the language of **PA** is too long to constitute the epistemic basis on which to perceive the truth of the statement. These include, for instance, transfinite induction claims for infinite ordinals, as well as consistency statements for theories of arithmetic weaker than **PA**. We then try to show that the way in which Isaacson, who had foreseen the reading that leads

I would like to thank Carlo Nicolai, Dan Isaacson, Luca Incurvati, Johannes Stern, Maciej Głowacki, Luca Dondoni and Amedeo Robiolio for reading and commenting on this paper. I am also greatly indebted to two anonymous reviewers, as their extensive feedback helped me improve this paper substantially. This work was made possible with the support of an LAHP-AHRC studentship.

to this impurity concern, tries to prevent the latter, is not entirely satisfactory. Finally, we explore a different route to restore purity: justifying the arithmeticality, in Isaacson's sense, of those claims that motivated the move in the first place. As a case study, we take transfinite induction claims for infinite ordinals; thus, we argue that shortened proofs of these statements can be shown to be arithmetical in Isaacson's sense, as they do not really contain higher-order notions. We end up by considering this case study as evidence in favour of a conjecture we advance, namely that the arithmeticality of all statements leading to the impurity concern can be justified.

2 Isaacson's thesis

Ever since at least Tarski, the mainstream conception of arithmetical truth has equated the latter with satisfiability in the standard model for the language of arithmetic, that we will call $\mathcal{L}_0$ and which includes the nonlogical constants $(\mathtt{S}, 0, +, \cdot, <)$.[1] We refer to this model simply as the standard model of arithmetic $\mathcal{N}$, and to the set of sentences true in this model as true arithmetic, or $\mathrm{Th}(\mathcal{N})$—see e.g. (Boolos et al., 2007, 295).

Contra this widespread view on arithmetical truth, Daniel Isaacson advances his own. For him, an arithmetical truth is a truth that is perceived as such 'from the purely arithmetical content of a categorical conceptual analysis of the notion of natural number' (1987/1996, 203). Accordingly, he defends that the way to determine what counts as an arithmetical truth is not only a formal matter but also an *epistemic* one, since what counts as arithmetical 'has to do with the way in which we are able to perceive [a] statement's truth or falsity' (1992, 95). In particular, he defends that the set of arithmetical truths is to be captured via a recursive definition. The base clause asserts that a true statement is arithmetical when its truth can be seen to follow directly from our understanding of the natural number structure; he seems to think that the axioms of **PA** (and perhaps those alone) are arithmetical in this sense. The recursive clause asserts that a true statement is arithmetical if its truth can be perceived as such via first-order logical inferences from known truths whose arithmetical nature has been granted.[2] Thus:

> [A] truth expressed in the (first-order) language of arithmetic is arithmetical just in case its truth is directly perceivable on the basis of our (higher-order) articulation of our grasp of the structure of the natural numbers or directly perceivable from truths in the language of arithmetic which are themselves arithmetical. The analysis of the number concept in §§2-4 seems to me to render the axioms of Peano Arithmetic arithmetical, in the sense that their truth is directly perceivable so expressed, and on this basis the second clause renders the theorems of PA arithmetical.(Isaacson, 1987/1996, 217)

Admittedly, Isaacson's recursive definition only accounts for arithmetical truths, that is, 'being arithmetical' is a property that applies only to true statements that can be expressed in the signature of **PA**. But one can easily account for arithmetical falsities by taking them to be all statements the negation of which is an arithmetical truth.[3] On the other hand, *higher-order* statements are

[1] For differing views, see Sayward (1990).

[2] One can question, in any case, the appropriateness of this recursive definition. The problem has to do with the base clause: while **PA** can be seen to capture the first-order content of second-order elementary number theory, so do the different theories that are mutually elementary reducible with standard **PA**, such as $\bigcup_n \mathrm{I}\Sigma_n$ (for a definition of elementary reducibility see, e.g., (Łełyk & Nicolai, 2022, p. 8 of 26)). It seems hence arbitrary to establish that one set of axioms, and not the other, can be directly perceived as the set of truths about the natural number structure. They all correspond to different axiomatizations of what we consider first-order elementary number theory with full induction to be. In sum, the base clause ought to allow for a wider range of applicability regarding what counts as following directly from our understanding of the concept of natural number.

[3] As Isaacson does in his (1992, 96).

truths and falsities that incorporate what Isaacson calls 'higher-order notions'. These include not only higher-order (in contrast with first-order) quantification, but also infinitary notions, 'in the sense of presupposing an infinite totality' (1987/1996, 210), as opposed to finitary notions.

Once this much is clear, we can fix the terminology we will use in the remainder of the paper, for the sake of readability. 'Elementary number theory' will here refer simply to the realm of mathematics that deals with natural numbers and their basic operations, as traditionally understood; and 'number-theoretic' will just be the corresponding adjective. When we write 'arithmetical', we mean 'arithmetical in Isaacson's sense' (i.e., whose truth or falsity is seen to follow from the purely number-theoretic content of a categorical conceptual analysis of the notion of natural number); when we talk about 'arithmeticality' or 'arithmetical nature', we mean 'the status of being arithmetical in Isaacson's sense'. We avoid the use of the term 'arithmetic' to prevent any confusion. The only exception will be the term 'theory of arithmetic', by which we mean a theory aimed at capturing, fully or partially, the content of Dedekind's analysis of the natural numbers, as Isaacson believes that **PA** does in a first-order setting (1987/1996, 207). Moreover, other potential uses of the word 'arithmetical' will be qualified as appropriate. For instance, an arithmetical statement as traditionally understood will be referred to as a 'statement expressible in $\mathcal{L}_0$' (where $\mathcal{L}_0$ is as above); and an arithmetical truth in the traditional, Tarskian sense, will be referred to as a 'statement of true arithmetic'.

Our current understanding of the natural number structure owes much to Dedekind's and Frege's studies of the principles of elementary number theory. Hence, theirs (and perhaps Dedekind's to a greater extent) are seen as the best categorical conceptual analyses of the notion of natural number. Admittedly, Dedekind's analysis contains higher-order concepts in the form of second-order quantification over subsets of natural numbers. But what remains when we strip this analysis of its second-order content—i.e., when we 'first-orderize' this second-order quantification—is just **PA**. As a result, according to Isaacson, **PA** enjoys a privileged position among all first-order axiomatizations of elementary number theory: not only does the analysis of the natural number structure allow us to perceive **PA** as true and strictly arithmetical, but it is also the case that **PA** captures all there is to *arithmetical*—as opposed to just *mathematical*—truth: if a statement expressible in $\mathcal{L}_0$ is not provable in **PA**, then some 'hidden' higher-order concept is needed either to directly perceive its truth or to carry out a proof of it.

With this in mind, we offer a precise formulation of Isaacson's thesis. There are a couple of different phrasings in the literature—see e.g. (Incurvati, 2008, 263); (Smith, 2008, 1); or (Horsten, 2001, 181), who instead calls it Isaacson's 'theory about arithmetical truth'. Isaacson's seemingly preferred way to put it is that Peano Arithmetic consists of those truths which can be perceived as truths either directly or via a proof from the purely number-theoretic content of the categorical conceptual analysis of the notion of natural number. However, and since we already know that 'those truths which...' is just short for Isaacson's notion of arithmetical truth, we offer the following, shorter wording:

> ***Isaacson's thesis*** Peano Arithmetic proves all and only arithmetical truths (in the sense of Isaacson).

As we see it, Isaacson's thesis gains a great deal of plausibility from the fact that it captures the long-standing mathematical intuition that our natural number system is at the heart of all finite mathematics, and that **PA**, as a set of axioms, is the best, natural approximation of such a system in first-order logic. Even so, the thesis must be tested, and its most pressing challenge is accommodating the kinds of sentences that show that **PA** is incomplete: statements of true arithmetic that are nonetheless independent of **PA**. The thesis predicts that all these sentences will present a common feature, namely their not being arithmetical in nature. Two clear examples

Isaacson examines are the Gödel sentence for **PA** and Goodstein's theorem. In the first case, the arithmeticality of the sentence is denied on the basis that seeing its truth requires the assumption that **PA** is consistent. That is, we can only come to see the truth of the sentence that says of itself 'This sentence is unprovable in **PA**' by first acknowledging the consistency of **PA**. But the latter is the kind of notion that, by Isaacson's thesis, and due to Godel's second incompleteness theorem, cannot be arithmetical—hence the Godel sentence will not be arithmetical either. As per Goodstein's theorem, the proof of the theorem relies on the well-ordering of ordinals (i.e. transfinite induction) for ε_0 ($\mathtt{TI}(\varepsilon_0)$ henceforth). The latter, however, is known to entail, over **PA**, the sentence Con(**PA**) (i.e., the sentence asserting the consistency of **PA**), and hence is also higher-order in nature. As a result, we should expect **PA** to prove neither the Gödel sentence nor Goodstein's theorem, so Isaacson's thesis stands.

Similar reasonings are given for two further well-known theorems independent of **PA**: the Paris-Harrington theorem and Friedman's finitization of Kruskal's theorem. Thus, although none of these arguments is conclusive enough to secure Isaacson's thesis—what happens, for instance, with the Kanamori-McAloon theorem or **PA**-unprovable versions of the graph minor theorem?[4]—they make it rather convincing. In other words, they seem to indicate that all arithmetically-expressible theorems that **PA** cannot prove aren't, after all, arithmetical truths.

3 The impurity concern

One of the key points behind Isaacson's thesis is that it lifts **PA** as *the* first-order axiomatization of elementary number theory, in the sense of proving all and only arithmetical truths. The 'all' part of the claim is established through completeness and it has certainly been the main focus of the literature, possibly due to its novelty after (and its defiance of) Gödel's incompleteness theorems (see Smith (2008); Tatton-Brown (2018)). But the 'only' side has not been thoroughly addressed so far. This section aims to show that, under a certain reading of Isaacson's original 1987/1996 paper, there is a real possibility of **PA** being an *impure* theory of arithmetic. Here, the notion of 'purity' has a precise meaning, in line with Isaacson's conception of arithmetical truth, that we will now explain.

3.1 The notion of purity

When we assert that, following Isaacson's thesis, **PA** is complete with respect to arithmetical truth, what we mean by completeness differs as much from the model-theoretic notion of completeness as Isaacson's understanding of arithmetical truth does from the Tarskian one. That is, we do not intend to say that **PA** proves all statements of true arithmetic, for this is plainly not the case. Rather, we just mean that there is no arithmetical truth in the sense of Isaacson that **PA** does not prove. To be fully unambiguous, we could have given this notion a new name—e.g. 'I-completeness'—since it is not what is usually meant by completeness alone.

We now intend to define the counterpart of this notion, which one can understand as the analogue of soundness under Isaacson's conception of arithmetical truth. Although, following the above, we could have called it 'I-soundness', we introduce a new term for it. Ordinarily, we say that a theory of arithmetic T is sound iff every theorem of T is a statement of true arithmetic; we will stick to this understanding of the term in what follows. Given this definition, when we read 'arithmetical truth' in the Tarskian sense, the question of whether every theorem of **PA** is an arithmetical truth is just implicit in the question of the soundness of **PA**. However, since 'statement of true arithmetic' and 'arithmetical truth' are concepts with different extensions under Isaacson's

[4] See e.g. Bovykin (2009).

thesis, there is a possibility that these two questions come apart in Isaacson's case. In other words: a theory of arithmetic being sound (i.e., proving only statements of true arithmetic) *does not entail* its proving only *arithmetical truths*. Accordingly, we distinguish between soundness and the fact of proving only arithmetical truths, a feature of theories of arithmetic that we label 'purity'.[5]

> ***Purity*** A theory of arithmetic T is *pure* iff every theorem of T is an *arithmetical* truth.

As can be understood from our discussion above, whether a theory is pure or not is relative to a given view on arithmetical truth. Hence, under the framework we work with (i.e., Isaacson's thesis), a theory of arithmetic T is pure iff every theorem of T is a true statement expressible in $\mathcal{L}_0$ that follows from the recursive definition proposed by Isaacson. Now, since the questions of soundness and purity have been separated, we must note that, for a theory to be impure, it need not be unsound, that is, it need not prove a false statement expressible in $\mathcal{L}_0$. It will suffice for it to prove a statement of true arithmetic that is not an arithmetical truth in the sense of Isaacson.

Clearly, purity is an essential feature for **PA** in the context of Isaacson's thesis. Should **PA** be impure, Isaacson's thesis, at least in the way we formulated it here, would simply be wrong. But even if the thesis was formulated in a way that is less liable to the threats of impurity, we understand that purity would still be an essential feature for **PA** in Isaacson's framework. The reason is that, in this framework, **PA** is meant to capture the boundaries of a well-defined region of mathematical truth, namely arithmetical truth or the truths of finite mathematics. Hence, if **PA** is to play that role, it should arguably be able to prove only *arithmetical* truths. And this is, precisely, what the idea of being pure amounts to. In fact, and in our view, the desirability of purity extends to *any* theory that purports to encapsulate a well-defined region of mathematical truth.[6]

The above does not imply, however, that a pure theory of arithmetic will prove *all* arithmetical truths—it suffices that all its theorems are arithmetical truths. For instance, **PRA**, or Robinson arithmetic, might perfectly well be pure theories of arithmetic if every theorem they prove is an *arithmetical* truth. This is in contrast to the case of **PA**, at least if Isaacson's thesis holds. For, if Isaacson's thesis is true, then **PA** (i) proves only true statements (soundness), (ii) proves *all* arithmetical truths in the sense of Isaacson (I-completeness), and (iii) proves *only* arithmetical truths in the sense of Isaacson (purity).

With this notion of purity in mind, we can now move on to see what it would mean for **PA** to be an impure theory under Isaacson's framework.

3.2 Impurity as a potential problem

We start by noticing, as Isaacson does, that some statements that are provable in **PA** seem to belong to the class of statements that Isaacson dubs higher-order, since they are about infinitary objects, or because they involve purportedly higher-order syntactic notions, such as consistency, for axiomatic systems of elementary number theory. An example of the former is transfinite induction

[5] Our main reason to employ the term 'purity' instead of simply calling this notion 'soundness in Isaacson's sense' or 'I-soundness' is to keep any possible confusion away. Thus, when we say that a theory is *unsound*, this is often associated with the theory proving a false statement. But this is by no means what goes on when we say that **PA** might be impure. Therefore, to avoid misleading claims, we leave the term 'soundness' and derivatives aside.

[6] What's more: some relaxed form of purity might also be deemed a desirable property of mathematical theories with a clearly defined and restricted domain (sometimes known as 'non-algebraic theories'), since it guarantees that they do not 'overshoot' in relation to that intended matter. To give an example, it would be rather unsettling if we were to show that, from the axioms for Euclidean geometry, one can prove the existence of a Mahlo cardinal.

for any ordinal $\alpha < \varepsilon_0$, that we will denote $\mathtt{TI}(< \varepsilon_0)$;[7] clearly, ordinals like ω^{ω^3} are infinite—but **PA** shows ω^{ω^3} is well-ordered! An example of the latter is Con(**PRA**), the sentence that formalizes the consistency of Primitive Recursive Arithmetic.

Why, Isaacson asks, when statements similar to these ones are not provable in **PA** (i.e., $\mathtt{TI}(\varepsilon_0)$ or Con(**PA**)), are we justified in taking their corresponding $\mathcal{L}_0$-formula to be a higher-order truth, and such move is not available when the statements are **PA**-provable? The reason, he argues, is that the very same tool that helped uncovering the higher-order nature of the former statements, namely coding (broadly understood), also reveals the arithmetical nature of the latter statements. The possibility of arithmetizing its syntax allows **PA** to speak about syntactic notions, 'coding' such notions with strings of number-theoretic constants; an ordinal notation system does the same in relation to infinite ordinals. And the application of coding, Isaacson argues, suffices to realise that these kinds of sentences are, after all, arithmetical in nature: as an auxiliary device, coding 'pulls the ostensibly higher-order truth into the arithmetical' (1987/1996, 221) and allows for a proof of the statement in strictly number-theoretic terms, which is all we need for the statement to count as arithmetical. Note that this is a consequence of Isaacson's epistemic approach to arithmetical truth: arithmeticality is not solely a feature of the statement in question but of the way we come to see its truth.

This cannot be taken, however, to be a conclusive answer, as Isaacson acknowledges and we shall now explain. The reason has to do with the length of certain proofs in **PA** when these are strictly formulated in $\mathcal{L}_0$. Thus, there are certain statements whose **PA**-proofs in the language $\mathcal{L}_0$ exhibit too many symbols (e.g., certain transfinite induction claims, or consistency statements), and hence the only way to present a proof that a human agent might realistically follow is by employing seemingly higher-order notions, e.g., infinite ordinals. Indeed, this is the reason we work with the latter and not their notations in proving, e.g., $\mathtt{TI}(\omega^{\omega^\omega})$ in **PA**. Given the correctness of our ordinal notation,[8] we know that there exists a corresponding proof with formulas that strictly belong to $\mathcal{L}_0$. But such a proof would be too long to be carried out in practice, so the deployment of uncoded infinite ordinals becomes indispensable for the presentation of the proof.

Now, if this is the case, someone could reason in the following way. First (1), as we have seen, under Isaacson's epistemic take on arithmetical truth, what allows us to establish the arithmeticality of a given statement expressible in $\mathcal{L}_0$ is the perceivability of a statement as true on the basis of a proof stripped of higher-order notions and consisting of arithmetical truths alone. Second (2), as just said, the proofs of some $\mathcal{L}_0$-expressible statements solely consisting in arithmetical truths formulated in $\mathcal{L}_0$ is too long to be surveyed. Third (3), the sort of proof that is surveyable for these statements employs seemingly higher-order notions. Then, the second and third claims lead to (4) the possibility, as Isaacson admits, that in these cases 'the higher-order perspective is essential for actual conviction as to truth of the arithmetically expressed sentence.' (1987/1996, 221). But this, together with the first claim, suggest that (5) we cannot establish the arithmeticality of these statements. Therefore, one might be inclined to conclude that Isaacson's understanding of arithmetical truth entails that these statements are not arithmetical truths, despite being provable in **PA**. Thus, for instance, $\mathtt{TI}(\omega^{\omega^\omega})$ might be a statement provable in **PA**, but not an arithmetical truth.

A key component of this problem concerns the idea of being provable in practice, that is, what follows from proofs that a human agent might realistically be able to check, versus what is simply provable in principle. The move, in the previous paragraph, from (1) and (4) to (5) relies on the idea that perceivability via proof consists somehow in being able to check the proof by oneself, i.e.,

[7] The expression is a little sloppy here: $\mathtt{TI}(< \varepsilon_0)$ is a schema, that is, needs to be instantiated by some formula. Let's take that for granted in what follows.

[8] See e.g. (Pohlers, 2009, Th.3.3.17) for a theorem establishing such correctness.

that the statement in question is provable in practice. Thus, Isaacson contends that someone who accepts that provability in principle in **PA** is sufficient to define the boundaries of arithmeticality need not worry further. Insofar as a statement is in principle provable in **PA** solely through formulae formulated in $\mathcal{L}_0$, the statement counts as arithmetical:

> If one is prepared to countenance a notion of being 'in principle' derivable in PA, then the present problem disappears. One might consider that this move is legitimate, as enabling one to define precisely a theoretical boundary, to which mathematical practice approximates. (Isaacson, 1987/1996, 221)

However, and as we have seen, Isaacson's thesis puts the emphasis on the *epistemic* character of arithmeticality. Therefore, there is a strong case for demanding that proofs be feasibly apprehensible, and not solely ideally apprehensible—an attitude we shall call the 'feasibility attitude'. Being arithmetical is here as much a product of our possibility to perceive the truth of the statement as it is a product of the language in which the statement can be expressed. Hence, it looks as if *followability* is a reasonable constraint on what counts as a proof that allows us to establish the arithmetical nature of a statement. This is something that Isaacson (1987/1996, 221-222) concedes: 'I have in my discussion been considering provability in terms of providing a basis for perceiving the truth of a given statement. In these terms, a proof in PA of a given proposition being infeasibly long has to be taken seriously.'

The problem is that the feasibility attitude, despite being a reasonable one, has an important implication. Since it gives us reasons to buy the argument above, and to conclude that some **PA**-provable statements are not arithmetical truths, it also leads to what we have called the *impurity concern*: the concern that ***PA** might be an impure theory of arithmetic*. Impurity here must be understood as above, i.e., as implying that some statements provable in the theory of arithmetic are not arithmetical truths in Isaacson's sense—roughly, that **PA** proves too much for a theory of arithmetic. Thus, under this implication, Isaacson's thesis as we presented it here collapses; for **PA** might still be I-complete, and hence prove all arithmetical truths, but it is no longer the case that it only proves arithmetical truths. The situation is depicted in figure 1 below: arithmetical truth would be a proper subset of the set of **PA**-provable truths, which is in turn a proper subset of true statements expressible in $\mathcal{L}_0$ (due to Gödel's theorem).

This is something Isaacson himself acknowledges, for he grants that, should one adopt the feasibility attitude,

> then within the arithmetically expressible truths of mathematics, we must think of the boundary between those which are purely arithmetical and those which are essentially higher-order as running somewhat inside the collection of those for which derivations in PA exist. (Isaacson, 1987/1996, 221)

In recent conversation, Isaacson has made clear to me that he favours an 'in-principle' take on provability. His opinion seems to be that the feasibility attitude puts one on the road of strict finitism, an undesirable philosophy of mathematics that Isaacson now, and unlike then, definitely rules out. Be that as it may, and as we have argued, we still think that the epistemic turn on arithmetical truth fostered by Isaacson makes a case for the feasibility reading. Thus, in the remaining of the paper, we follow that reading.

3.3 Isaacson's proposed way-out

As it happens, Isaacson offers a solution to the impurity concern on behalf of the advocate of the feasibility attitude. To follow his reasoning, let us recap the problem: there are true statements

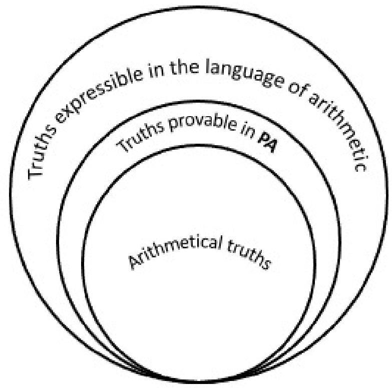

Figure 1: The relation between arithmetical truth, truths provable in **PA** and truths expressible in the language of arithmetic if Isaacson's thesis is true and **PA** is impure

expressible in $\mathcal{L}_0$, e.g., $\mathtt{TI}(\omega^{\omega^\omega})$, that can be proved in **PA** either employing seemingly higher-order notions embedded in a relatively short proof, or using solely formulae of $\mathcal{L}_0$ but with an unsurveyably long proof. Now, we could appeal to the mere existence of the latter proof in **PA** (even if it is humanly ungraspable) to argue that the higher-order notions are not indispensable. But, given his epistemic approach to arithmeticality, in which a proof has to be a vehicle to perceive the truth of a statement, the advocate of the feasibility attitude does not buy that argument, and will remain at best skeptical regarding the arithmeticality of such statements, leaving a door open for the impurity concern. Then, and possibly with the aim of avoiding the implications linked to this concern, Isaacson makes a move on behalf of such hypothetical advocate. According to Isaacson, one could reject extremely long proofs, such as the one for $\mathtt{TI}(\omega^{\omega^\omega})$ or the one for Con(**PRA**), as genuine proofs in **PA**. As a result, 'provable in **PA**' would acquire a new, more limited character, and the set of truths provable in **PA** would coincide with the set of arithmetical truths. This can be visualized by considering again figure 1: the circle that represents truths provable in **PA** 'shrinks' to the boundaries of the circle of arithmetical truths. In this case, the impurity concern no longer applies: all statements that we can consider as *genuine* proofs of **PA** are arithmetical.

Now, let me counter this move. There are at least two considerations as for why we might not want to reject very long proofs as genuine **PA**-proofs. In the first place, doing so deprives **PA** of its privileged proof-theoretic status among first-order axiomatizations of elementary number theory. After all, **PA** is widely considered as the *strongest* first-order theory of arithmetic that directly follows from our standard understanding of the natural number structure as exposed in the work of Dedekind. This is a key point underlying Isaacson's thesis: to a great extent, the proof-theoretic privilege buttresses the epistemic privilege that Isaacson defends for **PA**. Theories like $\mathrm{I}\Sigma_n$, for $n \in \omega$, can also be said to follow directly from our understanding of the Dedekian analysis; but, crucially, they are weaker than **PA**, and hence do not enjoy the same *proof-theoretic* privilege. Now, the standard measures of relative proof-theoretic strength between two theories of arithmetic

S and T in the style of **PA**, **PRA** or $\mathrm{I}\Sigma_n$ include the determination of their proof-theoretic ordinals, and whether $S \vdash \mathrm{Con}(T)$ or vice versa. Then, suppose that the length in symbols of the shortest **PA**-proof of, e.g., $\mathrm{Con}(\mathrm{I}\Sigma_1)$, is n, while the length of the shortest **PA**-proof of any instance of transfinite induction up to ω^ω in **PA** is m. Suppose further, following Isaacson's suggestion, that we only admit proofs in **PA** of symbol length less than $\min(n, m) = k$. In other words, and if $\vdash_k$ is the symbol we use for this restricted notion of provability in **PA** and $\rho(x)$ is a function that gives the length in symbols of the shortest proof in **PA** of the formula represented by x, we write $\mathbf{PA} \vdash_k \varphi$ iff $\mathbf{PA} \vdash \varphi$ and $\rho(\varphi) < k$. The result is then that $\mathbf{PA} \nvdash_k \mathrm{Con}(\mathrm{I}\Sigma_1)$ and $\mathbf{PA} \nvdash_k \mathtt{TI}(\omega^\omega)$. Therefore, it is no longer clear whether **PA** is in any standard way proof-theoretically stronger, and hence more privileged, than, in this case, $\mathrm{I}\Sigma_1$: **PA** understood in this new way does not prove the consistency of $\mathrm{I}\Sigma_1$, nor can it be said to have a larger proof-theoretic ordinal.[9]

In the second place, it seems likely that the downgrading of **PA** could happen not only at the proof-theoretic but also at the strictly number-theoretic level. That is, the issue is not only that the restrained view leaves out of **PA** statements that are of interest to the logician but only of relative interest to the number-theorist—statements of proof-theoretic nature, or syntactic statements, like Con(**PRA**)—but also that we might need to equally give up on certain important number-theoretic theorems from being considered **PA**-provable. For consider Theorem 2 in (Buss, 1994), an analogue of Godel's famous speed-up theorem on the length of proofs measured by number of steps (Gödel, 1936). The theorem shows that, for each computable function Φ, there are infinitely many different formulae x provable in **PA** (or in any first-order theory of arithmetic, for that matter) such that $\rho(x) > \Phi(\rho_2(x))$, where $\rho(x)$ is defined as above and $\rho_2(x)$ is the length of the shortest proof of x in $\mathbf{PA}_2$. Now, let's suppose that, among all instances of transfinite induction up to ω^ω, the instantiation with formula φ is the one whose shortest proof involves the greatest number of symbols, and that the proof is too long to be surveyed. Using upper corners ($\ulcorner\urcorner$) to indicate that what comes inside corresponds to the 'coded', $\mathcal{L}_0$ version of the formula, we write $\rho(\ulcorner\mathtt{TI}(\omega^\omega, \varphi)\urcorner)$ for the shortest proof of the instantiation with formula φ of the transfinite induction schema up to ω^ω. And, following Isaacson's suggestion, let's suppose that only proofs of length $< \rho(\ulcorner\mathtt{TI}(\omega^\omega, \varphi)\urcorner)$ are accepted. Then, we can find a computable function Ψ such that $\Psi(\rho_2(\ulcorner\mathtt{TI}(\omega^\omega, \varphi)\urcorner)) = \rho(\ulcorner\mathtt{TI}(\omega^\omega, \varphi)\urcorner)$. After that, it is not difficult to generate a countably infinite number of computable functions Ψ' that bound Ψ from above, i.e. such that

$$\Psi(n) \leq \Psi'(n), \text{ for all } n \in \mathbb{N}$$

For each of those Ψ', Buss' result tells us that there are infinitely many different formulas of $\mathcal{L}_0$ that are provable in **PA** and such that the length of their shortest proof is greater than $\Psi(\rho_2(\ulcorner\mathtt{TI}(\omega^\omega, \varphi)\urcorner))$. However, all these formulas need to be considered as unprovable in **PA**, or at least as formulas the proof of which are not genuine for **PA**. There are thus infinitely many different theorems of **PA** that we stop considering as such. And it might well be possible that relevant number-theoretic results (say, Fermat's last theorem) are included among these many formulae. Plus, this is not merely a speculative point: we *know* that there are relevant number-theoretic theorems of this sort. A well-known example includes the instances of Friedman's finitization of Kruskal's theorem. This finitization is a universal statement of the form $\forall k \exists n \psi(k, n)$ and is known to be independent of **PA**. Nonetheless, its particular instances, i.e., $\exists n \psi(m, n), m \in \omega$, are provable in **PA**—but, in most cases, their proofs incorporate a disproportionate number of symbols.[10].

[9]Should someone suggest that a weaker metatheory of arithmetic can still show that **PA** proves these claims, or even establish a relative consistency proof, we can insist that what the metatheory should prove, and obviously cannot, is that these results must be recovered in the new, restricted conception of **PA**.

[10]For instance, it is known that proving $\exists n \psi(10, n)$ requires at least the number of symbols represented with an exponential tower of one thousand 2's (Smoryński, 1982).

Of course, someone who accepts Isaacson's move might already be aware of this consequence, namely that **PA** would be very lacking as a formal axiomatization of number-theory, and willing to accept it. This might align them indeed with a strict finitist philosophy of mathematics, as we said Isaacson thinks. But since our discussion had so far been framed in terms of stripping **PA** of theorems like Con(**PRA**) or transfinite induction claims, the consequence just drawn might not have been evident to someone keen on giving up these more logical statements, but still thinking that **PA** should get most number-theory right.[11]

These two considerations suggest that provability in **PA** cannot be so freely adjusted to match the set of arithmetical truths, and we are left with the impurity concern under the feasibility reading of Isaacson's thesis. The remaining of the paper will now be devoted to showing how we can still avoid this concern with arguments different to those of Isaacson.

4 Resisting the impurity concern

In the previous section, we argued that, according to certain reading of Isaacson's work, **PA** could be impure, and that this would certainly be a blow to Isaacson's thesis. The reading in question epistemically favours proofs that can be feasibly apprehended, as opposed to unsurveyably long, humanly unapprehensible proofs. As we saw and objected to, Isaacson suggests that the advocate of the feasibility attitude may just do away in **PA** with all those statements the proof of which is too long to be carried out in practice. But we argued that such an advocate should not take the path delineated by Isaacson. As we pointed out, it also seems that Isaacson himself would accept this conclusion now, having identified that this path leads to strict finitism—and would discard the feasibility attitude altogether.

Nonetheless, this last move makes us think that Isaacson might have conflated two views that need to be distinguished: the feasibility attitude as regards arithmetical truth, and the feasibility attitude as regards derivability in a theory of arithmetic. That is: one can defend the view that feasible apprehensibility must be a criterion for actual perceivability of the truth of a statement and thus, following Isaacson, of its arithmetical nature; and one can defend the view that feasible apprehensibility must be a *formal* criterion for derivability over a theory of arithmetic. Only the latter seems to be related to strict finitism (sometimes also known as ultrafinitism). The former, on the contrary, just concerns what we can consider arithmetical in Isaacson's sense. Now, in what follows, we try to show that the as regards arithmetical truth is on safe grounds, so that even those statements that fall outside the scope of what is feasibly apprehensible with statements written in

[11] A reviewer of this paper has pointed out a further issue which I had overlooked, and for which I am grateful. As it happens, there are *two* kinds of advocates of the feasibility attitude. One such kind is the one I am assuming all along, namely an advocate for the view that any statement that can be proved in a feasible number of steps can rightly be called an arithmetical truth. Of course, what 'feasible' exactly means here is to be kept loose, as it might involve a lengthy discussion. Perhaps one can conjecture that a statement with Rayo's number-steps is already unfeasible. The other kind of advocate of the feasibility attitude understands 'in practice' as that which has or will be proved. And so the aforementioned issue stems out of this view: since, presumably, the totality of human agents that there was, there is, and there will be can only establish the arithmeticality of finitely many claims, this advocate must conclude that the class of arithmetical statements is finite. Further, if we buy Isaacson's proposal that provability in **PA** be restrained to what we can prove in practice, they would need to accept that the class of **PA**-provable statements must also be finite.

We take any of these consequences to be truly undesirable. But we also take the second kind of feasibility attitude to be extremely unpalatable. While marginal, the first feasibility attitude seems to have been held by certain finitists. To the best of our knowledge, no one has ever held anything like the second attitude. Among other things, this might have to do with the fact that any such advocate will be accused of not having understood the modality involved the notion of '*provable* in practice'. The substantial question is what can be considered a proof, not how many proofs are actually carried out. So, while a fully-fledged dismissal of this attitude is outside the scope of this paper, we will not consider it further.

the language $\mathcal{L}_0$ alone can, by other means, be considered arithmetical on Isaacson's understanding of the term. That is, we will argue that we can establish the arithmeticality of these statements in a way other than following the proof with only $\mathcal{L}_0$-formulae in **PA**. Or, to be more precise, what we propose is more of a conjecture—a conjecture whose establishment can, in a sense, be understood as complementary to Isaacson's original project, and in fact suggested by Isaacson himself in his remarks on $\mathtt{TI}(\omega^{\omega})$ (1987/1996, 221). The conjecture in question is to be summed up as follows:

Conjecture. *There is a way to justify the arithmetical nature of each statement whose proof in the language of* **PA** *is too long to be carried out in practice, but which is nevertheless provable in* **PA** *in principle.*

The idea behind the conjecture is that, for any statement S whose proof employing strictly $\mathcal{L}_0$-formulae is unsurveyable, but which we know to be in principle provable in **PA**, there is some argument that settles the arithmeticality of such statement. Some examples of argumentative strategies of this sort include, but might not be limited to, showing that S is equivalent to some other statement S' which is accepted as an arithmetical truth, or demonstrating that some proof of S which is not formulated in $\mathcal{L}_0$ is nonetheless based solely on arithmetical truths. In these cases, the feasibility attitude is respected: a surveyable proof is still needed to establish the arithmeticality of a statement. Still, if the conjecture holds, the threat of impurity for Isaacson's thesis fades away: suspected higher-order truths of the sort presented in section 3 could be shown to be arithmetical truths.

How can we defend this conjecture? The option we follow, in line with Isaacson's original paper, consists in examining some case studies. We look at two paradigmatic kinds of statements that may lead to the impurity problem: transfinite induction claims and consistency statements. Or rather: we will be looking at only one of these, transfinite induction claims, and, we believe, this will suffice to show that we can justify the arithmeticality of consistency claims too. The reason is that claims of the form $\mathrm{Con}(T)$, where T is a theory of arithmetic weaker than **PA**, can be proven equivalent to a transfinite induction claim up to a certain ordinal below ε_0, over a subsystem of **PA** proof-theoretically weaker than T itself.[12] This follows from the fact that each of these first-order subsystems, which are weaker than **PA**, has a proof-theoretic ordinal strictly smaller than ε_0. Hence, should we show that all transfinite induction statements up to ε_0 are, after all, arithmetical truths, we could conclude that all syntactic statements of this sort are arithmetical truths: epistemically, the truth of the syntactic statement would be perceivable insofar as the entailment can be established via a first-order derivation that only employs other established arithmetical truths.

Thus, we will try to provide evidence for our conjecture as follows. The problem of impurity with statements such as $\mathtt{TI}(\omega^{\omega^{\omega}})$ is that their not-so-long proofs make use of infinite ordinals and not their notations, which seem to be higher-order (infinitary) notions. Therefore, it might seem as if the only way we can feasibly carry out a proof in **PA** of certain features (i.e., well-orderings) of these ordinals is a proof which is essentially higher-order. We will show that this intuition is mistaken. In order to get there, we inspect the argument which constitutes the proof in **PA** of the transfinite induction claim in question, and argue that no higher-order resources are employed in such an argument. Furthermore, we later draw on the proof given to argue that the ordinals the proofs are about, i.e. the ordinals which **PA** proves well-ordered, are finitary in nature.

[12] An earlier version of this paper suggested that the reason had to do with the entailment being provable in **PA**. But clearly this does not suffice, and I thank the audience at the Konstanz Summer School on the Phil of Mathematics for fruitful discussion on this point. After all, *every* **PA**-provable statement is entailed by any other statement over **PA**. And we do not want to say that the arithmeticality of Con(**PRA**) is granted by the fact that $0 = 0$ is an arithmetical truth. It is the special connection between transfinite induction and consistency that must do the job.

Putting the pieces together, we will then conclude that the argument of the standard proof in **PA** of transfinite induction claims like $\mathtt{TI}(\omega^{\omega^{\omega}})$, even when given in terms of ordinals and not notations, involves no higher-order notions whatsoever, and are thus based solely on arithmetical truths. In other words, we will be showing that a proof of the claim can be given which is not formulated in $\mathcal{L}_0$ yet which is based solely on arithmetical truths. But this—and given the *epistemic* ideal of arithmetical truth that underlies Isaacson's thesis, by which a derivation in first-order logic from known arithmetical truths suffices to establish that statement as an arithmetical truth—will be enough to assert that $\mathtt{TI}(\omega^{\omega^{\omega}})$ and similar statements are arithmetical truths in Isaacson's sense, and hence to dispel the threat of impurity generated by transfinite induction claims.[13]

4.1 The proof of transfinite induction

The first question we address then is: how can **PA** prove transfinite induction claims, i.e., well-orderings, for *infinite* ordinals? How can we make sense of the fact that the theory of finite mathematics is able to deal, manipulate and establish properties of these infinite objects? We believe that the way to approach these questions relates to the nature of the supremum of all ordinals for which transfinite induction claims are provable in **PA**: ε_0. The point is that the way **PA** deals with ordered sets of order-type (or lists/sequences/proof-trees of length) less than ε_0 does not go beyond the strictly finite, as we will now see; therefore, they are somehow tractable in a finitary way.

In order to clarify what we mean here, we turn to the proof in **PA** of transfinite induction for all ordinals $\alpha < \varepsilon_0$. This result requires a primitive recursive well-ordering of the natural numbers of order-type $(\varepsilon_0, \prec)$ obtained, by coding, from the Cantor Normal Form Theorem for ordinals of base ω.[14] Whereas the original proof is due to Gentzen (1943), we consider a more up-to-date version by Halbach (2014, 204-7). The proof in question relies on two lemmas. The first of them is the following:

Lemma 1. $\mathbf{PA} \vdash \mathit{Prog}(\varphi) \to \mathit{Prog}(\mathcal{J}(\varphi))$

where Prog(φ) (that reads 'φ is progressive') is the formula $\forall\alpha(\forall\beta \prec \alpha\, \varphi(\beta) \to \varphi(\alpha))$, and $\mathcal{J}(\varphi)$ is the formula $\forall\alpha(\forall\xi(\forall\eta \prec \xi\, \varphi(\eta) \to \forall\eta \prec \xi + \omega^\alpha\, \varphi(\eta)))$.

And, as for the second lemma:

Lemma 2. *If*

$$\mathbf{PA} \vdash \mathit{Prog}(\varphi) \to \forall\xi \prec \alpha\, \varphi(\xi)$$

for all formulas φ of $\mathcal{L}_0$, then

$$\mathbf{PA} \vdash \mathit{Prog}(\varphi) \to \forall\xi \prec \omega^\alpha\, \varphi(\xi)$$

for all formulas φ of $\mathcal{L}_0$.
NB: these expressions correspond to $\mathtt{TI}(\alpha)$ and $\mathtt{TI}(\omega^\alpha)$, respectively.

[13] An anonymous reviewer has kindly pointed out that the strategy that we follow here, via Gentzen's proof of transfinite induction up to ε_0, is only necessary for ordinals $> \omega^\omega$. Transfinite induction for infinite ordinals up to, and including, ω^ω can be obtained in alternative fashions. For example, one can consider the set of finite sequences ordered by the so-called shortlex ordering (that is: any two sequences are ordered by first comparing their lengths and, if the latter are equal, employing the lexicographical order—see e.g.(Mancosu et al., 2021, ch.8)). This is indeed a well-order of type ω^ω.

[14] This theorem shows that any ordinal below ε_0 can be written as the sum of powers of ω with exponent $< \varepsilon_0$, whereas ε_0 itself and greater ordinals cannot.

Transfinite induction up to any ordinal below ε_0 can be reached by applying Lemma 2 finitely many times, and Lemma 2 is easily obtainable from Lemma 1. It is thus the latter that requires careful examination. And it is in fact the crux of the proof, for it is where the interweaving with infinite ordinals happens. The formula $\mathcal{J}(\varphi)$, sometimes known as Gentzen's jump formula, lies at the heart of this lemma. In all cases in which it is instantiated with $\alpha \succeq 1$, Gentzen's jump formula seems to announce the possibility of 'infinite jumps'. We can (very informally) understand the jump as stating that, when a given formula φ holds for all ordinals below a given one—finite or not—we can carry that formula along for ω^α-many more numbers above that ordinal. That is, it is as if we were indeed 'jumping' over powers of ω—taking an infinite leap the 'safety' of which (in the sense of well-foundedness) is guaranteed by Gentzen's formula. Notwithstanding these intuitions, we will now argue that these leaps are not infinite after all.

There is, however, a limit to these leaps. This limit is given by Cantor's Normal Form Theorem. Since Gentzen's jump formula works exclusively with towers of ω such that the next element of the tower is always smaller or equal than the previous one, ε_0 marks the boundary to the number of ordinals we can 'jump over'; hence, even if the jumps were infinite (contrary to what we argue below), they could not be of an arbitrarily big number of infinite ordinals. This is also why transfinite induction for ε_0 cannot be established with an argument in the style of lemmas 1 and 2: the inner structure of Gentzen's formula prevents us from reaching ε_0, and in this we see how pivotal this formula is for the proof. We will say more about this below.

Now, the other component of Lemma 1 is the notion of 'progressiveness', there abbreviated as *Prog*. To say that a formula is progressive is to say that, when it holds for all ordinals below a given one, it holds for that ordinal. Once we know that a formula is progressive, a transfinite induction claim for some ordinal α is just the assertion that, should the formula be satisfied by 0, progressiveness will carry the formula along the ordinal sequence all the way to α. This is all there is to transfinite induction, as Gentzen held (1943, 291); therefore, progressiveness is the cornerstone of transfinite induction. Yet the apparent mystery of Lemma 1 in relation to our project is that it shows that Gentzen's jump for a certain formula holds whenever the formula is progressive. That is, the formula is carried along 1 ordinal, and then ω ordinals, and then ω^2 ... and all the way to ω^ω and beyond. As such, the mystery lies in asking how it is possible that a finite, indeed unitary, increment in the satisfaction of a formula along the ordinal sequence can result in increments of the order of powers of ω.

The proof of Lemma 1 gives what we take to be a clear answer to this. If a formula φ is progressive, $\mathcal{J}(\varphi(0))$ holds trivially, for it just expresses that φ is carried one ordinal forward. Informally, **PA** 'sees' the unitary jump as safe (in the sense above, i.e., of well-foundedness).[15] Now, for $\mathcal{J}(\varphi)$ to be progressive, $\mathcal{J}(\varphi(1))$, i.e., $\forall\xi(\forall\eta \prec \xi\,\varphi(\eta) \to \forall\eta \prec \xi + \omega\,\varphi(\eta))$, must hold. The key then is that, although we seem to face an ω jump, it is after all a finite one. **PA** is given a certain ordinal ξ as input and has to carry that property for a number of ordinals *below* ω (for whatever η we pick, it will be strictly less than $\xi + \omega$). Hence, **PA** only needs to reiterate what it already 'sees' as a 'safe jump' (the unitary one) a given finite (hence, also safe) number of times. A very similar reasoning goes for $\mathcal{J}(\varphi(2))$: since **PA** 'sees' the ω-jump as safe now, it can perform it once and combine it with a finite number of steps (or perform it twice!) to leap just under ω^2-many ordinals. The same idea applies to any jump made over ω^n ordinals. Thus, in more formal terms, we are performing an outer or external induction on n for ω^n—allowing us to conclude that the jump must be safe, in the sense of being well-founded, for ω^ω ordinals.

Likewise, when we consider powers of ω of the form $\omega^\alpha, \omega^\omega \succ \alpha \succeq \omega$, the induction is happening at the next exponential level. That is, having been able to establish the safety of jumps over ω^ω-

[15]The reader need not interpret 'sees' here in anything like a model-theoretic sense, as a model that 'thinks' of itself in a certain way (e.g., as containing uncountable objects despite being countable, as given by Skolem's paradox). It is just a very informal way to describe the operations that are going on in **PA** to reach the desired results.

ordinals as above, we perform now an induction on n for ω^{ω^n}. This will allow us to conclude, in turn and by induction, that jumps over ω^{ω^ω} ordinals are also safe. Unsurprisingly, one will say argue in the same way for any ω with exponent $\prec \varepsilon_0$. Since induction is an entirely arithmetical task, in the sense that its correctness can be seen to follow from the number-theoretic content of our categorical conceptual analysis of the notion of natural number, **PA** can carry out these nested inductions, one after the other, to complete the transfinite induction. Even if the ordinals themselves are infinite, their structure is such that ordinary induction need only be performed a finite amount of times, and so **PA** can deal with it.

Hence, the mathematical procedures underlying the proof of transfinite induction for ordinals below ε_0 has an arithmetical nature: we need not invoke any proof resources other than number-theoretic induction to establish that these ordinals are well-ordered and, *a fortiori*, we need not invoke higher-order proof resources.

4.2 The finite nature of (some) infinite ordinals

Despite the above, here is a reason one may doubt that we have really shown the arithmetical nature, in Isaacson's sense, of transfinite induction claims like $\mathtt{TI}(\omega^{\omega^\omega})$. One can think that, since transfinite induction claims are about ordinals, the equivalent $\mathcal{L}_0$ statement will involve coding techniques for these ordinals. And does not the presence of coding threaten the arithmetical status of the $\mathcal{L}_0$-based formulation of $\mathtt{TI}(\omega^{\omega^\omega})$? The answer to this worry is: not necessarily. As we mentioned, under Isaacson's framework, coding is simply a device that, *in most cases*, allows us to discern whether a seemingly higher-order truth is arithmetical after all, or whether a seemingly arithmetical truth is higher-order. So the mere fact that the $\mathcal{L}_0$-based formulation of $\mathtt{TI}(\omega^{\omega^\omega})$ involves coding is not, per se, problematic. What was problematic, at least for the advocate of the feasibility attitude, was precisely the fact that here coding cannot directly serve as the vehicle to establish the arithmeticality of the $\mathcal{L}_0$-based formulation of $\mathtt{TI}(\omega^{\omega^\omega})$, because its application renders a proof too long to be surveyed. Accordingly, what we have been trying to show is that the uncoded version of the statement $\mathtt{TI}(\omega^{\omega^\omega})$—that is, the statement asserting that the *ordinal* (as opposed to the *code* for the ordinal) ω^{ω^ω} is well-founded—is not higher-order but arithmetical. And we have done it by verifying that the argument by which (the coded version of) such a claim is proved in **PA** employs no higher-order procedures. Hence, this would entail that the $\mathcal{L}_0$-based formulation of $\mathtt{TI}(\omega^{\omega^\omega})$ is also an arithmetical truth, given the correctness of our coding apparatus. Thus, this would be a way to verify that, in this case, and to use Isaacson's words, coding constitutes a 'linkage [that] pulls the ostensibly higher-order truth down into the arithmetical' (Isaacson, 1987/1996, 221).

There is however a second point that the reader may raise here. All we have shown is that the steps that constitute the argument by which **PA** can establish $\mathtt{TI}(< \varepsilon_0)$, and thus by which we come to see the arithmeticality of this statement, are of a finitary nature. That is, we have outlined the core of a proof that establishes such claims, and which can be fully formalised. But note that, in this outline, we have availed ourselves to *infinite* ordinals all along, instead of their notations in the language $\mathcal{L}_0$. And we could have not in fact employed the notations since these are, in many cases, likely to render such an argument unsurveyable; in other words, the argument above cannot be formalised in $\mathcal{L}_0$ without becoming unfeasibly long, and any feasible formalisation would seem to appeal to infinite ordinals. Yet these ordinals seem to be higher-order concepts, given that they might be considered infinitary in nature. Hence, insofar as the argument expounded is the core of the proof of $\mathtt{TI}(< \varepsilon_0)$, it does seem that the way by which we may convince ourselves of the arithmeticality of $\mathtt{TI}(< \varepsilon_0)$ *does* make use of higher-order concepts after all. That is, the possibility that 'the higher-order perspective is essential for *actual* conviction' (Isaacson, 1987/1996, 221) of the arithmeticality of the claim seems not to have vanished. Again, insisting that these ordinals are translatable into expressions in the language $\mathcal{L}_0$ is of no use, as the impossibility to do that without ending up with an unsurveyable proof is what brought us here in the first place. Thus, our aim now is to show that these ordinals are finitary in nature, and hence not higher-order, so as to conclude that the arguments used in the proofs of $\mathtt{TI}(\omega^{\omega^{\omega}})$ and similar claims involve no higher-order concepts at all. The key is to reflect on the proof of the transfinite induction claims just displayed.

We explained that the main argument for the proof in **PA** establishing that ordinals below ε_0 are well-ordered consists in exploiting a nested induction applied to the inner structure of these ordinals. But the possibility of this nested induction is only given in the first place because the structure of these ordinals is finitary. Indeed, what facilitates the nested induction is the fact that ordinals up to ε_0 are capable of being treated as finite objects, as is revealed by their specific Cantor Normal Form. And what do we mean by this? Cantor's Normal Form Theorem shows that we can see any ordinal below ε_0 as a sum of towers of ω of the form $\{\alpha_0, \alpha_1, ..., \alpha_n\}$ where $\alpha_i \leq \alpha_j$ when $i < j$ and $\alpha_i \leq \omega$ for each i. Due to this, one can think of any such ordinal as a finite list with two types of elements: further two-sorted lists, or individuals—symbolised, for instance, by $\lhd$. Thus, the ordinal ω^{ω^2} can be understood as a list with one element: another list, itself containing yet another list, which finally contains two elements: $\lhd, \lhd$. On the other hand, the ordinal $\omega^{\omega^2} + 2$ can be understood as the same list, now containing also $\lhd$ and $\lhd$. In following this idea, the theorem allows us to understand each ordinal below ε_0 as a finite list, the members of which are also finite lists, the members of which are also finite lists... and so on. What all of this reveals, in any case, is that the structure of the ordinal itself responds to a finitary nature. And, as we said, this makes possible the overall inductive procedure: **PA** 'sees as safe' (in the sense given in the previous subsection), i.e., apprehends as well-founded, each list in the construction of the ordinal, and it can also easily establish the well-foundedness of a finite sequence of individuals.

To be clear, this is a semi-informal picture that aims to uncover the finitary structure of the ordinals we are interested in.[16] Of course, alternative pictures are also possible. For example, we can see the ordinal as a finite tree, the nodes of which are finite trees, the nodes of which are finite trees, etc. Any such picture will hopefully lead to conviction as to the fact that these infinite ordinals $\omega \leq \alpha < \varepsilon_0$ can truly be said to belong to the realm of finite mathematics, and hence not to be higher-order concepts.

In fact, while this dividing line between the finitary and infinitary, to be located well into the infinite ordinals, may initially come as a surprise, it becomes increasingly less so as we learn of different situations where the link between infinite ordinals below ε_0 and finitary mathematics is made explicit. Some of these examples have been thoroughly investigated in the literature. The following are just two of them:

- The set of ordinals below ε_0, equipped with the usual well-ordering of ordinals, is isomorphic to the set $\mathbb{N}$ with the ordering induced by the so-called Matula numbers—see (Weiermann, 2005).
- Weiermann (2002) has shown that the behaviour regarding limit laws (roughly, the probability that any property holds in a structure of arbitrarily large size) for classes of infinite structures of order type up to ε_0 is continuous with that for classes of structures of finite size (and hence order type), assuming certain background conditions on the order. In particular, when seen as additive systems, these classes of structures meet the so-called zero-one law, that is, all properties have probability either 0 or 1 to be satisfied in structures of arbitrarily large size,

[16] I would like to thank the reviewers of this paper for pointing out that a previous argument I advanced to the effect that ordinals up to ε_0 followed from our understanding of natural numbers, based on Kreisel's notion of ordinal visualization (Kreisel, 1965), is not available here. First, because the argument relied on the positing of 'arithmetical concepts', which can be problematic in the context of Isaacson's thesis. Secondly, because a similar argument would establish the arithmetical nature of $\mathtt{TI}(\varepsilon_0)$, contradicting Isaacson's thesis.

whether finite or infinite, as long as the order type induced is less than ε_0. In loose terms: finite structures and infinite structures of order type up to, but not including, ε_0, show a certain 'decidability' when it comes to satisfying any property.

4.3 Additional remarks

In this section, we have argued that: (i) the argument behind the proofs in **PA** of transfinite induction claims up to ε_0 does not employ higher-order resources, and (ii) the ordinals with which these arguments are presented are inherently finitary, and not higher-order. Thus, we conclude that a surveyable proof can be given for these claims that does not appeal to higher-order notions overall—which, in the spirit of Isaacson's thesis, suffices to be convinced that the statements these proofs establish can be considered arithmetical truths. Therefore, these claims are not a counterexample to the purity of **PA** under the interpretation of arithmetical truth given by Isaacson. We take this to render further support and plausibility to Isaacson's thesis.

What's more: we believe that the explanation provided reinforces Isaacson's thesis with respect to two additional and different (but related) fronts. First, because it gives an answer to a question that underlay the specific case study: how can **PA**, which according to Isaacson's thesis coincides exactly with the truths of *finite* mathematics, prove that certain *infinite* ordinals, i.e., seemingly higher-order objects, are well-founded? Our response consists in pointing out that these objects are not really higher-order in nature but, as Cantor's Normal Form Theorem reveals, finitary; and that precisely because of this, their inner structure of blended finite strings can be proved well-founded by **PA** through the application of ordinary induction finitely many times.

The second front has to do with some remarks presented by Gentzen in his original proof of transfinite induction up to ε_0 in **PA**, for whom the situation was the opposite of the one we have presented here. According to him, for an important segment of the countable ordinals (including ordinals well beyond ε_0), '*transfinite induction* is a form of inference which, *in substance, belongs to elementary number theory*', (Gentzen, 1943, 307, italics in original) so that '[t]he fact that transfinite induction even up to the number ε_0 is no longer derivable from the remaining number-theoretical forms of inference therefore reveals from a new angle the *incompleteness* of the number-theoretical formalism' (ibid.). In other words, he seems to suggest that transfinite induction for $\alpha \geq \varepsilon_0$ is a genuine arithmetical truth—as opposed to some true statement cooked-up by logicians—that is nevertheless independent of **PA**. Following Isaacson's terminology, we could read him as saying that transfinite induction for such ordinals is not 'higher-order'. Recently, Saul Kripke (2022) has defended a very similar idea, arguing that $\mathtt{TI}(\varepsilon_0)$ is the first genuine arithmetical true statement that was shown independent from **PA**. For both Gentzen and Kripke, the unprovability of $\mathtt{TI}(\varepsilon_0)$ is yet another example of the incompleteness of **PA** with respect to arithmetical truth, constituting thus a challenge to Isaacson's thesis. We believe, however, that our account of what underlies transfinite induction for ordinals below ε_0 renders important support (even if perhaps not decisive) to a very different conclusion, namely that $\mathtt{TI}(\alpha)$, for $\alpha \geq \varepsilon_0$, unlike transfinite induction for smaller ordinals, is not really arithmetical. After all, the widely accepted strategy for proving transfinite induction claims in **PA** (namely, the Gentzen proof just outlined) is not applicable for ordinals $\alpha \geq \varepsilon_0$, insofar as the inner structure of these ordinals does not allow Gentzen jumps.

5 In search of more evidence

As a reminder, our driving conjecture here is that there is a way to justify the arithmeticality of each statement whose proof is too long to be proved in practice, but which **PA** can still prove in principle. These might include statements about infinitary concepts, or certain syntactic statements.

To provide some evidence for this conjecture, we have justified the arithmeticality of transfinite induction claims for ordinals up to ε_0.

The argument deployed seems to do well not only with transfinite induction claims but with many other statements involving infinitary objects and, in particular, infinite ordinals (for instance, results on ordinal arithmetic; see e.g., (Sommer, 1995, §3)). Likewise, it seems to us that it fares well with respect to *consistency* statements about theories weaker than **PA**. But these statements by no means exhaust the class of 'syntactic' statements that might involve unfeasibly long proofs. For instance, we find that statements that code provability in a theory of arithmetic are of an equally syntactic nature. If we are to defend the conjecture—and, with it, Isaacson's thesis—one will have to tell a convincing story on why these statements are also arithmetical.

As a paradigmatic case, take the following: what can we make of Henkin sentences, that is, formulae φ such that

$$\varphi \leftrightarrow \mathrm{Pr}_{\mathrm{PA}}(\ulcorner \varphi \urcorner)?$$

Some considerations come into play here. First of all, there is no one single formula expressing provability in a formal system. The formula in question will depend, among other things, on the choice of coding made, and on the conditions we believe a formula expressing provability in a system should meet. The last point is particularly relevant, and has been the object of some discussion—see e.g. Halbach & Visser (2014). Indeed, if some formula $\pi(x)$ that is intended to express provability is generally believed to be unsuccessful for that aim, we are (arguably) no longer talking about a *syntactic* statement, insofar as it fails to capture the relevant syntactic property. Hence, formulae like the ones Kreisel devised to answer Henkin's problem (i.e., whether Henkin sentences are provable in their relevant systems) (Kreisel, 1954) might not be strictly relevant when it comes to testing the conjecture: since most would argue that they fail to capture a syntactic property (as Henkin, and Halbach and Visser, have done), we can discard right away their containing higher-order notions.

Thus, one could argue that it all boils down to justifying the arithmeticality of Henkin sentences expressed with the 'canonical' formula capturing provability, which we denote as $\mathrm{Bew}(x)$. It is at this point where the defender of the conjecture must step in and try to explain in what sense these types of sentences are arithmetical. We shall not attempt to do that there. Nonetheless, we venture that one can accomplish this task for the formulae in question by identifying provability with the existence of a certain *finite* sequence and, in turn, justifying the arithmeticality of the notion of 'sequence'.

6 Conclusion

In this paper, we introduced the notion of purity for theories of arithmetic, and showed that there is a reading of Isaacson's thesis under which **PA** can be considered an impure theory of arithmetic, thus undermining Isaacson's thesis. As we see it, two possibilities stand out now if such a conclusion is to be avoided. Either we take this to be significant evidence in favour of retaining the Tarskian conception of arithmetical truth as truth in $\mathcal{N}$, thus going back to the incompleteness of **PA**, or we find a way to justify the arithmetical character of statements such as $\mathtt{TI}(\omega^{\omega^{\omega}})$, Con(**PRA**) and the like. Here, we tried to pursue the second path. As we said, our argument is just conjectural, based on a paradigmatic case study, and more may need to be done. But, if the conjecture holds, it is definitely a way to buttress the claim that **PA** proves all and only arithmetical truths—that is, Isaacson's thesis.

References

Boolos, G., Burgess, J. P., & Jeffrey, R. (2007). *Computability and logic*. Cambridge University Press.

Bovykin, A. (2009). Exact unprovability results for compound well-quasi-ordered combinatorial classes. *Annals of Pure and Applied Logic*, *157*(2-3), 77–84.

Buss, S. R. (1994). On Gödel's theorems on lengths of proofs I: Number of lines and speedup for arithmetics. *Journal of Symbolic Logic*, *59*(3), 737–756.

Gentzen, G. (1943). Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der Transfiniten Induktion in der Reinen Zahlentheorie. *Mathematische Annalen*, *119*(1), 140–161. English translation 'Provability and nonprovability of restricted transfinite induction in elementary number theory' in M.E. Szabo (Ed.) *The Collected Papers of Gerhard Gentzen*, North Holland, 287-308.

Gödel, K. (1936). Über die Länge von Beweisen. *Ergebnisse eines mathematischen Kolloquiums*, *7*, 23–24. Republished with English translation 'On the length of proofs', in Solomon Feferman et al. (Eds.) *Kurt Gödel Collected Works Volume I*, Oxford University Press, 1986, 394-399.

Halbach, V. (2014). *Axiomatic theories of truth*. Cambridge University Press.

Halbach, V., & Visser, A. (2014). Self-reference in arithmetic I. *The Review of Symbolic Logic*, *7*(4), 671–691.

Horsten, L. (2001). Platonistic formalism. *Erkenntnis*, *54*(2), 173–194.

Incurvati, L. (2008). Too naturalist and not naturalist enough: Reply to horsten. *Erkenntnis*, *69*(2), 261–274.

Isaacson, D. (1987/1996). Arithmetical truth and hidden higher order concepts. In The Paris Logic Group (Ed.) *Logic Colloquium '85*, (p. 147–169). North-Holland. Reprinted with revisions in Hart, W.D. (Ed.), *The Philosophy of Mathematics*, pages 203-224, Oxford Readings in Philosophy, Oxford University Press. Page references and quotations are from the 1996 edition.

Isaacson, D. (1992). Some considerations on arithmetical truth and the ω-rule. In M. Detlefsen (Ed.) *Proof, Logic and Formalization*, (p. 94–138). Routledge.

Kreisel, G. (1954). On a problem of Henkin's. *Indagationes mathematicae*, *15*, 405–406.

Kreisel, G. (1965). Mathematical logic. In T. L. Saaty (Ed.) *Lectures on Modern Mathematics*, vol. III, (p. 95–195). Wiley.

Kripke, S. A. (2022). Mathematical incompleteness results in first-order peano arithmetic: A revisionist view of the early history. *History and Philosophy of Logic*, *43*(2), 175–182.

Łełyk, M., & Nicolai, C. (2022). A theory of implicit commitment. *Synthese*, *200*(284), 1–26.

Mancosu, P., Galvan, S., & Zach, R. (2021). *An Introduction to Proof Theory: Normalization, Cut-Elimination, and Consistency Proofs*. Oxford: Oxford University Press.

Pohlers, W. (2009). *Proof Theory. The first step into impredicativity*. Springer.

Sayward, C. (1990). Four views of arithmetical truth. *The Philosophical Quarterly*, *40*(159), 155–168.

Smith, P. (2008). Ancestral arithmetic and Isaacson's thesis. *Analysis*, *68*(1), 1–10.

Smoryński, C. (1982). The varieties of arboreal experience. *The Mathematical Intelligencer*, *4*(4), 182–189.

Sommer, R. (1995). Transfinite induction within Peano Arithmetic. *Annals of Pure and Applied Logic*, *76*(3), 231–289.

Tatton-Brown, O. (2018). Primitive recursion and Isaacson's thesis. *Thought: A Journal of Philosophy*, *8*(1), 4–15.

Weiermann, A. (2002). Zero-one law characterizations of ε_0. In B. Chauvin (Ed.) *Mathematics and computer science II: Algorithms, trees, combinatorics and probabilities*, (p. 527–539). Birkhäuser.

Weiermann, A. (2005). Analytic combinatorics, proof-theoretic ordinals, and phase transitions for independence results. *Annals of Pure and Applied Logic*, *136*(1-2), 189–218.