Lower bounds for the provability of Herbrand consistency in weak arithmetics

Zofia Adamowicz Konrad Zdanowski

December 9, 2007

Abstract

We prove that for $i \geq 1$, the arithmetic $I\Delta_0 + \Omega_i$ does not prove its own Herbrand consistency restricted to the terms of the depth in $(1+\varepsilon)\log^{i+2}$, where ε is an arbitrary small constant greater than zero.

1 Introduction

One of the main methods of showing that one set of axioms, say T, is strictly stronger than the other one, say $S \subseteq T$, is to show that $T \vdash \mathrm{Con}_S$. However, as it was proved by Wilkie and Paris in [WP87] this method does not work for bounded arithmetic theories if we use the usual Hilbert style provability predicate. Indeed, they proved that even the strong arithmetic $\mathrm{I}\Delta_0 + \exp$ does not prove the Hilbert style consistency of Robinons's arithmetic Q, that is $\mathrm{I}\Delta_0 + \exp$ does not prove that there is no Hilbert prove of $0 \neq 0$ from Q. Thus, if we hope to differentiate various bounded arithmetics by provability of consistency we should use some other provability notions, like tableux or Herbrand provability. Indeed, for these notions it is usually easier to show that a given theory is consistent since, e.g., Herbrand proofs are of a bigger size than Hilbert ones. Thus, it may happen in a model of $\mathrm{I}\Delta_0 + \exp$ that a theory S is inconsistent in the Hilbert sense and consistent in the Herbrand sense. Only when we know that the superexponentiation function is total we can prove the equivalence of the above notions of provability. (The superexponentiation function is defined by inductive conditions: $\mathrm{supexp}(0) = 1$ and $\mathrm{supexp}(x + 1) = \exp(2, \mathrm{supexp}(x))$.) For some time it has been even unknown whether the second Gödel incompleteness theorem holds for arithmetics $\mathrm{I}\Delta_0 + \Omega_i$ and the Herbrand style provability predicate. Adamowicz and Zbierski in [AZ01] proved, for $i \geq 2$, the second incompleteness theorem for $\mathrm{I}\Delta_0 + \Omega_i$ and the Herbrand notion of consistency and later Adamowicz in [A01] proved this result for $I\Delta_0 + \Omega_1$. Recently, Kołodziejczyk showed in [K06] a strengthening of these results. He proved that there is a finite fragment S of $I\Delta_0 + \Omega_1$ such that no theory $I\Delta_0 + \Omega_i$ proves the Herbrand consistency of S. Thus, if one wants to differentiate bounded arithmetics by means of provability of Herbrand consistency one should consider a thinner notion, e.g., Herbrand proofs restricted to some cuts of a given model of a bounded arithmetic.

In our main result, we show that for each i, $I\Delta_0 + \Omega_i$ does not prove its Herbrand consistency, when it is suitably formulated, restricted to the terms of depth in $(1+\varepsilon)\log^{i+2}$ of a given model, where $\varepsilon > 0$, (see theorem 32). On the other hand it is known that for each i,

$$I\Delta_0 + \Omega_i \vdash \mathrm{HCons}(I\Delta_0 + \Omega_i, \log^{i+3})$$

that is $I\Delta_0 + \Omega_i$ proves its Herbrand consistency restricted to terms of depth $\log^{i+3}$ (see theorem 20).

It is tempting to close the gap by proving, at least for some $i \geq 1$, either that

$$I\Delta_0 + \Omega_i \vdash \mathrm{HCons}(I\Delta_0 + \Omega_i, \log^{i+2}) \tag{1}$$

or

$$I\Delta_0 + \Omega_i \nvdash \mathrm{HCons}(I\Delta_0 + \Omega_i, A\log^{i+3}),\ \text{for some } A \in \mathbb{N}. \tag{2}$$

Indeed both conjectures (1) and (2) have interesting consequences for bounded arithmetics. If (1) holds then $I\Delta_0 + \Omega_{i+1}$ would not be Π_1–conservative over $I\Delta_0 + \Omega_i$. This is so because $\log^{i+2}$ is closed under addition in the presence of Ω_{i+1}. Thus, in $I\Delta_0 + \Omega_{i+1}$ the cuts $\log^{i+2}$ and $(1+\varepsilon)\log^{i+2}$ are the same. It follows then from (1) that $I\Delta_0 + \Omega_{i+1} \vdash \mathrm{HCons}(I\Delta_0 + \Omega_i, A\log^{i+2})$, for each $A \in \mathbb{N}$.

On the other hand, if (2) holds this would mean that we cannot mimic the proof of theorem 20 for the cut $A\log^{i+3}$. But the only tool needed in that proof which is unavailable in this situation is the existence of a suitable truth definition for Δ_0 formulas. Thus, it would follow that there is no such truth definition for Δ_0 formulas whose suitable properties are provable in $I\Delta_0 + \Omega_i$.

2 Basic notions and facts

For a detailed treatment of bounded arithmetics we refer to [HP93]. We consider the bounded arithmetics theories $I\Delta_0 + \Omega_i$, for $i \geq 1$. $I\Delta_0$ is just the first order arithmetic with the induction axioms restricted to bounded formulas i.e. formulas with quantification of the form $Qx \leq t(\bar{z})$, where $Q \in \{\exists, \forall\}$ and $x \notin \{\bar{z}\}$. For $i \geq 1$, the axiom Ω_i states the totality of the function ω_i. The functions ω_i are defined as follows. Let $\log(x)$ be the logarithm with the base 2 . Let the length function $\mathrm{lh}(x)$ be the length of the binary representation of x,

$$\mathrm{lh}(x) = \ulcorner \log(x+1) \urcorner.$$

Now,

$$\omega_1(x) = \begin{cases} 0 & \text{if } x = 0 \\ 2^{(\mathrm{lh}(x)-1)^2} & \text{if } x > 0. \end{cases}$$

and

$$\omega_{i+1}(x) = \begin{cases} 0 & \text{if } x = 0 \\ 2^{\omega_i(\mathrm{lh}(x)-1)} & \text{if } x > 0. \end{cases}$$

Let $\exp(x) = 2^x$. The following relation between exp and ω_i will be important for us. For all $i \geq 1$, for all k

$$\omega_i^k(\exp^{i+2}(0)) = \exp^{i+2}(k). \tag{3}$$

This equation allows us to infer the existence of $(i+2)$-th iterated exp on a number k from the existence of the interpretation for a term $\omega_i^k(\exp^{i+2}(0))$. We will need also $\mathrm{supexp}(x)$ function defined by conditions $\mathrm{supexp}(0) = 1$ and $\mathrm{supexp}(x+1) = \exp(\mathrm{supexp}(x))$ and $\log^*(x)$ function, which is a kind of inverse of supexp, defined as

$$\log^*(x) = \max\{i \leq x : \mathrm{supexp}(i) \leq x\} \cup \{0\}\,.$$

We extend the language by adding a function symbol $s^{\exists x\varphi}$ of arity n for each formula $\exists x\varphi$ with n free variables. As a numeral for i we take $\underline{i} = 2i$. We take the tree depth of $\underline{i}$ as $\log^2(i)$. The tree depth of other terms is defined by the inductive condition:

$$\mathrm{tr}(f(t_1,\ldots,t_k)) = 1 + \max\{\mathrm{tr}(t_i) : i \leq k\}\,.$$

By the depth of a term t we define the maximum of its tree depth and the size of the greatest function symbol in t. That is

$$\mathrm{dp}(t) = \max\{f :\ f \text{ occurs in } t\} \cup \{\mathrm{tr}(t)\}\,.$$

For a set of terms Λ, the depth of Λ, $\mathrm{dp}(\Lambda) = \max\{\mathrm{dp}(t) : t \in \Lambda\}$.

We assume that our coding has the property that a code of a term $s^\varphi(t_1,\ldots,t_k)$ is not greater than $(\varphi \prod_{i\leq k} t_i)^{O(1)}$. The last expression should be read simply as a product of numbers coding the formula φ and terms t_i, for $i \leq k$. Let us remark that usual efficient codings possess this property. Indeed, for the length of a term $t = s^{\varphi}(t_1, \ldots, t_k)$ we have

$$\mathrm{lh}(t) \leq A(\mathrm{lh}(\varphi) + \sum_{i \leq k} \mathrm{lh}(t_i)). \tag{4}$$

for some integer $A \in \mathbb{N}$. Thus, $t \leq 2^{A(\mathrm{lh}(\varphi) + \sum_{i \leq k} \mathrm{lh}(t_i))}$ and

$$t \leq (\varphi \prod_{i \leq k} t_i)^A. \tag{5}$$

Later in the paper we will refer to the constant A from the above equations. However, let us note that the precise value of A depends on the coding method that one uses.

An evaluation p on a set of terms Λ is a boolean function from Λ^2 into $\{0, 1\}$. A given evaluation p tells which terms are equal under p. That is t equals t' under p if $p(t, t') = 1$. For an evaluation p on Λ we define a model $M(\Lambda, p)$. The equality relation of $M(\Lambda, p)$ is given by $p(t, t') = 1$. Then, for a function symbol f and terms $t_1, \ldots, t_k$ the value of f on $t_1, \ldots, t_k$ is just $f(t_1, \ldots, t_k)$. We define the ordering in M as: $t \leq t'$ if and only if there is $s \in \Lambda$ such that $p(t + s, t') = 1$. Thus, we adopt the standard method of defining the ordering relation.

Let us observe that $M(\Lambda, p)$ is a well defined model if and only if Λ has the property that if f is a function symbol in our language and $t_1, \ldots, t_{\mathrm{ar}(f)} \in \Lambda$, then $f(t_1, \ldots, t_{\mathrm{ar}(f)}) \in \Lambda$. To have the equality relation well defined we have to require that the relation on terms given by $p(t, t') = 1$ is reflexive, symmetric and transitive. Moreover, it should be a congruence relation with respect to an operation of the application of a function symbol, that is for each $t_1, \ldots, t_n$ and $s_1, \ldots, s_n$ and for each n-ary function symbol f,

$$\text{if for each } i \leq n,\ p(t_i, s_i) = 1 \text{ then } p(f(t_1, \ldots, t_n), f(s_1, \ldots, s_n)) = 1.$$

We assume that all considered evaluations satisfy the above conditions.

Let $M \models \mathrm{I}\Delta_0$ and let $\Lambda \in M$ be a set of terms. For $I \subseteq M$ by $\Lambda \restriction I$ we define the set of terms from Λ with depths in I that is

$$\Lambda \restriction I = \{t \in \Lambda : \mathrm{dp}(t) \in I\}.$$

If I is a cut in M (i.e. I is closed downward and closed on successor) then $M(\Lambda \restriction I, p)$ is a well defined model (where we restricted also an evaluation p from Λ to $\Lambda \restriction I$).

In the definition below and in the rest of this article we deal with formulas in a prenex normal form only. Thus, if we write $\neg\varphi$ for φ in a prenex normal form we assume that negation is pushed into the quantifier free part of φ using rules: $\neg\exists x\gamma \equiv \forall x\neg\gamma$ and $\neg\forall x\gamma \equiv \exists x\neg\gamma$.

Definition 1 *For a formula $\varphi(x_1, \ldots, x_n)$ and terms $t_1, \ldots, t_n$ we define a sequence of terms $s_1, \ldots, s_r$ by induction on the complexity of φ.*

If φ is quantifier free then the sequence for φ and $t_1, \ldots, t_n$ is just $t_1, \ldots, t_n$.

If $\varphi = \exists x\, \psi(x_1, \ldots, x_n, x)$ then the sequence for φ and $t_1, \ldots, t_n$ is the sequence for $\psi(x_1, \ldots, x_n, x)$ and terms $t_1, \ldots, t_n, s^{\exists x\, \psi(x_1,\ldots,x_n,x)}(t_1, \ldots, t_n)$.

If $\varphi = \forall x\, \psi(x_1, \ldots, x_n, x)$ then the sequence for φ and $t_1, \ldots, t_n$ is the sequence for $\neg\psi$ and terms $t_1, \ldots, t_n, s^{\exists x \neg\psi(x_1,\ldots,x_n,x)}(t_1, \ldots, t_n)$.

Sometimes we call $s_1, \ldots, s_r$ the terms needed to evaluate $\varphi(x_1, \ldots, x_n)$ on $t_1, \ldots, t_n$. As we will see, the definition of the relation $p \models \varphi[t_1, \ldots, t_n]$, in order to work properly, requires all terms $s_1, \ldots, s_r$ to be in Λ..

Definition 2 *Let $t_1, \ldots, t_n \in \Lambda$ and $\varphi(x_1, \ldots, x_n)$ be a formula. We say that $(\varphi, t_1, \ldots, t_n)$ is good enough (g.e. in short) for Λ if all terms from the sequence for φ and $t_1, \ldots, t_n$ are in Λ.*

Since the sequence of terms needed to evaluate φ on $t_1, \ldots, t_n$ is the same as the sequence needed to evaluate $\neg\varphi$ on $t_1, \ldots, t_n$ we have an obvious fact.

Fact 3 *For each Λ, φ and $t_1, \ldots, t_n \in \Lambda$, $(\varphi, t_1, \ldots, t_n)$ is g.e. for Λ if and only if $(\neg\varphi, t_1, \ldots, t_n)$is g.e. for Λ.*

Now, we define the notion of a satisfaction for evaluations. Later, we relate this notion to the satisfaction relation in a model $M(\Lambda, p)$.

Definition 4 *Let p be an evaluation on Λ. By induction on φ we define $p \models \varphi[\bar{t}]$, for $\bar{t} \in \Lambda$ such that $(\varphi, \bar{t})$ is g.e. for Λ:*

- *$p \models t = t'$ if $p(t, t') = 1$,*
- *$p \models t \leq t'$ if there is $s \in \Lambda$ such that $p \models (t + s = t')$,*
- *for φ quantifier free $p \models \varphi[\bar{t}]$ if p makes φ true in the sense of propositional logic,*
- *$p \models \exists x\varphi(\bar{x}, x)[\bar{t}]$ if $p \models \varphi(\bar{x}, x)[\bar{t}, s^{\exists\varphi}(\bar{t})]$,*
- *$p \models \forall x\varphi(\bar{x}, x)[\bar{t}]$ if for all terms $t \in \Lambda$ such that $(\varphi, \bar{t}, t)$ is g.e. for Λ, $p \models \varphi(\bar{x}, x)[\bar{t}, t]$.*

Of course, whenever we write $p \models \varphi[\bar{t}]$ we assume that $(\varphi, \bar{t})$ is g.e. for Λ.

One can easily prove by induction on the construction of a formula φ the following fact.

Fact 5 *Let p be an evaluation on Λ and let φ and $\bar{t}$ be g.e. for Λ. It holds that if $p \models \varphi[\bar{t}]$ then $p \not\models \neg\varphi[\bar{t}]$.*

Definition 6 *Let T be a theory and let p be an evaluation on Λ. We call p a T–evaluation if for all $\varphi \in T$ such that φ is g.e. for Λ, $p \models \varphi$.*

If T has Δ_0 definable set of axioms then the notion of T–evaluation is definable by a Δ_0 formula

We have the following relation between $p \models \varphi$ and $M(\Lambda, p) \models \varphi$.

Proposition 7 *Let p be an evaluation on Λ and let $M(\Lambda, p)$ be well defined. Then for a formula φ and $\bar{t} \in \Lambda$,*

$$\textit{if } p \models \varphi[\bar{t}] \textit{ then } M(\Lambda, p) \models \varphi[\bar{t}].$$

Proof. The proof of the above proposition is a straightforward induction on the complexity of φ. For the quantifier free formulas the thesis is obvious. If $\varphi = \exists y \psi(\bar{t}, y)$ then from $p \models \varphi[\bar{t}]$, we deduce that $p \models \psi[\bar{t}, s]$, where $s = s^{\exists y \psi(\bar{x}, y)}(\bar{t})$ and we may use our inductive assumption to conclude that $M(\Lambda, p) \models \psi[\bar{t}, s]$ and that $M \models \varphi[\bar{t}]$.

For $\varphi = \forall y \psi(\bar{t}, y)$ we observe that since $M(\Lambda, p)$ is well defined then for all $s \in \Lambda$, $(\psi, \bar{t}, s)$ is g.e. for Λ. It follows that for all $s \in \Lambda$, $p \models \psi[\bar{t}, s]$. Since the universe of $M(\Lambda, p)$ is made from terms in Λ, we obtain by the inductive assumption that for all $a \in M$, $M(\Lambda, p) \models \psi[\bar{t}, a]$ and $M(\Lambda, p) \models \varphi[\bar{t}]$. $\square$

Let us observe that it is possible that neither $p \models \varphi[\bar{t}]$ nor $p \models \neg\varphi[\bar{t}]$. This is the case when e.g. for some $\psi(x, y)$, $p \models \neg\psi[t, s^{\exists y \psi}(t)]$ and $p \models \psi[t, s]$, for some term s. In this case $p \not\models \exists y \psi[t]$ nor $p \models \forall y \neg\psi[t]$. This is why we need the following definition which describes the situation when p satisfies for a given formula $\varphi(\bar{x})$ the law of excluded middle.

Definition 8 *Let $(\varphi(x_1, \ldots, x_k)$, $t_1, \ldots, t_k)$ be g.e. for Λ.*

An evaluation p on Λ decides $(\varphi, t_1, \ldots, t_k)$ if

$$p \models \varphi[t_1, \ldots, t_k] \textit{ or } p \models \neg\varphi[t_1, \ldots, t_k].$$

An evaluation p decides a formula $\varphi(\bar{x})$ if for each terms $\bar{t} \in \Lambda$, such that $(\varphi, \bar{t})$ is g.e. for Λ, p decides $(\varphi, \bar{t})$.

For formulas which are decided by an evaluation p the satisfaction relation behaves in a way which is easy to handle.

Lemma 9 *Let $(\forall x\varphi, \bar{t})$ be g.e. for Λ and let p decide $\forall x\varphi$. Then,*

$$p \models \forall x\varphi[\bar{t}] \iff p \models \varphi[\bar{t}, s^{\exists x\neg\varphi}(\bar{t})].$$

Proof. The direction from the left to the right is obvious. So let us assume $p \models \varphi[\bar{t}, s^{\exists x\neg\varphi}(\bar{t})]$. Since p decides $\forall x\varphi$ we have either

$$p \models \forall x\varphi[\bar{t}]$$

or

$$p \models \exists x\neg\varphi[\bar{t}].$$

But if the latter is true then $p \models \neg\varphi[\bar{t}, s^{\exists x\neg\varphi}(\bar{t})]$ what $\square$
is impossible by our assumption and fact 5.

We have the following proposition

Proposition 10 *Let $\varphi = Q_1x_1 \ldots Q_nx_n\psi(\bar{z}, x_1, \ldots, x_n)$ for ψ quantifier free and let $(\varphi, \bar{t})$ be g.e. for Λ. Let p, an evaluation on Λ, decide $\varphi(\bar{t})$. Then,*

$$p \models \varphi[\bar{t}] \iff p \models \psi[\bar{t}, s_1/x_1, \ldots, s_n/x_n],$$

where $\bar{t}, s_1, \ldots, s_n$ is the sequence for $(\varphi, \bar{t})$.

Proof. The proof is an easy induction on the complexity of φ. For the only nontrivial step for universal quantifier one should use lemma 9. $\square$

The ralation $p \models \varphi[\bar{t}]$ is preserved while going to some subsets of the original set of terms Λ. As a consequence we obtain that we can deduce that if $\Lambda' \subseteq \Lambda$ and $M(\Lambda', p)$ is a well defined model, then its properties may be deduced from the properties of p considered as an evaluation on Λ.

Proposition 11 *Let $M \models \mathrm{I}\Delta_0$ and let $p \in M$ be an evaluation on a set of terms $\Lambda \in M$. Let $I \subseteq M$ be a cut in M and let $p{\restriction}I$ be an evaluation p restricted to $\Lambda{\restriction}I$. If $\bar{t} \in \Lambda{\restriction}I$ and $p \models \varphi[\bar{t}]$ then $p{\restriction}I \models \varphi[\bar{t}]$. In consequence $M(\Lambda{\restriction}I, p{\restriction}I) \models \varphi[\bar{t}]$.*

Proof. Let M, Λ, p and I be as in the assumptions of the proposition. We need to show that for all $\bar{t} \in \Lambda{\restriction}I$, if $p \models \varphi[\bar{t}]$ then $p{\restriction}I \models \varphi[\bar{t}]$.

For the quantifier free φ the thesis is obvious. For the case of $\varphi = \exists y\psi(\bar{t}, y)$ one should use the fact, that the term for skolem witness, $s^{\exists y\psi(\bar{x},y)}(\bar{t})$ is a member of $\Lambda{\restriction}I$ and use the inductive assumption. For the universal quantifier step one should use the fact that $\Lambda{\restriction}I$ is a subset of Λ. $\square$
In what follows we write p for an evaluation on a set of terms Λ as well as for the evaluation p restricted to any of subset of Λ. The last proposition

shows that they will satisfy the same formulas. In what follows we will use the fact that, in order to establish that $M(\Lambda\restriction I, p) \models \varphi$, it suffices to show that $p \models \varphi$, when we treat p as an evaluation on the whole Λ.

The next lemma shows that if an evaluation p decides a formula $\exists y\varphi(\bar{t})$ then, to check whether $p \models \exists y\varphi(\bar{t})$, it suffices to check whether $p \models \varphi(s, \bar{t})$ for some term $s \in \Lambda$. Indeed any s is as good as the canonical witness for $\exists y\varphi(\bar{t})$ which is $s^{\exists y\varphi}(\bar{t})$.

Lemma 12 *Let p be an evaluation on Λ and let p decide $\exists y\varphi(y, \bar{t})$. Then,*

$$p \models \exists y\varphi(y, \bar{t}) \textit{ if and only if}$$

$$\textit{there is } s \in \Lambda \textit{ such that } (\varphi, s, \bar{t}) \textit{ is g.e. for } \Lambda \textit{ and } p \models \varphi(s, \bar{t}).$$

Proof. Let p decide $\exists y\varphi(y, \bar{t})$. To prove the direction from the left to the right it suffices to take $s = s^{\exists y\varphi}(\bar{t})$. For the direction from the right to the left let us assume that there is $s_0 \in \Lambda$ such that $(\varphi, s_0, \bar{t})$ is g.e. for Λ and $p \models \varphi(s_0, \bar{t})$. By definition we have:

$$p \models \exists y\varphi(y, \bar{t}) \text{ if and only if } p \models \varphi(s^{\exists y\varphi}(\bar{t}), \bar{t}).$$

Thus let us assume, for the sake of contradiction, that

$$p \not\models \varphi(s^{\exists y\varphi}(\bar{t}), \bar{t}).$$

Since p decides $\exists y\varphi(y, \bar{t})$, it follows that

$$p \models \neg\exists y\varphi(y, \bar{t}).$$

This is equivalent to saying that for all $s' \in \Lambda$ such that $(\varphi, s', \bar{t})$ is g.e. for Λ,

$$p \models \neg\varphi(s', \bar{t}).$$

But this contradicts our assumption that $p \models \varphi(s_0, \bar{t})$. $\Box$

In the next lemma we show a kind of closure of the relation $p \models \varphi$ under the Hilbert notion of provability. This lemma will be useful in establishing that a given T–evaluation p will satisfy some consequences of T.

Lemma 13 *Let $T \vdash \varphi$, let $M \models I\Delta_0$ and let $p \in M$ be a T–evaluation on Λ, where Λ contains all standard terms. If p decides φ, then $p \models \varphi$.*

Proof. In the proof we use the fact that if $M \models I\Delta_0$ then p and Λ have all the properties proven above for evaluations.

Let $\Lambda' \subseteq \Lambda$ be a set of all terms in Λ of a standard depth. Then, $M(\Lambda', p)$ is a well defined model. Moreover, since Λ' contains all standard terms, each axiom of T is g.e. for Λ'. Then, by the fact that p is a T–evaluation, $M(\Lambda', p) \models T$. Now, if $p \models \neg\varphi$, then $M(\Lambda', p) \models \neg\varphi$ what is impossible. $\Box$

Let T_0 be the finite set of axioms of $\mathrm{I}\Delta_0$ which characterize the recursive properties of successor, addition and multiplication and basic properties of ordering. We put in T_0 all axioms which are used in the proof of lemma 14, e.g., such as $\forall x \forall y (x \leq y + 1 \Rightarrow (x \leq y \vee x = y + 1))$.

Lemma 14 *Let $M \models I\Delta_0$, let Λ be a set of terms from M such that $\{\underline{0}, \ldots, \underline{k}\} \subseteq \Lambda$ and let $p \in M$ be a T_0–evaluation on Λ. The following holds:*

1. *for each $t \in \Lambda$, $i \leq k$, if $p \models t \leq \underline{i}$ then there exists $j \leq i$, $p \models t = \underline{j}$;*
2. *for each $i, j \leq k$, $i \leq j$ if and only if $p \models \underline{i} \leq \underline{j}$;*
3. *for each $i \leq k$, $l \leq k$, $m \leq k$,*
   - $i + j = m \iff p \models \underline{i} + \underline{j} = \underline{m}$,
   - $ij = m \iff p \models \underline{ij} = \underline{m}$,

Proof. The proof of the first point is an easy induction on $i \leq k$. For $i = 0$ one should use the fact that p makes true the following axioms of T: $\forall x (0 \leq x)$ and $\forall x \forall y ((x \leq y \wedge y \leq x) \Rightarrow x = y)$. Thus, if $p \models \underline{i} \leq \underline{0}$ then $p \models \underline{i} = 0$. The induction step follows easily from the fact that p makes true the following axiom: $\forall x \forall y (x \leq y + 1 \Rightarrow (x \leq y \vee x = y + 1))$.

For the second and the third point one should use the inductive definitions of addition and multiplication and the properties of the ordering. $\Box$

The next lemma shows that if $\{\underline{0}, \ldots, \underline{k}\} \subseteq \Lambda$ then any T_0–evaluation on Λ has to reflect the truth on $\{\underline{0}, \ldots, \underline{k}\}$ for Δ_0 formulas.

Lemma 15 (Absoluteness lemma) *Let $M \models \mathrm{I}\Delta_0$, let $\Lambda \in M$ be a set of terms such that $\{\underline{0}, \ldots, \underline{k}\} \subseteq \Lambda$ and let $p \in M$ be a T_0–evaluation on Λ. Let φ be a Δ_0–formula with only variables as bounds of quantifiers, such that values of terms in $\varphi(\bar{x})$ are not greater than* $\max\{\bar{x}\}$*. We have that for each*

$i_1, \ldots, i_m \leq k$, *such that* $(\varphi, \underline{i_m}, \ldots, \underline{i_m})$ *is g.e. for* Λ, *the following holds in* M*:*

$$\begin{aligned} &\textit{If } p \textit{ decides } (\varphi, \underline{i_m}, \ldots, \underline{i_m}) \textit{ then} \\ &\qquad\qquad \varphi(i_1, \ldots, i_m) \iff p \models \varphi[\underline{i_1}, \ldots, \underline{i_m}]. \end{aligned}$$

Proof. The proof is by induction on the complexity of φ. The case for atomic formulas holds by points 2 and 3 of lemma 14. The bounded quantifier step can be carried out by point 1 of the same lemma. □

Now, we will estimate the size of terms which occur in the sequence for a given formula φ and $\bar{t}$.

Lemma 16 (Estimation lemma) *Let* $\varphi(x_1, \ldots, x_k)$ *be a formula, let* $t_1, \ldots, t_k$ *be arbitrary terms and let* $t_1, \ldots, t_k, w_1, \ldots, w_r$ *be the sequence of terms needed to evaluate* φ *on* $t_1, \ldots, t_k$*. Then, for all* $i \leq r$

$$w_i \leq \max\{t_j : j \leq r\}^{(\varphi^E)} \varphi^{(\varphi^E)},$$

where E *is a standard constant.*

Proof. First, we prove by induction on $i \leq r$,

$$w_i \leq (\varphi \prod_{j \leq k} t_j)^{(2A)^i}$$

Let A be, by equation (4), such that

$$\mathrm{lh}(s^{\varphi}(t_1, \ldots, t_k) \leq A(\mathrm{lh}(\varphi) + \sum_{j \leq k} \mathrm{lh}(t_j)).$$

Then,

$$\begin{aligned} w_1 &\leq 2^{A(\mathrm{lh}(\varphi) + \sum_{j \leq k} \mathrm{lh}(t_j))} \\ &\leq \varphi^A (\prod_{j \leq k} t_j)^A \\ &\leq \varphi^{2A} (\prod_{j \leq k} t_j)^{2A}. \end{aligned}$$

Now, let

$$w_i = s^{\psi}(t_1, \ldots, t_k, w_1, \ldots, w_{i-1})$$

and

$$w_{i+1} = s^{\psi'}(t_1, \ldots, t_k, w_1, \ldots, w_i),$$

for ψ and ψ' being subformulas of φ. Then,

$$\begin{aligned}
\mathrm{lh}(w_{i+1}) &\leq A(\mathrm{lh}(\psi') + \sum_{j\leq k} \mathrm{lh}(t_j) + \sum_{j\leq i} \mathrm{lh}(w_j)) \\
&\leq A(\mathrm{lh}(\psi) + \sum_{j\leq k} \mathrm{lh}(t_j) + \sum_{j\leq i-1} \mathrm{lh}(w_j)) + A\,\mathrm{lh}(w_i) \\
&\leq A(\mathrm{lh}(w_i) + \mathrm{lh}(w_i)) \\
&\leq 2A\,\mathrm{lh}(w_i).
\end{aligned}$$

So,

$$\begin{aligned}
w_{i+1} &\leq 2^{2A\,\mathrm{lh}(w_i)} \\
&\leq (w_i)^{2A} \\
&\leq ((\varphi \prod_{j\leq k} t_j)^{(2A)^i})^{2A} \\
&\leq (\varphi \prod_{j\leq k} t_j)^{(2A)^{i+1}}.
\end{aligned}$$

But $r, k \leq \log(\varphi)$, so

$$\begin{aligned}
w_i &\leq (\prod_{j\leq k} t_j)^{O(1)^{\log(\varphi)}} \varphi^{O(1)^{\log(\varphi)}} \\
&\leq (\max\{t_i : i \leq r\})^{\log(\varphi) O(1)^{\log(\varphi)}} \varphi^{O(1)^{\log(\varphi)}} \\
&\leq (\max\{t_i : i \leq r\})^{\log(\varphi)(\varphi^{O(1)})} \varphi^{(\varphi^{O(1)})} \\
&\leq (\max\{t_i : i \leq r\}^{(\varphi^{O(1)})} \varphi^{(\varphi^{O(1)})}.
\end{aligned}$$

□

The following theorem by Adamowicz is theorem 1.1 from [A02]

Theorem 17 (Adamowicz, [A02]) *For each $m, n \in \mathbb{N}$ there is a bounded formula $\theta(\bar{x})$ such that*

$$I\Delta_0 + \Omega_n + \exists \bar{x} \in \log^m \theta(\bar{x}) \text{ is consistent}$$

and

$$I\Delta_0 + \Omega_n + \exists \bar{x} \in \log^{m+1} \theta(\bar{x}) \text{ is inconsistent.}$$

Definition 18 *An evaluation p on Λ is a T–evaluation if for each $\varphi \in T$ such that φ is g.e. for Λ, $p \models \varphi$.*

Let N be an integer. An evaluation p on Λ is N–deciding p decides all formulas φ with codes less than N.

We define the following version of Herbrand consistency.

Definition 19 *Let N be a sufficiently large standard constant.*

$\mathrm{HCons}(T, i)$ *is a Π_1 arithmetical formula which states that for each set of terms Λ of depth not greater than i, there exists an N–deciding, T–evaluation on Λ.*

The Herbrand theorem can be stated in the following form: a theory T is consistent if and only if for each set of terms Λ in the skolemized language there is an evaluation p on pairs of terms from Λ such that p makes true all axioms of T which are g.e. for Λ.

Our definition of Herbrand consistency is parameterized by a constant N. We take this parameter for technical reasons to have evaluations better behaved. As we will see in theorem 20 this additional condition does not restrict the provability of some cases of Herbrand consistency while it allows us to have an interesting and still natural unprovability result.

We do not specify what is the size of the constant N. We do not need to fix it because for each i, $\mathrm{I}\Delta_0 + \Omega_i \vdash \mathrm{HCons}(\mathrm{I}\Delta_0 + \Omega_i, \log^{i+3})$, when we take an arbitrary constant N (theorem 20). On the other hand for our unprovability result one should take N so large that evaluations decide all relevant formulas which occur in the course of the proof of the unprovability of $\mathrm{HCons}(\mathrm{I}\Delta_0 + \Omega_i, (1+\varepsilon)\log^{i+2})$. It will be a large constant but its precise value is irrelevant for us.

3 Provability of Herbrand consistency

In this section we show a case for which a Herbrand consistency is provable in bounded arithmetic.

Theorem 20 $\mathrm{I}\Delta_0 + \Omega_i$ *proves its Herbrand consistency restricted to the terms of depth not greater than* $\log^{i+3}$ *that is*

$$\mathrm{I}\Delta_0 + \Omega_i \vdash \mathrm{HCons}(\mathrm{I}\Delta_0 + \Omega_i, \log^{i+3}).$$

Proof. We prove the theorem for the case of $i = 1$. The proof for $i > 1$ is essentially the same.

Let $T = \mathrm{I}\Delta_0 + \Omega_1$. The proof is a straightforward induction. Let $M \models T$ and let $\Lambda = \{t_1, \ldots, t_k\}$ be a set of terms of depth not greater than some $d \in \log^4(M)$. For simplicity we assume that if t_m is a subterm of t_j then $m \leq j$. We prove by induction on $m \leq k$ the following

$$\exists H_m = \{h_{t_1}, \ldots, h_{t_m}\} \forall j \leq m [\forall a \leq \Lambda (t_j = \underline{a} \Rightarrow h_{t_j} = a) \wedge$$
$$\forall r \forall \varphi \leq \Lambda \forall (n_1, \ldots, n_r) \leq \Lambda (t_j = s^{\exists y \varphi}(t_{n_1}, \ldots, t_{n_r}) \Rightarrow$$
$$h_{t_j} = \text{ the least witness } \exists y \varphi(h_{t_{n_1}}, \ldots, h_{t_{n_r}}) \text{ or } 0 \text{ otherwise})].$$

Since the theory T is Π_1, it is easy to see that the greatest element of H_i may be only for a term $\omega_1^m(\underline{0})$ and is less than $\exp^3(m) \in \log(M)$ because $m \in \log^4(M)$. Thus, we may use a universal formula to compute the witness for $\exists y \varphi(h_{n_1}, \ldots, h_{n_r})$. It is worth to mention that this is the only place where we use the relation between the rate of the growth of ω_1 function and the 4-th logarithm (or, more generally, between the rate of the growth of ω_i and the $(i+3)$-th logarithm).

It is also easy to see that H_m is small enough to be in M. The number of terms of depth d is not greater than $d^{\log(d)^d}$. Indeed, the number of nodes in the tree for a term of depth not greater than d is at most $\log(d)^d$ ($\log(d)$ is the branching of a tree and d is the depth of a tree). Since we have only d labels for these nodes, the number of terms is at most $d^{\log(d)^d}$. Thus,

$$\begin{aligned} \mathrm{card}(H_i) &\leq (\log^4(M))^{(\log^5(M))^{(\log^4(M))}} \\ &\leq 2^{2^{\log^6(M)(\log^4(M)+1)}} \\ &\leq 2^{2^{\log^3(M)}} \\ &\leq \log(M). \end{aligned}$$

It follows that the size of H_m, the set of $\log(M)$ elements of sizes in $\log(M)$, is not greater than

$$\log(M)^{\log(M)} \leq 2^{(\log(M))^2}$$

which is an element of M. Thus we can take an element of M to bound the quantifier $\exists H_m$ in the induction formula.

Now, we define an evaluation p on $\Lambda = \{t_1, \ldots, t_k\}$ according to $H_k = \{h_{t_1}, \ldots, h_{t_k}\}$:

$$p(t, t') = 1 \iff h_t = h_{t'}.$$

It suffices to show that p is an N–deciding, T–evaluation. By induction on the complexity of a formula we show that p decides all standard formulas. Indeed, we show something stronger: for each formula φ and for all terms $s_1, \ldots, s_r \in \Lambda$,

$$M \models \varphi[h_{s_1}, \ldots, h_{s_r}] \iff p \models \varphi[s_1, \ldots, s_r].$$

For atomic formulas the statement is obvious. So it is for all quantifier free formulas. Now, let us take the formula $\varphi = \exists y \psi(y, \bar{x})$ and $\bar{s} \in \Lambda$, $\bar{s} = s_1, \ldots, s_r$, such that $(\varphi, \bar{s})$ is g.e. for Λ. If $M \models \exists y \psi[h_{s_1}, \ldots, h_{s_r}]$, then for $s = s^{\exists y \psi}(s_1, \ldots, s_r)$, $M \models \psi[h_s, h_{s_1}, \ldots, h_{s_r}]$ and, by the inductive assumption, $p \models \psi[s, s_1, \ldots, s_r]$. So, $p \models \exists y \psi[s_1, \ldots, s_r]$.

On the other hand, if $M \models \neg \exists y \psi[h_{s_1}, \ldots, h_{s_r}]$, then for all $h \in H$, $M \models \neg \psi[h, h_{s_1}, \ldots, h_{s_r}]$. It easily follows by the inductive assumption that $p \models \neg \exists y \psi[s_1, \ldots, s_r]$.

Let us observe that the argument above works also for all nonstandard Δ_0 formulas if we change our statement to: for all terms $s_1, \ldots, s_r \in \Lambda$,

$$M \models \mathrm{Tr}_{\Delta_0}(\varphi, \langle h_{s_1}, \ldots, h_{s_r} \rangle) \iff p \models \varphi[s_1, \ldots, s_r].$$

Since for all $i \leq r$, $h_{s_i} \in \log(M)$, $\mathrm{Tr}_{\Delta_0}(\varphi, \langle h_{s_1}, \ldots, h_{s_r} \rangle)$ has in M the properties of a Δ_0–truth definition. Moreover, since $H \subseteq \log(M)$, we can bound the existential quantifier in Tr_{Δ_0} by an element of a model.

Now we show that p satisfies the bounded induction axioms. Let $\varphi(x)$ be a Δ_0 formula. We want to show that

$$p \models \forall z(\neg \varphi(\underline{0}) \vee \exists x \leq z(\varphi(x) \wedge \neg \varphi(x+1)) \vee \varphi(z)).$$

Let us assume that $M \models \mathrm{Tr}_{\Delta_0}(\varphi, 0)$ and $M \models \forall x \leq h_{s_i}(\mathrm{Tr}(\varphi, x) \Rightarrow \mathrm{Tr}(\varphi, x+1))$, where h_{s_i} is an arbitrary, fixed element of H. If not, then by the remark above, we could easily show that, either $p \models \neg \varphi(\underline{0})$ or $p \models \exists x \leq s_i(\varphi(x) \wedge \neg \varphi(x+1))$. Now, by the Δ_0 induction in M for $\mathrm{Tr}_{\Delta_0}(\varphi, x)$ we infer that $M \models \mathrm{Tr}_{\Delta_0}(\varphi, h_{s_i})$ (Tr_{Δ_0} is in fact a Δ_0 formula if we put inside a big parameter which is available in M). Thus, $p \models \varphi[s_i]$. Since s_i is arbitrary we showed that the induction axioms holds under p. $\square$

4 Unprovability of Herbrand consistency

In this section we prove that for $T_i = \mathrm{I}\Delta_0 + \Omega_i$, the arithmetic T_i does not prove its Herbrand consistency restricted to terms of depth in $(1+\varepsilon)\log^{i+2}$. However, for simplicity, we present the proof only for the most subtle case, for $\mathrm{I}\Delta_0 + \Omega_1$. Indeed, only in this case we should take care whether all the objects that we construct are inside a model and that the main inductive argument can be carried out in bounded induction. We encourage the reader to review the proof after reading it with an eye on how it behaves for $i > 1$. For such an $i > 1$ all the estimations become easier. One should only replace $\log^3$ with $\log^{i+2}$, $\log^4$ with $\log^{i+3}$ and insure that all elements needed in the proof are in the model.

Thus, from now on let $T = \mathrm{I}\Delta_0 + \Omega_1$. For the sake of contradiction, till the end of this section we assume that

$$T \vdash \mathrm{HCons}(T, (1+\varepsilon)log^3),$$

for some $\varepsilon > 0$.

Let us also fix a model $M \models T$ and an element $a \in \log^3(M)$. We will consider only evaluations for the set of Skolem terms for formulas $\varphi \leq \log^*(a)$. Since our result is about unprovability of $\mathrm{HCons}(T, (1+\varepsilon)\log^3)$ such a restriction only makes our result stronger.

The main idea of the proof is the following. Under the assumption that $T \vdash \mathrm{HCons}(T, (1+\varepsilon)\log^3)$, we show that for any model $M \models T$ and any $a \in \log^3(M)$ we can construct a model M' such that

$$M \restriction \{0, \ldots, a\} \cong M' \restriction \{0, \ldots, a\}$$

and

$$M' \models a \in \log^4.$$

Together with theorem 17 this will allow us to obtain a contradiction when we suitably choose an element $a \in M$ such that by its Δ_0–properties it cannot be (provably in T) in $\log^4$.

In order to construct M' we work in M. We construct a sequence of sets of terms and evaluations on them $\{(\Lambda_i, p_i)\}_{i \leq a/C^2}$, for some standard constant C. The key property of the sequence will be that under p_i the element $\exp^4((1+\varepsilon/2)^i a)$ exists. Then, the desired model M' will be, roughly saying, the model defined by $\Lambda_{a/C}$ and $p_{a/C}$.

We should say a word on how we choose the constant C. Again the reader should think about C as about a fixed large integer. Our construction is uniform in C thus the particular choice of C is not important for us. We only require that it so large that it satisfies some inequalities as: $C > 4E\log(C)$ and $C \geq (\models)\log(2A)+1$, where E is the constant from the estimation lemma, A is the constant from the equation (4) and $\models$ is understood here as a Gödel number for the formula $x \models y[z]$.

We start with a definition which will be used in the formulation of the main inductive argument.

Definition 21 *Let $M \models \mathrm{I}\Delta_0$, let $a \in \log^3(M)$, let $\varepsilon > 0$ be small standard constant and let $d \in M$ be such that $d > \mathbb{N}$ and $d < \min\left\{(\varepsilon/2)a, 2^{a-a/(C^2)(1+\log(1+\varepsilon/2))}\right\}$ (we comment on this choice of d below the definition). The elements a, ε and d are parameters of the definition which will be fixed during the whole proof.*

Let Λ be a set of terms, p be an evaluation on Λ and let $k, b \in M$. The sequence (Λ, p, k, b) is suitable when

1. *k and b are nonstandard parameters,*
2. *Λ is the set of terms of the form*

$$\Lambda = \left\{ t : \mathrm{Term}(t) \wedge \mathrm{dp}(t) \leq b \wedge t \leq 2^{2^k} \right\},$$

3. *p is a T–evalution on Λ and p is N–deciding,*
4. *$k + d < b$ and $bd < 2^k$.*

During the induction we will consider only suitable sequences $(\Lambda_i, p_i, k_i, b_i)$ where $k_i = a - iC$ and $b_i = (1 + \varepsilon/2)^{i+1}a$, for $i \leq a/C^2$. We chose the parameter d as above to ensure that the 4–th point of the definition will be always satisfied. It would suffice for our needs that only $k_i + \mathbb{N} < b_i$ and $b_i\mathbb{N} < 2^{k_i}$ however this condition is even not expressible by an arithmetical formula unless we cannot define the standard part of the model.

Fact 22 *Let (Λ, p, k, b) be suitable. Then $\left\{ \underline{0}, \ldots, \underline{2^{2^k-1}} \right\} \subseteq \Lambda$.*

Proof. For $i \leq 2^{2^k-1}$, we have $\underline{i} \leq 2i \leq 2^{2^k}$ and $\mathrm{dp}(\underline{i}) \leq k < b$. $\square$

In the next lemma we show which formulas with numerals as parameters are g.e. for a suitable sequence (Λ, p, k, b).

Lemma 23 *Let (Λ, p, k, b) be suitable and let $\varphi(x_1, \ldots, x_r)$ be a formula less than C. Then, for $m_1, \ldots, m_k \leq 2^{2^{k-C}}$ $(\varphi, \underline{m_1}, \ldots, \underline{m_r})$ is g.e. for Λ.*

Proof. The lemma follows from the estimation lemma and from the fact that, by our choice of C, $C > 4E\log(C)$, where E is the constant from the estimation lemma. Indeed, the size of the greatest term needed to evaluate $\varphi(\underline{m_1}, \ldots, \underline{m_r})$ is, by the estimation lemma, not greater than $(\max\{\underline{m_i} : i \leq r\})^{(C^E)} C^(C^E)$. It follows that terms are not greater than

$$\begin{aligned}
(22^{2^{k-C}})^{(C^E)} C^{(C^E)} &\leq 2^{(2^{k-C}+1)(C^E)+\log(C)C^E} \\
&\leq 2^{(2^{k-C+1+E\log(C)+E\log(C)+\log\log(C)})} \\
&\leq 2^{2^{k-C+4E\log(C)}} \\
&\leq 2^{2^k}.
\end{aligned}$$

Moreover, the depths of the terms are not greater than $k + \mathbb{N}$ which is less than b. $\square$

In the most important case φ from lemma 23 will be just $x \models y[z]$.

In the lemma below we show how much exponentiation is available under an evaluation p which occurs in a suitable sequence (Λ, p, k, b). We show that $p \models$ "$\exp^3(\underline{b-C})$ exists". Ideally, we would like to have $\exp^3(\underline{b})$. Unfortunately, to show that $\exp^3(\underline{i})$ exists we need a term $\omega_1^i(\underline{8})$ and we need that some formulas with this term (used in the course of a proof of lemma 24) are g.e. for Λ. This is why we restrict lemma 24 to $\exp^3(\underline{b-C})$.

Lemma 24 *Let (Λ, p, k, b) be suitable. Then,*

$$p \models \exists x(x = \exp^3(\underline{b-C})).$$

Proof. Let (Λ, p, k, b) be a suitable sequence. In order to show that $p \models \exists x(x = \exp^3(\underline{b-C}))$ we show that for each $i \leq b - C$,

$$p \models x = \exp^3(y)[\omega_1^i(\underline{8})/x, \underline{i}/y]$$

what clearly suffices. This is so, because the formula $\forall x \forall y(x = \omega_1^y(\underline{8}) \Rightarrow x = \exp^3(y))$ is provable in T. Thus, the following equation holds by lemma 13 and by the fact that N is chosen large enough,

$$p \models \forall y \forall x(x = \omega_1^y(\underline{8}) \Rightarrow x = \exp^3(y)). \tag{6}$$

Of course, in the formula $x = \omega_1^y(\underline{8})$, y is a free variable so $\omega_1^y(\underline{8})$ should not be read as a closed term but as a formula with free variables x and y.

By equation (6), to show that for all $i \leq b - C$,

$$p \models x = \exp^3(y)[\omega_1^i(\underline{8})/x, \underline{i}/y]$$

it suffices to show that for each $i \leq b - C$,

$$p \models x = \omega_1^y(\underline{8})[\omega_1^i(\underline{8})/x, \underline{i}/y].$$

For $i = 0$ there is nothing to prove. Indeed, $\underline{8} = \omega_1^0(\underline{8})$ is a true Δ_0 formula thus is has to be decided positively by p.

Now, let us assume that for some $i < b - C$, $p \models x = \omega_1^y(\underline{8})[\omega_1^i(\underline{8}), \underline{i}]$. Then, since

$$T \vdash \forall x \forall y[x = \omega_1^y(\underline{8}) \Rightarrow \omega_1(x) = \omega_1^{y+1}(\underline{8})]$$

we have, again by lemma 13,

$$p \models \omega_1(x) = \omega_1^{y+1}(\underline{8})[\omega_1^i(\underline{8})/x, \underline{i}/y].$$

Since $p \models z = y + 1[\underline{i+1}/z, \underline{i}/y]$, the last equation is nothing else than

$$p \models x = \omega_1^y(\underline{8})[\omega_1^{i+1}(\underline{8})/x, \underline{i+1}/y].$$

Thus, we have proved the induction step and this finishes the proof of the lemma. $\square$

In the next lemma we show how to construct from a suitable sequence (Λ, p, k, b) a new suitable sequence. The new sequence is of the form $(\tilde{\Lambda}, \tilde{p}, k - C, (1+\varepsilon/2)b)$. The lemma below is, essentially, the inductive step in our construction. The key property of the new constructed sequence $(\tilde{\Lambda}, \tilde{p}, k - C, (1+\varepsilon/2)b)$ is that we will have more exp available under $\tilde{p}$.

Lemma 25 *Let $\varepsilon' = \varepsilon/2$, $M \models \mathrm{I}\Delta_0$, $a \in \log^3(M)$ and let $i < a/(C^2)$. Let d be chosen according to definition 21.*

Let (Λ, p, k, b) be suitable, were $k = a - iC$ and $b = (1+\varepsilon')^{i+1}a$. Then, there are $\tilde{\Lambda}, \tilde{p}$ such that $(\tilde{\Lambda}, \tilde{p}, \tilde{k}, \tilde{b})$ is suitable, where $\tilde{k} = k - C$ and $\tilde{b} = (1+\varepsilon')b$.

Proof. Let $\Lambda, p, k, b, \tilde{k}, \tilde{b}$ satisfy the assumptions of the lemma. By our choice of the parameter d in the definition 21, $\tilde{k}$ and $\tilde{b}$ satisfy the 4-th point of definition 21.

definition 21 and $(1+\varepsilon/2) \geq 2^{k-2C}$ then $\tilde{k}$ and $\tilde{b}$ also satisfy the 4-th point of definition 21.

We define $\tilde{\Lambda}$ in the only possible way as

$$\tilde{\Lambda} = \left\{ t : \mathrm{Term}(t) \wedge \mathrm{dp}(t) \leq \tilde{b} \wedge t \leq 2^{2^{\tilde{k}}} \right\}.$$

Claim 26 *For each $t \leq 2^{2^{k-C}}$,*

$$t \in \tilde{\Lambda} \textit{ if and only if}$$

$$p \models \mathrm{Term}(\underline{t}),\ p \models \mathrm{dp}(\underline{t}) \leq \underline{(1+\varepsilon')b} \textit{ and } p \models \underline{t} \leq \underline{2^{2^{k-C}}}.$$

Proof. For $t \leq 2^{2^{k-C}}$, formulas $\mathrm{Term}(x)$, $\mathrm{dp}(x) \leq y$ and $x \leq 2^{2^y}$ with terms $\underline{t}$ and $\underline{(1+\varepsilon')b}$ are g.e. for Λ (see lemma 23). Since p decides these formulas the result follows from the absoluteness lemma. $\square$

Let $\Lambda(t, x, y)$ be a formula expressing that term t is such that $t \leq 2^{2^x}$ and $\mathrm{dp}(t) \leq y$. Claim 26 established that

$$\forall t (t \in \tilde{\Lambda} \iff p \models \Lambda[\underline{t}, \underline{k - C}, \underline{(1+\varepsilon')b}]. \tag{7}$$

Let $\Lambda(x, y)$ be the set of terms defined by $\Lambda(t, x, y)$. We can refer to this set by a term $s^{\exists z \forall t \leq x (t \in z \equiv \Lambda(t,x,y))}$. Let

$$\gamma(x, y) := \exists z (z \text{ is an } N\text{–deciding, } T\text{–evaluation on } \Lambda(x, y)).$$

Then, let

$$\hat{p} = s^{\exists z \gamma(x,y)}(\underline{k-C}, \underline{(1+\varepsilon')b}).$$

We assumed that $T \vdash \mathrm{HCons}(T, (1+\varepsilon)\log^3)$. Thus, by the fact that

$$(1+\varepsilon')b \leq (1+\varepsilon)(b-C)$$

and by lemmas 24 and 23, we have:

$$p \models \exists z(z \text{ is an } N\text{–deciding, } T\text{–evaluation on } \Lambda(x,y))[\underline{k-C}/x, (1+\varepsilon')b/y]$$

and, consequently,

$$p \models (z \text{ is an } N\text{–deciding, } T\text{–evaluation on } \Lambda(x,y))[\hat{p}/z, \underline{k-C}/x, (1+\varepsilon')b/y].$$

Of course, by our choice of N, p decides this formula.[1]

Since p decides the formula $x_1 \models x_2[x_3]$, then for each $t_1, t_2 \in \tilde{\Lambda}$,

$$p \models \text{“}\hat{p} \models \underline{t_1} = \underline{t_2}\text{” or } p \models \text{“}\neg\hat{p} \models (\underline{t_1} = \underline{t_2})\text{”},$$

and not both. Thus, we define $\tilde{p}$ on $\tilde{\Lambda}$ as follows: for each $t_1, t_2 \in \tilde{\Lambda}$,

$$\tilde{p} \models t_1 = t_2 \iff p \models \text{“}\hat{p} \models \underline{t_1} = \underline{t_2}\text{”}.$$

We claim that $(\tilde{\Lambda}, \tilde{p}, \tilde{k}, \tilde{b})$ is suitable. It suffices to show that $\tilde{p}$ is an N–deciding, T–evaluation on $\tilde{\Lambda}$. The other conditions from definition 21 are easily seen to be satisfied.

Now, we establish the relationship between $\tilde{p} \models \varphi$ and $p \models$ “$\hat{p} \models \underline{\varphi}$”. We need to show that it makes sense to ask whether $p \models$ “$\hat{p} \models \underline{\varphi}$” when we ask whether $\tilde{p} \models \varphi$.

Claim 27 *Let $\varphi \leq \log^*(a)$ and $t_1, \ldots, t_m \in \tilde{\Lambda}$. If $(\varphi, t_1, \ldots, t_m)$ is g.e. for $\tilde{\Lambda}$ then $(\models, \hat{p}, \underline{\varphi}, \langle \underline{t_1}, \ldots, \underline{t_m} \rangle)$ is g.e. for Λ.*

Proof. Let us assume that $(\varphi, t_1, \ldots, t_m)$ is g.e. for $\tilde{\Lambda}$. Under the usual coding the term $s^\varphi(t_1, \ldots, t_m)$ is greater than $\varphi \Pi_{i \leq m} t_i$. Thus, by the construction of $\tilde{\Lambda}$,

$$\varphi \Pi_{i \leq m} t_i \leq 2^{2^{k-C}}.$$

[1] One may object that we want N to be greater than a formula which uses N as a fixed parameter. However, this is easily possible if N has a short encoding. If N is of the form $\exp^n(\underline{2})$, for some n, than a formula $\varphi(N)$ can be written in a short but equivalent form $\exists z(z = \exp^n(\underline{2}) \wedge \varphi(z))$.

Let

$$\begin{aligned}
s_0 &= f_1(\hat{p}, \underline{\varphi}, \langle \underline{t_1}, \dots, \underline{t_m} \rangle), \\
s_1 &= f_2(\hat{p}, \underline{\varphi}, \langle \underline{t_1}, \dots, \underline{t_m} \rangle, s_0), \\
&\vdots \\
s_r &= f_r(\hat{p}, \underline{\varphi}, \langle \underline{t_1}, \dots, \underline{t_m} \rangle, s_0, \dots, s_{r-1})
\end{aligned}$$

be all terms in the sequence for a formula $\hat{p} \models \underline{\varphi}[\underline{t_1}, \dots, \underline{t_m}]$ besides parameters $\hat{p}$, $\underline{\varphi}$ and $\langle \underline{t_1}, \dots, \underline{t_m} \rangle$. (Here $\langle x_1, \dots, x_m \rangle$ is the m-th iteration of the pairing function.) The depth of s_i is not greater than $k + \mathbb{N}$ thus it is less than b. So, it suffices to show that $s_1, \dots, s_r \leq 2^{2^k}$. We estimate the size of s_r. Since $s_1, \dots, s_r$ are terms witnessing quantifiers in $\models$ we have $r \leq \mathrm{lh}(\models)$. Now we estimate lengths of terms s_i. We show that for $i \leq r$,

$$\mathrm{lh}(s_i) \leq 2^i A^i \, \mathrm{lh}(s_0)$$

where A is the constant from equation 4. Of course, there is nothing to prove for s_0 but we write the formula for $\mathrm{lh}(s_0)$ since it will be useful later.

$$\mathrm{lh}(s_0) \leq A(\mathrm{lh}(\models) + \mathrm{lh}(\hat{p}) + \mathrm{lh}(\underline{\varphi}) + m \, \mathrm{lh}(\langle\rangle) + \sum_{i \leq m} \mathrm{lh}(\underline{t_i})),$$

$$\begin{aligned}
\mathrm{lh}(s_1) &\leq A(\mathrm{lh}(\models) + \mathrm{lh}(\hat{p}) + \mathrm{lh}(\underline{\varphi}) + m \, \mathrm{lh}(\langle\rangle) + \sum_{i \leq m} \mathrm{lh}(\underline{t_i}) + \mathrm{lh}(s_0)) \\
&\leq 2A \, \mathrm{lh}(s_1)
\end{aligned}$$

$$
\begin{aligned}
\mathrm{lh}(s_2) &\le A(\mathrm{lh}(\models) + \mathrm{lh}(\hat{p}) + \mathrm{lh}(\underline{\varphi}) + m\,\mathrm{lh}(\langle\rangle) + \sum_{i\le m} \mathrm{lh}(\underline{t_i}) + \mathrm{lh}(s_0) + \mathrm{lh}(s_1))\\
&\le A(\mathrm{lh}(\models) + \mathrm{lh}(\hat{p}) + \mathrm{lh}(\underline{\varphi}) + m\,\mathrm{lh}(\langle\rangle) + \sum_{i\le m} \mathrm{lh}(\underline{t_i})) + A(\mathrm{lh}(s_0) + \mathrm{lh}(s_1)))\\
&\le A\,\mathrm{lh}(s_0) + A(2^0A^0\,\mathrm{lh}(s_0) + 2^1A^1\,\mathrm{lh}(s_0))\\
&\le (A + \sum_{j<2} 2^jA^j)\,\mathrm{lh}(s_0)\\
&\le (1 + \sum_{j<2} 2^j)A^2\,\mathrm{lh}(s_0)\\
&\le 2^2A^2\,\mathrm{lh}(s_0)\\
&\vdots\\
\mathrm{lh}(s_r) &\le A(\mathrm{lh}(\models) + \mathrm{lh}(\hat{p}) + \mathrm{lh}(\underline{\varphi}) + m\,\mathrm{lh}(\langle\rangle) + \sum_{i\le m} \mathrm{lh}(\underline{t_i}) + \sum_{i<r} \mathrm{lh}(s_i))\\
&\le A(\mathrm{lh}(\models) + \mathrm{lh}(\hat{p}) + \mathrm{lh}(\underline{\varphi}) + m\,\mathrm{lh}(\langle\rangle) + \sum_{i\le m} \mathrm{lh}(\underline{t_i})) + A\sum_{i<r} \mathrm{lh}(s_i))\\
&\le A\,\mathrm{lh}(s_0) + A\sum_{i<r} 2^iA^i\,\mathrm{lh}(s_0)\\
&\le (1 + \sum_{i<r} 2^i)A^r\,\mathrm{lh}(s_0)\\
&\le 2^rA^r\,\mathrm{lh}(s_0).
\end{aligned}
$$

The parameter m is less than $\mathrm{lh}(\varphi) \le \log^*(a)$. So, we can estimate the elements in the sum for $\mathrm{lh}(s_0)$ as follows:

$$
\begin{aligned}
\mathrm{lh}(\hat{p}) &\le A(2\,\mathrm{lh}(\underline{k-C}) + \mathrm{lh}(s^\gamma)),\\
\mathrm{lh}(\varphi) &\le \log^*(a),\\
m\,\mathrm{lh}(\langle\rangle) &\le \log^*(a)\,\mathrm{lh}(\langle\rangle),
\end{aligned}
$$

Thus, the length of s_0 can be estimated by

$$2A\log(k) + \log^*(a) + \log^*(a)\log^*(a) + \sum_{i\le m} \mathrm{lh}(t_i) + \mathbb{N}.$$

Since $a \le 2^{2^k}$, it follows that

$$\mathrm{lh}(s_0) \le 3A\log(k) + \sum_{i\le m} \mathrm{lh}(t_i).$$

Since the length of the greatest term s_r is not greater than $2^r A^r \operatorname{lh}(s_0)$ we can bound the size of s_r by

$$\begin{aligned} s_r &\leq 2^{(2A)^r \operatorname{lh}(s_0)} \\ &\leq 2^{(2A)^r(3A\log(k)+\sum_{i\leq m}\operatorname{lh}(t_i))} \\ &\leq 2^{(2A)^r 3A\log(k)}(\prod_{i\leq m} t_i)^{(2A)^r} \end{aligned}$$

and, because $\varphi \prod_{i\leq m} t_i \leq 2^{2^{k-C}}$,

$$\begin{aligned} &\leq 2^{(2A)^r 3A\log(k)}(2^{2^{k-C}})^{(2A)^r} \\ &\leq 2^{(2A)^r 3A\log(k)} 2^{2^{k-C}(2A)^r} \\ &\leq 2^{3\log(k)(2A)^r A} 2^{2^{k-C+r\log(2A)}} \\ &\leq 2^{2^{k-C+r\log(2A)}+3\log(k)(2A)^r A} \end{aligned}$$

and, since $k > \mathbb{N}$ we have $2^{k-C+r\log(2A)} > 3\log(k)(2A)^r A$,

$$\begin{aligned} &\leq 2^{2^{k-C+r\log(2A)+1}} \\ &\leq 2^{2^k}. \end{aligned}$$

The last inequality is true since r is not greater than $\log(\models)$ and we have chosen C so that $C \geq r\log(2A)+1$.

It is also easy to see that $\operatorname{dp}(s_r) \leq k + \mathbb{N} \leq b$. This completes the proof of claim 27. □

Now, we need to show that $\tilde{p}$ reflects $\hat{p}$ not only for equality but for all formulas of size $\log^*(a)$.

Claim 28 *For each $\varphi \leq \log^*(a)$, for each $t_1, \ldots, t_m \in \tilde{\Lambda}$ such that $(\varphi, t_1, \ldots, t_m)$ are g.e. for $(\tilde{\Lambda}, \tilde{p})$,*

$$\tilde{p} \models \varphi[t_1, \ldots, t_m] \iff p \models \text{“}\hat{p} \models \underline{\varphi}[\underline{t_1}, \ldots, \underline{t_m}]\text{”}.$$

Proof. The proof is by induction on φ. For atomic φ the statement follows from the definition of $\tilde{p}$.

Since,

$$p \models \text{"}\hat{p} \text{ satisfies Tarski conditions for propositional connectives"},$$

it follows that the equivalence holds for all quantifier free formulas. Now, let us consider a case where φ is of the form $\exists y\psi(y,\bar{z})$. Let us assume that $\tilde{p} \models \exists y\psi(y, t_1, \dots, t_m)$. Then,

$$\tilde{p} \models \psi[s^{\exists y\psi}(t_1, \dots, t_m), t_1, \dots, t_m].$$

By the inductive assumption,

$$p \models \text{“}\hat{p} \models \underline{\psi}[\underline{s^{\exists y\psi}(t_1, \dots, t_m)}, \underline{t_1}, \dots, \underline{t_m}]\text{”}.$$

Since all terms $\underline{s^{\exists y\psi}(t_1, \dots, t_m)}, \underline{t_1}, \dots \underline{t_m}$ are in $\left\{\underline{0}, \dots, \underline{2^{2^k}}\right\}$ it follows that p decides that $\underline{s^{\exists y\psi}(t_1, \dots, t_m)}$ is a witnessing term for $\underline{\exists y\psi}$ and $\underline{t_1}, \dots, \underline{t_m}$. Thus,

$$p \models (\text{“}\hat{p} \models \underline{\psi}[\underline{s^{\exists y\psi}(t_1, \dots, t_m)}, \underline{t_1}, \dots, \underline{t_m}]\text{”} \iff \text{“}\hat{p} \models \underline{\exists y\psi}[\underline{t_1}, \dots, \underline{t_m}]\text{”}).$$

Let us observe, that in the above formula ψ and $\exists y\psi$ are given as terms. Thus, p does not need to decide these formulas to decide positively the above equivalence. The proof of the implication from $p \models \text{“}\hat{p} \models \underline{\exists y\psi}[\underline{t_1}, \dots, \underline{t_m}]\text{”}$ to $p \models \exists y\psi[t_1, \dots, t_m]$ goes exactly in the same way.

Finally, let $\tilde{p} \models \forall x\psi[t_1, \dots, t_m]$. This is equivalent to

$$\forall t \in \tilde{\Lambda}\, \tilde{p} \models \psi[t, t_1, \dots, t_m]$$

and

$$\forall t \in \tilde{\Lambda}\, p \models \text{“}\hat{p} \models \underline{\psi}[\underline{t}, \underline{t_1}, \dots, \underline{t_m}]\text{”}.$$

But, by equation (7),

$$\forall t(t \in \tilde{\Lambda} \iff p \models \Lambda[\underline{t}, \underline{k-C}, \underline{(1+\varepsilon')b}]).$$

Thus

$$p \models \forall z(z \in \Lambda(\underline{k-C}, \underline{(1+\varepsilon')}b) \Rightarrow \hat{p} \models \underline{\psi}[\underline{t_1}, \dots, \underline{t_m}, z])$$

and this is just the definition of

$$p \models \text{“}\hat{p} \models \underline{\forall x\psi}[\underline{t_1}, \dots, \underline{t_m}]\text{”}.$$

Again, we skip the proof of the other implication. $\square$

Since $\hat{p}$ is (under p) a T–evaluation, $p \models \text{“}\hat{p} \models \underline{\varphi}\text{”}$, for $\varphi \in T$ ($\varphi \leq \log^*(a)$). By claim 28, for each $\varphi \in T$ such that φ is g.e. for $\tilde{\Lambda}$,

$$\tilde{p} \models \varphi.$$

Thus $\tilde{p}$ is a T–evaluation. Moreover, $\tilde{p}$ decides formulas less or equal N because

$$p \models \text{“}\hat{p} \text{ decides formulas less or equal } \underline{N}\text{”}$$

and this property is also transferred to $\tilde{p}$ by claim 28. So, $(\tilde{\Lambda}, \tilde{p}, \tilde{k}, \tilde{b})$ satisfies the third condition for being suitable. This completes the proof of lemma 25. $\square$

Now, we are ready to formulate the induction.

Proposition 29 *Assume that* $T \vdash \mathrm{HCons}((1+\varepsilon)\log^3)$, $M \models T$, $a \in \log^3(M)$ *and that* d *is chosen according to definition 21. Then, for all* $i \leq a/(C^2)$ *there exist* Λ_i, p_i *such that*

$$(\Lambda_i, p_i, a - iC, (1+\varepsilon/2)^{i+1)}a)$$

is suitable.

Proof. We prove the conclusion by induction on $i \leq a/C^2$. To carry on the induction we should take sufficiently large parameter to express the induction formula as bounded. To see that such a parameter exists we should estimate the size of Λ_0 and p_0 which are the greatest ones among Λ_i, p_i for $i \leq a/C^2$. Λ_0 is a set of terms less or equal 2^{2^a}. Thus,

$$\Lambda_0 \leq 2^{2^{2^a}}.$$

Since $a \in \log^3$, $\Lambda_0 \in M$. An evaluation p_0 is just a 0–1 function from the set of pairs of terms from Λ and we can bound p_0 as

$$p_0 \leq 2^{\mathrm{card}(\Lambda)^2} = 2^{(2^{2^a})^2} = 2^{2^{22^a}} = 2^{2^{2^{a+1}}}.$$

In T, the third logarithm of the model is closed on successor. Thus,

$$p_0 \in M$$

and we can bound the induction formula by an element from M.

To start the induction, the existence of a suitable $(\Lambda_0, p_0, a, (1+\varepsilon/2)a)$ is guaranteed by our assumption that $T \vdash \mathrm{HCons}((1+\varepsilon)\log^3)$.

To prove the induction step let us assume that for some $i \leq a/C^2$ there are Λ_i, p_i such that the sequence $(\Lambda_i, p_i, a - iC, (1+\varepsilon/2)^{i+1}a)$ is suitable. Then, $(\Lambda_i, p_i, a - iC, (1+\varepsilon)^{i+1}a)$ and chosen parameters satisfy the assumptions of lemma 25. It follows that there exists $(\tilde{\Lambda}, \tilde{p}, a - (i+1)C, (1+\varepsilon/2)^{i+2}a)$ which is suitable. And this is just the $(i+1)$-th sequence. $\square$

Lemma 30 *Assume that* $T \vdash \mathrm{HCons}((1+\varepsilon)\log^3)$. *Then for each model* $M \models T$ *and each* $a \in \log^3(M)$ *there exists a model* $M' \models T$ *such that* $M'{\upharpoonright}a = M{\upharpoonright}a$ *and* $a \in (C^2/\log(1+\varepsilon/2))\log^4(M')$.

Proof. By proposition 29, for $i = a/C^2$, there exists a suitable sequence

$$(\Lambda, p, a(1-1/C), (1+\varepsilon/2)^{a/C^2}a).$$

Now, for any cut $I < (1+\varepsilon/2)^{a/C^2}a$ if we take terms from Λ of depths in I we can define from p a model M' for T. By lemma 24 and proposition 11,

$$M' \models \exists z(z = \exp^3((1+\varepsilon/2)^{a/C^2}a - C)).$$

But

$$(1+\varepsilon/2)^{a/C^2}a - C = 2^{\log(1+\varepsilon/2)C^{-2}a+\log(a)} - C \geq 2^{\log(1+\varepsilon/2)C^{-2}a}.$$

It follows that

$$M' \models 2^{\log(1+\varepsilon/2)C^{-2}a} \in \log^3$$

what means that $a \in (C^2/\log(1+\varepsilon/2))\log^4(M')$. Moreover, since $\{\underline{0},\ldots,\underline{a}\} \subseteq \Lambda$, we obtain, by lemma 14, that $M'{\upharpoonright}a = M{\upharpoonright}a$. $\Box$

Theorem 31 *Assume that* $T \vdash \mathrm{HCons}((1+\varepsilon)\log^3)$. *Then for each model* $M \models T$ *and each* $a \in \log^3(M)$ *there exists a model* $M' \models T$ *such that* $M'{\upharpoonright}a = M{\upharpoonright}a$ *and* $a \in \log^4(M')$.

Proof. It suffices to use lemma 30 twice to obtain the thesis. Indeed, let $M \models T$, $a \in \log^3(M)$ and let $D = C^2/\log(1+\varepsilon/2) \in \mathbb{N}$. Then, by lemma 30, there exists $M'' \models T$ such that $M''{\upharpoonright}a = M{\upharpoonright}a$ and $a \in D\log^4(M'')$. Thus, $2^{a/D} \in \log^3(M'')$ and, again by lemma 30, there exists $M' \models T$ such that $M'{\upharpoonright}(2^{a/D}) = M{\upharpoonright}(2^{a/D})$ and $2^{a/D} \in D\log^4(M')$. So, $2^{a/D-\log(D)} \in \log^4(M')$. Since $a > \mathbb{N}$ and $D \in \mathbb{N}$, $a < 2^{a/D-\log(D)}$. Thus, we have that $a \in \log^4(M')$. This ends the proof of the theorem. $\Box$

Theorem 32 *Let* $T = I\Delta_0 + \Omega_1$. *Then, for any* $\varepsilon > 0$, T *does not prove its Herbrand consistency restricted to terms of depth not greater than* $(1+\varepsilon)\log^3$, *that is* $T \nvdash \mathrm{HCons}(T, (1+\varepsilon)\log^3)$.

Proof. The thesis is an easy consequence of theorems 31 and 17. $\Box$

References

[A01] Z. Adamowicz, *On tableaux consistency in weak theories*, preprint 618, Institute of Mathematics of the Polish Academy of Sciences, 2001.

[A02] Z. Adamowicz, *Herbrand consistency and bounded arithmetic*, in Fundamenta Mathematicae 171(2002), pp. 279–292

[AZ01] Z. Adamowicz and P. Zbierski, *On Herbrand consistency in weak arithmetics*, in Archive for Mathematical Logic, 40(2001), pp. 399–413.

[HP93] P. Hájek and P. Pudlák, *Metamathematics of first–order arithmetic*, Springer–Verlag, 1993.

[K06] L. A. Kołodziejczyk, *On the Herbrand notion of consistency for finitely axiomatizable fragments of bounded arithmetic theories*, in Journal of Symbolic Logic 71(2006), pp. 624–638.

[WP87] A. J. Wilkie and J. B. Paris, *On the scheme of induction for bounded arithmetical formulas*, in Annals of Pure and Applied Logic, 35(1987), pp. 261–302.