A NOTE ON PARAMETER FREE $\Pi_1$–INDUCTION AND RESTRICTED EXPONENTIATION

A. Cordón-Franco, A. Fernández-Margarit, and F. F. Lara-Martín

Dpto. Ciencias de la Computación e Inteligencia Artificial.
Facultad de Matemáticas, Universidad de Sevilla.
C/ Tarfia s/n, 41012 Sevilla, Spain
{acordon, afrmargarit, fflara}@us.es

Abstract. We characterize the sets of all $\Pi_2$ and all $B(\Sigma_1)$ (= boolean combinations of $\Sigma_1$) theorems of $I\Pi_{n+1}$ in terms of restricted exponentiation, and use these characterizations to prove that both sets are not deductively equivalent. We also discuss how these results generalize to $n > 0$. As an application, we prove that a conservation theorem of L.D. Beklemishev stating that $I\Pi_{n+1}$ is conservative over $I\Sigma_n$ with respect to $B(\Sigma_{n+1})$ sentences cannot be extended to $\Pi_{n+2}$ sentences.

1. Introduction

A natural question in studying the strength of an arithmetic theory $T$, is to characterize the sets $Th_{\Gamma}(T)$ of all theorems of $T$ of a specific logical complexity $\Gamma$. In this paper we shall deal with this question for the scheme of parameter free $\Pi_1$–induction $I\Pi_1^-$. $I\Pi_1^-$ is the first order theory in the language of Peano Arithmetic $L = \{0, 1, +, \cdot, \leq\}$ given by finitely many algebraic axioms $P^-$ together with the induction scheme

$$(I\varphi) \quad \varphi(0) \land \forall x (\varphi(x) \rightarrow \varphi(x + 1)) \rightarrow \forall x \varphi(x)$$

for each $\varphi(x) \in I\Pi_1^-$. As usual, we write $\varphi(x) \in I\Pi_1^-$ to mean that $\varphi$ is in $\Gamma$ and contains no other free variables but $x$ (we refer the reader to [7], [9] for the fundamental notions and results of first order Arithmetic, and to [6], [8] for properties of $I\Pi_1^-$).

Our interest in $I\Pi_1^-$ is motivated by the fact that, in a sense, this theory lies in between weak and strong fragments of Arithmetic. First, observe that $I\Pi_1^-$ is axiomatized by a set of true $\Sigma_2$ sentences and hence follows from the set of all true $\Pi_1$ arithmetic sentences, $Th_{\Pi_1}(\mathcal{N})$. So, $I\Pi_1^-$ cannot prove the totality of functions of more than polynomial growth and in particular exponentiation is not necessarily total in a model of $I\Pi_1^-$. From this point of view, $I\Pi_1^-$ can be considered a weak theory close to the scheme of induction for bounded formulas $I\Delta_0$. In fact, it is a theorem of T. Bigorajska that $I\Pi_1^-$ satisfies Parikh’s theorem “asymptotically.”

Theorem 1 (Bigorajska, [2]). Suppose $\varphi(x, y) \in \Sigma_1$ and $\Pi_1^- \vdash \forall x \exists y \varphi(x, y)$. Then, there is a term $t(x)$ such that $\Pi_1^- \vdash \exists u \forall x > u \exists y \leq t(x) \varphi(x, y)$.

Since [8] it has been known, however, that there is a tight relationship between $\Pi_1^-$ and the strong system $I\Delta_0 + exp$ (as usual $exp$ denotes a $\Pi_2$ axiom asserting “exponentiation is total”). Although both theories are incomparable with respect to inclusion, they share the same $B(\Sigma_1)$ consequences, where $B(\Sigma_1)$ denotes the set of all boolean combinations of $\Sigma_1$ sentences.

Theorem 2 (Kaye–Paris–Dimitracopoulos, [8]).

1. $I\Delta_0 + exp$ is conservative over $\Pi_1^-$ with respect to $\Sigma_2$ sentences.
2. $\Pi_1^-$ is conservative over $I\Delta_0 + exp$ with respect to $\Pi_2$ sentences.
3. So, $Th_{\Pi_1}(\Pi_1^-) = Th_{\Pi_1}(I\Delta_0 + exp)$ and $Th_{B(\Sigma_1)}(\Pi_1^-) = Th_{B(\Sigma_1)}(I\Delta_0 + exp)$.

As a result $\Pi_1^-$ proves that there are unboundedly many primes and implies the $\Delta_0$ pigeonhole principle, for they are $\Pi_1$ consequences of $I\Delta_0 + exp$ (whether or not these two properties can be proved from $I\Delta_0$ are important open problems). Besides, the following hierarchy of theories between $I\Delta_0$ and $I\Delta_0 + exp$ arises:

$$I\Delta_0 \subseteq Th_{\Pi_1}(\Pi_1^-) \subseteq Th_{B(\Sigma_1)}(\Pi_1^-) \subseteq Th_{\Pi_2}(\Pi_1^-) \subseteq I\Delta_0 + exp$$

In [12], [13] A.J. Wilkie and J.B. Paris studied the $\Pi_1$ theorems of $\Pi_1^-$ under its equivalent formulation of $Th_{\Pi_1}(I\Delta_0 + exp)$. In [12] Wilkie characterized the $\Pi_1$ sentences provable in $I\Delta_0 + exp$ as those $\Pi_1$ sentences which are interpretable in Robinson’s Arithmetic $Q$. In [13] the authors showed that a $\Pi_1$ sentence expressing the Herbrand consistency of $Q$ is provable in $I\Delta_0 + exp$ but not in $I\Delta_0$ and hence $Th_{\Pi_1}(\Pi_1^-)$ is strictly stronger than $I\Delta_0$. In contrast, no similar results were known for the $B(\Sigma_1)$ and the $\Pi_2$ theorems of $\Pi_1^-$. In particular, it was not known whether $Th_{\Pi_2}(\Pi_1^-)$ follows from $Th_{B(\Sigma_1)}(\Pi_1^-)$.

The aim of this paper is to fill these gaps in our understanding of the hierarchy above by obtaining informative axiomatizations of $Th_{B(\Sigma_1)}(\Pi_1^-)$ and $Th_{\Pi_2}(\Pi_1^-)$, and by proving that the inclusions are all strict. To this end, the key idea is to consider certain forms of restricted exponentiation. We shall show that $\Pi_1^-$ can prove that $x^y$ exists whenever $x, y$ are restricted to be in certain initial segments determined by the $\Sigma_1$ definable elements of a model. We say that $a$ is $\Sigma_n$ definable in $\mathfrak{A}$ if there is $\varphi(x) \in \Sigma_n$ such that $a$ is the unique element satisfying $\varphi(x)$. We write $\mathcal{K}_n(\mathfrak{A}, X)$ to denote the set of all $\Sigma_n$ definable elements of $\mathfrak{A}$ (possibly with a parameter from $X$); or simply $\mathcal{K}_n(\mathfrak{A})$ if $X = \emptyset$. We write $I_n(\mathfrak{A}, X)$ to denote the least initial segment of $\mathfrak{A}$ containing $\mathcal{K}_n(\mathfrak{A}, X)$. Given $\mathfrak{A} \models I\Delta_0$ we shall deal with the following hierarchy of submodels of $\mathfrak{A}$ based on iterated $\Sigma_n$ definability\footnote{The use of these structures goes back to Paris–Kirby’s [10], where the authors used them to construct a model of $B\Sigma_{n+1} + \neg I\Sigma_{n+1}$.}.

- $\mathcal{K}_0(\mathfrak{A}, X) = \mathcal{K}_0(\mathfrak{A})$, i.e. the submodel of $\mathfrak{A}$ whose domain consists of all $\Sigma_n$ definable elements; $I_0(\mathfrak{A}) = I_n(\mathfrak{A})$, i.e. the least initial segment of $\mathfrak{A}$ containing $\mathcal{K}_n(\mathfrak{A})$.
- For each $k \geq 0$, $\mathcal{K}_{k+1}(\mathfrak{A}) = \mathcal{K}_n(\mathfrak{A}, I_k(\mathfrak{A}))$, i.e. the submodel of $\mathfrak{A}$ whose domain consists of all elements which are $\Sigma_n$ definable with a parameter from $I_k(\mathfrak{A})$; $I_{k+1}(\mathfrak{A}) = I_n(\mathfrak{A}, I_k(\mathfrak{A}))$, i.e. the least initial segment of $\mathfrak{A}$ containing $\mathcal{K}_{k+1}(\mathfrak{A})$. 

2
By considering this hierarchy for $n = 1$, we can isolate the exact amount of exponentiation available in models of $\Pi^+_1$: $x^y$ exists whenever the base $x$ is in $I^+_1(\mathfrak{A})$ and the exponent $y$ is in $I_1(\mathfrak{A})$. In addition, the $\Pi_2$ and the $B(\Sigma_1)$ theorems of $\Pi^+_1$ can be characterized over $I \Delta_0$ by principles of the form “For certain $x, y$, $x^y$ exists.” More precisely, next theorem collects together the main results of the paper.

**Theorem 3.**

1. $Th_{\Pi_2}(\Pi^+_1) = I \Delta_0 + \forall a \in I_1(\mathfrak{A}) \forall b \in I^+_1(\mathfrak{A}) \exists y (y = b^a)$.
2. $Th_{B(\Sigma_1)}(\Pi^+_1) = I \Delta_0 + \forall a, b \in I_1(\mathfrak{A}) \exists y (y = b^a)$.
3. $Th_{\Pi_1}(\Pi^+_1) \not\models Th_{B(\Sigma_1)}(\Pi^+_1)$, and $Th_{B(\Sigma_1)}(\Pi^+_1) \not\models Th_{\Pi_2}(\Pi^+_1)$.

Let us observe that the characterizations above are to be seen as axiomatizations in the original language $\mathcal{L} = \{0, 1, +, \cdot, \leq\}$. Indeed, a quantifier of the form “$\forall x \in I_1(\mathfrak{A})$” in front of a formula $\Psi(x)$ is intended as a shorthand for the scheme:

$$\{ \forall x, y \, (\delta(y) \land \forall y' (\delta(y') \rightarrow y' = y) \land x \leq y \rightarrow \Psi(x)) : \delta \in \Sigma_1 \}$$

In a similar vein, a quantifier of the form “$\forall x \in I^+_1(\mathfrak{A})$” in front of a formula $\Psi(x)$ unravels to:

$$\{ \forall x, y, z, u \left( \begin{array}{c} \delta_1(y) \land \forall y' (\delta_1(y') \rightarrow y' = y) \land z \leq y \\ \delta_2(u, z) \land \forall u' (\delta_2(u', z) \rightarrow u' = u) \land x \leq u \end{array} \rightarrow \Psi(x) \right) : \delta_1, \delta_2 \in \Sigma_1 \}$$

Our proofs are model–theoretic and are based on the proofs of Theorem 2 and other related results in [6] and [8]. It will be important, however, to make explicit a property that appeared only implicitly in [6] or [8]: although $\Pi^+_1$ is much weaker than the full scheme of induction for $\Sigma_1$ formulas $I \Sigma_1$, $\Pi^+_1$ does allow for induction up to $I_1(\mathfrak{A})$ for $\Sigma_1$ formulas with parameters in $I_1(\mathfrak{A})$. This property clarifies why a certain amount of exponentiation is available in $\Pi^+_1$ and, in a sense, explains the particular behavior of $\Pi^+_1$ mixing features of weak and strong fragments of Peano Arithmetic.

To close this section, we state some basic notions and results on definable elements that we shall use throughout the paper (for proofs we refer the reader to [7] and [9]). As usual, we write $\mathfrak{A} \prec_n \mathfrak{B}$, $\mathfrak{A} \subseteq^e \mathfrak{B}$, and $\mathfrak{A} \subseteq^{cf} \mathfrak{B}$ to mean that $\mathfrak{A}$ is an $n$–elementary substructure of $\mathfrak{B}$, $\mathfrak{A}$ is an initial segment of $\mathfrak{B}$, and $\mathfrak{A}$ is cofinal in $\mathfrak{B}$ (i.e. unbounded above), respectively. We say that $a$ is in $\Pi_n$ minimal in $\mathfrak{A}$ if there is $\varphi(x) \in \Pi_n$ such that $a$ is the least element satisfying $\varphi(x)$. We write $\mathcal{M}_n(\mathfrak{A}, X)$ to denote the set of all $\Pi_n$ minimal elements of $\mathfrak{A}$ (possibly with a parameter from $X$). Finally, $(x)_0, (x)_1$ denote the inverse functions of Cantor’s paring function.

**Proposition 4.** $(n > 0)$ Suppose $\mathfrak{A} \models I \Sigma_{n-1}$ and $X \subseteq \mathfrak{A}$.

1. $K_n(\mathfrak{A}, X) \prec_n \mathfrak{A}$; and $K_n(\mathfrak{A}, X) \prec^{cf}_n I_n(\mathfrak{A}, X) \prec^{e}_{n-1} \mathfrak{A}$. In particular, every nonempty $\Sigma_n$ definable set contains some $\Sigma_n$ definable element.
2. $\mathcal{M}_{n-1}(\mathfrak{A}, X)$ is cofinal in $K_n(\mathfrak{A}, X)$. Even more, for each $a$ in $K_n(\mathfrak{A}, X)$ there is $b$ in $\mathcal{M}_{n-1}(\mathfrak{A}, X)$ such that $a = (b)_0$.
3. If $X = \emptyset$, $\mathfrak{A} \models I \Sigma_{n-1}$ suffices in 1 and 2.
2. Restricted Exponentiation in Models of $\Pi^1_1$

Although $\Pi^1_1$ is known to be much weaker than the scheme of induction for $\Sigma_1$ formulas $I\Sigma_1$, our first result shows that $\Pi^1_1$ allows for $\Sigma_1$ induction restricted to the initial segment determined by the $\Sigma_1$ definable elements (we write it for $\Pi^1_{n+1}$ and $n \geq 0$ since we will need this general version in a subsequent section).

**Proposition 5.** Suppose $\mathfrak{A} \models \Pi^1_{n+1}$. Then, for each $\varphi(x, v) \in \Sigma_{n+1}$ and $a, b \in I_{n+1}(\mathfrak{A})$, it holds that:

$$\mathfrak{A} \models \varphi(0, b) \land \forall x (\varphi(x, b) \rightarrow \varphi(x + 1, b)) \rightarrow \varphi(a, b)$$

**Proof.** Assume $\mathfrak{A} \models \varphi(0, b) \land \forall x (\varphi(x, b) \rightarrow \varphi(x + 1, b))$. We wish to prove $\mathfrak{A} \models \varphi(a, b)$. The idea is to show that $\varphi(x, b)$ is equivalent to a bounded formula for elements $x \leq a$ and then apply (parametric) $\Delta_0$ induction. Pick $c \in K_{n+1}(\mathfrak{A})$ such that $a, b < c$.

**Claim 1.** If $\theta(x) \in \Pi_{n+1}$ and there is $d \in K_{n+1}(\mathfrak{A})$ such that $\mathfrak{A} \models \theta(d)$, then $\mathfrak{A} \models \exists x (x = (\mu t) (\theta(t)))$.

Towards a contradiction, assume $\mathfrak{A} \models \forall x (\theta(x) \rightarrow \exists x' < x \theta(x'))$. Put $\theta(x) \equiv \forall y \theta_0(x, y)$, where $\theta_0 \in \Sigma_n$.

$n = 0$: Define $\theta'(z)$ to be $\exists u \forall x \leq z \exists y \leq u \neg \theta_0(x, y)$.

$n > 0$: Define $\theta'(z)$ to be $\exists s \forall x \leq z \neg \theta_0(x, s(x))$ (observe that $\mathfrak{A} \models exp$ and so we can make use of a standard coding for sequences in $\mathfrak{A}$).

Clearly, $\theta'(z)$ is in $\Sigma_{n+1}$ and implies $\forall x \leq z \neg \theta(x)$. So,

$$\mathfrak{A} \models \theta'(0) \land \forall z (\theta'(z) \rightarrow \theta'(z + 1)) \land \neg \theta'(d).$$

Now consider $\theta''(z, d) \equiv \neg \theta'(d - z)$. Then, $\theta'' \in \Pi_{n+1}$ and $\mathfrak{A} \models \theta''(0, d) \land \forall z (\theta''(z, d) \rightarrow \theta''(z + 1, d))$. If $\delta(v)$ is a $\Sigma_{n+1}$ definition of $d$, we can get rid of the parameter in $\theta''$ by considering the formula $\forall v (\delta(v) \rightarrow \theta''(z, v))$. Hence $\mathfrak{A} \models \forall z \theta''(z, d)$ by $\Pi^1_{n+1}$. In particular, $\mathfrak{A} \models \theta''(d, d)$ and so $\mathfrak{A} \models \theta(0)$, which is a contradiction.

Observe that $2^x$ exists in $\mathfrak{A}$ for each $\Sigma_{n+1}$ definable exponent (for $n > 0$ it is trivial; for $n = 0$ it follows from the claim by considering the least $t$ satisfying $\neg \exists y (y = 2^t)$). Hence, $2^x$ also exists for each exponent in $I_{n+1}(\mathfrak{A})$. As a consequence, we can use a $\Delta_0$ formula $x \in u$ expressing “the $x$-th digit of the binary expansion of $u$ is 1” to codify finite sets included in $I_{n+1}(\mathfrak{A})$. Now define $\theta(u, c)$ to be the $\Pi_{n+1}$ formula:

$$\forall x, v < c (\varphi(x, v) \rightarrow \langle x, v \rangle \in u)$$

Clearly, $\mathfrak{A} \models \theta(2(c, c) - 1)$. By the claim, there exists $e \in I_{n+1}(\mathfrak{A})$ satisfying $e = (\mu u) (\theta(u, c))$ (again we can get rid of the parameter $c$ in $\theta$ by using a $\Sigma_{n+1}$ definition of it). Now it follows from the definition of $\theta$ that $\mathfrak{A} \models \forall x, v < c (\varphi(x, v) \rightarrow \langle x, v \rangle \in e)$. So, we can obtain $\mathfrak{A} \models \varphi(a, b)$ by applying $\Delta_0$ induction to the formula $x < c \rightarrow \langle x, b \rangle \in e$.

\[\square\]

It immediately follows from the previous result that $\Pi^1_1$ proves that $b^a$ exists whenever $a, b \in I_1(\mathfrak{A})$ (it is sufficient to apply induction up to $a$ to the $\Sigma_1$ formula $\exists y (y = b^x)$). However, a stronger result actually holds:
Theorem 6. Suppose $\mathfrak{A} \models \text{II}^+_1$. Then, for each $a \in I_1(\mathfrak{A})$ and $b \in I_1(\mathfrak{A})$, $b^a$ exists in $\mathfrak{A}$.

Proof. Pick $c \in K^1_1(\mathfrak{A})$ such that $b \leq c$ and let $\delta(u, d)$ be a $\Sigma_1$ definition of $c$, where $d \in I_1(\mathfrak{A})$. By applying induction up to $a$ to the $\Sigma_1$ formula $\exists y, u (\delta(u, d) \land y = u^c)$ (available in $\mathfrak{A}$ thanks to Proposition 5), we get $\mathfrak{A} \models \exists y (y = c^a)$. But it is straightforward to check that $\exists y (y = z^e) \land u \leq z \rightarrow \exists y (y = u^c)$ is provable in $I_\Delta_0$. Consequently, $b^a$ exists in $\mathfrak{A}$.

It is natural to ask whether Theorem 6 captures the amount of exponentiation available in every model of $\text{II}^+_1$. We close this section by showing that this is the case: Theorem 6 is best possible in terms of the hierarchy $I_1^k(\mathfrak{A})$ (and hence so is Proposition 5). First, we need the following two lemmas.

Lemma 7. Suppose $\mathfrak{A} \models \text{II}^+_1$, $\mathfrak{B} \prec_0 \mathfrak{A}$ and $K^1_1(\mathfrak{A}) \subseteq \mathfrak{B}$. Then, $\mathfrak{B}$ also satisfies $\text{II}^+_1$.

Proof. Let $\varphi(x) \in \Pi_1$. We need to prove that $\mathfrak{B} \models I_{\varphi(x)}$. Put $\varphi(x) \equiv \forall y \varphi_0(x, y)$, where $\varphi_0 \in \Delta_0$. The induction axiom for $\varphi$ can be reexpressed as

$$\exists x, y, z \, [\neg \varphi_0(0, y) \lor (\varphi(x) \land \neg \varphi_0(x + 1, z)) \lor \forall x \varphi(x)]$$

Let $\theta(x, y, z)$ denote the $\Pi_1$ formula in brackets above. It suffices to show that $\theta$ can be witnessed in $K^1_1(\mathfrak{A})$, for it follows from $\mathfrak{B} \prec_0 \mathfrak{A}$ and $K^1_1(\mathfrak{A}) \subseteq \mathfrak{B}$ that $\mathfrak{B} \models \exists x, y, z \, \theta(x, y, z)$.

Case 1: $\mathfrak{A} \models \forall x \varphi(x)$. Then, $\mathfrak{A} \models \theta(0, 0, 0)$.

Case 2: $\mathfrak{A} \models \neg \varphi(0)$. There is $a \in K_1(\mathfrak{A})$ such that $\mathfrak{A} \models \neg \varphi_0(0, a)$ and then $\mathfrak{A} \models \theta(0, 0, 0)$.

Case 3: $\mathfrak{A} \models \varphi(0) \lor \exists x \, \neg \varphi(x)$. Pick $a \in K_1(\mathfrak{A})$ such that $\mathfrak{A} \models \neg \varphi(a)$. Then, there is $b \leq a$ such that $\mathfrak{A} \models \varphi(b) \land \neg \varphi(b + 1)$ (otherwise, it would follow from the induction axiom for $x \leq a$ that $\mathfrak{A} \models \varphi(a)$). Since $b \in I_1(\mathfrak{A})$, there is $c \in K_1(\mathfrak{A})$ satisfying $\mathfrak{A} \models \neg \varphi_0(b + 1, c)$. Thus, $\mathfrak{A} \models \theta(b, 0, c)$.

Lemma 8. Suppose $\mathfrak{A} \models I_\Sigma_1$ with nonstandard $\Sigma_1$ definable elements. Then, $\{I^k(\mathfrak{A}) : k \in \omega\}$ form a proper hierarchy, i.e. $K^{k+1}_1(\mathfrak{A}) \neq I^k(\mathfrak{A})$ for each $k \in \omega$.

Proof. Assume that there is $k \in \omega$ such that $K^{k+1}_1(\mathfrak{A}) = I^k(\mathfrak{A})$ (we may also assume that $K^{m+1}_1(\mathfrak{A}) \neq I^m(\mathfrak{A})$ for all $m < k$). Then, $I^k(\mathfrak{A})$ is a nonstandard model of $I_\Sigma_1$. To see this, observe that if $I^k(\mathfrak{A}) = \mathfrak{A}$, then the result is immediate; otherwise, it follows from $I^k(\mathfrak{A}) \prec_1 \mathfrak{A} \models I_\Delta_0$ and $I^k(\mathfrak{A}) \neq \mathfrak{A}$ that $I^k(\mathfrak{A})$ satisfies the $\Sigma_2$ collection scheme $B\Sigma_2$ which is well-known to imply $I_\Sigma_1$. In addition, there is $a \in I^k(\mathfrak{A})$ satisfying that $K_1(I^k(\mathfrak{A}), (\leq a))$ is cofinal in $I^k(\mathfrak{A})$ (if $k = 0$, consider $a = 0$; otherwise, pick $a \in K^k_1(\mathfrak{A}) - I^{k-1}(\mathfrak{A})$). But this is a contradiction, for the $\Sigma_1$ definable elements (even with parameters from a proper cut) of a nonstandard model of $I_\Sigma_1$ are not cofinal in it (see, e.g., chapter IV, lemma 1.37 of [9]).

Proposition 9.

(1) There are $\mathfrak{A} \models \text{II}^+_1$ and $a \in K^1_1(\mathfrak{A})$ such that $2^a$ does not exist in $\mathfrak{A}$.

(2) There are $\mathfrak{A} \models \text{II}^+_1$, $a \in I_1(\mathfrak{A})$ and $b \in K^1_1(\mathfrak{A})$ such that $b^a$ does not exist in $\mathfrak{A}$.
Proof. (1): Let $\mathfrak{B}$ be a model of $\Pi_1^− + \neg \text{con}(I\Delta_0)$, where $\text{con}(I\Delta_0)$ is a $\Pi_1$ sentence expressing the consistency of $I\Delta_0$. Consider $\mathfrak{A} = \mathcal{K}_1(\mathfrak{B})$. Then, $\mathcal{K}_1(\mathfrak{A}) = \mathfrak{A}$, $\mathfrak{A} \models \neg \text{con}(I\Delta_0)$, and $\mathfrak{A} \models \Pi_1^−$ by Lemma 7.

Claim 2. $\mathfrak{A} \not\models I\Delta_0 + \text{exp}$. Towards a contradiction, assume $\mathfrak{A} \models I\Delta_0 + \text{exp}$. Let $z = \text{supexp}(x, y)$ be a $\Delta_0$ formula defining the graph of the superexponential function and satisfying the usual recursion laws of superexponentiation in $I\Delta_0 + \text{exp}$. Pick $a, b \in \mathcal{K}_1(\mathfrak{A})$. By applying induction up to $a$ to the $\Sigma_1$ formula $\exists z (z = \text{supexp}(b, y))$ (available in $\mathfrak{A}$ thanks to Proposition 5), we get $\mathfrak{A} \models (\exists z (z = \text{supexp}(b, y))$. Hence, $\mathcal{K}_1(\mathfrak{A})$ is a model of $I\Delta_0 + \text{supexp} \vdash \text{con}(I\Delta_0)$. So, $\mathfrak{A} \models \text{con}(I\Delta_0)$, which is impossible. Consequently, $\mathfrak{A}$ satisfies $\Pi_1^−$ and there is $a \in \mathcal{K}_1(\mathfrak{A})$ such that $2^a$ does not exist in $\mathfrak{A}$.

(2): Let $\mathfrak{B}$ be a model of $I\Sigma_1$ containing nonstandard $\Sigma_1$ definable elements. Pick $b \in \mathcal{M}(\mathfrak{B}, \mathcal{I}_1(\mathfrak{B}) \setminus \mathcal{I}_1(\mathfrak{B})$ (such an element exists by Lemma 8) and define $\mathfrak{A}$ to be the initial segment of $\mathfrak{B}$ determined by the standard powers of $b$, i.e. $\mathfrak{A} = \{ a \in \mathfrak{B} : \exists k \in \omega, a \leq b^k \}$. Then, $\mathfrak{A} \models \Pi_1^−$ by Lemma 7, $b$ is in $\mathcal{K}_1(\mathfrak{A})$, and $b^a$ does not exist in $\mathfrak{A}$ for any nonstandard exponent $a$. □

3. The $\Pi_2$ and the $B(\Sigma_1)$ Theorems of $\Pi_1^−$

We have shown that $\Pi_1^−$ proves “$\forall a \in \mathcal{I}_1(\mathfrak{A}) \forall b \in \mathcal{I}_1(\mathfrak{A}) b^a$ exists.” This principle can be reexpressed as a set of $\Pi_2$ sentences of $\mathcal{L}$ and hence follows from $Th_{\Pi_2}(\Pi_1^−)$. Interestingly, the converse is also true. Our first characterization theorem states that, over $I\Delta_0$ declaring “$\forall a \in \mathcal{I}_1(\mathfrak{A}) \forall b \in \mathcal{I}_1(\mathfrak{A}) b^a$ exists” allows us to recover all $\Pi_2$ theorems of $\Pi_1^−$. As a by–product, we also obtain a Parsons–like characterization of $Th_{\Pi_2}(\Pi_1^−)$. A well–known result of C. Parsons [11] states that the $\Pi_2$ theorems of $I\Sigma_1$ can be characterized as the closure of $I\Delta_0$ under the $\Sigma_1$ induction rule $I\Delta_0 + \Sigma_1$–IR. Likewise, the $\Pi_2$ theorems of $\Pi_1^−$ can be captured by using a variant of that inference rule: the $\Sigma_1$ induction rule up to $\mathcal{I}_1(\mathfrak{A})$ with parameters in $\mathcal{I}_1(\mathfrak{A})$, $(\Sigma_1, \mathcal{I}_1, \mathcal{I}_1^1)$–IR. This inference rule is given by

$$
\varphi(0, v) \land \forall x \varphi(x, v) \rightarrow \varphi(x + 1, v)
\forall v \in \mathcal{I}_1(\mathfrak{A}) \forall x \in \mathcal{I}_1(\mathfrak{A}) \varphi(x, v),
$$

where $\varphi(x, v)$ runs over $\Sigma_1$. Then, $I\Delta_0 + (\Sigma_1, \mathcal{I}_1, \mathcal{I}_1^1)$–IR denotes the closure of $I\Delta_0$ under $(\Sigma_1, \mathcal{I}_1, \mathcal{I}_1^1)$–IR; whereas $[I\Delta_0, (\Sigma_1, \mathcal{I}_1, \mathcal{I}_1^1)$–IR] denotes the closure of $I\Delta_0$ under unnested applications of the rule.

Theorem 10. The following theories are equivalent.

1. $Th_{\Pi_2}(\Pi_1^−)$.
2. $I\Delta_0 + (\Sigma_1, \mathcal{I}_1, \mathcal{I}_1^1)$–IR.
3. $[I\Delta_0, (\Sigma_1, \mathcal{I}_1, \mathcal{I}_1^1)$–IR].
4. $I\Delta_0 + \forall a \in \mathcal{I}_1(\mathfrak{A}) \forall b \in \mathcal{I}_1(\mathfrak{A}) \exists y (y = b^a)$.

Proof. (1$\Rightarrow$2): Since $I\Delta_0 + (\Sigma_1, \mathcal{I}_1, \mathcal{I}_1^1)$–IR $\subseteq$ $\Pi_2$, it suffices to show that it can be proved from $\Pi_1^−$. Indeed, we show the stronger result that $\Pi_1^−$ is closed under $(\Sigma_1, \mathcal{I}_1, \mathcal{I}_1^1)$–IR.
To this end, suppose $\mathfrak{A} \models \Pi \Pi^1_1^-$ and $\Pi \Pi^1_1^+ \vdash \forall v \varphi(0, v) \land \forall v, x (\varphi(x, v) \rightarrow \varphi(x + 1, v))$, with $\varphi(x, v) \in \Sigma_1$. Let $a \in I_1(\mathfrak{A})$ and $b \in I_1^1(\mathfrak{A})$. Pick $c \in K_1^1(\mathfrak{A})$ such that $b \leq c$ and let $\delta(v, d)$ be a $\Sigma_1$ definition of $c$, where $d \in I_1(\mathfrak{A})$. Write $\varphi(x, v) \equiv \exists y \varphi_0(x, y, v)$, with $\varphi_0 \in \Delta_0$, and define $\theta(x, d)$ to be the $\Sigma_1$ formula

$$\exists v (\delta(v, d) \land \exists u \forall w \leq v \exists y \leq u \varphi_0(x, y, w))$$

Clearly, $\mathfrak{A} \models \forall x (\theta(x, d) \rightarrow \theta(x + 1, d))$. Since $\Pi \Pi^1_1^-$ proves $\forall v \exists y \varphi_0(0, y, v)$ and $\Pi \Pi^1_1^+$ is closed under $\Sigma_1$ collection rule, $\Pi \Pi^1_1^+$ also proves $\forall v \exists u \forall w \leq v \exists y \leq u \varphi_0(0, y, w)$ and so $\mathfrak{A} \models \theta(0, d)$. Thus, it follows from Proposition 5 that $\mathfrak{A} \models \theta(a, d)$ and hence $\mathfrak{A} \models \varphi(a, b)$.

Case 1: $I_1^1(\mathfrak{A}) \neq \mathfrak{A}$

Then, $\mathfrak{A}$ satisfies the least number axiom scheme for parameter free $\Sigma_1$ formulas $L\Sigma_1^-$, which is known to be equivalent to $\Pi \Pi^1_1^-$. To see that, let $\varphi(x) \in \Sigma_1$ and $c \in K_1(\mathfrak{A})$ such that $\mathfrak{A} \models \varphi(c)$. Put $\varphi(x) \equiv \exists y \varphi_0(x, y)$, where $\varphi_0 \in \Delta_0$, and pick $d \in \mathfrak{A}$ satisfying $I_1^1(\mathfrak{A}) < d$. Then, $\mathfrak{A} \models \forall x \leq c (\exists y \varphi_0(x, y) \rightarrow \exists y \leq d \varphi_0(x, y))$ and so $\varphi(x)$ is equivalent to the $\Delta_0$ formula $\exists y \leq d \varphi_0(x, y)$ for elements $x \leq c$. By applying $\Delta_0$–minimization in $\mathfrak{A}$, we get that there is a least element satisfying $\varphi(x)$.

Case 2: $I_1^1(\mathfrak{A}) = \mathfrak{A}$

Then, we can assume that $b^o$ exists in $\mathfrak{A}$ for all $a \in I_1(\mathfrak{A})$ and all $b \in \mathfrak{A}$. Now we essentially follow the proof of the $\Pi_2$ conservativity of $\Pi \Pi^1_1^-$ over $I \Delta_0 + \text{exp}$ included in Kaye's Ph.D. thesis (theorem 10.8 of [6]). Towards a contradiction, assume that there is a $\Pi_2$ sentence $\theta$ such that $\Pi \Pi^1_1^- \vdash \theta$ and $\mathfrak{A} \models \neg \theta$. By compactness, there are $\varphi_1(x), \ldots, \varphi_n(x) \in \Pi_1$ such that $I \Delta_0 + I \varphi_1 + \ldots + I \varphi_n \vdash \theta$. For the sake of simplicity, we first observe that

Claim 3. There is a single $\Pi_1$ formula $\varphi(x)$ such that $Th_{\Sigma_2}(\mathfrak{A}) + I \varphi \vdash \theta$.

By reordering the formulas $\varphi_i(x)$ if necessary, we can assume that there is a $m \leq n$ such that: a) none of $I \varphi_1, \ldots, I \varphi_m$ is true in $\mathfrak{A}$ (hence $\exists x \neg \varphi_i(x)$ is in $Th_{\Sigma_2}(\mathfrak{A})$ for $1 \leq i \leq m$); and b) $I \varphi_{m+1}, \ldots, I \varphi_n$ are all true in $\mathfrak{A}$ (and hence included in $Th_{\Sigma_2}(\mathfrak{A})$). It is well-known that the least number axiom for a formula $\neg \delta(x)$, $L \neg \delta(x)$, is equivalent to the induction axiom for $\forall y \leq x \delta(y)$. In addition, it easy to check that $Th_{\Sigma_2}(\mathfrak{A}) + L \neg \varphi_1((x_1) \land \ldots \land \neg \varphi_m((x)_m) \vdash L \neg \varphi_1(x) \land \ldots \land L \neg \varphi_m(x)$, where $(x)_1, \ldots, (x)_m$ denote the inverse functions for a coding for $m$–tuples. Thus, it is sufficient to define $\varphi(x)$ to be the $\Pi_1$ formula $\forall y \leq x (\varphi_1(y_1) \lor \ldots \lor \varphi_m((y)_m))$.

Let $\mathfrak{B}$ be a recursively saturated elementary extension of $\mathfrak{A}$. It follows from the saturation of $\mathfrak{B}$ that there exists $c \in \mathfrak{B}$ such that, for each $\delta(x) \in \Pi_2$, $\mathfrak{B} \models \exists x \delta(x) \rightarrow \exists x \leq c \delta(x)$. As a result, $I_1(\mathfrak{B}) < c$ and every initial segment of $\mathfrak{B}$ containing $c$ is a model of $Th_{\Sigma_2}(\mathfrak{A})$.

Since $\mathfrak{A}$ and $\mathfrak{B}$ are elementary equivalent, $b^o$ also exists in $\mathfrak{B}$ for $a \in I_1(\mathfrak{B})$ and all $b \in \mathfrak{B}$. In addition, $I \varphi$ must fail in $\mathfrak{B}$ and so:

i) $\mathfrak{B} \models \varphi(0) \land \forall x (\varphi(x) \rightarrow \varphi(x + 1))$, and
ii) there is $a \in K_1(\mathfrak{B})$ such that $\mathfrak{B} \models \neg \varphi(a)$.

Put $\varphi(x) \equiv \exists y \varphi_0(x, y)$, with $\varphi_0 \in \Delta_0$. Then, we have:
Claim 4. \( \exists a' \leq a \exists b \forall k \in \omega : \mathcal{B} \models \forall z \leq (b + c)^k \varphi_0(a', z) \land \neg \varphi_0(a' + 1, b) \).

By the saturation of \( \mathcal{B} \), it is sufficient to show that

\[ \forall k \in \omega \exists a' \leq a \exists b : \mathcal{B} \models \forall z \leq (b + c)^k \varphi_0(a', z) \land \neg \varphi_0(a' + 1, b) \]

Assume not. Then, there is some \( k \in \omega \) such that:

\[ \mathcal{B} \models \forall x \leq a \forall y (\forall z \leq (c + y)^k \varphi_0(x, z) \rightarrow \varphi_0(x + 1, y)) \]

Since \( \mathcal{B} \models \neg \varphi(a) \), there is \( d \in K_1(\mathcal{B}) \) such that \( \mathcal{B} \models \neg \varphi_0(a, d) \). Define \( \psi(x) \) to be the \( \Delta_0 \) formula

\[ \forall z \leq (c + d)^{k(2^n + 1) - 2x} \varphi_0(x, z) \]

Observe that \( k(2^n + 1) - 2x \) is in \( I_1(\mathcal{B}) \) and hence \( (c + d)^{k(2^n + 1) - 2x} \) exists. Clearly, \( \mathcal{B} \models \psi(0) \) and it is easy to check that \( \mathcal{B} \models \forall x < a (\psi(x) \rightarrow \psi(x + 1)) \). Thus, by applying \( I\Delta_0 \) in \( \mathcal{B} \), we get that \( \mathcal{B} \models \psi(a) \) and so \( \mathcal{B} \models \varphi_0(a, d) \), which is a contradiction.

Let \( a' \leq a \) and \( b \) as in the Claim above and consider \( \mathcal{C} = \{ d \in \mathcal{B} : \forall k \in \omega, \ d \leq (c + b)^k \} \).

Since \( \mathcal{C} \) is an initial segment of \( \mathcal{B} \) containing \( c \), \( \mathcal{C} \) satisfies \( Th_{\Sigma_2}(\mathcal{A}) \) and so \( \mathcal{C} \models \neg \varphi \).

However, by the Claim, \( \mathcal{C} \models \varphi(a') \land \neg \varphi(a' + 1) \) and hence \( \mathcal{C} \) satisfies \( Th_{\Sigma_2}(\mathcal{A}) + I_\varphi \models \varphi \), which gives the desired contradiction.

As a first application of Theorem 10, we obtain that

**Corollary 11.** \( Th_{\Pi_1}(\Pi_1) \) is not finitely axiomatizable.

**Proof.** Let \( T_0 \) be a finite part of \( Th_{\Pi_1}(\Pi_1) \). Since \( Th_{\Pi_1}(\mathcal{N}) + Th_{\Pi_2}(\Pi_1) \), there is a true \( \Pi_1 \) sentence \( \varphi \) satisfying \( I\Delta_0 + \varphi \models T_0 \). Let \( \mathcal{A} \) be a model of \( I\Sigma_1 + \varphi \) with nonstandard \( \Sigma_1 \) definable elements. Since \( I_1(\mathcal{A}) \neq I_1(\mathcal{A}) \) by Lemma 8, there is \( b \in M_0(\mathcal{A}, I_1(\mathcal{A})) - I_1(\mathcal{A}) \).

Define \( \mathcal{B} \) to be the initial segment of \( \mathcal{A} \) determined by the standard powers of \( b \). Clearly, \( \mathcal{B} \models I\Delta_0 + \varphi \models T_0 \). In addition, \( \mathcal{B} \models Th_{\Pi_2}(\Pi_1) \), for \( b \) is in \( I_1(\mathcal{B}) \) and \( b^a \) does not exist in \( \mathcal{B} \) for any nonstandard exponent \( a \). \( \square \)

We can also derive from Theorem 10 a characterization of the \( \mathcal{B}(\Sigma_1) \) consequences of \( \Pi_1 \). First, we need the following simple but useful observation (recall that \( \mathcal{A} \) and \( \mathcal{B} \) are said to be \( 1 \)-elementary equivalent, written \( \mathcal{A} \equiv_1 \mathcal{B} \), if for each \( \Sigma_1 \) sentence \( \varphi \), \( \mathcal{A} \models \varphi \) iff \( \mathcal{B} \models \varphi \)).

**Lemma 12.** Suppose \( \mathcal{A} \models I\Delta_0 \). The following are equivalent:

1. \( \mathcal{A} \models Th_{\Sigma_1}(T) \).
2. There exists \( \mathcal{B} \models T \) such that \( \mathcal{A} \equiv_1 \mathcal{B} \).
3. \( \mathcal{K}_1(\mathcal{A}) \models Th_{\Pi_2}(T) \).

**Proof.** (1\( \Rightarrow \)2): It suffices to show that \( T + Th_{\Pi_1}(\mathcal{A}) + Th_{\Sigma_1}(\mathcal{A}) \) is a consistent theory. Assume not, then there are \( \varphi_1 \in \Pi_1 \) and \( \varphi_2 \in \Sigma_1 \) satisfying that \( T + \varphi_1 + \varphi_2 \) is inconsistent and \( \mathcal{A} \models \varphi_1 \land \varphi_2 \). Define \( \theta \) to be \( \neg \varphi_1 \lor \neg \varphi_2 \). Then, \( \theta \) is a \( \mathcal{B}(\Sigma_1) \) sentence provable in \( T \) and \( \mathcal{A} \models \neg \theta \), which is impossible since \( \mathcal{A} \models Th_{\Sigma_1}(T) \).

(2\( \Rightarrow \)3): Assume \( \mathcal{B} \) is a model of \( T \) satisfying that \( \mathcal{A} \equiv_1 \mathcal{B} \). Then, \( \mathcal{K}_1(\mathcal{B}) \) satisfies \( Th_{\Pi_2}(T) \) since \( \mathcal{K}_1(\mathcal{B}) \) is a canonical model of \( \Pi_1 \). It follows from \( \mathcal{A}, \mathcal{B} \models I\Delta_0 \) and \( \mathcal{A} \equiv_1 \mathcal{B} \) that there exists a canonical
isomorphism between $\mathcal{K}_1(\mathfrak{A})$ and $\mathcal{K}_1(\mathfrak{B})$ (if $\varphi(x) \in \Sigma_1$ defines $a \in \mathfrak{A}$, map $a$ to the unique $b \in \mathfrak{B}$ satisfying $\mathfrak{B} \models \varphi(b)$). So, $\mathcal{K}_1(\mathfrak{A}) \models Th_{\Pi_1}(T)$.

(3$\Rightarrow$1): Since $\mathcal{K}_1(\mathfrak{A}) \preceq_1 \mathfrak{A} \models Th_{\Sigma_2}(Th_{\Pi_1}(T))$. But clearly $Th_{\mathcal{B}(\Sigma_1)}(T) \subseteq Th_{\Sigma_2}(Th_{\Pi_1}(T))$. \hfill $\Box$

**Corollary 13.**

1. If $T$ implies $I\Delta_0$ then $Th_{\Sigma_2}(Th_{\Pi_1}(T)) \equiv Th_{\mathcal{B}(\Sigma_1)}(T)$.
2. $Th_{\Sigma_2}(I\Delta_0 + exp) \equiv Th_{\mathcal{B}(\Sigma_1)}(I\Delta_0 + exp) \equiv Th_{\mathcal{B}(\Sigma_1)}(I\Pi_1^-)$.

We are now ready to give our characterization of the $\mathcal{B}(\Sigma_1)$ consequences of $I\Pi_1^-$ (and hence also of the $\Sigma_2$ consequences of $I\Delta_0 + exp$) in terms of restricted exponentiation. As in Theorem 10, we also obtain a Parsons–like characterization of this class. To this end, we introduce the parameter free $\Sigma_1$ induction rule up to $I_1(\mathfrak{A})$, $(\Sigma_1^- , I_1)$–IR. This inference rule is given by

$$
\varphi(0) \land \forall x (\varphi(x) \rightarrow \varphi(x + 1))
\forall x \in I_1(\mathfrak{A}) \varphi(x)
$$

where $\varphi(x) \in \Sigma_1^-$. 

**Theorem 14.** The following theories are equivalent.

1. $Th_{\mathcal{B}(\Sigma_1)}(I\Pi_1^-)$.
2. $I\Delta_0 + (\Sigma_1^- , I_1)$–IR.
3. $[I\Delta_0, (\Sigma_1^- , I_1)$–IR$]$.
4. $I\Delta_0 + \forall a, b \in I_1(\mathfrak{A}) \exists y (y = b^a)$.

**Proof.** (1$\Rightarrow$2): We prove the stronger fact that $Th_{\mathcal{B}(\Sigma_1)}(I\Pi_1^-)$ is closed under $(\Sigma_1^- , I_1)$–IR. Assume that $\mathfrak{A} \models Th_{\mathcal{B}(\Sigma_1)}(I\Pi_1^-)$ and $Th_{\mathcal{B}(\Sigma_1)}(I\Pi_1^-) \models \varphi(0) \land \forall x (\varphi(x) \rightarrow \varphi(x + 1))$, with $\varphi(x) \in \Sigma_1^-$. Let $a \in I_1(\mathfrak{A})$ and let $b \in K_1(\mathfrak{A})$ such that $a \leq b$. By Lemma 12, there exists $\mathfrak{B} \models I\Pi_1^-$ such that $\mathfrak{A} \equiv_1 \mathfrak{B}$ (and so $K_1(\mathfrak{A}) = K_1(\mathfrak{B})$). Write $\varphi(x) \equiv \exists y \varphi_0(x,y)$, where $\varphi_0 \in \Delta_0$, and define $\theta(z)$ be the $\Sigma_1$ formula $\exists u \forall x \leq z \exists y \leq u \varphi_0(x,y)$. By applying $\Sigma_1$ induction up to $b$ (available in $\mathfrak{B}$ thanks to Proposition 5), we get $\mathfrak{B} \models \theta(b)$ and so $\mathfrak{A} \models \varphi(a)$.

(2$\Rightarrow$3): Immediate.

(3$\Rightarrow$4): Assume $\mathfrak{A} \models [I\Delta_0, (\Sigma_1^- , I_1)$–IR$]$. It follows that $2^a$ exists for each exponent $a \in I_1(\mathfrak{A})$. So, $I_1(\mathfrak{A}) \models I\Delta_0 + exp$ and hence (4) follows.

(4$\Rightarrow$1): Assume $\mathfrak{A} \models I\Delta_0$ and $b^a$ exists for all $a, b \in I_1(\mathfrak{A})$. Clearly, $K_1(\mathfrak{A}) \models I\Delta_0 + exp$ and so $K_1(\mathfrak{A}) \models Th_{\Pi_2}(I\Pi_1^-)$ by Theorem 10. Thus, $\mathfrak{A} \models Th_{\mathcal{B}(\Sigma_1)}(I\Pi_1^-)$ by Lemma 12. \hfill $\Box$

Reasoning as in Corollary 11, we also get that

**Corollary 15.** $Th_{\mathcal{B}(\Sigma_1)}(I\Pi_1^-)$ is not finitely axiomatizable.
4. Separation Results

We start by proving that \( Th_{\Pi_1}(\Pi_1^-) \) is strictly weaker than \( Th_{B(\Sigma_1)}(\Pi_1^-) \). Roughly speaking, given a nonstandard \( \Delta_0 \) minimal element \( a \), we consider the function

\[
f_a(x) = \begin{cases} 
2^x & \text{if } x = a \\
0 & \text{if } x \neq a
\end{cases}
\]

Then, \( Th_{B(\Sigma_1)}(\Pi_1^-) \) can prove that \( f_a \) is a total function since \( a \in I_1(\mathfrak{A}) \) but \( Th_{\Pi_1}(\Pi_1^-) \) cannot: for otherwise \( 2^a \) would be bounded by a polynomial in \( a \) by Parikh’s theorem.

More formally:

**Definition 16.** Let \( \delta(x) \) be a \( \Delta_0 \) formula. Define \( y = f_\delta(x) \) to be the following \( \Delta_0 \) formula:

\[
(x = (\mu z)(\delta(z)) \land y = 2^x) \lor (x \neq (\mu z)(\delta(z)) \land y = 0)
\]

**Theorem 17.** Let \( \delta(x) \) be a \( \Delta_0 \)-formula such that \( \mathcal{N} \models \forall x \neg \delta(x) \) and \( \Pi_1^- \not\vdash \forall x \neg \delta(x) \) (con(PA) \( \equiv \forall x \neg \delta(x) \), say). Then:

1. \( Th_{B(\Sigma_1)}(\Pi_1^-) \vdash \forall x \exists! y (y = f_\delta(x)) \).
2. \( Th_{\Pi_1}(\Pi_1^-) \not\vdash \forall x \exists y (y = f_\delta(x)) \).
3. Hence, \( Th_{B(\Sigma_1)}(\Pi_1^-) \) does not follow from \( Th_{\Pi_1}(\Pi_1^-) \).

**Proof.** (1): Let \( \mathfrak{A} \) be a model of \( Th_{B(\Sigma_1)}(\Pi_1^-) \) and \( a \in \mathfrak{A} \). If \( a \) is not the least element satisfying \( \delta(x) \) then \( \mathfrak{A} \models 0 = f_\delta(a) \). Otherwise, \( a \) is in \( K_1(\mathfrak{A}) \) and hence \( \mathfrak{A} \models \exists y (y = 2^a) \) by Theorem 14.

(2): Towards a contradiction, assume that \( Th_{\Pi_1}(\Pi_1^-) \) proves \( \forall x \exists y (y = f_\delta(x)) \). By Parikh’s theorem there is a term \( t(x) \) such that \( Th_{\Pi_1}(\Pi_1^-) \vdash \forall x \exists y \leq t(x) (y = f_\delta(x)) \). Let \( \mathfrak{A} \) be a model of \( \Pi_1^- + \exists x \delta(x) \) and \( a = (\mu x)(\delta(x)) \). Then, \( y = f_\delta(x) \) defines a total function in \( \mathfrak{A} \) and \( 2^a = f_\delta(a) \). From \( \mathfrak{A} \models \exists y \leq t(a) (y = f_\delta(a)) \) it follows that \( 2^a \leq t(a) \), which is a contradiction since \( a \) is nonstandard. \( \square \)

Let us observe that the proof of the previous result actually shows the following more general fact:

**Corollary 18.** \( Th_{\Pi_1}(\mathcal{N}) \) is the only consistent \( \Pi_1 \) theory to imply \( Th_{B(\Sigma_1)}(\Pi_1^-) \).

In what follows we show that \( Th_{B(\Sigma_1)}(\Pi_1^-) \) is strictly weaker than \( Th_{\Pi_2}(\Pi_1^-) \). Following the ideas in Theorem 17, given two \( \Delta_0 \) minimal elements \( a, c \) and a \( \Delta_0 \) formula \( \delta(z,v) \), we consider the function:

\[
g(x) = \begin{cases} 
(x)_3^{(x)_0} & \text{if } (x)_0 = a \land (x)_1 = c \land (x)_2 \leq (x)_1 \land (x)_3 = (\mu z)(\delta(z,(x)_2)) \\
0 & \text{otherwise}
\end{cases}
\]

Then, \( Th_{\Pi_2}(\Pi_1^-) \) proves that \( g \) defines a total function (observe that if an element \( x \) satisfies the first condition in the definition of \( g \) then \( (x)_3 \) is in \( I_1(\mathfrak{A}) \) and \( (x)_0 \) is in \( I_1(\mathfrak{A}) \)).
However, since $Th_{B(\Sigma_1)}(\Pi_1^\Gamma)$ is not $\Pi_1$ axiomatizable, now we cannot make use of Parikh’s theorem to obtain that the totality of $g$ is not provable in $Th_{B(\Sigma_1)}(\Pi_1^\Gamma)$. Instead, we need to show that a (finer) version of Bigorajska’s theorem holds for $Th_{B(\Sigma_1)}(\Pi_1^\Gamma)$. First of all, it is worth noticing that Bigorajska’s theorem can be derived from Parikh’s theorem as follows (Bigorajska’s proof was based on an ultrapower construction).

**Proposition 19.** Suppose $\varphi(x, y) \in \Sigma_1$ and $I \Delta_0 + \exists u \forall \theta(u) \vdash \forall x \exists y \varphi(x, y)$, where $\theta(u) \in \Pi_1$.

1. There is a term $t(x)$ such that $I \Delta_0 \vdash \forall x \exists y \varphi(x, y)$.
2. So, $I \Delta_0 + \exists u \forall \theta(u) \vdash \exists u \forall x > u \exists y \leq t(x) \varphi(x, y)$.

**Proof.** Clearly $I \Delta_0 \vdash \forall x, u \exists y (\theta(u) \rightarrow \varphi(x, y))$. By Parikh’s theorem there is a term $t'(x, u)$ such that $I \Delta_0 \vdash \forall x, u \exists y \leq t'(x, u) (\theta(u) \rightarrow \varphi(x, y))$. By the monotonicity of terms, $I \Delta_0 \vdash x > u \rightarrow t'(x, u) \leq t'(x, x)$ and so $I \Delta_0 \vdash \forall u (\theta(u) \rightarrow \forall x > u \exists y \leq t'(x, x) \varphi(x, y))$.

Now it is sufficient to put $t(x) = t'(x, x)$.

Our proof makes clear that Bigorajska’s theorem actually holds for every $\Sigma_2$ extension of $I \Delta_0$. Besides, it gives us a useful piece of information: if a $\Sigma_2$ theory $T$ proves the totality of a $\Sigma_1$ function $f$ and the axioms of $T$ can be “witnessed” in a certain subset $\mathcal{A}$, then there is an element $u$ already in $\mathcal{A}$ such that $f$ is bounded by a polynomial for all $x > u$. Thus, we have:

**Proposition 20.** Suppose $\varphi(x, y) \in \Sigma_1$ and $T \vdash \forall x \exists y \varphi(x, y)$. Let $\mathfrak{A}$ be a model of $T$.

1. If $T \subseteq B(\Sigma_1)$, there are a term $t(x)$ and $a \in \mathcal{I}_1(\mathfrak{A})$ such that $\mathfrak{A} \models \forall x > a \exists y \leq t(x) \varphi(x, y)$.
2. If $T = \Pi_1^\Gamma$, there are a term $t(x)$ and $a \in \mathcal{I}_1^1(\mathfrak{A})$ such that $\mathfrak{A} \models \forall x > a \exists y \leq t(x) \varphi(x, y)$.

**Proof.** If $T \subseteq B(\Sigma_1)$, every axiom of $T$ can be “witnessed” in $\mathcal{K}_1(\mathfrak{A})$ since $\mathcal{K}_1(\mathfrak{A}) \prec_1 \mathfrak{A}$. If $T = \Pi_1^\Gamma$, we know from the proof of Lemma 7 that every axiom of $\Pi_1^\Gamma$ can be “witnessed” in $\mathcal{K}_1^1(\mathfrak{A}) \subseteq \mathcal{I}_1^1(\mathfrak{A})$.

We are now ready to obtain the desired non collapse result. First, we write down a formal definition of the function $g$ described above.

**Definition 21.** Let $D = (\delta_1(x), \delta_2(x), \delta_3(x, v))$ be a triple of $\Delta_0$ formulas. Consider $\theta_D(x)$ to be

$$(x)_{\theta} = (\mu z)(\delta_1(z)) \land (x)_{\delta_2} = (\mu z)(\delta_2(z)) \land (x)_{\delta_3} = (\mu z)(\delta_3(z, x))_{\delta_2}$$

and define $y = g_D(x)$ to be the $\Delta_0$ formula: $(\theta_D(x) \land y = (x)_{\delta_3}(x)) \lor (\neg \theta_D(x) \land y = 0)$.

**Theorem 22.** There is a triple of $\Delta_0$ formulas $D = (\delta_1(x), \delta_2(x), \delta_3(x, v))$ satisfying that:

1. $Th_{\Pi_2}(\Pi_1^\Gamma) \vdash \forall x \exists y (y = g_D(x))$.
2. $Th_{B(\Sigma_1)}(\Pi_1^\Gamma) \not\vdash \forall x \exists y (y = g_D(x))$.
3. Hence, $Th_{\Pi_2}(\Pi_1^\Gamma)$ does not follow from $Th_{B(\Sigma_1)}(\Pi_1^\Gamma)$.
Proof. (1): Assume $\mathfrak{A} \models Th_{\Pi_2}(\Pi_1^-)$ and $a \in \mathfrak{A}$. If $\mathfrak{A} \models \neg \theta_D(a)$ then $\mathfrak{A} \models 0 = g_D(a)$. Otherwise, $(a)_0$ is in $I_1(\mathfrak{A})$ and $(a)_3$ is in $I_1'\mathfrak{(A)}$ and hence $(a)_3^{(a)_0}$ exists by Theorem 10. (2): Let $\mathfrak{A}$ be a model of $I\Sigma_1$ with nonstandard $\Sigma_1$ definable elements. By Lemma 8, $K_1\mathfrak{A} \neq I_1(\mathfrak{A})$ and so there are $c \in M_0(\mathfrak{A})$ and $d \leq c$ such that $M_0(\mathfrak{A}, d) - I_1(\mathfrak{A}) \neq \emptyset$. Pick $a \in M_0(\mathfrak{A})$ nonstandard and $b \in M_0(\mathfrak{A}, d) - I_1(\mathfrak{A})$. Let $\delta_1(x)$, $\delta_2(x)$ and $\delta_3(x, v)$ be $\Delta_0$ formulas such that $a$ is the least element satisfying $\delta_1(x)$, $c$ is the least element satisfying $\delta_2(x)$, and $b$ is the least element satisfying $\delta_3(x, d)$. Finally, let $D = (\delta_1, \delta_2, \delta_3)$ and let $y = g_D(x)$ be as in Definition 21. Towards a contradiction, assume that $Th_{\mathfrak{B}(\Sigma_1)}(\Pi_1^-)$ proves $\forall x \exists y (y = g_D(x))$. Then, it follows from Proposition 20 that there are a term $t(x)$ and $a' \in I_1(\mathfrak{A})$ such that $\mathfrak{A} \models \forall x > a' \exists y \leq t(x) (y = g_D(x))$. Since $e = (a, c, d, b) \geq a'$, $\mathfrak{A} \models \exists y \leq t(e) (y = g_D(e))$ and so $\mathfrak{A} \models \exists y \leq t(e) (y = b^a)$. Since $a, c, d \leq b$, $t(e) \leq b^k$ for some $k \in \omega$. Hence, $b^a \leq b^k$ for some $k \in \omega$, which is a contradiction since $a$ is nonstandard. □

Combining Theorem 22 and Proposition 12 we obtain

**Corollary 23.** $Th_{\Pi_2}(\Pi_1^-)$ is not $\Sigma_2$-axiomatizable.

We close this section with some remarks.

(a) Our results give us a description of what total functions of $\Pi_1^1$ look like. Suppose $\Pi_1^1 \vdash \forall x \exists y (y = f(x))$, where $y = f(x)$ is a $\Sigma_1$ formula. Let $\mathfrak{A}$ be a nonstandard model of $\Pi_1^1$. By Proposition 20 there are $b \in I_1(\mathfrak{A})$ and a polynomial $p_1(x)$ such that $f(x) \leq p_1(x)$ for all $x > b$. If we define $\mathfrak{B}$ to be the submodel given by $\{ c \in \mathfrak{A} : \forall z \in M_0(\mathfrak{A}) - \omega, c \leq z \}$, then it is easy to check that $\mathfrak{B}$ is the largest initial segment of $\mathfrak{A}$ satisfying $Th_{\Pi_1}(\mathcal{N})$. By applying Parikh’s theorem for $Th_{\Pi_1}(\mathcal{N})$, we get that there is a polynomial $p_2(x)$ such that $f(x) \leq p_2(x)$ for all $x \in \mathfrak{B}$. By $\Delta_0$-overspill there is $a \in M_0(\mathfrak{A})$ such that $f(x) \leq p_2(x)$ for all $x \leq a$. Put $p(x) = p_1(x) + p_2(x)$. Thus, $f$ is bounded by $p$ for all $x$ in $[0, a] \cup [b, +\infty)$; whereas $f$ may present an “exponential interference” in the interval $(a, b) \subseteq I_1(\mathfrak{A})$.

(b) Our separation results also inform us of the limits of Parikh’s and Bigorajka’s theorems. Let $y = f_b(x)$ and $y = g_D(x)$ be as in Theorem 17 and Theorem 22, respectively. Firstly, by Theorem 17, $Th_{\mathfrak{B}(\Sigma_1)}(\Pi_1^-)$ proves $\forall x \exists y (y = f_b(x))$ but does not prove $\forall x \exists y \leq t(x) (y = f_b(x))$ for any term $t(x)$. Secondly, it follows from Theorem 22 and Corollary 13 that $Th_{\mathfrak{B}(\Pi_1^-)}(\Pi_1^-)$ proves $\forall x \exists y (y = g_D(x))$ but does not prove $\exists u \forall x > u \exists y \leq t(x) (y = g_D(x))$ for any term $t(x)$. Thus, we have:

**Proposition 24.**

1. $Th_{\mathfrak{B}(\Sigma_1)}(\Pi_1^-)$ satisfies Bigorajka’s theorem and does not satisfy Parikh’s theorem.
2. $Th_{\Pi_2}(\Pi_1^-)$ does not satisfy Bigorajka’s theorem.

This result solves in the negative a question posed in [3], where the authors studied the quantifier complexity of the induction schema for the class of $\Delta_{n+1}$ formulas in an arithmetic theory $T$, $I\Delta_{n+1}(T)$. Whereas for $n > 0$ this scheme is not $\Sigma_{n+2}$ axiomatizable for any theory $T$, for $n = 0$ the authors gave a number of examples for which $I\Delta_1(T)$ is $\Sigma_2$ axiomatizable (e.g. $T = I\Delta_0$, or $B\Sigma_1$). Indeed, for all of those examples $I\Delta_1(T)$ is also a $\Pi_1$ axiomatizable theory. This fact motivated the following question.
Question 1 (Problem 7.2 of [3]). Suppose $I\Delta_0 \subseteq T \subseteq I\Sigma_1$ and $T$ is closed under $\Sigma_1$ collection rule. Are the following conditions equivalent?

1. $I\Delta_1(T)$ is $\Pi_1$ axiomatizable.
2. $I\Delta_1(T)$ satisfies Parikh’s theorem.
3. $I\Delta_1(T)$ is $\Sigma_2$ axiomatizable.
4. $I\Delta_1(T)$ satisfies Bigorajska’s theorem.

It follows from theorem 2.4 of [3] that, under the assumptions of Question 1 $I\Delta_1(T) \equiv \text{Th}_{\Pi_2}(T)$. So, it is clear that (1) and (2) are equivalent; and it is easy to check that (3) and (4) are equivalent too. However, if we put $T = \text{Th}_{B(S_1)}(\Pi_1^\omega)$ then $I\Delta_1(T) \equiv \text{Th}_{B(S_1)}(\Pi_1^\omega)$ and so $I\Delta_1(T)$ is $\Sigma_2$ axiomatizable and not $\Pi_1$, for it does not satisfy Parikh’s theorem. Consequently, (1) and (3) are not equivalent.

Let us note, in passing, that in the statement of Question 1 in [3] the assumption $T \subseteq I\Sigma_1$ is missing; as well as in the remarks following that problem it is erroneously claimed that $\text{Th}_{\Pi_2}(\Pi_1^\omega)$ satisfies Bigorajska’s theorem.

5. Concluding Remarks

It is natural to ask how the results in the previous sections generalize to the induction scheme for parameter free $\Pi_{n+1}$ formulas $\Pi_{n+1}$ for $n > 0$. Whereas in the case $n = 0$ we have the anomalous situation $\Pi_1^\omega + I\Delta_0 \equiv I\Delta_0$, in the case $n > 0$ the theories $\Sigma_n$ and $I\Sigma_n$ cease to be equivalent and indeed $I\Sigma_n + \Pi_{n+1}$ and $\Pi_{n+1}$ differ significantly in strength. It turns out that the ideas in the previous sections apply equally well to theories $I\Sigma_n + \Pi^{-n}_{n+1}$ for $n > 0$. The key point is to think of the exponential function as the iteration of a polynomial. Indeed, building on a well–known construction of Kaye, in [5] it is shown that, for each $n \geq 0$, there is a $\Pi_n$ formula $y = \mathbb{K}_n(x)$ satisfying:

- (a) $\exists x \forall y (y = \mathbb{K}_n(x))$; and
- (b) initial segments of $\mathcal{N} \models I\Sigma_n$ closed under function $y = \mathbb{K}_n(x)$ are $\Sigma_n$-elementary substructures of $\mathcal{N}$.

Observe that for $n = 0$ it suffices to put $\mathbb{K}_0(x) = (x + 2)^2$; and indeed for $n > 0$ functions $y = \mathbb{K}_n(x)$, with $k \in \omega$, will play the role of polynomials. Moreover, in section 3 of [3] it is presented a $\Pi_n$ formula $y = \mathbb{K}_n(x)$ that expresses the iteration of the function $y = \mathbb{K}_n(x)$. It is not hard to check that our characterization and non collapse results remain true for $n > 1$ and for $I\Sigma_{n-1} + \Pi_{n}^-$ when restricted exponentiation is replaced with restricted iterations of functions $y = \mathbb{K}_n(x)$. Namely,

Theorem 25.

1. $\text{Th}_{\Pi_{n+2}}(I\Sigma_n + \Pi_{n+1}) \equiv I\Sigma_n + \forall a \in I_{n+1}(\mathcal{A}) \forall b \in I_{n+1}(\mathcal{A}) \exists y (y = \mathbb{K}_n^a(b))$.
2. $\text{Th}_{B(S_{n+1})}(I\Sigma_n + \Pi_{n+1}) \equiv I\Sigma_n + \forall a, b \in I_{n+1}(\mathcal{A}) \exists y (y = \mathbb{K}_n^a(b))$.
3. $\text{Th}_{\Pi_{n+2}}(I\Sigma_n + \Pi_{n+1}^-)$ does not follow from $\text{Th}_{B(S_{n+1})}(I\Sigma_n + \Pi_{n+1}^-)$, which, in turn, does not follow from $\text{Th}_{\Pi_{n+1}}(I\Sigma_n + \Pi_{n+1}^-)$.

With respect to theories $\Pi_{n+1}^-$ for $n > 0$, the following questions are in order.
Question 2. Does $\text{Th}_{B(\Sigma_{n+1})}(\Pi_{n+1}^-)$ follow from $\text{Th}_{\Pi_{n+1}}(\Pi_{n+1}^-)$?

Question 3. Does $\text{Th}_{\Pi_{n+2}}(\Pi_{n+1}^-)$ follow from $\text{Th}_{B(\Sigma_{n+1})}(\Pi_{n+1}^-)$?

The negative answer to Question 2 can be derived from axiomatization properties of $\Sigma_n^-$. In [4] it is proved that, for $n > 0$, $\text{Th}_{\Pi_{n+1}}(N)$ is the only $\Pi_{n+1}$ theory to imply $\Sigma_n^-$. Thus, $\text{Th}_{B(\Sigma_{n+1})}(\Pi_{n+1}^-)$ does not follow from $\text{Th}_{\Pi_{n+1}}(\Pi_{n+1}^-)$ since $\text{Th}_{B(\Sigma_{n+1})}(\Pi_{n+1}^-)$ implies $\Sigma_n^-$. Question 3 is related to a conservation theorem of L.D. Beklemishev:

Theorem 26 (Beklemishev, [1]). $(n > 0)$ $\Pi_{n+1}^-$ is conservative over $\Sigma_n^-$ w.r.t. $B(\Sigma_{n+1})$ sentences.

As a result, $\text{Th}_{B(\Sigma_{n+1})}(\Pi_{n+1}^-) \equiv \Sigma_n^-$ and Question 3 is equivalent to asking whether Beklemishev’s conservation theorem can be extended to $\Pi_{n+2}$ sentences. We conclude by proving that Question 3 also has the negative answer and hence Beklemishev’s conservation theorem is best possible (observe that it cannot be extended to $\Sigma_{n+2}$ sentences either, for the least number principle for parameter free $\Delta_{n+1}$ formulas $L\Delta_{n+1}$ is a $\Sigma_{n+2}$ theory that can be proved from $\Pi_{n+1}^-$ but not from $\Sigma_n^-$.)

The idea to separate $\Sigma_n^-$ and $\text{Th}_{\Pi_{n+2}}(\Pi_{n+1}^-)$ is to consider the following model-theoretic property:

$$\begin{align*}
A & \equiv \forall a \in K_n^1(\mathfrak{A}) \; " \mathcal{K}_n(\mathfrak{A},a) \text{ is not cofinal in } \mathfrak{A}" \\
A' & \equiv \forall a \in K_n^1(\mathfrak{A}) \; " \mathcal{M}_{n-1}(\mathfrak{A},a) \text{ is not cofinal in } \mathfrak{A}." 
\end{align*}$$

Recall that in models of $\Sigma_{n-1}$ the set of the $\Pi_{n-1}$ minimal elements $\mathcal{M}_{n-1}(\mathfrak{A},a)$ is cofinal in $\mathcal{K}_n(\mathfrak{A},a)$. Hence, (A) can be reexpressed as:

$$A' \equiv \forall a \in K_n^1(\mathfrak{A}) \; " \mathcal{M}_{n-1}(\mathfrak{A},a) \text{ is not cofinal in } \mathfrak{A}."$$

Let $\text{Sat}_{n-1}(x,y,v)$ denote a truth predicate for $\Pi_{n-1}$ formulas (which is available in $\Sigma_{n-1}$) expressing “$x$ satisfies the $\Pi_{n-1}$ formula $y$ with a parameter $v$. Then, it is straightforward to write down a $\Delta_n$ formula $\psi(x,y,v)$ formalizing the property “$x$ is the least element satisfying the $\Pi_{n-1}$ formula $y$ with a parameter $v."$ We can now formalize over $\Sigma_{n-1}$, (a stronger version of) property $(A')$ using the following set of $\Pi_{n+2}$ sentences of $\mathcal{L}$:

$$B \equiv \forall a \in I_{n+1}(\mathfrak{A}) \; \forall b \in I_{n+1}(\mathfrak{A}) \; \exists u \forall x \forall y \leq b(\psi(x,y,a) \rightarrow x \leq u)$$

Theorem 27 $(n > 0)$.

1. $\text{Th}_{\Pi_{n+2}}(\Pi_{n+1}^-) \vdash (B)$.
2. $\text{Th}_{B(\Sigma_{n+1})}(\Pi_{n+1}^-) \not\vdash (B)$.
3. Hence, $\text{Th}_{\Pi_{n+2}}(\Pi_{n+1}^-)$ does not follow from $\text{Th}_{B(\Sigma_{n+1})}(\Pi_{n+1}^-) \equiv \Sigma_n^-.$

Proof. (1): Since $(B) \subseteq \Pi_{n+2}$, it suffices to show that $\Pi_{n+1}^- \vdash (B)$. Suppose $\mathfrak{A} \models \Pi_{n+1}^-$, $a \in K_n^1(\mathfrak{A})$ and $b \in K_{n+1}(\mathfrak{A})$. There exist $\delta(v,w) \in \Sigma_{n+1}$ and $c \in I_{n+1}(\mathfrak{A})$ such that $\delta(v,c)$ defines $a$. Let $\varphi(z,c)$ be the $\Sigma_{n+1}$ formula

$$\exists v [\delta(v,c) \land \exists u \forall y \leq z(\exists x \text{Sat}_{n-1}(x,y,v) \rightarrow \exists x \leq u \text{Sat}_{n-1}(x,y,v))].$$
Clearly, $\mathfrak{A} \models \varphi(0, c)$ and $\mathfrak{A} \models \forall z \ (\varphi(z, c) \rightarrow \varphi(z + 1, c))$. So, $\mathfrak{A} \models \varphi(b, c)$ by Proposition 5. Hence, property (B) holds in $\mathfrak{A}$, for it follows from the definition of $\psi$ that $I\Sigma_{n-1} + exp \vdash \psi(x, y, v) \land Sat_{n-1}(x', y, v) \rightarrow x \leq x'$.

(2): Let $\mathfrak{A}$ be a model of $I\Sigma_{n+1}$ with nonstandard $\Sigma_{n+1}$ definable elements. Reasoning as in the proof of Lemma 8, we get that $\mathcal{K}_{n+1}(\mathfrak{A}) \neq I_{n+1}(\mathfrak{A})$ and so there is $a \in \mathcal{M}_n(\mathfrak{A}, b) - I_{n+1}(\mathfrak{A})$, with $b \in I_{n+1}(\mathfrak{A})$. Pick $c \in \mathcal{M}_n(\mathfrak{A})$ such that $b \leq c$. Put $d = \langle a, b, c \rangle$ and define $\mathcal{C}$ to be $I_n(\mathfrak{A}, d)$, i.e. the least initial segment of $\mathfrak{A}$ containing $K_n(\mathfrak{A}, d)$. Then, $I_{n+1}(\mathfrak{A}) \subset \mathcal{C}$ and $\mathcal{C} \prec_{\mathfrak{A}} n - 1$ by Proposition 4. In addition, it follows from theorem 1.33, Chapter IV of [9] that $\mathcal{C} \models B\Sigma_n$ and $\mathcal{C}$ satisfies $Th_{I_{n+1}}(\mathfrak{A}; d)$, i.e. the set of all $\Pi_{n+1}$ sentences (possibly containing the parameter $d$) which are true in $\mathfrak{A}$. Then, we have:

i) $\mathcal{C} \models Th_{B(\Sigma_{n+1})}(\Pi_{n+1}^-)$.

Since $\mathfrak{A} \models \Pi_{n+1}^-$, it is sufficient to prove that $\mathcal{C} \models Th_{\Pi_{n+1}^-}(\mathfrak{A})$. Clearly, $\mathcal{C} \models Th_{\Pi_{n+1}^-}(\mathfrak{A})$. To prove that $\mathcal{C} \models Th_{\Sigma_{n+1}}(\mathfrak{A})$, suppose $\mathfrak{A} \models \exists x \theta(x)$, where $\theta(x)$ is in $\Pi_n$. It follows from $I_{n+1}(\mathfrak{A}) < a$ that $\mathfrak{A} \models \exists x \leq a \theta(x)$ and so $\mathcal{C} \models \exists x \leq a \theta(x)$ since $\mathcal{C} \prec_{\mathfrak{A}} n - 1$ and $\mathfrak{A}$ and $|0| \subseteq \mathcal{C}$.

ii) $\mathcal{C} \not\models (B)$.

By Proposition 4, $K_n(\mathfrak{A}, d) \prec \mathcal{C}$ and so $K_n(K_n(\mathfrak{A}, d), d) = K_n(\mathcal{C}, d)$. It follows from $K_n(\mathfrak{A}, d) \prec \mathfrak{A}$ that $K_n(K_n(\mathfrak{A}, d), d) = K_n(\mathfrak{A}, d)$. So, $K_n(\mathcal{C}, d) = K_n(\mathfrak{A}, d)$ and hence $K_n(\mathcal{C}, d)$ is cofinal in $\mathcal{C}$. Thus, in order to prove that (B) fails in $\mathcal{C}$ we only need to show that $d \in K_{n+1}^1(\mathcal{C})$. Indeed, it is sufficient to show that $d = \langle a, b, c \rangle \in \mathcal{M}_n(\mathfrak{A}, \mathcal{C}, I_{n+1}(\mathcal{C}))$, for it follows from $\mathcal{C} \models B\Sigma_n$ that $\mathcal{M}_n(\mathcal{C}, I_{n+1}(\mathcal{C})) \subseteq K_{n+1}^1(\mathcal{C})$.

First, observe that $c \in \mathcal{M}_n(\mathcal{C})$ since $c \in \mathcal{M}_n(\mathfrak{A})$, $\mathfrak{A} \models Th_{\Pi_{n+1}}(\mathfrak{A}, d)$ and $c = (d)_2$. Second, it follows from $\mathcal{C} \models B\Sigma_n$ that $\mathcal{M}_n(\mathcal{C}, I_{n+1}(\mathcal{C})) \subseteq K_{n+1}^1(\mathcal{C})$ and so $b \in I_{n+1}(\mathcal{C})$.

Finally, $a \in \mathcal{M}_n(\mathcal{C}, \mathcal{I}_{n+1}(\mathcal{C}))$ since $a \in \mathcal{M}_n(\mathfrak{A}, b)$, $\mathfrak{A} \models Th_{\Pi_{n+1}}(\mathfrak{A}, d)$ and $a = (d)_0, b = (d)_1$. Hence, $a, b, c$ are in $\mathcal{M}_n(\mathcal{C}, I_{n+1}(\mathcal{C}))$ and so is $d$. \qed

Let us observe that $I\Sigma_n$ does prove property (B). Indeed, the following question is open.

**Question 4.** Does $Th_{\Pi_{n+2}}(\Pi_n^-)$ follow from $I\Sigma_n$?

**References**


15


