THE JOURNAL OF SYMBOLIC LOGIC
Volume 78, Number 4, Dec. 2013

ON THE OPTIMALITY OF CONSERVATION RESULTS FOR LOCAL REFLECTION IN ARITHMETIC

A. CORDÓN-FRANCO, A. FERNÁNDEZ-MARGARIT, AND F. F. LARA-MARTÍN

Abstract. Let T be a recursively enumerable theory extending Elementary Arithmetic EA. L. D. Beklemishev proved that the Σ_2 local reflection principle for T, $\mathsf{Rfn}_{\Sigma_2}(T)$, is conservative over the Σ_1 local reflection principle, $\mathsf{Rfn}_{\Sigma_1}(T)$, with respect to boolean combinations of Σ_1-sentences; and asked whether this result is best possible. In this work we answer Beklemishev's question by showing that Π_2-sentences are not conserved for $T = \mathrm{EA} + $ "*f is total*," where f is any nondecreasing computable function with elementary graph. We also discuss how this result generalizes to $n > 0$ and obtain as an application that for $n > 0$, $I\Pi_{n+1}^{-}$ is conservative over $I\Sigma_n$ with respect to Π_{n+2}-sentences.

§1. Introduction. This work was motivated by a question of L. D. Beklemishev on the optimality of a conservation result for reflection principles in first order arithmetic. Reflection principles for a given theory T are axiom schemes expressing the soundness of T. More precisely, if T is a recursively enumerable (r.e.) arithmetic theory extending Elementary Arithmetic EA and $\Box_T(x) = \exists y\ \mathrm{Prf}_T(x, y)$ denotes a standard provability predicate for T, the *local reflection principle* for T is the axiom scheme given by

$$\mathsf{Rfn}(T)\colon \Box_T(\ulcorner\varphi\urcorner) \to \varphi,$$

where φ ranges over all sentences of the language of T and $\ulcorner\varphi\urcorner$ denotes (the numeral of) the Gödel number of φ. The term *local* refers to the fact that the scheme is restricted to sentences, in contrast with the *uniform reflection principle* for T, where formulas with free variables are allowed:

$$\mathsf{RFN}(T)\colon \Box_T(\ulcorner\varphi(\dot{x})\urcorner) \to \varphi(x).$$

In [3], using provability logic techniques, Beklemishev showed that over T, full local reflection $\mathsf{Rfn}(T)$ is Γ-conservative over local reflection restricted to Γ-sentences, denoted $\mathsf{Rfn}_\Gamma(T)$, for $\Gamma = \Sigma_n$ or Π_n. More precisely,

THEOREM 1.1 (Beklemishev, [3]). 1. *For $n > 1$, $T + \mathsf{Rfn}(T)$ is conservative over $T + \mathsf{Rfn}_{\Sigma_n}(T)$ with respect to Σ_n-sentences (and dually for Π_n).*
2. *$T + \mathsf{Rfn}(T)$ is conservative over $T + \mathsf{Rfn}_{\Sigma_1}(T)$ with respect to $\mathscr{B}(\Sigma_1)$ (= boolean combinations of Σ_1) sentences.*

Received September 12, 2011.
This work was supported by grant MTM2008–06435, Ministerio de Ciencia e Innovación, Spain. Co-financed with FEDER funds, EU.

0022-4812/13/7804-0001/$2.10
DOI:10.2178/jsl.7804010

In [3] Beklemishev noted that part (1) in the above theorem is best possible with respect to arithmetic complexity and in [6] he raised the problem of determining the optimality of part (2). In general, $T + \mathsf{Rfn}(T)$ is neither Π_2 nor Σ_2 conservative over $T + \mathsf{Rfn}_{\Sigma_1}(T)$, for otherwise $T + \mathsf{Rfn}_{\Pi_2}(T)$ (resp. $T + \mathsf{Rfn}_{\Sigma_2}(T)$) would collapse to $T + \mathsf{Rfn}_{\Sigma_1}(T)$ and it is known that the local reflection hierarchy is proper unless its Σ_1 level is already inconsistent. Nevertheless, it follows from part (2) that both $T + \mathsf{Rfn}_{\Pi_2}(T)$ and $T + \mathsf{Rfn}_{\Sigma_2}(T)$ are $\mathscr{B}(\Sigma_1)$-conservative over $T + \mathsf{Rfn}_{\Sigma_1}(T)$ and now there could be room for improvement. In fact, by Proposition 4.5 in [5] conservativity of $T + \mathsf{Rfn}_{\Pi_2}(T)$ over $T + \mathsf{Rfn}_{\Sigma_1}(T)$ can be extended to Σ_2-sentences if $T \subseteq \Pi_2$. In contrast, the corresponding question for $T + \mathsf{Rfn}_{\Sigma_2}(T)$ seems to have been open.

QUESTION 1 (Beklemishev). Is $T + \mathsf{Rfn}_{\Sigma_2}(T)$ conservative over $T + \mathsf{Rfn}_{\Sigma_1}(T)$ with respect to Π_2-sentences?

Typically, non Π_2-conservation can be shown by exhibiting a computable function provably total in a theory and not in the other. However, $\mathsf{Rfn}_{\Sigma_2}(T)$ has quantifier complexity Σ_2 and so $T + \mathsf{Rfn}_{\Sigma_2}(T)$ and $T + \mathsf{Rfn}_{\Sigma_1}(T)$ share the same class of provably total functions. This motivates the following problem on general arithmetic theories, of which Beklemishev's Question 1 for $T \subseteq \mathscr{B}(\Sigma_1)$ is a particular case ($\mathrm{Th}_\Gamma(S)$ denotes the Γ-consequences of a theory S):

QUESTION 2. Suppose S contains EA. Are $\mathrm{Th}_{\Pi_2}(\mathrm{Th}_{\Sigma_2}(S))$ and $\mathrm{Th}_{\mathscr{B}(\Sigma_1)}(S)$ deductively equivalent?

In this work we solve in the negative both questions for a wide class of theories and apply the proof ideas to obtain some new results on local reflection. Our methods are model-theoretic and we exploit the connections between reflection principles and induction schemes in the spirit of Kreisel and Lévy' [13].

The paper is organized as follows. Sections 1 and 2 are introductory. Section 3 presents a preliminary result that puts in context the negative answer to Question 2. In Section 4 we show that Question 1 has the negative answer for $T = \mathrm{EA} +$ "*f is total*," where f is any nondecreasing computable function with elementary graph; and, similarly, we solve in the negative Question 2 for every consistent, r.e. extension of $\mathrm{EA} + \mathsf{Rfn}_{\Sigma_2}(\mathrm{EA})$. Both results are derived from Theorem 4.4, which is an unboundness theorem of independent interest. In Section 5, we deal with results à la Kreisel–Lévy relating local reflection and various forms of induction. We improve known results for $\mathsf{Rfn}_{\Sigma_2}(T)$ and $\mathrm{Th}_{\Pi_1}(T + \mathsf{Rfn}(T))$ as well as we fill an obvious gap in our understanding of partial local reflection by obtaining a Kreisel and Lévy-like theorem for $\mathsf{Rfn}_{\Sigma_1}(T)$. Finally, in Section 6 we discuss how our results generalize to $n > 0$ and obtain a new conservation theorem for parameter free induction: $I\Pi^-_{n+1}$ is Π_{n+2}-conservative over $I\Sigma_n$.

§2. Basic notions and notation.

Our notation is standard and we assume that the reader is familiar with the basic notions of first order arithmetic (we recommend [11] and [10] for a detailed introduction to the subject; [12] for results on parameter free schemes; and [6] for information on reflection principles). To a large extent, our results are independent of the language we are working in. Nevertheless, for the sake of definiteness we assume that we work in the language $\mathscr{L}_{\exp} = \{0, 1, +, \cdot, <, \exp\}$

extending that of Peano Arithmetic PA with a symbol for the function 2^x. Also, we assume that all the theories we shall deal with are extensions of Elementary Arithmetic EA, which is axiomatized by a finite set of defining axioms for the symbols in $\mathscr{L}_{\exp}$ plus the scheme of induction for bounded (elementary) formulas of $\mathscr{L}_{\exp}$. Finally, we also assume, sometimes without explicit mention, that every theory T for which we consider reflection principles is an elementary presented extension of EA; the set of its axioms being represented by an elementary formula of the form $Ax_{\mathrm{EA}}(x) \vee Ax_T(x)$.

We recall some notation on parameter free induction schemes from [12]. If Γ is a class of formulas, the scheme $I\Gamma^-$ is given by

$$\varphi(0) \wedge \forall x\,(\varphi(x) \rightarrow \varphi(x+1)) \rightarrow \forall x\,\varphi(x),$$

where $\varphi(x) \in \Gamma^-$ (we write $\varphi(x) \in \Gamma^-$ to mean that φ is in Γ and contains no other free variables than x).

We will be also concerned with a number of inference rules associated to induction principles. Given an inference rule R and a theory T, $T + R$ denotes the closure of T under R and first order logic; while $[T, R]$ denotes the closure of T under *non-nested* applications of R and first order logic. A rule R_1 is *reducible* to R_2 if $[T, R_1] \subseteq [T, R_2]$ for every theory T extending EA; two rules R_1 and R_2 are *congruent*, written $R_1 \cong R_2$, if they are mutually reducible to each other.

Finally, if $\mathfrak{A}$ and $\mathfrak{B}$ are $\mathscr{L}_{\exp}$-structures we write $\mathfrak{A} \prec_n \mathfrak{B}$ to mean that $\mathfrak{A}$ is a Σ_n-elementary substructure of $\mathfrak{B}$, i.e., for all $\varphi(x) \in \Sigma_n$ and $a \in \mathfrak{A}$, $\mathfrak{A} \models \varphi(a)$ iff $\mathfrak{B} \models \varphi(a)$. Natural examples of Σ_n-elementary substructures are provided by submodels of Σ_n-definable elements, $\mathscr{K}_n(\mathfrak{A})$. Recall that a is a Σ_n-definable element of a model $\mathfrak{A}$ if there is $\varphi(x) \in \Sigma_n$ such that a is the unique element satisfying $\varphi(x)$ in $\mathfrak{A}$, and

- $\mathscr{K}_n(\mathfrak{A})$ denotes the set of all Σ_n-definable elements of $\mathfrak{A}$; $\mathscr{I}_n(\mathfrak{A})$ denotes the initial segment of $\mathfrak{A}$ determined by $\mathscr{K}_n(\mathfrak{A})$.
- $\mathscr{K}_n^1(\mathfrak{A}) = \mathscr{K}_n(\mathfrak{A}, \mathscr{I}_n(\mathfrak{A}))$ denotes the set of all elements of $\mathfrak{A}$ which are Σ_n-definable with a parameter from $\mathscr{I}_n(\mathfrak{A})$.

Note that for $n \geq 1$, if $\mathfrak{A} \models \mathrm{EA} + I\Sigma_{n-1}^-$ then $\mathscr{K}_n(\mathfrak{A}) \prec_n \mathfrak{A}$, see e.g., [10].

§3. **A preliminary result.** At first glance the negative answer to Question 2 might seem to be rather obvious. However, it should be compared with the following result stating that a dual version of that question has a positive answer. Thus, our work also brings into evidence the different behavior of the dual classes $\mathrm{Th}_{\Pi_2}(\mathrm{Th}_{\Sigma_2}(S))$ and $\mathrm{Th}_{\Sigma_2}(\mathrm{Th}_{\Pi_2}(S))$ for a general arithmetic theory S.

PROPOSITION 3.1. *If S implies* EA *then* $\mathrm{Th}_{\Sigma_2}(\mathrm{Th}_{\Pi_2}(S)) \equiv \mathrm{Th}_{\mathscr{B}(\Sigma_1)}(S)$.

PROOF. It is obvious that $\mathrm{Th}_{\mathscr{B}(\Sigma_1)}(S) \subseteq \mathrm{Th}_{\Sigma_2}(\mathrm{Th}_{\Pi_2}(S))$. For the opposite direction, assume that $\mathfrak{A}$ is a model of $\mathrm{Th}_{\mathscr{B}(\Sigma_1)}(S)$. Consider the theory

$$S' = \mathrm{Th}_{\Sigma_1}(\mathfrak{A}) + \mathrm{Th}_{\Pi_1}(\mathfrak{A}) + S,$$

where $\mathrm{Th}_\Gamma(\mathfrak{A})$ denotes the set of all Γ-sentences which are true in the model $\mathfrak{A}$. Clearly, S' is consistent. Let $\mathfrak{B}$ be a model of S'. Then $\mathscr{K}_1(\mathfrak{A}) \prec_1 \mathfrak{A}$ and $\mathscr{K}_1(\mathfrak{B}) \prec_1 \mathfrak{B}$; and, as a consequence, $\mathscr{K}_1(\mathfrak{B}) \models \mathrm{Th}_{\Pi_2}(S)$ and $\mathfrak{A} \models \mathrm{Th}_{\Sigma_2}(\mathscr{K}_1(\mathfrak{A}))$. But it follows from the fact that $\mathfrak{A}$ and $\mathfrak{B}$ satisfy the same Σ_1-sentences that $\mathscr{K}_1(\mathfrak{A})$ and $\mathscr{K}_1(\mathfrak{B})$

are isomorphic (if $\varphi(x)$ defines $a \in \mathfrak{A}$, map a to the unique $b \in \mathfrak{B}$ satisfying $\varphi(x)$ in $\mathfrak{B}$). Consequently, $\mathfrak{A} \models \mathrm{Th}_{\Sigma_2}(\mathrm{Th}_{\Pi_2}(S))$. ⊣

Note that the presence of the exponentiation in S is irrelevant and the result easily generalizes to $\mathrm{Th}_{\Sigma_{n+2}}(\mathrm{Th}_{\Pi_{n+2}}(S)) \equiv \mathrm{Th}_{\mathscr{B}(\Sigma_{n+1})}(S)$ if S contains $I\Sigma_n^-$.

Proposition 3.1 provides us with a recipe for lifting a $\mathscr{B}(\Sigma_1)$-conservation result to Σ_2-conservation. Namely, if T_1 is a Π_2-theory extending Δ_0-induction and T_1 is $\mathscr{B}(\Sigma_1)$-conservative over T_2, then Σ_2-sentences are also conserved. In particular, this gives us a simple proof of the known fact that the dual version of Question 1 has a positive answer for Π_2-theories, i.e., $T + \mathsf{Rfn}_{\Pi_2}(T)$ is Σ_2-conservative over $T + \mathsf{Rfn}_{\Sigma_1}(T)$ if $T \subseteq \Pi_2$.

Despite its simplicity it seems that Proposition 3.1 has not been widely known, although some cases of theories that turn out to have axiomatizations of unexpectedly low quantifier complexity have been observed in the literature. For example, in [12] it is proved that, for $n > 0$, the $\mathscr{B}(\Sigma_{n+1})$-consequences of $I\Sigma_n$ axiomatize the Σ_{n+2}-consequences of $I\Sigma_n$. Also, in [5] it is shown that over T, the Σ_2-consequences of $\mathsf{RFN}_{\Sigma_1}(T)$ (uniform reflection restricted to Σ_1 formulas) can be axiomatized by $\mathsf{Rfn}_{\Sigma_1}(T) \subseteq \mathscr{B}(\Sigma_1)$. Both results can be seen as particular cases of Proposition 3.1.

§4. An unboundedness theorem for $\mathrm{Th}_{\Pi_2}(T + \mathsf{Rfn}_{\Sigma_2}(T))$. The so-called unboundedness theorems due to Kreisel and Lévy [13] state that $\mathsf{Rfn}_{\Pi_n}(T)$ is not contained in any consistent finite extension of T of complexity Σ_n (and dually for $\mathsf{Rfn}_{\Sigma_n}(T)$). Here we obtain a variant of these results for the Π_2-consequences of $T + \mathsf{Rfn}_{\Sigma_2}(T)$. To this end, a crucial fact is that, somewhat surprisingly, $\mathsf{Rfn}_{\Sigma_2}(T)$ allows for a modicum of parametric Σ_1-induction. In [8], we proved that $I\Pi_n^-$ is equivalent to the following *local* variants of the Σ_n-induction scheme, where the conclusion of the induction axiom is relativized to definable elements:

- The scheme $I(\Sigma_n^-, \mathscr{K}_n)$ is given by

$$\varphi(0) \wedge \forall x\,(\varphi(x) \to \varphi(x+1)) \to \forall x \in \mathscr{K}_n\, \varphi(x),$$

where $\varphi(x) \in \Sigma_n^-$ and a quantifier of the form $\forall x \in \mathscr{K}_n$ in front of a formula $\Psi(x)$ is intended as a shorthand for the scheme:

$$\{\forall x\,[\mathrm{Def}_\delta(x) \to \Psi(x)] \colon \delta \in \Sigma_n\}$$

(we write $\mathrm{Def}_\delta(x,v)$ to denote $\delta(x,v) \wedge \forall x, x'\,(\delta(x,v) \wedge \delta(x',v) \to x = x')$, for a formula $\delta(x,v)$, or $\mathrm{Def}_\delta(x)$ if $\delta(x)$ does not contain any parameters).

- The scheme $I(\Sigma_n, \mathscr{I}_n, \mathscr{K}_n^1)$ is given by:

$$\forall v \in \mathscr{K}_n^1\,[\varphi(0,v) \wedge \forall x\,(\varphi(x,v) \to \varphi(x+1,v)) \to \forall x \in \mathscr{I}_n\, \varphi(x,v)],$$

where $\varphi(x,v) \in \Sigma_n$, a quantifier $\forall x \in \mathscr{I}_n\, \Psi(x)$ unravels to

$$\{\forall x, y\,[\mathrm{Def}_\delta(y) \wedge x \le y \to \Psi(x)] \colon \delta \in \Sigma_n\},$$

and a quantifier of the form $\forall x \in \mathscr{K}_n^1\, \Psi(x)$ unravels to

$$\{\forall x, y, z\,[\mathrm{Def}_{\delta_1}(z) \wedge y \le z \wedge \mathrm{Def}_{\delta_2}(x,y) \to \Psi(x)] \colon \delta_1, \delta_2 \in \Sigma_n\}.$$

The equivalence between $I\Pi_n^-$ and the above schemes was essentially proved in [8], see (the proofs of) Proposition 2.1 and Theorem 2.3 there.

PROPOSITION 4.1. *Over* $\mathrm{EA} + I\Sigma_{n-1}^-$, $I\Pi_n^- \equiv I(\Sigma_n^-, \mathscr{K}_n) \equiv I(\Sigma_n, \mathscr{I}_n, \mathscr{K}_n^1)$.

Since $T + \mathsf{Rfn}_{\Sigma_2}(T)$ contains $I\Pi_1^-$ by Theorem 1 of [4], it follows that

PROPOSITION 4.2. *$T + \mathsf{Rfn}_{\Sigma_2}(T)$ implies $I(\Sigma_1, \mathcal{I}_1, \mathcal{K}_1^1)$.*

Equipped with $I(\Sigma_1, \mathcal{I}_1, \mathcal{K}_1^1)$ a proof of the desired unboundness theorem proceeds as follows. Suppose $T = \mathrm{EA} +$ "*f is total*," where f is a computable function. Then a natural candidate for "a hardest Π_2-problem" for the theory $T + \mathsf{Rfn}_{\Sigma_2}(T)$ is: $\forall x \in \mathcal{I}_1\, \forall z \in \mathcal{K}_1^1$ "*$f^x(z)$ exists*," where f^x is the x-th iterate of f. In what follows, we shall show that this is actually the case whenever the graph of f is defined by an EA-*honest* formula:

DEFINITION 4.3. We say that a formula $y = f(x)$ is EA-*honest* if

1. $y = f(x)$ is elementary,
2. $\mathrm{EA} \vdash y = f(x) \to y \geq 2^x$, and
3. $\mathrm{EA} \vdash x_1 \leq x_2 \wedge y_1 = f(x_1) \wedge y_2 = f(x_2) \to y_1 \leq y_2$.

By Proposition 5.4 of [2] a theory T can be written as $\mathrm{EA} +$ "*f is total*," where $y = f(x)$ is EA-honest, if and only if T is a finite Π_2-extension of EA closed under the Σ_1-collection rule

$$\Sigma_1\text{-CR}: \ \frac{\forall x\, \exists y\, \varphi(x, y)}{\forall z\, \exists u\, \forall x \leq z\, \exists y \leq u\, \varphi(x, y)},$$

where $\varphi(x, y) \in \Sigma_1$. In addition, if $y = f(x)$ is an EA-honest formula then, using elementary coding of sequences, the iteration of f can be expressed by the following elementary formula (denoted by $y = f^x(z)$):

$$\exists s \leq bt\, [\mathrm{length}(s) = x + 1 \wedge (s)_0 = z \wedge \forall i < x\, ((s)_{i+1} = f((s)_i)) \wedge y = (s)_x],$$

where bt is a bounding term for the code of a sequence consisting of $(x + 1)$-many y's. We are now in a position to state the main result of this section:

THEOREM 4.4 (Unboundedness). *Suppose T is a finite Π_2-extension of* EA *closed under Σ_1-CR. Then, $\mathrm{Th}_{\Pi_2}(T + \mathsf{Rfn}_{\Sigma_2}(T))$ is not contained in any consistent, r.e. extension of T by $\mathcal{B}(\Sigma_1)$-sentences.*

PROOF. Put $T \equiv \mathrm{EA} + \forall x\, \exists! y\, (y = f(x))$, where $y = f(x)$ is EA-honest. Let Γ be an r.e. set of $\mathcal{B}(\Sigma_1)$-sentences satisfying that $T + \Gamma$ is consistent. We shall construct a model of $T + \Gamma$ in which $\mathrm{Th}_{\Pi_2}(T + \mathsf{Rfn}_{\Sigma_2}(T))$ fails. First of all, note that $T + \Gamma$ does not imply the set of all true Π_1 sentences $\mathrm{Th}_{\Pi_1}(\mathbb{N})$, for otherwise it would follow that $\mathrm{Th}_{\Pi_1}(T + \Gamma) = \mathrm{Th}_{\Pi_1}(\mathbb{N})$ and this is impossible since the first set is r.e. and the second one is Π_1^0-complete. Thus, there is $\mathfrak{A} \models T + \Gamma$ in which $\mathrm{Th}_{\Pi_1}(\mathbb{N})$ fails and so $\mathcal{K}_1(\mathfrak{A})$ is nonstandard. By considering an elementary extension of $\mathfrak{A}$ if necessary, we may also assume that $\mathcal{K}_1(\mathfrak{A})$ is bounded above in $\mathfrak{A}$. Then, it holds that $\mathcal{K}_1^1(\mathfrak{A}) \neq \mathcal{I}_1(\mathfrak{A})$, for otherwise $\mathcal{I}_1(\mathfrak{A})$ would have a proper Σ_1-elementary end extension and hence $\mathcal{I}_1(\mathfrak{A}) \models B\Sigma_2$ by Theorem B of [16]. But $I\Sigma_1 \subseteq B\Sigma_2$ and $I\Sigma_1$ is well-known to fail in $\mathcal{I}_1(\mathfrak{A})$ whenever $\mathcal{I}_1(\mathfrak{A})$ is nonstandard. It thus follows that there is $a \in \mathcal{K}_1^1(\mathfrak{A})$ such that $\mathcal{I}_1(\mathfrak{A}) < a$. Consider $\varphi(x, y, v)$ elementary and $b \in \mathcal{I}_1(\mathfrak{A})$ such that $\exists y\, \varphi(x, y, b)$ defines a in $\mathfrak{A}$. Since φ is bounded, there is a minimal c satisfying $\varphi((z)_0, (z)_1, b)$ in $\mathfrak{A}$. It is clear that $(c)_0 = a$ and so $\mathcal{I}_1(\mathfrak{A}) < a < c$. Now define $\mathfrak{B}$ to be initial segment of $\mathfrak{A}$ determined by the standard iterations $f^k(c)$, i.e.,

$$\mathfrak{B} = \{d \in \mathfrak{A} \colon \exists k \in \omega,\ \mathfrak{A} \models \exists y\, (y = f^k(c) \wedge d \leq y)\}.$$

It follows that $\mathfrak{B}$ is an initial substructure of $\mathfrak{A}$, $\mathfrak{B} \prec_0 \mathfrak{A}$, and $\mathscr{K}_1(\mathfrak{A}) \subseteq \mathfrak{B}$.

In addition, we have $\mathfrak{B} \models T + \Gamma$. Note that $\mathfrak{B} \models T$, since $y = f(x)$ defines a total function in $\mathfrak{B}$. To see $\mathfrak{B} \models \Gamma$, consider $\theta \in \Gamma$. By logical operations,

$$\theta \equiv (\theta_1^0 \vee \exists x\, \theta_2^0(x)) \wedge \cdots \wedge (\theta_1^n \vee \exists x\, \theta_2^n(x)),$$

with $\theta_1^i \in \Pi_1$ and $\theta_2^i(x) \in \Delta_0$. Fix $i \leq n$. Since $\mathfrak{A}$ satisfies Γ, $\mathfrak{A} \models \theta_1^i \vee \exists x\, \theta_2^i(x)$.

Case 1: $\mathfrak{A} \models \theta_1^i$. Then θ_1^i is also valid in $\mathfrak{B}$ since $\mathfrak{B} \prec_0 \mathfrak{A}$.

Case 2: $\mathfrak{A} \models \exists x\, \theta_2^i(x)$. Then there is $a_i \in \mathscr{K}_1(\mathfrak{A})$ such that $\mathfrak{A} \models \theta_2^i(a_i)$. But $a_i \in \mathfrak{B}$ since $a_i \leq c$. So, $\mathfrak{B} \models \exists x\, \theta_2^i(x)$.

Hence, $\mathfrak{B} \models \theta$, as required.

Finally, we prove that $\mathfrak{B} \not\models \mathrm{Th}_{\Pi_2}(T + \mathsf{Rfn}_{\Sigma_2}(T))$.

First, let us observe that

$$\mathrm{Th}_{\Pi_2}(T + \mathsf{Rfn}_{\Sigma_2}(T)) \vdash \forall x \in \mathscr{I}_1\, \forall z \in \mathscr{K}_1^1\, \exists y\, (y = f^x(z)). \tag{$\dagger$}$$

Indeed, by Proposition 4.2, $T + \mathsf{Rfn}_{\Sigma_2}(T)$ proves $\forall x \in \mathscr{I}_1\, \forall z \in \mathscr{K}_1^1\, \exists y\, (y = f^x(z))$, and a quantifier of the form $\forall x \in \mathscr{I}_1$ or $\forall z \in \mathscr{K}_1^1$ in front of a Π_2-formula unravels to a scheme of Π_2-formulas; hence, ($\dagger$) follows.

Now observe that $c \in \mathscr{K}_1^1(\mathfrak{B})$. In fact, $c = (\mu z)\,(\varphi((z)_0, (z)_1, b))$ in $\mathfrak{B}$, with $b \in \mathscr{I}_1(\mathfrak{A})$. But it follows from $\mathfrak{B} \prec_0 \mathfrak{A}$ and $\mathscr{K}_1(\mathfrak{A}) \subseteq \mathfrak{B}$ that $\mathscr{K}_1(\mathfrak{B}) = \mathscr{K}_1(\mathfrak{A})$. So, $b \in \mathscr{I}_1(\mathfrak{B})$ and hence $c \in \mathscr{K}_1^1(\mathfrak{B})$.

Pick $e \in \mathscr{K}_1(\mathfrak{A})$ nonstandard. Since $f(x) \geq 2^x \geq x$, EA proves that $y = f^x(z)$ defines a nondecreasing function in the variable x. So, $f^e(c)$ does not exist in $\mathfrak{B}$. Consequently, $\mathrm{Th}_{\Pi_2}(T + \mathsf{Rfn}_{\Sigma_2}(T))$ fails in $\mathfrak{B}$ by ($\dagger$). ⊣

We can now derive the answers to Question 1 and 2 as direct corollaries.

COROLLARY 4.5 (Answer to Question 1). *Suppose T_1 is a finite Π_2-extension of* EA *closed under Σ_1-CR, T_2 is an r.e. set of $\mathscr{B}(\Sigma_1)$-sentences and $T = T_1 + T_2$. Then $T + \mathsf{Rfn}_{\Sigma_2}(T)$ is not Π_2-conservative over $T + \mathsf{Rfn}_{\Sigma_1}(T)$ provided $T + \mathsf{Rfn}_{\Sigma_1}(T)$ is consistent.*

PROOF. It follows from Theorem 4.4 for T_1, because $T + \mathsf{Rfn}_{\Sigma_1}(T)$ is a consistent, r.e. extension of T_1 by $\mathscr{B}(\Sigma_1)$-sentences. ⊣

COROLLARY 4.6 (Answer to Question 2). *Suppose S is a consistent, r.e. extension of* $\mathrm{EA} + \mathsf{Rfn}_{\Sigma_2}(\mathrm{EA})$. *Then* $\mathrm{Th}_{\Pi_2}(\mathrm{Th}_{\Sigma_2}(S))$ *is strictly stronger than* $\mathrm{Th}_{\mathscr{B}(\Sigma_1)}(S)$.

PROOF. Since EA is closed under Σ_1-CR, by Theorem 4.4 for $T = \mathrm{EA}$,

$$\mathrm{EA} + \mathrm{Th}_{\mathscr{B}(\Sigma_1)}(S) \nvdash \mathrm{Th}_{\Pi_2}(\mathrm{EA} + \mathsf{Rfn}_{\Sigma_2}(\mathrm{EA}))$$

and so $\mathrm{Th}_{\mathscr{B}(\Sigma_1)}(S) \nvdash \mathrm{Th}_{\Pi_2}(\mathrm{Th}_{\Sigma_2}(S))$. ⊣

§5. Some results *à la* Kreisel–Lévy. In [13] Kreisel and Lévy showed that PA is equivalent to the full uniform reflection principle over primitive recursive arithmetic. D. Leivant and H. Ono [14, 15] sharpened that result by showing that $\mathrm{EA} + \mathsf{RFN}_{\Sigma_{n+1}}(\mathrm{EA}) \equiv I\Sigma_n$ for each $n > 0$. Later Beklemishev [2, 4] extended this correspondence to parameter free induction schemes and to theories described in terms of induction rules. In particular, he proved the following result over the base theory EA^+ (which is EA plus a Π_2-axiom declaring the totality of the superexponentiation function). EA^+ is needed because the proof uses formalized Cut-elimination Theorem.

THEOREM 5.1. [2, 4] *Suppose T is a finite Π_2-extension of* EA^+.

1. $T + \mathsf{Rfn}_{\Sigma_2}(T) \equiv T + I\Pi_1^-$.
2. $T + \mathrm{Th}_{\Pi_1}(T + \mathsf{Rfn}(T)) \equiv T + \Pi_1\text{-}IR \equiv T + T_\omega$.

In the above theorem T_ω denotes the so-called ω times iterated consistency assertion for T, i.e., $Con(T) + Con(T + Con(T)) + \cdots$

In this section we improve Theorem 5.1 in two ways. On the one hand, we show that it also holds over the weaker base theory EA. On the other hand, we obtain a new Kreisel and Lévy-like result for $\mathsf{Rfn}_{\Sigma_1}(T)$.

We begin by showing that EA^+ can be dropped in part 1 in the above theorem thanks to a simple but useful trick. A similar result has been obtained by A. Visser (private communication).

THEOREM 5.2. *If T is a finite Π_2-extension of* EA, $T + \mathsf{Rfn}_{\Sigma_2}(T) \equiv T + I\Pi_1^-$.

PROOF. We only have to prove that $T + I\Pi_1^-$ implies $\mathsf{Rfn}_{\Sigma_2}(T)$, as the opposite direction is already proved in [2] over EA. Suppose $\mathfrak{A} \models T + I\Pi_1^-$. It suffices to show that local Σ_2-reflection for Predicate Calculus PC is valid in $\mathfrak{A}$, for $T + \mathsf{Rfn}_{\Sigma_2}(T) \equiv T + \mathsf{Rfn}_{\Sigma_2}(\mathrm{PC})$ by the formalized Deduction theorem. Assume $\mathfrak{A} \models \exists y\ \mathrm{Prf}_{\mathrm{PC}}(\ulcorner\varphi\urcorner, y)$. Since $\mathscr{K}_1(\mathfrak{A}) \prec_1 \mathfrak{A}$, $\mathscr{K}_1(\mathfrak{A}) \models \exists y\ \mathrm{Prf}_{\mathrm{PC}}(\ulcorner\varphi\urcorner, y)$. But it follows from $\mathfrak{A} \models \mathrm{EA} + I\Pi_1^-$ and Proposition 4.1 that $\mathscr{K}_1(\mathfrak{A}) \models \mathrm{EA}^+$. Hence, Cut-elimination is available and $\mathscr{K}_1(\mathfrak{A}) \models \exists y\ \mathrm{Prf}^{\mathrm{cf}}_{\mathrm{PC}}(\ulcorner\varphi\urcorner, y)$, where $\mathrm{Prf}^{\mathrm{cf}}_{\mathrm{PC}}(x, y)$ denotes cut-free provability. So, $\mathfrak{A} \models \exists y\ \mathrm{Prf}^{\mathrm{cf}}_{\mathrm{PC}}(\ulcorner\varphi\urcorner, y)$ as well and the result follows since the proof of Theorem 1 in [4] (see also Theorem 10 in [6]) shows that $I\Pi_1^-$ implies local Σ_2-reflection for PC w.r.t. cut-free provability. ⊣

Now we turn to the promised theorem characterizing $\mathsf{Rfn}_{\Sigma_1}(T)$ by some form of induction. To this end, let Γ-IR and Γ-IR_0 denote the inference rules

$$\mathrm{IR}\colon \frac{\varphi(0) \wedge \forall x\,(\varphi(x) \to \varphi(x+1))}{\forall x\,\varphi(x)} \qquad \mathrm{IR}_0\colon \frac{\forall x\,(\varphi(x) \to \varphi(x+1))}{\varphi(0) \to \forall x\,\varphi(x)},$$

where $\varphi(x) \in \Gamma$. In his detailed analysis of inference rules in arithmetic [2] Beklemishev proved that, if parameters are allowed, IR_0 is congruent with the usual formulation of the induction rule: Σ_n-$\mathrm{IR}_0 \cong \Pi_n$-$\mathrm{IR}_0 \cong \Sigma_n$-IR. The parameter free version of IR_0 was not considered there, however. It turns out that, whereas Σ_n^--IR_0 is easily seen to be congruent with Σ_n-IR_0, Π_n-IR_0 and Π_n^--IR_0 cease to be equivalent. In fact, we have

THEOREM 5.3. *If T is a finite Π_2-extension of* EA, *then* $T + \mathsf{Rfn}_{\Sigma_1}(T) \equiv T + \Pi_1^-\text{-}IR_0 \equiv [T, \Pi_1^-\text{-}IR_0]$.

PROOF. Write S for $T + \mathsf{Rfn}_{\Sigma_1}(T)$, R_ω^- for $T + \Pi_1^-$-IR_0, and R_1^- for $[T, \Pi_1^-\text{-}\mathrm{IR}_0]$. First, observe that R_ω^- is an extension of T by $\mathscr{B}(\Sigma_1)$-sentences. Hence, $S \vdash R_\omega^-$ follows, since $T + \mathsf{Rfn}_{\Sigma_2}(T)$ implies $I\Pi_1^-$ and is $\mathscr{B}(\Sigma_1)$-conservative over S. Second, the implication $R_\omega^- \vdash R_1^-$ is trivial. Finally, we prove that $R_1^- \vdash S$. Suppose $\mathfrak{A} \models R_1^-$. It follows that $\mathscr{K}_1(\mathfrak{A}) \models [T, \Sigma_1\text{-IR}]$. To see this, let $\varphi(x) \in \Sigma_1$ and assume $T \vdash \varphi(0) \wedge \forall x\,(\varphi(x) \to \varphi(x+1))$ and $\varphi(a)$ fails with $a \in \mathscr{K}_1(\mathfrak{A})$. Let $\delta(z)$ be a Σ_1-formula defining a. Applying Π_1^--IR_0 to $\forall z\,(\delta(z) \to \neg\varphi(z - x))$ we get a contradiction and the result follows. But $[T, \Sigma_1\text{-IR}]$ implies $T + \mathsf{RFN}_{\Sigma_1}(T)$ by Theorem 2 of [2]. Hence, $\mathfrak{A} \models T + \mathrm{Th}_{\Sigma_2}(T + \mathsf{RFN}_{\Sigma_1}(T))$ and so $\mathfrak{A} \models S$. ⊣

Now we show that EA^+ can also be dropped in part 2 of Theorem 5.1. This is interesting because in [2] Beklemishev remarked that, for characterizing Π_1-IR by iterated consistency modulo EA one should replace $Con(T)$ with a weaker notion of consistency based on cut-free provability. Nonetheless, Theorem 5.4 below will show that this result is true even for the usual notion of consistency.

THEOREM 5.4. *If T is a finite Π_2-extension of* EA, $T + \mathrm{Th}_{\Pi_1}(T + \mathsf{Rfn}(T)) \equiv T + \Pi_1\text{-}IR \equiv T + T_\omega$.

PROOF. Write S for $T + \mathsf{Rfn}(T)$ and R_ω for $T + \Pi_1$-IR. First, note that $T + \mathrm{Th}_{\Pi_1}(S) \equiv T + T_\omega$ is proved over EA in Proposition 2.33 of [6] (a result that goes back to S. V. Goryachev's [9]). Second, $T + \mathrm{Th}_{\Pi_1}(S) \vdash R_\omega$ because R_ω is an extension of T by Π_1-sentences and $S \vdash I\Pi_1^-$ by Theorem 5.2. Finally, we must prove that $R_\omega \vdash \mathrm{Th}_{\Pi_1}(S)$. By Theorem 1.1, S is Π_1-conservative over $T + \mathsf{Rfn}_{\Sigma_1}(T)$ and so, by Theorem 5.3, it is sufficient to show that $[T, \Pi_1^-\text{-IR}_0]$ is Π_1-conservative over R_ω.

To this end, we use a model-theoretic method inspired by J. Avigad's [1] who, in turn, builds on some previous ideas of Visser (unpublished) and D. Zambella, [17]. The key notion is that of a Σ_1*-closed model.* We say that $\mathfrak{A}$ is Σ_1-closed with respect to (w.r.t.) a theory U if $\mathfrak{A} \models U$ and, for every $\mathfrak{B} \models U$, $\mathfrak{A} \prec_0 \mathfrak{B}$ implies $\mathfrak{A} \prec_1 \mathfrak{B}$. By a union of chain argument every model of $U \subseteq \Pi_2$ can be Σ_0-elementarily extended to a Σ_1-closed model w.r.t. U. Thus, if every Σ_1-closed model w.r.t. $U \subseteq \Pi_2$ is a model of a theory W, then W is Π_1-conservative over U (see [7] for further details and applications of this method).

Turning back to the proof of the theorem, suppose $\mathfrak{A}$ is Σ_1-closed w.r.t. R_ω. Assume $T \vdash \forall x\,(\varphi(x) \to \varphi(x+1))$, with $\varphi(x) \in \Pi_1^-$, and $\varphi(0)$ holds in $\mathfrak{A}$. It follows from the Σ_1-closedness condition that there is $\delta(a)$ in the Π_0-diagram of $\mathfrak{A}$ such that $R_\omega \vdash \delta(v) \to \varphi(0)$. Put $\varphi'(x) \equiv \exists v\,\delta(v) \to \varphi(x)$. Clearly, $\varphi'(x) \in \Pi_1^-$ and R_ω proves the antecedent of the induction axiom for φ'. Hence, $\mathfrak{A} \models \forall x\,\varphi'(x)$ and so $\mathfrak{A} \models \forall x\,\varphi(x)$ since $\mathfrak{A} \models \delta(a)$. Therefore, $\mathfrak{A} \models [T, \Pi_1^-\text{-IR}_0]$. ⊣

§6. The case $n > 0$.

It is natural to ask ourselves how our results generalize to $\mathsf{Rfn}_{\Sigma_{n+2}}(T)$ and $I\Pi_{n+1}^-$ for an arbitrary $n > 0$. First of all, it should be noticed that, in order to characterize $I\Pi_{n+1}^-$ in terms of reflection principles, one needs to consider *relativized local reflection*. For each $n > 0$, the relativized local reflection principle for T is the scheme given by

$$\mathsf{Rfn}^n_\Gamma(T)\colon [n]_T(\ulcorner\varphi\urcorner) \to \varphi,$$

where φ ranges over all sentences in Γ and $[n]_T(x)$ denotes a Σ_{n+1}-formula expressing "x is provable from $T + \mathrm{Th}_{\Pi_n}(\mathbb{N})$." That is $[n]_T(x) \equiv \Box_U(x)$, where

$$Ax_U(x) = (Ax_T(x) \vee \mathrm{True}_{\Pi_n}(x))$$

and $\mathrm{True}_{\Pi_n}(x)$ is a truth-definition for Π_n-sentences in EA (see Section 2.3 of [6] for details). It is a theorem of Beklemishev (see [4]) that over EA, $I\Pi_{n+1}^- \equiv \mathsf{Rfn}^n_{\Sigma_{n+2}}(\mathrm{EA})$ and $I\Sigma_n^- \equiv \mathsf{Rfn}^n_{\Sigma_{n+1}}(\mathrm{EA})$ for each $n > 0$. In addition, a relativization of Theorem 1.1 holds, i.e., $T + \mathsf{Rfn}^n_{\Sigma_{n+2}}(T)$ is $\mathscr{B}(\Sigma_{n+1})$-conservative over $T + \mathsf{Rfn}^n_{\Sigma_{n+1}}(T)$. Thus, a relativized version of Question 1 is in order.

QUESTION 3 ($n > 0$). Is $T + \mathsf{Rfn}^n_{\Sigma_{n+2}}(T)$ conservative over $T + \mathsf{Rfn}^n_{\Sigma_{n+1}}(T)$ with respect to Π_{n+2}-sentences?

It turns out that the ideas in the previous sections apply equally well to the case $n > 0$ for theories T extending the Σ_n-induction scheme $I\Sigma_n$ and closed under Σ_{n+1}-collection rule. Such a theory T can be reformulated as $I\Sigma_n$ + "f is total", where f is a nondecreasing function with a Π_n-graph. Thus, considering the Π_{n+2} separation property $\forall x \in \mathscr{I}_{n+1}\, \forall z \in \mathscr{K}^1_{n+1}$ "$f^x(z)$ *exists*," allows us to obtain an analog of the unboundedness Theorem 4.4 for the class $\mathrm{Th}_{\Pi_{n+2}}(T + \mathsf{Rfn}^n_{\Sigma_{n+2}}(T))$ and, in turn, a negative answer to Question 3.

It will be useful, however, to obtain an answer to Question 3 for theories T extending EA rather than extending $I\Sigma_n$. To this end we have to use a different separation property. Now the key point is the following local variant of the *finite axiom of choice* for Π_{n-1}-formulas, denoted $FAC(\Pi_{n-1}, \mathscr{K}_1, \mathscr{K}^1_{n+1})$,

$$\forall v \in \mathscr{K}^1_{n+1}\, \forall z \in \mathscr{K}_1\, (\forall x \leq z\, \exists y\, \varphi(x, y, v) \to \exists s\, \forall x \leq z\, \varphi(x, (s)_x, v)),$$

where φ is in Π_{n-1}.

It is easy to check that $FAC(\Pi_{n-1}, \mathscr{K}_1, \mathscr{K}^1_{n+1})$ can be reexpressed as a set of Π_{n+2}-sentences which are provable from $I\Pi^-_{n+1}$ by Proposition 4.1 (and hence also from $T + \mathsf{Rfn}^n_{\Sigma_{n+2}}(T)$). By using this separation property we get:

Theorem 6.1 ($n > 0$, Unboundedness). *Suppose T is a $\mathscr{B}(\Sigma_{n+1})$-extension of* EA. *Then,* $\mathrm{Th}_{\Pi_{n+2}}(\mathsf{Rfn}^n_{\Sigma_{n+2}}(T))$ *is not contained in any consistent, r.e. extension of T by $\mathscr{B}(\Sigma_{n+1})$-sentences.*

Proof. It is similar to that of Theorem 4.4 so we skip some details. Towards a contradiction, assume that there is an r.e. set of $\mathscr{B}(\Sigma_{n+1})$-sentences Γ such that $T + \Gamma$ is consistent and contains $\mathrm{Th}_{\Pi_{n+2}}(\mathsf{Rfn}^n_{\Sigma_{n+2}}(T))$. In particular, $T + \Gamma + I\Sigma_n$ is consistent, for $I\Sigma_n$ is Σ_{n+2}-conservative over $I\Sigma^-_n$ by Theorem 2.1 of [12]. Let $\mathfrak{A} \models T + \Gamma + I\Sigma_n$ with $\mathscr{K}_1(\mathfrak{A}) \neq \omega$ and $\mathscr{I}_{n+1}(\mathfrak{A}) \neq \mathfrak{A}$. Pick $a \in \mathscr{K}^1_{n+1}(\mathfrak{A}) - \mathscr{I}_{n+1}(\mathfrak{A})$. By overspill inside $\mathscr{K}^1_{n+1}(\mathfrak{A})$ there is a sequence coded by $b \leq a$ such that,

$$\mathscr{K}^1_{n+1}(\mathfrak{A}) \models \exists x, y \leq a\ \mathrm{Sat}_n(\ulcorner\theta\urcorner, x, y) \to (b)_{\ulcorner\theta\urcorner} = (\mu t)\, (\mathrm{Sat}_n(\ulcorner\theta\urcorner, (t)_0, (t)_1)),$$

for all $\theta(x, y) \in \Pi^-_n$, where Sat_n is a truth predicate for Π_n-formulas in EA and $(t)_0, (t)_1$ denote the inverse of the Cantor pairing function. Note that $\mathscr{K}_{n+1}(\mathfrak{A})$ is included in any substructure of $\mathfrak{A}$ containing b, for in models of $I\Sigma_n$ every Σ_{n+1}-definable element can be obtained as the projection of a Π_n-minimal one. Since $b \in \mathscr{K}^1_{n+1}(\mathfrak{A})$, there are $b_1 \in \mathscr{K}^1_{n+1}(\mathfrak{A})$ and $b_2 \in \mathscr{I}_{n+1}(\mathfrak{A})$ satisfying that $b_1 = (\mu t)\, (\varphi((t)_0, (t)_1, b_2))$ and $b = (b_1)_0$ for some $\varphi \in \Pi_n$. Put $c = \langle b_1, b_2 \rangle$ and define $\mathfrak{B}$ to be $\mathscr{K}_n(\mathfrak{A}, c)$. Then $\mathfrak{B} \models T + \Gamma$ and $c \in \mathscr{K}^1_{n+1}(\mathfrak{B})$.

Finally $\mathfrak{B} \not\models FAC(\Pi_{n-1}, \mathscr{K}_1, \mathscr{K}^1_{n+1})$. We argue as in Paris–Kirby' proof that the Σ_n-collection scheme fails in $\mathscr{K}_n(\mathfrak{A})$ (see Proposition 7 of [16]). Let $\mathrm{Min}_{n-1}(z, y, v)$ formalize "z is the least element satisfying the Π_{n-1}-formula y with a parameter v." If $\mathfrak{B}$ were to satisfy $FAC(\Pi_{n-1}, \mathscr{K}_1, \mathscr{K}^1_{n+1})$ then, for any nonstandard Σ_1-definable element d, there would be an element e such that

$$\mathfrak{B} \models \forall x \leq d\, \exists y < d\, [\exists z \leq e\, (\mathrm{Min}_{n-1}(z, y, c) \wedge x = (z)_0)],$$

violating the pigeon-hole principle for $\Sigma_0(\Sigma_{n-1})$-functions in models of EA $+ I\Sigma_{n-1}$. It follows that $T + \Gamma$ does not imply $FAC(\Pi_{n-1}, \mathscr{K}_1, \mathscr{K}^1_{n+1})$, which contradicts our assumption that $T + \Gamma$ contains $\mathrm{Th}_{\Pi_{n+2}}(\mathsf{Rfn}^n_{\Sigma_{n+2}}(T))$. ⊣

As a consequence, we obtain the corresponding negative answer to Question 3.

COROLLARY 6.2 ($n > 0$). *Suppose T is a $\mathscr{B}(\Sigma_{n+1})$-extension of* EA. *$T + \mathsf{Rfn}^n_{\Sigma_{n+2}}(T)$ is not Π_{n+2}-conservative over $T + \mathsf{Rfn}^n_{\Sigma_{n+1}}(T)$ provided that the latter theory is consistent.*

In addition, taking $T = \mathrm{EA}$, Theorem 6.1 gives us that

COROLLARY 6.3 ($n > 0$). *$I\Pi^-_{n+1}$ is not Π_{n+2}-conservative over $I\Sigma^-_n$.*

In [8] we gave another proof of Corollary 6.3 that uses a different separation property, namely, a formalized version of the model-theoretic property

$$(A) \equiv \forall a \in \mathscr{K}^1_{n+1} \text{ “}\mathscr{K}_n(\mathfrak{A}, a) \text{ is bounded above in } \mathfrak{A}.\text{”}$$

This approach has the small advantage that it allows one to prove that $I\Pi^-_{n+1}$ is not Π_{n+2}-conservative over $I\Sigma^-_n + B\Sigma_n$ for each $n > 0$ (see the proof of Theorem 5.5 in [8] for details). Note that both $FAC(\Pi_{n-1}, \mathscr{K}_1, \mathscr{K}^1_{n+1})$ and property (A) above are provable from $I\Sigma_n$. This is no accident, as we shall infer from the following conservation result.

Let $\mathsf{Rfn}^0_\Gamma(T)$ coincide, by definition, with its non-relativized analog $\mathsf{Rfn}_\Gamma(T)$ and let $\mathsf{Rfn}^n(T)$ denote the full relativized local reflection principle for T.

THEOREM 6.4 ($n \geq 0$).

1. *$T + \mathsf{Rfn}^n(T)$ is Π_{n+2}-conservative over $T + \mathsf{RFN}_{\Sigma_{n+1}}(T)$.*
2. *$T + I\Pi^-_{n+1}$ is Π_{n+2}-conservative over $[T, \Sigma_{n+1}\text{-}IR]$.*

PROOF. (1): Let φ be a Π_{n+2}-sentence such that $T + \mathsf{Rfn}^n(T) \vdash \varphi$. Then, by a relativized version of Theorem 1.1, $T + \mathsf{Rfn}^n_{\Pi_{n+2}}(T) \vdash \varphi$. Using formalized Deduction theorem for $n > 0$, we get $T + \mathsf{RFN}_{\Pi_{n+2}}(T) \vdash \varphi$. Finally, note that over EA, $\mathsf{RFN}_{\Pi_{n+2}}(T) \equiv \mathsf{RFN}_{\Sigma_{n+1}}(T)$ and hence the result follows.

(2): Let φ be a Π_{n+2}-sentence provable in $T + I\Pi^-_{n+1}$. Since $I\Pi^-_{n+1} \subseteq \Sigma_{n+2}$, by compactness there is a finite Π_{n+2}-axiomatized subtheory $T_0 \subseteq T$ satisfying that $T_0 + I\Pi^-_{n+1} \vdash \varphi$. Then, by a relativized version of Theorem 1.1, $T_0 + \mathsf{Rfn}^n_{\Sigma_{n+2}}(T_0) \vdash \varphi$. Then, trivially, $T_0 + \mathsf{Rfn}^n(T_0) \vdash \varphi$ and, by part (1), $T_0 + \mathsf{RFN}_{\Sigma_{n+1}}(T_0) \vdash \varphi$. But $[T_0, \Sigma_{n+1}\text{-IR}] \equiv T_0 + \mathsf{RFN}_{\Sigma_{n+1}}(T_0)$ by Theorem 4 of [2] and the result follows. ⊣

By using a model-theoretic construction Kaye, Paris and Dimitracopoulos proved that, when formulated in the usual language of arithmetic, $I\Pi^-_1$ is Π_2-conservative over $I\Delta_0 + \exp \equiv [I\Delta_0, \Sigma_1\text{-IR}]$ (see Theorem 2.9 of [12]). Thus, Theorem 6.4 can be seen as a reflection principle counterpart of Kaye–Paris–Dimitracopoulos' result. Since $[\mathrm{EA}, \Sigma_{n+1}\text{-IR}] \equiv \mathrm{EA} + \mathsf{RFN}_{\Sigma_{n+1}}(\mathrm{EA}) \equiv I\Sigma_n$, taking $T = \mathrm{EA}$ in Theorem 6.4 yields the following new conservation result for induction schemes.

COROLLARY 6.5 ($n > 0$). *$I\Pi^-_{n+1}$ is Π_{n+2}-conservative over $I\Sigma_n$.*

Corollary 6.5 explains why any Π_{n+2}-property separating $I\Pi^-_{n+1}$ and $I\Sigma^-_n$ turns out to be provable from $I\Sigma_n$ as well as, combined with Corollary 6.3, settles the question of the optimality of conservativity between $I\Pi^-_{n+1}$ and $I\Sigma^-_n$ for $n > 0$. On the one hand, $I\Pi^-_{n+1}$ is $\mathscr{B}(\Sigma_{n+1})$-conservative over $I\Sigma^-_n$ and this result is best possible; on the other hand, conservativity of $I\Pi^-_{n+1}$ over $I\Sigma_n$ can be extended to Π_{n+2}-sentences and this is, again, best possible.

Acknowledgements. The authors are grateful to the anonymous referees for their helpful suggestions for improving the presentation of the paper.

REFERENCES

[1] J. AVIGAD, *Saturated models of universal theories*, ***Annals of Pure and Applied Logic***, vol. 118 (2002), pp. 219–234.

[2] L. D. BEKLEMISHEV, *Induction rules, reflection principles, and provably recursive functions*, ***Annals of Pure and Applied Logic***, vol. 85 (1997), pp. 193–242.

[3] ———, *Notes on local reflection principles*, ***Theoria***, vol. 63 (1997), pp. 139–146.

[4] ———, *Parameter free induction and provably total computable functions*, ***Theoretical Computer Science***, vol. 224 (1999), pp. 13–33.

[5] ———, *Proof-theoretic analysis by iterated reflection*, ***Archive for Mathematical Logic***, vol. 42 (2005), pp. 515–552.

[6] ———, *Reflection principles and provability algebras in formal arithmetic*, ***Russian Mathematical Surveys***, vol. 60 (2005), pp. 197–268.

[7] A. CORDÓN-FRANCO, A. FERNÁNDEZ-MARGARIT, and F. F. LARA-MARTÍN, *On conservation results for parameter–free π_n–induction*, ***Studies in weak arithmetics*** (P. Cégielski, editor), CSLI Lecture Notes, vol. 196, The University of Chicago Press, 2010, pp. 49–97.

[8] ———, *A note on parameter free π_1–induction and restricted exponentiation*, ***Mathematical Logic Quarterly***, vol. 57 (2011), pp. 444–455.

[9] S. V. GORYACHEV, *On the interpretability of some extensions of arithmetic*, ***Rossiĭskaya Akademiya Nauk. Matematicheskie Zametki***, vol. 40 (1986), pp. 561–571.

[10] P. HÁJEK and P. PUDLÁK, ***Metamathematics offirst–order arithmetic***, Perpectives in Mathematical Logic, Springer–Verlag, 1993.

[11] R. KAYE, ***Models of Peano Arithmetic***, Oxford Logic Guides, vol. 15, Oxford University Press, 1991.

[12] R. KAYE, J. PARIS, and C. DIMITRACOPOULOS, *On parameter free induction schemas*, this JOURNAL, vol. 53 (1988), pp. 1082–1097.

[13] H. KREISEL and A. LÉVY, *Reflection principles and their use for establishing the complexity of axiomatic systems*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 14 (1968), pp. 97–142.

[14] D. LEIVANT, *The optimality of induction as an axiomatization of arithmetic*, this JOURNAL, vol. 48 (1983), pp. 182–184.

[15] H. ONO, *Reflection principles in fragments of peano arithmetic*, ***Zeitschrift für mathematische Logik und Grundlagen der Mathematik***, vol. 33 (1987), pp. 317–333.

[16] J. PARIS and L. KIRBY, *Σ_n-collection schemas in arithmetic*, ***Logic colloquium** 77* (A. Macintyre, L. Pacholski, and J. Paris, editors), North–Holland, 1978, pp. 285–296.

[17] D. ZAMBELLA, *Notes on polynomial bounded arithmetic*, this JOURNAL, vol. 61 (1996), pp. 942–966.

DEPARTAMENTO CIENCIAS DE LA COMPUTACIÓN E INTELIGENCIA ARTIFICIAL
FACULTAD DE MATEMÁTICAS. UNIVERSIDAD DE SEVILLA
C/ TARFIA, S/N, 41012 SEVILLA, SPAIN
E-mail: acordon@us.es
E-mail: fflara@us.es