On Strong Fragments of Peano Arithmetic

Vítězslav Švejdar
Charles University in Prague

Appeared in P. Arazim and M. Dančák eds., The Logica Yearbook 2014, pp. 281–291, College Publications, London, 2015.

Abstract

We analyze the classical proof of Paris and Kirby, showing that Σ_{n+1}-collection is not provable using Σ_n-induction. We then mention another principle that is also violated in the model of Paris and Kirby and that might be weaker than Σ_{n+1}-collection: there is no Σ_{n+1} definable bounded one-to-one function.

1 Robinson and Peano arithmetics, induction

Robinson arithmetic Q is a weak axiomatic theory of natural numbers. Its language contains the symbols $+$ and $\cdot$ for addition and multiplication of natural numbers, the symbol S for the successor function (adding the number one), and the constant 0. The original version of Robinson arithmetic, as defined in [TMR53], has seven simple axioms Q1–Q7, where for example the axioms Q4 and Q5 are $\forall x(x + 0 = x)$ and $\forall x \forall y(x + \mathrm{S}(y) = \mathrm{S}(x + y))$. Because of the bounded quantifiers (to be discussed soon), it nowadays seems more practical to add the symbols $\leq$ and $<$ for unstrict and strict ordering to the language of Q, and formulate two additional axioms about these symbols:

Q8: $\quad \forall x \forall y(x \leq y \equiv \exists v(v + x = y))$,

Q9: $\quad \forall x \forall y(x < y \equiv \exists v(\mathrm{S}(v) + x = y))$.

The resulting version of Q, with nine axioms, is conservative over the original version defined in [TMR53].

Despite the fact that the two sample axioms Q4 and Q5 mentioned above look like a recursive definition of uniquely defined addition, expected properties of addition, like associativity and commutativity, are not provable in Q. The same is true about multiplication. Expected properties of operations can be proved in *Peano arithmetic* PA, which is a theory obtained from Q by adding the *induction schema*: the sentence

$$\text{Ind:} \qquad \forall \underline{y}(\varphi(0,\underline{y}) \mathbin{\&} \forall x(\varphi(x,\underline{y}) \to \varphi(\mathrm{S}(x),\underline{y})) \to \forall x \varphi(x,\underline{y})),$$

where $\underline{y}$ stands for $y_1,\ldots,y_k$, is an axiom for any choice of a formula φ and a variable x. The variable x can be called *induction variable*, or *variable of induction*, the remaining variables $y_1,\ldots,y_k$ are *parameters*. It is understood that all free variables of φ are among $x, y_1,\ldots,y_k$. When using Ind to show that $\forall x \varphi(x,\underline{y})$, we say that $\forall x \varphi(x,\underline{y})$ is proved *by induction on x for (fixed) parameters $y_1,\ldots,y_k$*. As an example or an exercise, the reader may want to write down proofs of the sentences $\forall x(0+x=x)$, $\forall x \forall y(\mathrm{S}(x)+y=\mathrm{S}(x+y))$ and $\forall x \forall y(x+y=y+x)$. Peano arithmetic is incomplete because Gödel incompleteness theorems apply to it. However, it is a strong theory: natural examples of independent sentences are difficult to find. Peano arithmetic can also prove the schema

$$\text{B:} \qquad \forall \underline{y} \forall x(\forall u{<}x\, \exists v \varphi(u,v,\underline{y}) \to \exists z \forall u{<}x\, \exists v{<}z\, \varphi(u,v,\underline{y})),$$

where again $y_1,\ldots,y_k$ are parameters. This schema is called *collection schema*, or *boundedness schema*. It prevents the existence of an arithmetically definable relation that is *unbounded on some x*, where a relation unbounded on x is an "impossible" relation whose characteristic function is such as that in Fig. 1: for every bound z there are columns $u < x$ that contain nothing but zeros in all lines $v < z$, but, what might not be apparent in Fig. 1, every column $u < x$ contains some ones.

2 Hierarchy of formulas and bounded induction

To prove the expected properties of operations, like the commutativity of addition mentioned above, it is often sufficient to use the induction schema for very simple or even open (quantifier free) formula φ. The study of fragments of PA is motivated by the following question: in which situations, if ever, does one need induction (or collection) for a complicated φ, with alternating quantifiers?

Let $\forall v{<}z\, \varphi$ and $\exists v{<}z\, \varphi$ stand for $\forall v(v<z \to \varphi)$ and $\exists v(v<z \mathbin{\&} \varphi)$ respectively, and similarly for $\forall v{\leq}z\, \varphi$ and $\exists v{\leq}z\, \varphi$. The expressions $\forall v{<}z$, $\exists v{<}z$, $\forall v{\leq}z$, and $\forall v{\leq}z$, where v and z are different variables, are called *bounded quantifiers*. A formula is a *bounded formula*, or a *Δ_0-formula*, if all quantifiers in it are bounded. A formula is Σ_n if it has the form $\exists v_1 \forall v_2 \exists \ldots v_n \varphi$ where φ is bounded. Thus a *Σ_n-formula* has a prefix of n alternating quantifiers the first of which is existential, followed by a Δ_0-formula. Symmetrically, a *Π_n-formula* has the form $\forall v_1 \exists v_2 \forall \ldots v_n \varphi$ where φ is Δ_0-formula.

The class of all relations definable by Δ_0-formulas in the structure $\mathbb{N}$ of natural numbers is an easily defined (proper) subclass of the class of all recursive

$$
\begin{array}{|cccccccccccccc|c}
\vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \\
0&0&0&0&0&0&0&0&0&0&0&0&1&0 & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&0&0&0&1&0&0 & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&1&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&1&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&1&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&1&0&0&0&0&0 & \\
0&0&0&0&0&0&0&0&0&0&0&0&0&0 & \\
\hline
\multicolumn{1}{c}{0} & \multicolumn{2}{c}{\ldots} & & & & & \ldots & & & & & \multicolumn{2}{c}{\ldots} & x
\end{array}
$$

Figure 1: An unbounded relation on x

relations. It is quite inclusive in the sense that all r.e. relations can be obtained from Δ_0 relations by the projection operation (i.e. by existential quantification). It is also inclusive in the sense that it contains many important sets like the divisibility relation, the set of all primes, and the set of all powers of two. It also contains the graph of the function $n \mapsto 2^n$; that is, there is a Δ_0-formula $\varepsilon(x, y)$ such that $m = 2^n$ if and only if $\mathbb{N} \models \varepsilon(\overline{n}, \overline{m})$, where $\overline{n}$ and $\overline{m}$ are the numerals for n and m, i.e. the closed terms $\mathrm{S}(\mathrm{S}(\ldots(0)\ldots))$, having the corresponding number of occurrences of the symbol S. This fact has a not so trivial proof, see [Pud83] or [Ben62].

$\mathsf{I}\Delta_0$ is a theory like Peano arithmetic, but with the induction schema restricted to Δ_0-formulas. The *theory* $\mathsf{I}\Delta_0$ is also called *bounded arithmetic*, but this name is sometimes also used for similar theories that can be even weaker and that can have a somewhat different language.

In $\mathsf{I}\Delta_0$, i.e. using induction for Δ_0-formulas only, one can prove various facts about addition and multiplication: both operations are associative and commutative, multiplication is distributive over addition, for each pair x and y there is a unique z such that $z + x = y$ or $z + y = x$. It is also possible to develop coding of finite sets and sequences. Thus writing $x \in w$ has a good sense in $\mathsf{I}\Delta_0$. We also have

$$\mathsf{I}\Delta_0 \vdash \varepsilon(0, \overline{1}) \ \& \ \forall x \forall y \forall z (\varepsilon(x, y) \rightarrow \varepsilon(x + \overline{1}, z) \equiv z = \overline{2} \cdot y) \tag{1}$$

where ε is the formula mentioned above, that defines the graph of the exponen-

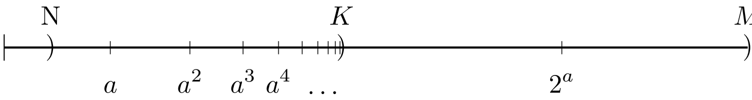

Figure 2: A simple construction of a model of $\mathsf{I}\Delta_0$

tiation function. Also, one can prove in $\mathsf{I}\Delta_0$ that every two formulas ε satisfying (1) are equivalent; thus exponentiation with base 2 has a unique meaning in $\mathsf{I}\Delta_0$. However, $\forall x \exists y \varepsilon(x,y)$ is *not* provable in $\mathsf{I}\Delta_0$. That is, the function $x \mapsto 2^x$ is a uniquely defined but possibly partial function. Let Exp be the sentence $\forall x \exists y \varepsilon(x,y)$, and let $\mathsf{I}\Delta_0{+}\mathsf{Exp}$ be $\mathsf{I}\Delta_0$ enhanced with the sentence Exp as an additional axiom. We take $\mathsf{I}\Delta_0{+}\mathsf{Exp}$ as the *base theory*, although we know that this choice is somewhat arbitrary. For some purposes, like coding of logical syntax, a weaker axiom would be sufficient, while in some other situations a stronger additional axiom would be needed. In $\mathsf{I}\Delta_0{+}\mathsf{Exp}$, coding of finite sets and sequences works as expected; in particular, there are sets and sequences of arbitrary (finite) lengths. The structure $\mathbb{N}$ of natural numbers is a model (the *standard model*) of the theories PA, $\mathsf{I}\Delta_0$, and $\mathsf{I}\Delta_0{+}\mathsf{Exp}$. It is also a model of all the remaining theories we will consider. However, one should not forget that all these theories have nonstandard models as well. So for example, when we speak, inside a theory, about the length of a (finite) sequence, it is good to see a nonstandard element of some model behind that length. We also have Δ_0-comprehension in $\mathsf{I}\Delta_0{+}\mathsf{Exp}$: for each Δ_0-formula $\varphi(v)$, possibly with parameters, $\mathsf{I}\Delta_0{+}\mathsf{Exp} \vdash \forall z \exists w \forall v(v \in w \equiv v < z \;\&\; \varphi(v))$. That is, for each z there exists a number w that codes the set $\{\, v\,;\, v < z \;\&\; \varphi(v)\,\}$.

Consider a model $\mathcal{M} \models \mathsf{PA}$ like in Fig. 2, with domain M, the standard part N and a nonstandard element a. Put $K = \{\, b \in M\,;\, \exists m{\in}\mathrm{N}(b < a^m)\,\}$. The set K is evidently closed under $+$ and $\cdot$. Thus the structure $\mathcal{K}$ with domain K and operations induced from $\mathcal{M}$ is a substructure of $\mathcal{M}$. Induction on complexity of $\varphi(x_1, .., x_k) \in \Delta_0$ shows that if $b_1, .., b_k \in K$, then $\mathcal{K} \models \varphi(\underline{b}) \Leftrightarrow \mathcal{M} \models \varphi(\underline{b})$. From this, absoluteness of Δ_0-formulas, it follows that $\mathcal{K} \models \mathsf{I}\Delta_0$. For each standard m, there exists a standard k such that $2^x > x^m$ for all $x > k$. For example, if $m = 3$, then $2^x > x^m$ for all $x > 9$. If $m = 4$, then $2^x > x^m$ for all $x > 16$. Since a is nonstandard, all elements a^2, a^3, a^4, ... are less than 2^a. It follows that $2^a \notin K$. From this we have that 2^a does not exist in $\mathcal{K}$, i.e. $\mathcal{K} \not\models \exists y \varepsilon(a,y)$. Thus we have $\mathcal{K} \not\models \mathsf{Exp}$.

3 The hierarchy of strong fragments of PA

Let $\mathrm{I}\Gamma$, for $\Gamma = \Sigma_n$ or $\Gamma = \Pi_n$, be the theory obtained from $\mathsf{I}\Delta_0$ by adding the schema $\mathrm{Ind}(\Gamma)$, the induction schema restricted to Γ-formulas. Thus $\mathsf{I}\Sigma_0$ and $\mathsf{I}\Pi_0$

are the same theories as $\mathsf{I}\Delta_0$. For the formula $\varepsilon(x,y)$ defining the graph of the function $x \mapsto 2^x$, one can show by induction on x that $\forall x \exists y \varepsilon(x,y)$. Since $\exists y \varepsilon(x,y)$ is a Σ_1-formula, the sentence $\forall x \exists y \varepsilon(x,y)$, i.e. the sentence Exp, is an example of a sentence provable in $\mathsf{I}\Sigma_1$ but not in $\mathsf{I}\Delta_0$. Using dummy quantifiers, it is easy to show that a Σ_n- or a Π_n-formula is (already in predicate logic, and thus in all theories considered here), equivalent to a Σ_m-formula for each $m > n$, and also to a Π_m-formula for each $m > n$. Thus the sentence Exp, a useful axiom possibly added to $\mathsf{I}\Delta_0$, brings nothing new if added to $\mathsf{I}\Sigma_n$ for $n > 0$ or to $\mathsf{I}\Pi_n$ for $n > 1$.

Besides induction and collection, we can consider yet another axiom schema, the *least number principle*

LNP: $\quad \forall \underline{y}(\exists x \varphi(x,\underline{y}) \rightarrow \exists x(\varphi(x,\underline{y}) \,\&\, \forall v{<}x\, \neg\varphi(v,\underline{y})))$.

Similarly as in the case of the induction schema, let LNP(Γ) and B(Γ) be the restriction of the schema LNP or B to Γ-formulas, where again Γ is Σ_n or Π_n. Let $\mathsf{L}\Gamma$ and $\mathsf{B}\Gamma$ be the theories obtained from $\mathsf{I}\Delta_0$ by adding the schema LNP(Γ) or B(Γ). We will see that $\mathsf{B}\Pi_0$ is not a subtheory of $\mathsf{I}\Delta_0{+}\mathsf{Exp}$. However, it is not difficult to check that the model $\mathcal{K}$ from Fig. 2 is a model of $\mathsf{B}\Pi_0$.

Assume that $\varphi(u,v,s)$ is a Π_n-formula, with possible parameters that are not shown. Assume further that $\exists s \varphi(u,v,s)$ defines an unbounded relation on x, i.e. that $\forall u{<}x\, \exists v \exists s \varphi(u,v,s)$, but $\neg\exists z \forall u{<}x\, \exists v{<}z\, \exists s \varphi(u,v,s)$. Then we have $\forall u{<}x\, \exists w \exists v{<}w\, \exists s{<}w\, \varphi(u,v,s)$. This is so because if $\exists v \exists s \varphi(u,v,s)$, then one can pick an arbitrary $w > \max\{v,s\}$ to obtain $\exists v{<}w\, \exists s{<}w\, \varphi(u,v,s)$. Furthermore,

$$\neg\exists z \forall u{<}x\, \exists w{<}z\, \exists v{<}w\, \exists s{<}w\, \varphi(u,v,s).$$

Indeed, from $\forall u{<}x\, \exists w{<}z\, \exists v{<}w\, \exists s{<}w\, \varphi(u,v,s)$ we can evidently conclude that $\forall u{<}x\, \exists v{<}z\, \exists s \varphi(u,v,s)$, which would contradict our assumption. Therefore, if $\exists s \varphi(u,v,s)$ defines an unbounded relation on x, then $\exists v{<}w\, \exists s{<}w\, \varphi(u,v,s)$ defines an unbounded relation (in x and w) on the same x. The whole argument shows that B(Π_n) implies B(Σ_{n+1}), provided $\exists v{<}w\, \exists s{<}w\, \varphi(u,v,s)$ is a Π_n-formula. Which it is, as it follows from (a) of the following theorem.

Theorem 1 *The theory $\mathsf{B}\Pi_n$ proves the following.*
(a) Σ_{n+1}- and Π_{n+1}-formulas are closed under bounded quantification. That is, a formula obtained from Σ_{n+1}- or Π_{n+1}-formula by bounded quantification is in $\mathsf{B}\Pi_n$ equivalent to some Σ_{n+1}- or Π_{n+1}-formula respectively.
(b) Σ_{n+2}-formulas are closed under $\exists$, Π_{n+2}-formulas are closed under $\forall$.
(c) Both Σ_{n+2}- and Π_{n+2}-formulas are closed under conjunctions and disjunctions.

Σ_0- and Π_0-formulas are trivially closed under bounded quantification and Boolean connectives—recall that $\Sigma_0 = \Pi_0 = \Delta_0$. It is also true that Σ_1-formulas are closed under $\exists$ and Π_1-formulas are closed under $\forall$, and that both Σ_1- and Π_1-formulas are closed under conjunctions and disjunctions. These facts are not listed in the previous theorem because their proofs do not require the collection schema (do not require anything above $\mathsf{I}\Delta_0$).

Theorem 2 *(a)* $\mathsf{B}\Sigma_{n+1}$ *and* $\mathsf{B}\Pi_n$ *are equivalent theories.*
(b) All $\mathsf{I}\Sigma_n$, $\mathsf{I}\Pi_n$, $\mathsf{L}\Sigma_n$, *and* $\mathsf{L}\Pi_n$ *are equivalent theories.*
(c) $\mathsf{I}\Sigma_{n+1}$ *proves* $\mathsf{B}\Pi_n$, *and* $\mathsf{B}\Pi_n$ *proves* $\mathsf{I}\Sigma_n$.

We already gave the proof of (a). The remaining proofs are similarly natural. Theorems 1 and 2 contain the basic relationships between strong fragments of PA and are well known. Details can be found in [PK78] and [HP93]. With Theorem 2 in mind, we have the following hierarchy of theories:

$$
\begin{array}{ccccccccccc}
\mathsf{I}\Delta_0+\mathsf{Exp} & \Leftarrow & \mathsf{B}\Pi_0+\mathsf{Exp} & \Leftarrow & \mathsf{I}\Sigma_1 & \Leftarrow & \mathsf{B}\Pi_1 & \Leftarrow & \dots & \Leftarrow & \mathsf{PA} \\
 & & \Updownarrow & & \Updownarrow & & \Updownarrow & & & & \\
 & & \mathsf{B}\Sigma_1+\mathsf{Exp} & & \mathsf{I}\Pi_1 & & \mathsf{B}\Sigma_2 & & & & \\
 & & & & \Updownarrow & & & & & & \\
 & & & & \mathsf{L}\Sigma_1 & & & & & & \\
 & & & & \Updownarrow & & & & & & \\
 & & & & \mathsf{L}\Pi_1 & & & & & &
\end{array}
$$

A formula is $\Sigma_0(\Sigma_n)$ if it is obtained from Σ_n formulas using Boolean connectives and bounded quantification. It is known that $\mathsf{I}\Sigma_n$ for $n \geq 1$ proves $\Sigma_0(\Sigma_n)$-comprehension: for each $\Sigma_0(\Sigma_n)$-formula $\varphi(v)$ the theory $\mathsf{I}\Sigma_n$ proves that for each z there exists a number w that codes the set $\{\, v\,;\, v < z \;\&\; \varphi(v)\,\}$. It is clear that $\Sigma_0(\Delta_0) = \Delta_0$. We have already noted that $\mathsf{I}\Delta_0+\mathsf{Exp}$ proves Δ_0-comprehension.

A *universal* Σ_{n+1}*-formula* is a formula $\exists v\gamma(z, x, v)$ such that $\gamma \in \Pi_n$ and for any Σ_{n+1}-formula $\psi(x)$ there exists a number $e \in \mathrm{N}$ such that $\psi(x)$ is in $\mathsf{I}\Sigma_1$ equivalent to $\exists v\gamma(\overline{e}, x, v)$. Universal Σ_{n+1}-formulas exist for each $n \geq 1$. Using universal Σ_{n+1}-formulas, it is possible to show that all theories $\mathsf{I}\Sigma_n$ and $\mathsf{B}\Pi_n$ for $n \geq 1$ are finitely axiomatizable.

4 The hierarchy of fragments does not collapse

We now rephrase the proof of [PK78] that $\mathsf{B}\Pi_n$ is strictly stronger than $\mathsf{I}\Sigma_n$. We will use the notion of sparse relation. We say that a *relation defined by a formula* $\theta(u, v)$ *is sparse with respect to* t if $\forall u \exists v \theta(u, v)$ but, for each s, the set $\{\, [u, v]\,;\, u < s \;\&\; v < s \;\&\; \theta(u, v)\,\}$ has less than t elements. In other words, if a relation is sparse, then every square $s \times s$ contains less than t elements (pairs) of that relation. The definition of sparse relation only makes sense in the presence of comprehension for the formula θ: only then the number of pairs in a square $s \times s$ is uniquely determined. Thus let the notion of sparse relation be made more precise as follows: in $\mathsf{I}\Delta_0{+}\mathsf{Exp}$, we speak about sparse relations defined by formulas $\theta \in \Delta_0$, while in $\mathsf{I}\Sigma_n$ for $n \geq 1$ we speak about sparse relations defined by formulas $\theta \in \Sigma_0(\Sigma_n)$. The following theorem says that a sparse relation is something even more weird than a relation unbounded on x. The lemma is easily proved using Theorem 2(a): a $\Sigma_0(\Sigma_n)$-formula is Σ_{n+1} in $\mathsf{I}\Sigma_n$.

Theorem 3 *In $\mathsf{I}\Sigma_n{+}\mathsf{Exp}$, if a relation defined by $\theta \in \Sigma_0(\Sigma_n)$ is sparse with respect to t, then it is unbounded on any $x \geq t$. Thus $\mathsf{B}\Pi_n{+}\mathsf{Exp}$ proves that there is no $\Sigma_0(\Sigma_n)$ sparse relation with respect to any t.*

Let n be given and let $\mathcal{M}$ be a model of PA containing nonstandard Σ_{n+1}-definable elements, where an element $a \in M$ is Σ_{n+1}-definable if it is the only element of $\mathcal{M}$ satisfying certain Σ_{n+1}-formula. To obtain $\mathcal{M}$, we can use Gödel incompleteness theorems: a witness for a false Σ_1-sentence (if 1st Gödel incompleteness theorem us used) or the least proof of contradiction in formalized PA (if 2nd incompleteness theorem us used) is Δ_0-definable and nonstandard. Thus we can begin with the same $\mathcal{M}$ regardless of the number $n \in \mathrm{N}$. Let $K = \{\, a \in M \,;\, a \text{ is } \Sigma_{n+1}\text{-definable in } \mathcal{M} \,\}$.

First, we claim that K contains all standard elements of $\mathcal{M}$, also some nonstandard elements, and is closed under addition and multiplication. Indeed, if $\psi_1(x) \in \Sigma_{n+1}$ defines a and $\psi_2(x) \in \Sigma_{n+1}$ defines b, then the formula

$$\exists u \exists v(\psi_1(u) \;\&\; \psi_2(v) \;\&\; x = u + v) \tag{2}$$

is satisfied by exactly one element of the model $\mathcal{M}$, the sum $a + b$, and so it defines the element $a + b$. The argument for $a \cdot b$ is analogical. The formula (2) is Σ_{n+1} because $\mathcal{M}$ is a model of full PA (in which Σ_{n+1}-formulas are closed under $\exists$ and &). Thus $\mathcal{K}$, i.e. the set K with operations inherited from $\mathcal{M}$, is a substructure of $\mathcal{M}$. A difference to the model in Fig. 2 is that $\mathcal{K}$ might be not downwards closed.

Second, we claim that if $\varphi(x, y_1, .\,., y_k)$ is a Σ_{n+1}-formula, $a_1, .\,., a_k$ elements of K and $\mathcal{M} \models \exists x \varphi(x, \underline{a})$, then there exists a $b \in K$ such that $\mathcal{M} \models \varphi(b, \underline{a})$. Indeed, let $\varphi(x, \underline{y})$ be $\exists v \theta(v, x, \underline{y})$ where $\theta \in \Pi_n$, and let $\psi_1(x), .\,., \psi_k(x)$ be Σ_{n+1}-formulas that define the elements $a_1, .\,., a_k$ respectively. Let $[v, x] \mapsto (v, x)$ be the pairing function, i.e. the function defined as $(v, x) = \frac{1}{2}(v{+}x)(v{+}x{+}\overline{1}){+}v$, and let l and r be its inverse functions, i.e. the functions satisfying the equations $(\mathrm{l}(w), \mathrm{r}(w)) = w$, $\mathrm{l}((v, x)) = v$, and $\mathrm{r}((v, x)) = x$ for every w, v, and x. Then

$$\begin{aligned} \exists u_1 .\,. \exists u_k \exists w(\psi_1(u_1) \;\&\; .\,. \;\&\; \psi_k(u_k) \;\&\; \theta(\mathrm{l}(w), \mathrm{r}(w), \underline{u}) \;\&\; \\ \&\; \forall z{<}w \neg\theta(\mathrm{l}(z), \mathrm{r}(z), \underline{u}) \;\&\; x = \mathrm{l}(w)) \end{aligned}$$

is a Σ_{n+1} formula satisfied by exactly one element of $\mathcal{M}$, the left part $\mathrm{l}(w)$ of the least w satisfying $\theta(\mathrm{l}(w), \mathrm{r}(w), \underline{u})$ with respect to the uniquely defined elements $a_1, .., a_k$. Recall that the trick of working with the left part of the least w satisfying $\theta(\mathrm{l}(w), \mathrm{r}(w), \underline{u})$ rather than the least x satisfying $\exists v \theta(v, x, \underline{u})$ is taken from recursion theory. The least x satisfying $\exists v \theta(v, x, \underline{u})$ is not Σ_{n+1}-definable.

Third, a usual argument from model theory shows that Σ_{n+1}-formulas are absolute between $\mathcal{K}$ and $\mathcal{M}$: if $\varphi(\underline{y}) \in \Sigma_{n+1}$ and $a_1, .., a_k \in K$, then $\mathcal{K} \models \varphi(\underline{a})$ if and only if $\mathcal{M} \models \varphi(\underline{a})$. We can describe this by saying that $\mathcal{K}$ is a Σ_{n+1}-*elementary substructure of* $\mathcal{M}$ and denote it as $\mathcal{K} \preceq_{n+1} \mathcal{M}$.

Fourth, from $\mathcal{K} \preceq_{n+1} \mathcal{M}$ it follows that any Π_{n+2} sentence valid in $\mathcal{M}$ is also valid in $\mathcal{K}$. Since all axioms of $\mathsf{I}\Sigma_n$ are Π_{n+2}, we have $\mathcal{K} \models \mathsf{I}\Sigma_n$.

Fifth, it remains to show what is violated in $\mathcal{K}$. Let $\exists v \gamma(z, x, v)$ where $\gamma \in \Pi_n$ be a universal Σ_{n+1}-formula. Since every $a \in K$ is Σ_{n+1}-definable in $\mathcal{M}$ and every Σ_{n+1} formula $\psi(x)$ is equivalent to $\exists v \gamma(\overline{e}, x, v)$ for some $e \in \mathrm{N}$, we see that for every $a \in K$ there exists an $e \in \mathrm{N}$ such that a is the only element x of $\mathcal{M}$ satisfying $\exists v \gamma(\overline{e}, x, v)$ in $\mathcal{M}$. Let a nonstandard $t \in K$ be fixed. Since Σ_{n+1}-formulas are absolute between $\mathcal{K}$ and $\mathcal{M}$, we have

$$\mathcal{K} \models \forall x \exists z{<}t \forall u (\exists v \gamma(z, u, v) \equiv u = x), \tag{3}$$

which says that, in $\mathcal{K}$, every number x is the only number u (in the entire universe) satisfying $\exists v \gamma(z, u, v)$ for a suitable $z < t$. Work in $\mathcal{K}$ and consider the following relation between z, x, and w:

$$\begin{aligned} x \leq w \ \& \ \forall w'{<}w \forall x'{<}w \neg\gamma(z, x', w') \ \& \\ \& \ \forall x'{\leq}w (x' \neq x \rightarrow \neg\gamma(z, x', w)) \ \& \\ \& \ \exists w'{\leq}w \, \gamma(z, x, w'). \end{aligned} \tag{4}$$

Let the formula (4) be denoted $\delta(z, x, w)$. To understand its meaning, think of z as fixed, imagine a characteristic function of a binary relation like in Fig. 1, with columns x or x' and lines w', and think of w as a size of a square. Then the three lines in (4) say that (i) the square $(w + 1) \times (w + 1)$ contains zeros everywhere except possibly in the right and upper borders, i.e. in column w or in line w, (ii) there are zeros in all columns $x' \neq x$ (even in line w), but (iii), there are some ones in column x. Put otherwise, either $x = w$ and the column x is the only column containing some ones, or $x < w$ and the coordinates of the only one in the entire square are $[x, w]$. Evidently, there exists at most one pair $[x, w]$ such that $\delta(z, x, w)$. For such a pair it is the case that $x \leq w$. Consider now all pairs $[x, w]$ such that $\exists z{<}t \, \delta(z, x, w)$, now with t fixed as said above. Then $\exists z{<}t \, \delta(z, x, w)$ can be seen as a union of t binary relations each of which is at most single-element. Thus the relation $\exists z{<}t \, \delta(z, x, w)$ satisfies the condition from the definition of sparse relation with respect to $t + 1$: every square $s \times s$ contains less than $t + 1$ elements of this relation. Let x be given and consider z like in (3), satisfying $z < t$ and $\forall u(\exists v\gamma(z,u,v) \equiv u = x)$. If w is the least number such that $\exists w'{\le}w\,\gamma(z,x,w')$, then $\delta(z,x,w)$. This shows that $\forall x\exists w\exists z{<}t\,\delta(z,x,w)$, and thus $\exists z{<}t\,\delta(z,x,w)$ satisfies also the remaining condition from the definition of sparse relation.

We can summarize that $\exists z{<}t\,\delta(z,x,w)$ is a sparse relation; hence by Theorem 3, $\mathsf{B}\Pi_n$ is violated in $\mathcal{K}$.

Theorem 4 *The following conditions are equivalent over* $\mathsf{I}\Sigma_n{+}\mathsf{Exp}$.
(i) There exists a $\Sigma_0(\Sigma_n)$*-definable sparse relation.*
(ii) There exists a Σ_{n+1}*-definable one-to-one function which is bounded.*
(iii) There exists a Σ_{n+1}*-definable one-to-one function the range of which is the interval* $\{\, y\,;\, y < t \,\}$ *for some* t.

We omit the proofs of this theorem, but we give some remarks and hints. if $\varphi(x,y,v) \in \Sigma_0(\Sigma_n)$ defines a one-to-one function bounded by z, then the relation defined by

$$\exists w\exists y{\le}w\,\exists v{\le}w\,(\varphi(x,y,v)\ \&\ \forall y'{<}w\forall v'{<}w\,\neg\varphi(x,y',v'))$$

is $\Sigma_0(\Sigma_n)$ and sparse with respect to any $t > z$. The proof of (ii) $\Rightarrow$ (iii) is due to Jeff Paris and is probably unpublished. Actually, the notion of sparse relation is extracted from Paris's proof, communicated privately, of (ii) $\Rightarrow$ (iii).

5 Final remarks

There are several variants of the pigeon hole principle (PHP), the most common being that there is no one-to-one function from $t+1$-element set to t-element set. The negation of the condition in (ii) of Theorem 4 says that there is no Σ_{n+1}-definable one-to-one function *from the entire universe* to some t-element set. This principle, i.e. the negation of the condition (ii) of Theorem 4, can be called a *weak PHP principle* and denoted WPHP(Σ_{n+1}). I conjecture (and have some ideas how to prove) that this principle is weaker than $\mathsf{B}\Pi_n$, i.e. that $\mathsf{I}\Sigma_n$+WPHP(Σ_{n+1}) does not prove $\mathsf{B}\Pi_n$. There may also be an open problem connected with WPHP(Σ_{n+1}): I do not know whether $\mathsf{I}\Delta_0$+WPHP(Σ_{n+1}) proves $\mathsf{I}\Sigma_n$.

References

[Ben62] J. H. Bennet. *On Spectra.* Dissertation, Princeton University, Princeton, NJ, 1962.

[HP93] P. Hájek and P. Pudlák. *Metamathematics of First Order Arithmetic.* Springer, 1993.

[PK78] J. B. Paris and L. A. S. Kirby. Σ_n-collection schemas in arithmetic. In A. Macintyre, L. Pacholski, and J. Paris, editors, *Logic Colloquium '77*, number 96 in Studies in Logic and the Foundations of Mathematics, pages 199–209. North-Holland, Amsterdam, 1978.

[Pud83] P. Pudlák. A definition of exponentiation by a bounded arithmetical formula. *Comm. Math. Univ. Carolinae*, 24(4):667–671, 1983.

[TMR53] A. Tarski, A. Mostowski, and R. M. Robinson. *Undecidable Theories*. North-Holland, Amsterdam, 1953.