Two Proofs of Parsons' Theorem

Joost J. Joosten

November 28, 2002

Abstract

It is well-known that $I\Sigma_1$ is Π_2-conservative over PRA. This fact is often referred to as Parsons' theorem. In this paper we provide two proofs of the theorem. One is proof-theoretic and easily formalizable. The other proof is model-theoretic. The main ingredient of the second proof is a characterization of $I\Sigma_1$ in terms of provable closure under the iteration operation on functions.

1 Introduction

This note is dedicated to a theorem that is proved independently by C. Parsons ([Par70], [Par72]), G. Mints ([Min72]) and G. Takeuti ([Tak75]). The theorem is usually referred to as Parsons' theorem and says that $I\Sigma_1$ is Π_2-conservative over PRA. Often, PRA is associated with finitism ([Sko67], [HB68], [Tai81]). In this light, Parsons' theorem can be considered of great importance as a partial realization of Hilbert's programme.

The first proofs of Parsons' theorem were all of proof-theoretical nature. Parsons' first proof, [Par70], is based upon Gödel's Dialectica interpretation. His second proof, [Par72], merely relies on a Cut-elimination. Mints' proof, [Min72], employs the no-counterexample interpretation of a special sequent calculus. The proof by Takeuti, [Tak75], employs an ordinal analysis in the style of Gentzen.

Over the years, many more proofs of Parsons' theorem have been published. In many accounts Herbrand's theorem plays a central role in providing primitive recursive Skolem functions for Π_2-statements provable in $I\Sigma_1$. (Cf. Sieg's method of Herbrand analysis [Sie91], Avigad's proof by his notion of Herbrand saturated models [Avi02], Buss' proof by means of his witness predicates [Bus98], and Ferreira using Herbrand's theorem for Σ_3 and Σ_1-formulas [Fer02].) A first model-theoretic proof is due to Paris and Kirby. They employ semi-regular cuts in their proof (cf. [Sim99]:373-381).

This note adds two more proofs to the long list. The first proof, due to L. Beklemishev, is a proof-theoretic one. It is explicitly written down here for the first time and can be seen as a modern version of Parsons' second proof. The main ingredient is the Cut-elimination theorem for Tait's sequent calculus.

The second proof is essentially contained in an unpublished note of A. Visser ([Vis90]). In that note a model-theoretic proof is sketched. A central ingredient is an analysis of the difference between PRA and $\mathrm{I}\Sigma_1$ in terms of iteration of total functions. The very same note inspired Zambella in his [Zam96] for a proof of a conservation result of Buss' S^1_2 over PV.

I am grateful to Lev Beklemishev and Albert Visser for making their proof-sketches available to me. Also the discussion with them helped the writing of this paper a lot.

2 A Proof-theoretic Proof of Parsons' Theorem

The first proof we give of Parsons' theorem is proof-theoretic. It is due to L. Beklemishev. It will become evident that the whole argument is easily formalizable as soon as the superexponential function is provably total. This is because the proof only uses the standard Cut-elimination theorem.

We denote by $\mathrm{I}\Sigma_1^R$ the system that arises from adding to some minimal arithmetical theory, for example PA^- from [Kay91], the Σ_1 induction rule. The Σ_1 induction rule allows one to conclude $\forall x\, \sigma(x)$ from $\sigma(0)$ and $\forall x\, (\sigma(x) \to \sigma(x+1))$ for $\sigma(x) \in \Sigma_1$. It is well known that $\mathrm{I}\Sigma_1^R$ is equi-interpretable with PRA and certainly that $\mathrm{I}\Sigma_1^R$ is Π_2 conservative over PRA. In Lemma 2.2 we shall make this conservation more precise. In this section we use $\mathrm{I}\Sigma_1^R$ to state and prove Parsons' theorem.

Theorem 2.1 $\mathrm{I}\Sigma_1$ *is a* Π_2 *conservative extension of* $\mathrm{I}\Sigma_1^R$.

PROOF OF Π_2 CONSERVATIVITY. In this proof we will use the Tait sequent calculus of first order logic which is presented in Schwichtenberg's contribution to the Handbook of Mathematical Logic. (See [Sch77].) It works with sequents which are finite sets and should be read disjunctively in the sense that $\Gamma = \{\varphi_1, \ldots, \varphi_n\}$ stands for $\varphi_1 \vee \ldots \vee \varphi_n$. Often we will omit the set-brackets $\{\}$. All formulas are written in a form that uses only $\wedge, \vee, \forall, \exists$ and literals, that is, atoms or negations of atoms. Negation of composed formulas is an operation defined by the de Morgan laws. The axioms of the Tait calculus are:

$$\Gamma, \varphi, \neg\varphi \quad \text{for atomic } \varphi.$$

The rules are:

$$\frac{\Gamma, \varphi \quad \Gamma, \psi}{\Gamma, \varphi \wedge \psi}, \qquad \frac{\Gamma, \varphi}{\Gamma, \varphi \vee \psi}, \qquad \frac{\Gamma, \psi}{\Gamma, \varphi \vee \psi},$$

$$\frac{\Gamma, \varphi(a)}{\Gamma, \forall x\ \varphi(x)}, \qquad \frac{\Gamma, \varphi(t)}{\Gamma, \exists x\ \varphi(x)},$$

plus the cut rule

$$\frac{\Gamma, \varphi \quad \Gamma, \neg\varphi}{\Gamma}.$$

In the rule for the universal quantifier introduction it is necessary that the a does not occur free anywhere else in Γ. And in the rule for the introduction of the existential quantifier one requires t to be substitutable for x in φ.

So, our aim is to prove that if $\mathrm{I}\Sigma_1 \vdash \pi$ then $\mathrm{I}\Sigma_1^R \vdash \pi$ whenever π is a Π_2-sentence. We reason as follows. If $\mathrm{I}\Sigma_1 \vdash \pi$, we have by the compactness and deduction theorem that $\vdash \sigma \to \pi$ where σ is the conjunction of a finite number of axioms of $\mathrm{I}\Sigma_1$. Or equivalently $\vdash \neg\sigma \vee \pi$. As the Tait calculus is complete this amounts to the same as saying that the sequent $\neg\sigma, \pi$ is derivable within the calculus. By the Cut-elimination theorem for the Tait calculus we know that there exists a cut-free derivation of the sequent. Cut-free proofs have all sorts of pleasant properties like the sub-formula property (modulo substitution of terms).

The proof is concluded by showing by induction on the length of cut-free derivations that if a sequent of the form Σ, Π is derivable then $\mathrm{I}\Sigma_1^R \vdash \bigvee \Pi$. Here Σ is a finite set of negations of induction axioms of Σ_1-formulas written in a specific form and Π is a finite non-empty set of strict Π_2-formulas.[1] $\bigvee \Pi$ denotes the disjunction of all elements in Π.

The basis case of the proof is trivial as $\mathrm{I}\Sigma_1^R \vdash \sigma \vee \neg\sigma \vee \bigvee \Gamma$ for any σ.

So, suppose we have a cut-free proof of Σ, Π. What can be the last step in the proof of this sequent? Either the last rule yielded something in the Π part of the sequent or in the Σ part of it. In the first case nothing interesting happens and we almost automatically obtain the desired result by the induction hypothesis.

So, suppose something had happened in the Σ part. We can assume that the Σ part only contains formulas of the form $\exists a\ [\varphi(0) \wedge \forall x\ (\varphi(x) \to \varphi(x+1)) \wedge \neg\varphi(a)]$, with $\varphi \in \Sigma_1$. With

[1] To see that hereby the proof is indeed completed it suffices to remark that we can take an axiomatization of $\mathrm{I}\Sigma_1$ that consists of Σ_2 (actually Π_1 is enough) -sentences that are provable in $\mathrm{I}\Sigma_1^R$ plus induction axioms for all Σ_1 formulas.

$\forall x\ (\varphi(x) \to \varphi(x{+}1))$ we actually denote the formula in prenex normal form in the calculus that is predicate-logically equivalent to it.[2]

The last deduction step thus must have been the introduction of the existential quantifier and we can by a one step shorter proof derive for some term t the following sequent

$$\Sigma', \varphi(0) \wedge \forall x\ (\varphi(x) \to \varphi(x+1)) \wedge \neg\varphi(t), \Pi$$

By the inversion property of the Tait calculus (for a proof and precise formulation of the statement consult [Sch77] page 873) we obtain proofs of the same length of the following sequents

$$\Sigma', \varphi(0), \Pi\ , \quad \Sigma', \forall x\ (\varphi(x) \to \varphi(x+1)), \Pi \quad \text{and} \quad \Sigma', \neg\varphi(t), \Pi.$$

As all of $\varphi(0), \forall x\ (\varphi(x) \to \varphi(x+1))$ and $\neg\varphi(t)$ are Π_2-formulas, we can apply the induction hypothesis to conclude that we have

$$\begin{array}{lll} \mathrm{I}\Sigma_1^R & \vdash & \varphi(0) \vee \bigvee \Pi, \\ \mathrm{I}\Sigma_1^R & \vdash & \forall x\ (\varphi(x) \to \varphi(x+1)) \vee \bigvee \Pi, \\ \mathrm{I}\Sigma_1^R & \vdash & \neg\varphi(t) \vee \bigvee \Pi. \end{array}$$

Recall that Π consists of Π_2-statements. So, Π is of the form (with some abuse of notation) $\forall \vec{u}\ \exists \vec{v}\ \Pi_0(\vec{u}, \vec{v})$. In our context we can omit the outer universal quantifiers.

If we now define $\varphi'(x, \vec{u}) := \varphi(x) \vee \bigvee \exists \vec{v}\ \Pi_0(\vec{u}, \vec{v})$, we obtain a Σ_1-formula to which we can apply the induction rule to obtain $\forall x\ \varphi'(x, \vec{u})$ and thus also $\varphi'(t, \vec{u})$. Combining this with $\mathrm{I}\Sigma_1^R \vdash \neg\varphi(t) \vee \bigvee \exists \vec{v}\ \Pi_0(\vec{u}, \vec{v})$ yields $\mathrm{I}\Sigma_1^R \vdash \bigvee \exists \vec{v}\ \Pi_0(\vec{u}, \vec{v})$ by one application of the cut rule (in $\mathrm{I}\Sigma_1^R$) and thus we have the desired result, $\mathrm{I}\Sigma_1^R \vdash \bigvee \Pi$. QED

In order to get the proof of Theorem 2.1 started we had to switch to a cut-free proof in $\mathrm{I}\Sigma_1$ of the Π_2-sentence. This makes that the corresponding proof in $\mathrm{I}\Sigma_1^R$ of the same Π_2-sentence is superexponentially

[2] In the case that the proof in $\mathrm{I}\Sigma_1$ of Π uses an induction axiom with parameters actually a little coding trick is needed here. The negation of the induction formula can be written then as $\exists x \exists y\ (\varphi(0, y) \wedge \forall z\ (\varphi(z, y) \to \varphi(z+1, y)) \wedge \neg\varphi(x, y))$ which is equivalent by the coding of pairs to $\exists x\ (\varphi(0, (x)_1) \wedge \forall z\ (\varphi(z, (x)_1) \to \varphi(z+1, (x)_1)) \wedge \neg\varphi((x)_0, (x)_1))$. The proof then runs the same if one just drags the parameter $(t)_1$ along. Of course we should have this coding machinery available for example by adding the pairing and projection functions to our language together with some axioms stating their basic properties. It is doubtful whether starting with "sufficiently many" negations of induction axioms instead of enriching the language also suffices. Anyway we are not dealing with an essential problem here as we could have easily set up the proof with Σ containing sentences of the form $\exists y_1 \ldots \exists y_m\ \exists a\ [\varphi(0, y_1, \ldots, y_m) \wedge \forall x\ (\varphi(x, y_1, \ldots, y_m) \to \varphi(x+1, y_1, \ldots, y_m)) \wedge \neg\varphi(a, y_1, \ldots, y_m)]$ with $\varphi \in \Sigma_1$. Our assumption is just to simplify the presentation of the argument.

larger. In Ignjatovic ([Ign90]) it is shown that this growth of proofs is essential.

In [Bek99] a generalization of Parsons' theorem is stated in Corollary 4.8: For $n \geq 1$, $\mathrm{I}\Sigma_n$ is Π_{n+2}-conservative over $\mathrm{I}\Sigma_n^R$. This is a corollary of his Reduction property, Theorem 2, which is also formalizable in the presence of the superexponential function. The proof of Parsons' theorem we have presented here is very close to the proof of the reduction property.

We conclude this section by showing that PRA as it is formulated usually contains all theorems of $\mathrm{I}\Sigma_1^R$. Often one defines PRA in a language that contains for every primitive recursive function a function symbol plus its defining axioms. In this extended language PRA allows for induction over Δ_0-formulas. It is also well-known that PRA is interpretable in $\mathrm{I}\Sigma_1^R$ in the expected way, that is, every function symbol is replaced by its definition in terms of sequences. The interpretability is in this section to a lesser extend of our concern.

Lemma 2.2 *$I\Sigma_1^R \subseteq$ PRA.*

Proof of Lemma 2.2. The proof goes by induction on the length of a proof in $\mathrm{I}\Sigma_1^R$. If $\mathrm{I}\Sigma_1^R \vdash \varphi$ without any applications of the Σ_1 induction rule, it is clear that $\mathrm{PRA} \vdash \varphi$.

So, suppose that the last step in the $\mathrm{I}\Sigma_1^R$-proof of φ were an application of the Σ_1 induction rule. Thus φ is of the form $\forall x \exists y\ \varphi_0(x,y,\vec{z})$ [3] and we obtain shorter $\mathrm{I}\Sigma_1^R$-proofs of the Σ_1-statements $\exists y\ \varphi_0(0,y,\vec{z})$ and $\exists y'\ (\varphi_0(x,y,\vec{z}) \rightarrow \varphi_0(x+1,y',\vec{z}))$. The induction hypothesis tells us that these statements are also provable in PRA. Herbrand's theorem for PRA provides us[4] with primitive recursive functions $g(\vec{z})$ and $h(x,y,\vec{z})$ such that

$$\mathrm{PRA} \vdash \varphi_0(0, g(\vec{z}), \vec{z}) \qquad (1)$$

and

$$\mathrm{PRA} \vdash \varphi_0(x,y,\vec{z}) \rightarrow \varphi_0(x+1, h(x,y,\vec{z}), \vec{z}) \qquad (2)$$

Let $f(x,\vec{z})$ be the primitive recursive function defined by

$$\begin{cases} f(0,\vec{z}) = g(\vec{z}), \\ f(x+1,\vec{z}) = h(x, f(x,\vec{z}), \vec{z}). \end{cases}$$

By (1) and (2) it follows from Δ_0-induction in PRA that $\mathrm{PRA} \vdash \forall x\ \varphi_0(x, f(x,\vec{z}), \vec{z})$ whence $\mathrm{PRA} \vdash \forall x \exists y\ \varphi_0(x,y,\vec{z})$. QED

[3]We treat here the case where φ contains only one unbounded existential quantifier. The more general case can be proved completely analogously.

[4]It is well known that we can take an open axiomatization of PRA to which we indeed can apply Herbrand's theorem. Cf. [Sch87] or [Joo01].

We note that our proof of Parsons' theorem also involves Herbrand's theorem for PRA. In our case however Herbrand's theorem only occurs in showing the conservation of $\mathrm{I}\Sigma_1^R$ over PRA. Also we note that one application of of the Σ_1 induction rule in $\mathrm{I}\Sigma_1^R$ corresponds to one application of primitive recursion in PRA. The correspondence can be stretched further as is carried out in [Bek97].

3 A Model-theoretic Proof of Parsons' Theorem

In this section we shall work out the proof sketch of Visser's note ([Vis90]). The precise date in which this note was finished is not clear. We (and A. Visser too) estimate that it was written around 1990.

Theorem 3.1 *(C. Parsons [Par70], [Par72]) $\mathrm{I}\Sigma_1$ is a Π_2-conservative extension of* PRA.

If Parsons' theorem is formulated in this way, we find it convenient if the language of $\mathrm{I}\Sigma_1$ contains the language of PRA. Therefore we choose some suitable formulations of these theories. PRA will be just $\mathrm{I}\Delta_0$ in a language that contains function symbols for some special primitive recursive functions and $\mathrm{I}\Sigma_1$ will be formulated in that same language with the induction schema for all Σ_1 formulas.

Theorem 3.13 describes the difference between PRA and $\mathrm{I}\Sigma_1$ in terms of totality statements of recursive functions. Lemma 3.14 tells us what it takes for a model $\mathcal{M}$ of PRA to also be a model of $\mathrm{I}\Sigma_1$: A class of functions of this model should be majorizable by another class of functions. This lemma is at the heart of our model-theoretic proof of Parsons' theorem. We will show that any countable model $\mathcal{N}$ of PRA falsifying $\pi \in \Pi_2$ can be extended to a countable model $\mathcal{N}'$ of $\mathrm{I}\Sigma_1 + \neg\pi$ whence $\mathrm{I}\Sigma_1 \nvdash \pi$. In extending the model we will, having Lemma 3.14 in the back of our mind, repeatedly majorize functions to finally obtain a model of $\mathrm{I}\Sigma_1 + \neg\pi$.

3.1 Introducing a new function symbol

In our discussion we shall like to work with a theory that arises as an extension of PRA by a definition. We will add a new function symbol f to the language of PRA together with the axiom φ that defines f. Moreover we would like to employ induction that involves this new function symbol, possibly also in the binding terms of the bounded quantifiers. We will see that if the function f allows for a simple definition and has some nice properties we have indeed access to the extended form of induction.

Essentially the justification boils down to a theorem of Gaifman and Dimitracopoulos [GD82] a proof of which can also be found in [HP93] (Theorem 1.48 on page 45 and Proposition 1.3 on page 271). We will closely follow here a proof of Beklemishev from [Bek97] which we slightly modified for our purposes.

Throughout this section we will adhere to the following notational convention. Arithmetical formulas defining the graph of a function are denoted by lowercase Greek letters. The corresponding lower case Roman letter is reserved to be the symbol that refers to the function described by its graph. By the corresponding upper case Roman letter we will denote the very short formula that defines the graph using the lower case Roman letter and the identity symbol only. Context, like indices and so forth, is inherited in the expected way.

For example, if $\chi_n(x,y)$ is an arithmetical formula describing a function, in a richer language this function will be referred to by the symbol g_n. The corresponding G_n will refer to the simple formula $g_n(x) = y$.

Definition 3.2 *Let $\{g_i\}_{i\in I}$ be a set of function symbols. The $\Delta_0(\{g_i\}_{i\in I})$-formulas are the bounded formulas in the language of* PA *enriched with the function symbols $\{g_i\}_{i\in I}$. The new function symbols are also allowed to occur in the binding terms of the bounded quantifiers. By $I\Delta_0(\{g_i\}_{i\in I})$ we mean the theory that comprises*

- *some open axioms describing some minimal arithmetic*[5]*,*
- *induction axioms for all $\Delta_0(\{g_i\}_{i\in I})$-formulas and*
- *(possibly) defining axioms of the symbols $\{g_i\}_{i\in I}$.*

The defining axioms of the symbols $\{g_i\}_{i\in I}$ are denoted by $\mathcal{D}(\{g_i\}_{i\in I})$.

Definition 3.3 *Let $\varphi(x,y)$ be a $\Delta_0(\{g_i\}_{i\in I})$ formula. By $\mathsf{Tot}(\varphi)$ we shall denote the formula $\forall x\, \exists !y \; \varphi(x,y)$*[6] *stating that φ can be regarded as a total function. By $\mathsf{Mon}(\varphi)$ we shall denote the formula $\forall x, x', y, y' \; (x \leq x' \wedge \varphi(x,y) \wedge \varphi(x',y') \rightarrow y \leq y') \wedge \mathsf{Tot}(\varphi)$ stating the monotonicity of the total φ.*

Definition 3.4 *Let φ be such that $I\Delta_0(\{g_i\}_{i\in I}) \vdash \mathsf{Tot}(\varphi)$. Recall that the uppercase letter F paraphrases the formula $f(x){=}y$. A $\Delta_0(\{g_i\}_{i\in I}, F)$-formula is a $\Delta_0(\{g_i\}_{i\in I})$-formula possibly containing occurrences of F. By $I\Delta_0(\{g_i\}_{i\in I}, F)$ we denote the theory $I\Delta_0(\{g_i\}_{i\in I})$ where we now also have induction for $\Delta_0(\{g_i\}_{i\in I}, F)$ formulas. The defining axiom of f, in our case φ, is also in $I\Delta_0(\{g_i\}_{i\in I}, F)$.*

[5] For example the open part of Robinson's arithmetic.

[6] That is, $\forall x\, \exists y \; \varphi(x,y) \wedge \forall x\, \forall y\, \forall y' \; (\varphi(x,y) \wedge \varphi(x,y') \rightarrow y = y')$.

Note that f cannot occur in a bounding term in an induction axiom of $\mathrm{I}\Delta_0(\{g_i\}_{i\in I}, F)$. Also note that F is nothing but a formula containing f stating $f(x) = y$ and consisting of just six symbols. Of course later we will substitute for F an arithmetical definition of the graph of f, that is, $\varphi(x, y)$.

The main interest of the extension of $\mathrm{I}\Delta_0(\{g_i\}_{i\in I})$ by a definition of f is in Theorem 3.7 and in its Corollary 3.8. The latter says that we can freely use $f(x)$ as an abbreviation of $\varphi(x, y)$ and have access to $\Delta_0(\{g_i\}_{i\in I}, f)$-induction whenever f is provably total and monotone in $\mathrm{I}\Delta_0(\{g_i\}_{i\in I})$ and has a $\Delta_0(\{g_i\}_{i\in I})$ graph.

First we prove some technical but rather useful lemmata. They are minor variations of Beklemishev's Lemma 5.12 and 5.13 from [Bek97].[7] From now on we will work under the assumptions of Lemma 3.7 so that $\mathrm{I}\Delta_0(\{g_i\}_{i\in I})$ is such that any term t in its language is provably majorizable by some other term $\tilde{t}$ that is strictly increasing in all of its arguments. Throughout the forthcoming proofs we will for any term t denote by $\tilde{t}$ such a term that is provably strictly monotone (in all of its arguments) and majorizing t.

Lemma 3.5 *For every term $s(\vec{a})$ of $I\Delta_0(\{g_i\}_{i\in I}, f)$ and every $R \in \{\leq, \geq, =, <, >\}$ there are terms t_s^R and $\tilde{s}(a)$ strictly increasing in all of their arguments and a $\Delta_0(\{g_i\}_{i\in I}, F)$-formula $\psi_s^R(\vec{a}, b, y)$ such that $I\Delta_0(\{g_i\}_{i\in I}, F) + \mathsf{Mon}(\varphi) \vdash \forall\, y{\geq}t_s^R(\vec{a})\ (s(\vec{a})Rb \leftrightarrow \psi_s^R(\vec{a}, b, y))$ and $I\Delta_0(\{g_i\}_{i\in I}, F) + \mathsf{Mon}(\varphi) \vdash \forall\, \vec{x}\ (s(\vec{x}) \leq \tilde{s}(\vec{x}))$.*

Proof of Lemma 3.5. The proof proceeds by induction on $s(\vec{a})$. In the basis case nothing has to be done as $x_i Rb$, $0Rb$ and $1Rb$ are all atomic $\Delta_0(\{g_i\}_{i\in I}, F)$-formulas. Moreover all of the x_i, 0 and 1 are (provably) strictly monotone in all of their arguments. For the induction case consider $s(\vec{a}) = h(s_1(\vec{a}))$, where h is either one of the g_i or $h = f$. For simplicity we assume here that h is a unary function.

The induction hypothesis provides us with a $\Delta_0(\{g_i\}_{i\in I}, F)$-formula $\psi_{s_1}^{=}(\vec{a}, b, y)$ and terms $t_{s_1}^{=}(\vec{a})$ and $\tilde{s}_1(\vec{a})$ such that

$$\mathrm{I}\Delta_0(\{g_i\}_{i\in I}, F) + \mathsf{Mon}(\varphi) \vdash \forall\, y{\geq}t_{s_1}^{=}(\vec{a})\ (s_1(\vec{a}) = b \leftrightarrow \psi_{s_1}^{=}(\vec{a}, b, y)),$$

and

$$\mathrm{I}\Delta_0(\{g_i\}_{i\in I}, F) + \mathsf{Mon}(\varphi) \vdash \forall\, \vec{x}\ (s_1(\vec{x}) \leq \tilde{s}_1(\vec{x})).$$

We now want to say that $h(s_1(\vec{a}))Rb$ in a $\Delta_0(\{g_i\}_{i\in I}, F)$ way. This can be done by $\exists\, y', y''{\leq}y\ (\psi_{s_1}^{=}(\vec{a}, y', y) \wedge h(y') = y'' \wedge y''Rb)$ whenever

[7] One can also consult the same lemmata of [Bek96]. This reference can be found online at the homepage of the Philosophy Department of Utrecht University at `http://www.phil.uu.nl/` or `http://preprints.phil.uu.nl/lgps/`.

$y \geq t^{=}_{s_1}(\vec{a})+\tilde{s}(\vec{a})$. Here we define $\tilde{s}(\vec{a})$ to be just $f(\tilde{s}_1(\vec{a}))$ in case $h=f$ and $\tilde{g}_i(\tilde{s}_1(\vec{a}))$ in case $h=g_i$. Clearly $\mathrm{I}\Delta_0(\{g_i\}_{i\in I},F)+\mathsf{Mon}(\varphi)\vdash \forall\,\vec{x}\;(s(\vec{x})\leq\tilde{s}(\vec{x}))$. Indeed one easily sees that

$$\begin{aligned}\mathrm{I}\Delta_0(\{g_i\}_{i\in I},F)+\mathsf{Mon}(\varphi)\vdash \forall\,y{\geq}t^{=}_{s_1}(\vec{a})+\tilde{s}(\vec{a})\;\;&[h(s_1(\vec{a}))Rb\leftrightarrow\\ &\exists\,y',y''{\leq}y\;(\psi^{=}_{s_1}(\vec{a},y',y)\wedge h(y')=y''\wedge y''Rb)].\end{aligned}$$

It is also easy to see that $t^{=}_{s_1}(\vec{a})+\tilde{s}(\vec{a})$ is indeed monotone in all of its arguments. In case $h=f$ we need $\mathsf{Mon}(\varphi)$ here.

A similar reduction applies to the case when the function h has more than one argument. QED

It is possible to simplify the above reduction a bit by distinguishing between $h=f$ and $h\neq f$ and also $R==$ and $R\neq=$, or by proving the lemma just for $R==$ and showing that all the other cases can be reduced to this. We are not very much interested in optimality at this point though.

Lemma 3.6 *For every $\Delta_0(\{g_i\}_{i\in I},f)$-formula $\theta(\vec{a})$ there is a $\Delta_0(\{g_i\}_{i\in I},F)$-formula $\theta_0(\vec{a},y)$ and a provably monotonic term $t_\theta(\vec{a})$ such that*
$I\Delta_0(\{g_i\}_{i\in I},F)+\mathsf{Mon}(\varphi)\vdash \forall\,y{\geq}t_\theta(\vec{a})\;\;(\theta(\vec{a})\leftrightarrow\theta_0(\vec{a},y))$.

PROOF OF LEMMA 3.6. The lemma is proved by induction on θ.

- Basis. In this case $\theta(\vec{a})$ is $s_1(\vec{a})R\,s_2(\vec{a})$. Applying Lemma 3.5 we see that $s_1(\vec{a})R\,s_2(\vec{a})\leftrightarrow\exists\,b{\leq}y\;(\psi^{=}_{s_2}(\vec{a},b,y)\wedge\psi^{R}_{s_1}(\vec{a},b,y))$[8] whenever $y\geq t_{s_1}(\vec{a})+t_{s_2}(\vec{a})$.
- The only interesting induction case is where a bounded quantifier is involved. We consider the case when $\theta(\vec{a})$ is $\exists\,x{\leq}s(\vec{a})\;\xi(\vec{a},x)$. The induction hypothesis yields a provably monotone term $t_\xi(\vec{a},x)$ and a $\Delta_0(\{g_i\}_{i\in I},F)$-formula $\xi_0(\vec{a},x,y)$ such that provably

$$\forall\,y{\geq}t_\xi(\vec{a},x)\;(\xi(\vec{a},x)\leftrightarrow\xi_0(\vec{a},x,y))$$

. Combining this with Lemma 3.5 we get that provably

$$\exists\,x{\leq}s(\vec{a})\;\xi(\vec{a},x)\leftrightarrow\exists\,x'{\leq}y\;(\psi^{=}_{s}(\vec{a},x',y)\wedge\exists\,x{\leq}x'\;\xi_0(\vec{a},x,y))^{9}$$

whenever $y\geq\tilde{s}(\vec{a})+t^{=}_{s}(\vec{a})+t_\xi(\vec{a},\tilde{s}(\vec{a}))$.

QED

Lemma 3.7 *Let $I\Delta_0(\{g_i\}_{i\in I})$ be such that any term t in its language is provably majorizable by some other term $\tilde{t}$ that is strictly increasing in all of its arguments. We have that $I\Delta_0(\{g_i\}_{i\in I},F)+\mathsf{Mon}(\varphi)\vdash I\Delta_0(\{g_i\}_{i\in I},f)$.*

[8] If we only want to use Lemma 3.5 with R being $=$ we can observe that $s_1(\vec{a})R\,s_2(\vec{a})\leftrightarrow\exists\,b,c{\leq}y\;(\psi^{=}_{s_1}(\vec{a},b,y)\wedge\psi^{=}_{s_2}(\vec{a},c,y)\wedge bRc)$ whenever $y\geq t_{s_1}(\vec{a})+t_{s_2}(\vec{a})$.

[9] Alternatively, one could take $\exists\,x{\leq}y\;(\psi^{\geq}_{s}(\vec{a},x,y)\wedge\xi_0(\vec{a},x,y))$ for $y\geq t^{\geq}_{s}(\vec{a})+t_\xi(\vec{a},\tilde{s}(\vec{a}))$.

Proof of Lemma 3.7. We will prove the least number principle[10] for $\Delta_0(\{g_i\}_{i\in I}, f)$-formulas in $\mathrm{I}\Delta_0(\{g_i\}_{i\in I}, F)+\mathsf{Mon}(\varphi)$ as this is equivalent to induction for $\Delta_0(\{g_i\}_{i\in I}, f)$-formulas. So, let $\theta(x, \vec{a})$ be a $\Delta_0(\{g_i\}_{i\in I}, f)$-formula and reason in $\mathrm{I}\Delta_0(\{g_i\}_{i\in I}, F) + \mathsf{Mon}(\varphi)$. By Lemma 3.6 we have a strict monotone term $t_\theta(x, \vec{a})$ and a $\Delta_0(\{g_i\}_{i\in I}, F)$-formula $\theta_0(x, \vec{a}, y)$ such that $\theta(x, \vec{a}) \leftrightarrow \theta_0(x, \vec{a}, y)$ whenever $y \geq t_\theta(x, \vec{a})$.

Now assume $\exists\, x\ \theta(x, \vec{a})$. We will show that

$$\exists\, x\ (\theta(x, \vec{a}) \wedge \forall\, x'{<}x\ \neg\theta(x', \vec{a})).$$

Let x be such that $\theta(x, \vec{a})$. We now fix some $y \geq t_\theta(x, \vec{a})$. Thus we have $\theta_0(x, \vec{a}, y)$. Applying the least number principle to the $\Delta_0(\{g_i\}_{i\in I}, F)$-formula $\theta_0(x, \vec{a}, y)$ we get a minimal x_0 such that $\theta_0(x_0, \vec{a}, y)$. As $x_0 < x$ and t_θ is monotone we have $y \geq t_\theta(x, \vec{a}) \geq t_\theta(x_0, \vec{a})$ and thus $\theta(x_0, \vec{a})$. If now $x'{<}x_0$ such that $\theta(x', \vec{a})$ then also $\theta_0(x', \vec{a}, y)$ which would conflict the minimality of x_0 for θ_0. Thus x_0 is the minimal element such that $\theta(x_0, \vec{a})$. QED

As in [Bek97] (Remark 5.14) we note here that Lemma 3.7 shows that $\Delta_0(\{g_i\}_{i\in I}, f)$ induction is actually provable from $\Delta_0(\{g_i\}_{i\in I}, F)$ induction where the bounding terms are just plain variables. Also we note that Lemma 3.5 and Lemma 3.6 do not use the full strength of $\mathrm{I}\Delta_0(\{g_i\}_{i\in I}, F)$.

Corollary 3.8 *Suppose $I\Delta_0(\{g_i\}_{i\in I})$ and φ are as above and $I\Delta_0(\{g_i\}_{i\in I}) \vdash \mathsf{Mon}(\varphi)$. If $\varphi \in \Delta_0(\{g_i\}_{i\in I})$ then we have that $I\Delta_0(\{g_i\}_{i\in I}, f) \vdash \psi \Rightarrow I\Delta_0(\{g_i\}_{i\in I}) \vdash \psi^*$ where ψ^* is any non-pathological translation of ψ where f is somehow replaced by φ.*

Proof of Corollary 3.8. As $\mathrm{I}\Delta_0(\{g_i\}_{i\in I}) \vdash \mathsf{Tot}(\varphi)$ we can form an extension by a definition of $\mathrm{I}\Delta_0(\{g_i\}_{i\in I})$ by adding an additional function symbol f to the language and by adopting an additional axiom $\forall x\ \varphi(x, f(x))$. For this definitional extension we know that if $\mathrm{I}\Delta_0(\{g_i\}_{i\in I}) + \mathcal{D}(f) \vdash \psi$ then $\mathrm{I}\Delta_0(\{g_i\}_{i\in I}) \vdash \psi^*$ for any non-pathological translation of ψ where f is somehow replaced by φ. (See for example [End72]:164-173, Section 2.7.) But

$$\mathrm{I}\Delta_0(\{g_i\}_{i\in I}) + \mathcal{D}(f) \vdash \mathrm{I}\Delta_0(\{g_i\}_{i\in I}, F)$$

as F is a $\Delta_0(\{g_i\}_{i\in I})$ formula. Theorem 3.7 tells us that $\mathrm{I}\Delta_0(\{g_i\}_{i\in I}, F) \vdash \mathrm{I}\Delta_0(\{g_i\}_{i\in I}, f)$ since $\mathrm{I}\Delta_0(\{g_i\}_{i\in I}) \vdash \mathsf{Mon}(\varphi)$. QED

It is noteworthy that the above treatment can be carried out in a far more general setting where we consider the results relative to

[10] The equivalence of the least number principle and induction indeed holds for our classes of $\Delta_0(\{g_i\}_{i\in I}, F)$ and $\Delta_0(\{g_i\}_{i\in I}, f)$-formulas. The standard proof for the equivalence can just be applied in our case.

some base theory T with possibly more function symbols than just the $\{g_i\}_{i\in I}$. For example, the generalization of Lemma 3.7 would be formulated as follows:

Corollary 3.9 *Let $I\Delta_0(\{g_i\}_{i\in I})$ and T be such that any term t in the language of $I\Delta_0(\{g_i\}_{i\in I})$ is provably majorizable in $T + I\Delta_0(\{g_i\}_{i\in I})$ by some other term $\tilde{t}$ that is strictly increasing in all of its arguments. We have that $T + I\Delta_0(\{g_i\}_{i\in I}, F) + \mathsf{Mon}(\varphi) \vdash I\Delta_0(\{g_i\}_{i\in I}, f)$.*

3.2 The difference between $\mathrm{I}\Sigma_1$ and PRA in terms of iteration of $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-definable functions

In this subsection we will use some definitions of PRA and $\mathrm{I}\Sigma_1$ slightly different from the usual ones.

Definition 3.10 *The* language *of* PRA *is the language of* PA *plus a family of new function symbols* $\{\mathsf{Sup}_n \mid n{\in}\omega\}$. *The* non-logical axioms *of* PRA *come in three sorts.*

- *Defining axioms for $+$, $\cdot$, and $<$,*[11]
- *Defining axioms for the new symbols*
  - $\forall x\ \mathsf{Sup}_0(x) = 2x$,
  - $\{\mathsf{Sup}_{n+1}(0) = 1\}$,
  - $\{\forall x\ \mathsf{Sup}_{n+1}(x+1) = \mathsf{Sup}_n(\mathsf{Sup}_{n+1}(x)) \mid n \in \omega\}$,
- *Induction axioms for $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-formulas in the following form: $\forall x\ (\varphi(0) \wedge \forall y{<}x\ (\varphi(y) \to \varphi(y+1)) \to \varphi(x))$.*

The logical axioms *and* rules *are just as usual.*

The functions Sup_i describe on the standard model a well-known hierarchy; Sup_0 is the doubling function, Sup_1 is the exponentiation function, Sup_2 is superexponentiation, Sup_3 is superduperexponentiation and so on. It is also known that the Sup_i form an envelope for PRA, that is, every provably total recursive function of PRA gets eventually majorized by some Sup_i. (Essentially this is Parikh's theorem.) Evidently all terms of PRA are majorizable by a strictly monotone one.

PRA proves all the expected properties of the Sup_i functions like

$$
\begin{aligned}
&\mathsf{Sup}_n(1) = 2,\\
&1 \leq \mathsf{Sup}_{n+1}(y),\\
&x \leq y \to \mathsf{Sup}_n(x) \leq \mathsf{Sup}_n(y),\\
&(n{\leq}m \wedge x{\leq}y) \to \mathsf{Sup}_n(x){\leq}\mathsf{Sup}_m(y),
\end{aligned}
$$

[11] We can take for example Kaye's system PA^- from [Kay91] where in Ax 13 we replace the unbounded existential quantifier by a bounded one.

and so on. Of course PRA proves in a trivial way the totality of all the Sup_i as these symbols form part of our language. We have chosen an equivalent variant of the usual induction axiom so that we end up with a Π_1 axiomatization of PRA. It is easy to see that our definition of PRA is equivalent, or more precisely equi-interpretable, to any other of our definitions of PRA.

Definition 3.11 *The theory $I\Sigma_1$ is the theory that is obtained by adding to* PRA *induction axioms $\varphi(0) \wedge \forall x\, (\varphi(x) \to \varphi(x+1)) \to \forall x\, \varphi(x)$ for all $\Sigma_1(\{\mathsf{Sup}_i\}_{i\in\omega})$-formulas $\varphi(x)$ that may contain additional parameters.*

We will in the sequel often make use of function iteration. If f denotes a function we will denote by f^{it} the (unique) function satisfying the following primitive recursive schema: $f^{\mathsf{it}}(0)=1$, $f^{\mathsf{it}}(x+1)=f(f^{\mathsf{it}}(x))$.

Definition 3.12 *Let $\varphi(x,y)$ be some formula. By $\varphi^{\mathsf{it}}(x,y)$ we denote $\exists \sigma\, \tilde{\varphi}^{\mathsf{it}}(\sigma,x,y)$ where $\tilde{\varphi}^{\mathsf{it}}(\sigma,x,y)$ is the formula*
$\mathsf{Finseq}(\sigma) \wedge (\mathsf{lh}(\sigma) = x+1) \wedge (\sigma_0 = 1) \wedge (\sigma_x = y) \wedge \forall i{<}x\; \varphi(\sigma_i,\sigma_{i+1})$.

Note that $\tilde{\varphi}^{\mathsf{it}}$ is a $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ formula if φ is so. Also note that if PRA proves the functionality of a $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-formula φ, it also proves the functionality of $\tilde{\varphi}^{\mathsf{it}}$, for example by proving by induction on σ that $\forall \sigma \forall x,y,y',\sigma'{\leq}\sigma\; (\tilde{\varphi}^{\mathsf{it}}(\sigma,x,y) \wedge \tilde{\varphi}^{\mathsf{it}}(\sigma',x,y') \to \sigma = \sigma' \wedge y = y')$.

As we will need upper-bounds on sequences of numbers a short remark on coding is due here. By $[a_0,\dots,a_n]$ we will denote the code of the sequence $a_0,\dots,a_n$ of natural numbers provided by the coding technique that we will fix below. By $[a_0,\dots,a_n]\sqcap[b_0,\dots,b_m]$ we will denote the code of the sequence $a_0,\dots,a_n,b_0,\dots,b_m$ that arises from concatenating $b_0,\dots,b_m$ to $a_0,\dots,a_n$ (to the right). Instead of $[a_0,\dots,a_n]\sqcap[b]$ we will often write $[a_0,\dots,a_n]\sqcap b$ if the context excludes any possible confusion. The projection functions are referred to by sub-indexing. So, σ_i will be a_i if $\sigma = [a_0,\dots,a_n]$ and $i \leq n$ and zero if $i > n$, and $n+1$ is called the length of σ. We use the same notation for the projection functions for pairs and trust that no confusion will arise from this. We say that σ is an initial subsequence of σ' if $\sigma = [a_0,\dots a_n]$ and $\sigma' = [a_0,\dots a_n,\dots a_m]$ and $m \geq n$. We denote this by $\sigma \sqsubseteq \sigma'$.

The binary representation of a number a is denoted by $\{a\}_2$. Note that the length of the binary representation can be estimated by $\lceil \log_2(a) \rceil + 1$. We will code the sequence $a_0,\dots,a_n$ by the unique number that has as ternary representation the following sequence:

$$2\{a_n\}_2 2 \dots 2\{a_0\}_2.$$

We define the code of the empty sequence to be zero. The sequence 2,1 will for example be coded by the ternary number 21210 which is the decimal number 210. It is clear that all of the coding, length and projection functions are primitive recursive.

The following theorem tells us what is the difference between PRA and $\mathrm{I}\Sigma_1$ in terms of totality statements of $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-definable functions.

Theorem 3.13 $I\Sigma_1 \equiv \mathrm{PRA}+\{\mathsf{Tot}(\varphi)\to\mathsf{Tot}(\varphi^{\mathsf{it}}) \mid \varphi\in\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})\}$.

Proof of Theorem 3.13. For one inclusion we only need to show that $\mathrm{I}\Sigma_1 \vdash \mathsf{Tot}(\varphi)\to\mathsf{Tot}(\varphi^{\mathsf{it}})$ but this follows easily from a Σ_1 induction on x in $\exists\,\sigma\,\exists\,y\;\tilde{\varphi}^{\mathsf{it}}(\sigma,x,y)$ under the assumption that $\forall\,x\exists\,y\;\varphi(x,y)$. We shall thus concentrate on the harder direction
$\mathrm{PRA}+\{\mathsf{Tot}(\varphi)\to\mathsf{Tot}(\varphi^{\mathsf{it}}) \mid \varphi\in\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})\}\vdash \mathrm{I}\Sigma_1$.

To this end we work in $\mathrm{PRA}+\{\mathsf{Tot}(\varphi)\to\mathsf{Tot}(\varphi^{\mathsf{it}}) \mid \varphi\in\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})\}$ and assume $\exists\,y\;\psi(0,y)\wedge\forall\,x\;(\exists\,y\;\psi(x,y)\to\exists\,y'\;\psi(x+1,y'))$ for some $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-formula $\psi(x,y)$. Our aim is to obtain $\forall\,x\exists\,y\;\psi(x,y)$. First we define a function $\varphi(x,y)$ as follows:

$$\varphi(x,y):=\begin{cases}
(i) & (x=0\wedge y=0)\;\;\vee\\
(ii) & (x=1\wedge\exists\,y_0{\leq}y\\
 & (y=[\langle 2^1,y_0\rangle]\wedge\psi(0,y_0)\wedge\forall\,y'{<}y_0\;\neg\psi(0,y')))\;\;\vee\\
(iii) & (x>1\wedge\mathsf{Finseq}(x)\wedge\\
 & \forall\,i{<}\mathsf{lh}(x)\;((x_i)_0=2^{i+1}\wedge\psi(i,(x_i)_1)\wedge\\
 & \forall\,y'{<}(x_i)_1\;\neg\psi(i,(x_i)_1))\;\wedge\\
 & \exists\,y_0{\leq}y\;(\psi(\mathsf{lh}(x),y_0)\wedge\forall\,y'{<}y_0\;\neg\psi(\mathsf{lh}(x),y')\wedge\\
 & y=x\sqcap[\langle 2^{\mathsf{lh}(x)+1},y_0\rangle]))\;\;\vee\\
(iv) & (x>1\wedge\neg(\mathsf{Finseq}(x)\wedge\\
 & \forall\,i{<}\mathsf{lh}(x)\;((x_i)_0=2^{i+1}\wedge\psi(i,(x_i)_1)\wedge\\
 & \forall\,y'{<}(x_i)_1\;\neg\psi(i,(x_i)_1)))\;\wedge\\
 & y=0).
\end{cases}$$

So, the function f that is described by $\varphi(x,y)$ is always zero unless $x=1$ or x is of the form $[\langle 2^1,y_0\rangle,\ldots\langle 2^{l+1},y_l\rangle]$ for some $l\geq 0$ where we denote here by y_j the minimal number such that $\psi(j,y_j)$. In the latter case f yields $[\langle 2^1,y_0\rangle,\ldots\langle 2^{l+1},y_l\rangle,\langle 2^{l+2},y_{l+1}\rangle]$. Of course the function $\varphi(x,y)$ is tailored for $\varphi^{\mathsf{it}}(x,y)$ to yield on input $x\geq 1$ the sequence $[\langle 2^1,y_0\rangle,\ldots\langle 2^x,y_{x-1}\rangle]$ where again y_j is the minimal number such that $\psi(j,y_j)$.[12]

[12] We could as well have defined φ so that φ^{it} would yield sequences of the form $[\langle 0,y_0\rangle,\ldots,\langle x,y_x\rangle]$. With our current choice we get a nicer growth-rate. Also do we see openings for generalizations of the lemma to more restrictive situations, for example where we work in $\mathrm{EA}+\{\mathsf{Tot}(\varphi)\to\mathsf{Tot}(\varphi^{\mathsf{it}}) \mid \varphi{\in}\Delta_0(\exp)\}$ or on a cut on which exponentiation is a total function. Here EA is Kalmar elementary arithmetic.

Note that we have used 2^x in our definition of φ but this is nothing but an abbreviation of $\mathsf{Sup}_1(x)$. We see that there is no overlap in the clauses (i)-(iv) when it comes to the x values.

By a simple case distinction we see that $\forall x \exists y\ \varphi(x,y)$. If $x = 0$, that is case (i), this is easy. Case (ii) is guaranteed by our assumption $\exists y\ \psi(0,y)$ and the $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ least number principle. If x falls under clause (iii) it follows from our assumption that $\forall x\,(\exists y\ \psi(x,y) \rightarrow \exists y'\ \psi(x+1,y'))$ and the fact that concatenation, pairing and exponentiation are all total functions in our setting. Also in this case the $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ least number principle is used. Case (iv) again is easy.

Again by a simple case distinction we see that φ is functional (one-valued), that is, $\forall x, y, y'\ (\varphi(x,y) \wedge \varphi(x,y') \rightarrow y = y')$. So, $\varphi(x,y)$ denotes a total function. Thus we apply $\mathsf{Tot}(\varphi) \rightarrow \mathsf{Tot}(\varphi^{\mathsf{it}})$ to conclude $\mathsf{Tot}(\varphi^{\mathsf{it}})$.

The idea is to show that $\forall x\ \psi(x, ((f^{\mathsf{it}}(x+1))_x)_1)$ (and thus $\forall x \exists y\ \psi(x,y)$) by means of $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega}, f^{\mathsf{it}})$-induction. To see that we have access to $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega}, f^{\mathsf{it}})$-induction we should, by Corollary 3.8 convince ourselves of two more things: $\mathsf{Mon}(\varphi^{\mathsf{it}})$ and the fact that f^{it} has a $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ graph.

The monotonicity of f^{it} is intuitively clear but we have to show that we can catch this intuition using only $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-induction.

For example, we can first prove by induction on x that all of the $f^{\mathsf{it}}(x+1)$ are "good sequences" where by a good sequence we mean one of the form $[\langle 2^1, y_0\rangle, \ldots, \langle 2^{x+1}, y_x\rangle]$. To make this a $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-induction we should reformulate the statement as
$\forall z \forall \sigma, x, y{\leq}z\ (\tilde{\varphi}^{\mathsf{it}}(\sigma, x+1, y) \rightarrow \mathsf{Goodseq(y)})$ or even
$\forall \sigma \forall x, y{\leq}\sigma\ (\tilde{\varphi}^{\mathsf{it}}(\sigma, x+1, y) \rightarrow \mathsf{Goodseq(y)})$ is sufficient.

Now assume $\tilde{\varphi}^{\mathsf{it}}(\sigma', x', y')$. We will show by induction on x that

$$\forall x{\leq}x'\ \exists \sigma{\leq}\sigma'\ \exists y{\leq}y'\ \tilde{\varphi}^{\mathsf{it}}(\sigma, x'-x, y) \qquad (+)$$

from which monotonicity follows. For if $\varphi^{\mathsf{it}}(x', y')$ and $x_0 \leq x'$ then $\tilde{\varphi}^{\mathsf{it}}(\sigma', x', y')$ for some σ' and $x_0 = x' - x$ some $x \leq x'$, whence by $(+)$ $\tilde{\varphi}^{\mathsf{it}}(\sigma, x'-x, y)$ for some $\sigma \leq \sigma'$ and $y \leq y'$.

Thus we prove $(+)$ under the assumption that $\tilde{\varphi}^{\mathsf{it}}(\sigma', x', y')$.
If $x = 0$ we take $\sigma' = \sigma$ and $y = y'$. For the inductive step, let $\sigma \leq \sigma'$ and $y \leq y'$ be such that $\tilde{\varphi}^{\mathsf{it}}(\sigma, x'-x, y)$. We assume that $x+1 \leq x'$ hence $\mathsf{lh}(\sigma) > 1$, for if not the solution is trivial. By σ_{-1} we denote the sequence that is obtained from σ by deleting the last element. Clearly $\tilde{\varphi}^{\mathsf{it}}(\sigma_{-1}, x'-x-1, (\sigma_{-1})_{x'-x-1})$ and $\varphi((\sigma_{-1})_{x'-x-1}, y)$. Thus $(\sigma_{-1})_{x'-x-1}$ is a good sequence which implies that clause (iii) in the definition of φ is used to determine y. Consequently $(\sigma_{-1})_{x'-x-1} \sqsubseteq y$ and thus $(\sigma_{-1})_{x'-x-1} \leq y \leq y'$. Moreover we note that $\sigma_{-1} \sqsubseteq \sigma$ and thus $\sigma_{-1} \leq \sigma \leq \sigma'$.

We now want to show the $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-ness of $\varphi^{\mathsf{it}}(x,y)$ by providing an upper-bound on the σ in $\tilde{\varphi}^{\mathsf{it}}(\sigma,x,y)$. If $\tilde{\varphi}^{\mathsf{it}}(\sigma,x,y)$ we have $\mathsf{lh}(\sigma) = x+1$, $(\sigma)_x = y \geq 2^x$ and moreover f^{it} is, as we have just showed, monotone. As $(\sigma)_x = y \geq 2^x$ we have that $x \leq \lfloor \log_2(y) \rfloor \leq \lceil \log_2(y) \rceil$. We will estimate the length of the binary representation of a number y by $\lceil \log_2(y) \rceil + 1$. So, σ can be roughly estimated by the following numbers written in base 3 :

$$\begin{array}{ccl}
\overbrace{2\ \{y\}_2 2\ldots 2\ \{y\}_2}^{x+1 \text{ times}} & \leq & \overbrace{2\ \{y\}_2 2\ldots 2\ \{y\}_2}^{\lceil \log_2(y)\rceil+1 \text{ times}} \leq \\
\overbrace{22\ldots 22}^{(\lceil \log_2(y)\rceil+1)^2 \text{ times}} & \leq & 100\ \overbrace{00\ldots 00}^{2\cdot\lceil \log_2(y)\rceil^2 \text{ times}} \quad .
\end{array}$$

This latter number can bounded from above by the following numbers which we will write now in base 10 notation $3^{2\cdot\lceil \log_2(y)\rceil^2+2} \leq \mathsf{Sup}_2(y+3)$. Thus $\varphi^{\mathsf{it}}(x,y)$ can be written as $\exists\, \sigma \leq \mathsf{Sup}_2(y+3)\ \tilde{\varphi}^{\mathsf{it}}(\sigma,x,y)$ being $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$. By a simple $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega}, f^{\mathsf{it}})$ induction on x we now prove $\forall\, x\ \psi(x, ((f^{\mathsf{it}}(x+1))_x)_1)$.

QED

We note that we filled the gap between PRA and $\mathrm{I}\Sigma_1$ by transforming an admissible rule of PRA to axiom form. Indeed $\mathsf{Tot}(\varphi) \mathrel{|\!\sim} \mathsf{Tot}(\varphi^{\mathsf{it}})$ is an admissible rule of PRA. For if $\mathrm{PRA} \vdash \mathsf{Tot}(\varphi)$ then f is a primitive recursive function as is well known. But f^{it} is constructed from f by a simple recursion. Thus f^{it} is primitive recursive and hence provably total in PRA. The same phenomenon occurs in passing from $\mathrm{I}\Sigma_1^R$ to $\mathrm{I}\Sigma_1$ where the (trivially) admissible Σ_1 induction rule is added in axiom form to PRA to obtain $\mathrm{I}\Sigma_1$.

The fact that we allow for variables in Theorem 3.13 is essential. For if not, the logical complexity of $\mathrm{PRA} + \{\mathsf{Tot}(\varphi) \to \mathsf{Tot}(\varphi^{\mathsf{it}}) \mid \varphi \in \Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})\}$ would be[13] Δ_3 and so would be the logical complexity of $\mathrm{I}\Sigma_1$. But the well-known fact that $\mathrm{I}\Sigma_1 \vdash \mathrm{RFN}_{\Pi_3}(\mathrm{EA})$ (where EA stands for elementary arithmetic, that is, $I\Delta_0$ plus the totality of the exponential function) would contradict the fact that $\mathrm{RFN}_{\Pi_3}(\mathrm{EA})$ is not proved by any consistent Σ_3 extension of EA.[14] A parameter-free version of $\mathrm{PRA} + \{\mathsf{Tot}(\varphi) \to \mathsf{Tot}(\varphi^{\mathsf{it}}) \mid \varphi \in \Delta_0^-(\{\mathsf{Sup}_i\}_{i\in\omega})\}$ will be equivalent to parameter-free Σ_1-induction, $\mathrm{I}\Sigma_1{}^-$.

[13] Actually we should be more careful here as we work in a richer language. However this makes no essential difference as all the Sup_n are Δ_1 definable over EA.

[14] Let S be some collection of Σ_3 sentences such that $\mathrm{EA}+S$ extends $\mathrm{EA}+\mathrm{RFN}_{\Pi_3}(\mathrm{EA})$. We also have $\mathrm{EA}+S \vdash \forall\, x\ (\Box_{\mathrm{EA}}(\mathsf{Tr}_{\Pi_3}(\dot{x})) \to \mathsf{Tr}_{\Pi_3}(x))$. By compactness we have for some particular Σ_3-sentence σ that $\mathrm{EA}+\sigma \vdash \forall\, x\ (\Box_{\mathrm{EA}}(\mathsf{Tr}_{\Pi_3}(\dot{x})) \to \mathsf{Tr}_{\Pi_3}(x))$. Consequently, $\mathrm{EA}+\sigma \vdash \Box_{\mathrm{EA}}(\mathsf{Tr}_{\Pi_3}(\ulcorner\neg\sigma\urcorner)) \to \mathsf{Tr}_{\Pi_3}(\ulcorner\neg\sigma\urcorner)$ and thus $\mathrm{EA} \vdash \sigma \to (\Box_{\mathrm{EA}}(\ulcorner\neg\sigma\urcorner) \to \neg\sigma)$. But we also have $\mathrm{EA} \vdash \neg\sigma \to (\Box_{\mathrm{EA}}(\ulcorner\neg\sigma\urcorner) \to \neg\sigma)$ hence $\mathrm{EA} \vdash (\Box_{\mathrm{EA}}(\ulcorner\neg\sigma\urcorner) \to \neg\sigma$. Löb's rule gives us $\mathrm{EA} \vdash \neg\sigma$ in which case $\mathrm{EA}+S$ is inconsistent.

We now come to prove a lemma that tells us when a model of PRA is also a model of $\mathrm{I}\Sigma_1$. This lemma is formulated in terms of majorizability behavior of some total functions. A total function of a model M is a relation $\varphi(x, y)$ (possibly with parameters from M) for which $M \models \mathsf{Tot}(\varphi)$. Often we will write $f \leq g$ as short for $\forall\, x(\exists\, y\ \varphi(x, y) \to \exists\, y'\ (\chi(x, y') \wedge y \leq y'))$ and say that f is majorized by g. Thus if $f \leq g$ we automatically have $\mathsf{Tot}(\varphi) \to \mathsf{Tot}(\chi)$.

Lemma 3.14 *Let $\mathcal{M}$ be a model of* PRA*. If every $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ definable total function (with parameters) of $\mathcal{M}$ is majorized by $m + \mathsf{Sup}_n$ for some $m \in \mathcal{M}$ and some $n \in \omega$, then $\mathcal{M}$ is also a model of $I\Sigma_1$.*

Proof of Lemma 3.14. Let $\mathcal{M}$ be satisfying our conditions. To see that $\mathcal{M} \models \mathrm{I}\Sigma_1$ we need in the light of Theorem 3.13 to show that $\mathcal{M} \models \mathsf{Tot}(\varphi) \to \mathsf{Tot}(\varphi^{\mathsf{it}})$ for any $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ function φ with parameters in $\mathcal{M}$. So, we consider some $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ function f such that $\mathcal{M} \models \mathsf{Tot}(\varphi)$. We choose $m \in \mathcal{M} \setminus \{0\}$ and $n \in \omega$ large enough so that

(*a.*) $\mathcal{M} \models f \leq m + \mathsf{Sup}_n$,

(*b.*) $\mathcal{M} \models \forall\, x\ (m + \mathsf{Sup}_{n+1}(mx + m + 1) \leq \mathsf{Sup}_{n+1}(mx + m + m))$.

The second condition is automatically satisfied if m is a non-standard element.

An easy $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-induction shows that $(m + \mathsf{Sup}_n)^{\mathsf{it}}(x) \leq \mathsf{Sup}_{n+1}(mx + m)$. (Remember that we have excluded $m = 0$.) The case $x = 0$ is trivial as $1 \leq \mathsf{Sup}_{n+1}(m)$. For the inductive step we see that[15]

$$\begin{array}{ll}
(m + \mathsf{Sup}_n)^{\mathsf{it}}(x + 1) & = \\
(m + \mathsf{Sup}_n)((m + \mathsf{Sup}_n)^{\mathsf{it}}(x)) & \leq_{\text{i.h.}} \\
m + \mathsf{Sup}_n(\mathsf{Sup}_{n+1}(mx + m)) & \leq_{\text{def.}} \\
m + \mathsf{Sup}_{n+1}(mx + m + 1) & \leq_{(b.)} \\
\mathsf{Sup}_{n+1}(mx + m + m) = \mathsf{Sup}_{n+1}(m(x + 1) + m). &
\end{array}$$

We can use the obtained bounds to show the totality of f^{it} by estimating the size of σ that witnesses $\tilde{\varphi}^{\mathsf{it}}(\sigma, x, y)$. We know (outside PRA) that σ is of the form

[15] This looks like a legitimate induction but remember that $(m + \mathsf{Sup}_n)^{\mathsf{it}}$ has an a priori $\Sigma_1(\{\mathsf{Sup}_i\}_{i\in\omega})$ definition. The argument should thus be encapsulated in a $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-induction, for example by proving $\forall\, z \forall\, \sigma, x, y{\leq}z\ \widetilde{(m + \mathsf{Sup}_n)}^{\mathsf{it}}(\sigma, x, y) \to y \leq \mathsf{Sup}_{n+1}(mx + m)$. The essential reasoning though boils down to the argument given here.

$$
\begin{array}{ll}
[1, f(1), f(f(1)), \ldots, f^x(1)] & \leq \\
[1, m + \mathsf{Sup}_n(1), m + \mathsf{Sup}_n(f(1)), \ldots, m + \mathsf{Sup}_n(f^{x-1}(1))] & \leq \\
[1, m + \mathsf{Sup}_n(1), (m + \mathsf{Sup}_n)^2(1), \ldots, (m + \mathsf{Sup}_n)^2(f^{x-2}(1))] & \leq \\
\vdots & \vdots \\
[1, m + \mathsf{Sup}_n(1), (m + \mathsf{Sup}_n)^2(1), \ldots, (m + \mathsf{Sup}_n)^x(1)] & \leq \\
[(m + \mathsf{Sup}_n)^x(1), \ldots, (m + \mathsf{Sup}_n)^x(1)] & \leq \\
[\mathsf{Sup}_{n+1}(mx + m), \ldots, \mathsf{Sup}_{n+1}(mx + m)] &
\end{array}
$$

Every time we used dots here in our informal argument, some $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-induction should actually be applied. To neatly formalize our reasoning we choose some $k \in \omega$ large enough for our n and m such that (in $\mathcal{M}$)

(c.) $[1] \leq \mathsf{Sup}_{n+k}(2m)$

(d.) $\mathsf{Sup}_{n+k}(m(x+1)+m) \sqcap [\mathsf{Sup}_{n+1}(m(x+1)+m)] \leq \mathsf{Sup}_{n+k}(m(x+2)+m)$[16]

With these choices for m, n and k it is easy to prove by $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-induction that

$$\forall\, x\, \exists\, \sigma \leq \mathsf{Sup}_{n+k}(m(x+1)+m)\, \exists\, y \leq \mathsf{Sup}_{n+1}(mx+m)\ \tilde{\varphi}^{\mathsf{it}}(\sigma, x, y).$$

If $x = 0$ then $\tilde{\varphi}^{\mathsf{it}}([1], 0, 1)$ and by (c.) we have $[1] \leq \mathsf{Sup}_{n+k}(m(0+1)+m)$. Also $1 \leq \mathsf{Sup}_{n+1}(m)$. Now suppose $\tilde{\varphi}^{\mathsf{it}}(\sigma, x, y)$ with σ and y below their respective bounds. We have by the definition of $\tilde{\varphi}^{\mathsf{it}}$ that $\tilde{\varphi}^{\mathsf{it}}(\sigma \sqcap [f(y)], x+1, f(y))$ (again we do as if we had f available in our language). We need to show that the new values do not grow too fast. But,

$$
\begin{array}{cccc}
f(y) & \leq_{\text{i.h.}} & f(\mathsf{Sup}_{n+1}(mx+m)) & \leq_{(a.)} \\
m + \mathsf{Sup}_n(\mathsf{Sup}_{n+1}(mx+m)) & \leq_{(b.)} & \mathsf{Sup}_{n+1}(m(x+1)+m) &
\end{array}
$$

[16]It is not hard to convince oneself that such a k does exist. As a feature of our coding we can easily estimate the value of $\sigma \sqcap [y]$. It will be just $\sigma+$ "something else". This "something else" is the contribution of $\{y\}_2$ whose length can be bounded by $\lceil \log_2(y) \rceil + 1 + 1$ (the extra $+1$ is because in our coding protocol we put a 2 in front). The place where this contribution appears is dependent on σ whose length can be estimated by $\lceil \log_3(\sigma) \rceil + 1$. The value of $\{y\}_2$ with a 2 in front interpreted as a base three number can be bounded by $3^{\lceil \log_2(y) \rceil + 3}$. The place where it occurs yields a bound on that "something else" of $3^{\lceil \log_2(y) \rceil + \lceil \log_3(\sigma) \rceil + 4}$. Thus $\sigma \sqcap [y] \leq \sigma + 3^{\lceil \log_2(y) \rceil + \lceil \log_3(\sigma) \rceil + 4}$. In turn $\sigma + 3^{\lceil \log_2(y) \rceil + \lceil \log_3(\sigma) \rceil + 4} \leq \sigma + 3^{\sigma + y + 4}$. If $k = 1$ then we get our concatenation situation with $\sigma(x) = y(x+1)$ where we denote by $\sigma(x)$ the binding function on σ, that is, $\mathsf{Sup}_{n+1}(m(x+1)+m)$. But $f(x+1) \geq f(x) + 3^{2f(x)+4}$ is a recursion that is satisfied if m and n are large enough and $f(x) = \mathsf{Sup}_{n+1}(m(x+1)+m)$.

as we have seen before. By $(d.)$ we get that

$$\begin{array}{lll} \sigma \sqcap [f(y)] & \leq_{\text{i.h.}} & \mathsf{Sup}_{n+k}(m(x+1)+m) \sqcap [\mathsf{Sup}_{n+1}(m(x+1)+m)] \\ & \leq_{(d.)} & \mathsf{Sup}_{n+k}(m(x+2)+m). \end{array}$$

QED

3.3 The actual proof of Parsons' theorem

Before we give the proof of Parsons' theorem we first agree on some model theoretic notation.

We recall the definition of M' being a 1-elementary extension of M, denoted by $M \prec_1 M'$. This means that $M \subseteq M'$ and that for $\vec{m} \in M$ and $\sigma(\vec{y}) \in \Sigma_1$ we have $M \models \sigma(\vec{m}) \quad \Leftrightarrow M' \models \sigma(\vec{m})$. In this case we also say that M is a 1-elementary submodel of M'. It is easy to see that

$$M \prec_1 M' \;\Leftrightarrow\; [M \models \sigma(\vec{m}) \Rightarrow M' \models \sigma(\vec{m})] \;\text{ for all } \sigma(\vec{y}) \in \Sigma_2.$$

A 1-elementary chain is a sequence $M_0 \prec_1 M_1 \prec_1 M_2 \prec_1 \ldots$. It is well known that the union of a 1-elementary chain is a 1-elementary extension of every model in the chain. It is worthy to note that in a 1-elementary chain the truth of Σ_2-sentences (with parameters) is preserved from left to right and the truth of Π_2-sentences (without parameters) is preserved from right to left.

By $\mathsf{Th}(M, C)$ we denote the first-order theory of M with all constants from C added to the language. This makes sense if we know how to interpret the constants of C in M.
We also recall the definition of the collection principle.

$$B\Gamma := \{\forall\, x{<}t\, \exists\, y\; \varphi(x,y) \to \exists\, s\, \forall\, x{<}t\, \exists\, y{<}s\; \varphi(x,y) \mid \varphi \in \Gamma\}$$

together with a minimum of arithmetical axioms, e.g. PA^-.

We now come to the actual proof of Theorem 3.1.

Proof of Parsons' Theorem, Theorem 3.1. Let a countable model $M \models \mathrm{PRA} + \sigma$ be given with $\sigma \in \Sigma_2$. We will construct a countable model M' of $\mathrm{I}\Sigma_1 + \sigma$ using Lemma 3.14.

Our strategy will be to make any $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ definable total function of M that is not bounded by any of the $m + \mathsf{Sup}_n$ ($n \in \omega$, $m \in M$) either bounded by some $m + \mathsf{Sup}_n$ ($n \in \omega$, $m \in M'$) or not total in the PRA-model M'. The model M' will be the union of a Σ_1-elementary chain of models $M = M_0 \prec_1 M_1 \prec_1 M_2 \ldots \prec_1 M' = \cup_{i\in\omega} M_i$.

At each stage either the boundedness of a total $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ definable function is guaranteed (a Π_1-sentence: $\forall\, x, y\; (\varphi(x,y) \to y \leq$

$m + \mathsf{Sup}_n(x)))$ or its non-totality (a Σ_2-sentence: $\exists x \forall y\, \neg\varphi(x, y)$). As we shall work with a 1-elementary chain of models, functions that are dealt with need no more attention further on in the chain. Their interesting properties, that is boundedness or non-totality, are stable. By choosing the order in which functions are dealt with in a good way, eventually all total funtions of all models M_i will be considered. We shall see that as a result of this process every total function in M' that is $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ definable is bounded by some $m+\mathsf{Sup}_n$ with $m \in M'$ and $n \in \omega$.

To properly order the functions that we shall deal with, we fix a bijective pairing function in this proof satisfying $x, y \leq \langle x, y\rangle$. We write $f_{n0}, f_{n1}, f_{n2}, \ldots$ for the list of the (countably many) total $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ definable functions of M_n. We emphasize that we allow the functions f_{ni} to contain parameters from M_n. Furthermore we define g_n to be f_{ab} for the unique $a, b \in \omega$ such that $\langle a, b\rangle = n$.

We define $M_0 := M$.

We will define M_{n+1} to be such that g_n becomes (or remains) either bounded or non-total in it and $M_n \prec_1 M_{n+1}$. If we can do so, we are done. For suppose $M = M_0 \models \mathrm{PRA} + \sigma$. As PRA is Π_1 axiomatizable in the language containing the $\{\mathsf{Sup}_i\}_{i\in\omega}$[17] we get that $M' \models \mathrm{PRA}$ and likewise $M' \models \sigma$.

If now $M' \models \mathsf{Tot}(\varphi)$ for some $\varphi \in \Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$, we see that for some n, $M_n \models \mathsf{Tot}(\varphi)$ as soon as M_n contains all the parameters that occur in φ. Thus $f = g_m$ for some $m \geq n$. Thus in M_{m+1} the function f will be surely majorized, for $M_{m+1} \models \neg\mathsf{Tot}(\varphi) \Rightarrow M' \models \neg\mathsf{Tot}(\varphi)$. Consequently $M' \models f \leq m' + \mathsf{Sup}_k$ for some $m' \in M_{m+1} \subseteq M'$, $k \in \omega$. By Lemma 3.14 we see that $M' \models \mathrm{I}\Sigma_1$.

If $M_n \models g_n \leq m + \mathsf{Sup}_k$ for some $m \in M_n$ and $k \in \omega$ we set $M_{n+1} := M_n$. Clearly $M_n \prec_1 M_{n+1}$ and g_n is bounded in M_{n+1}.

So, suppose that g_n is total in M_n and that $M_n \models \neg(g_n \leq m + \mathsf{Sup}_k)$ for all $m \in M_n$ and all $k \in \omega$. We obtain our required model M_{n+1} in

[17] By Matiyasevich's theorem we know that we also have a purely universal formulation of PRA. (We assume that the theorem extends to a theory formulated in a richer language.) We can consider PRA^M, the theory that is axiomatized by the pure Π^0_1 formulas that are provided by Matiyasevich's theorem. A natural question to ask is whether PRA^M and PRA are equivalent. We think this is highly unlikely. For PRA^M gives us an open formulation of PRA. By Herbrand's theorem we would then be able to write the graph of any provably total recursive function of PRA, that is, the graph of any primitive recursive function, as a disjunction of terms of PRA. However it seems unplausible to expect that for example $y = x!$, the graph of the factorial function, can be written as such a disjunction.

two steps.

Step 1.

First we go from $M_n \prec_1 M_{n1} \models \mathsf{B}\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})(+\text{PRA})$. To this purpose we temporarily expand our language. For every Σ_2 formula $\exists x\ \varphi(x)$ such that $M_n \models \exists x\ \varphi(x)$ we add a fresh constant $c_{\varphi(x)}$ to our language and call the set of all these contants C. We will fix some interpretation of the $c_{\varphi(x)}$ in M_n such that $M_n \models \varphi(c_{\varphi(x)})$. Let d be a new constant symbol. It is clear that $\mathsf{Th}(M_n, \{m\}_{m\in M_n} \cup \{c_\varphi\}_{c_\varphi\in C}) \cup \{d > \mathsf{Sup}_k(c_\varphi) \mid k \in \omega,\ c_\varphi \in C\}$ is finitely satisfiable, namely in M_n. Thus there is some countable model M_{n0} containing M_n in a natural way satisfying the whole set.

Let M_{n1} be the (initial) submodel of M_{n0} with domain $\{m \in M_{n0} \mid \exists\, k{\in}\omega\, \exists\, c_\varphi{\in}C\ m \leq \mathsf{Sup}_k(c_\varphi)\}$. Clearly M_{n1} is indeed a submodel, that is, it is closed under all the Sup_k. For if $m \leq \mathsf{Sup}_l(c_\varphi)$ then $\mathsf{Sup}_k(m) \leq \mathsf{Sup}_k(\mathsf{Sup}_l(c_\varphi)) \leq \mathsf{Sup}_{k+l+2}(c_\varphi)$. We see that M_{n1} is a model of PRA as PRA is Π_1 axiomatized.[18]

As all the c_φ live in M_{n1} and $M_n \subseteq M_{n1}$ we get $M_n \prec_1 M_{n1}$. For[19],

$$\begin{array}{lll}
M_n & \models & \exists x\ \varphi(x) \quad ,\varphi(x) \in \Pi_1 \quad \Rightarrow \\
M_n & \models & \varphi(c_\varphi) \quad \Rightarrow \\
M_{n0} & \models & \varphi(c_\varphi) \quad \Rightarrow \\
M_{n1} & \models & \varphi(c_\varphi) \quad \Rightarrow \\
M_{n1} & \models & \exists x\ \varphi(x).
\end{array}$$

We now see that $M_{n1} \models \mathsf{B}\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$.
So, suppose $M_{n1} \models \forall\, x{<}t\, \exists\, y\ \varphi(x,y)$ for some $t \in M_{n1}$ and $\varphi \in \Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$. Clearly $M_{n0} \models \forall\, x{<}t\, \exists\, y{<}d\ \varphi(x,y)$ for some $d \in M_{n0}$, actually for any $d \in M_{n0} \setminus M_{n1}$. Now by the $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ minimal number principle we get a minimal d_0 such that $M_{n0} \models \forall\, x{<}t\, \exists\, y{<}d_0\ \varphi(x,y)$. If d_0 were in $M_{n0} \setminus M_{n1}$, then $d_0 - 1$ would also suffice as a bound on the y's. The minimality of d_0 thus imposes that $d_0 \in M_{n1}$. Consequently $M_{n1} \models \exists\, d_0 \forall\, x{<}t\, \exists\, y{<}d_0\ \varphi(x,y)$ and $M_{n1} \models \mathsf{B}\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$.[20]

[18] We can also see the validity of the induction axioms by elementary means. For suppose $M_{n1} \models \psi(0) \wedge \forall\, x\ (\psi(x) \rightarrow \psi(x+1))$ for some $\psi(x) \in \Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$.

If $M_{n0} \models \forall\, x\ (\psi(x) \rightarrow \psi(x+1))$ we have by $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-induction in M_{n0} that $M_{n0} \models \forall\, x\ \psi(x)$, which is Π_1, so, $M_{n1} \models \forall\, x\ \psi(x)$. If $M_{n0} \not\models \forall\, x\ (\psi(x) \rightarrow \psi(x+1))$ then by the least number principle for $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-formulas there is a minimal $a \in M_{n0}$ such that $M_{n0} \models \psi(a) \wedge \neg\psi(a+1)$. Because of $M_{n1} \models \forall\, x\ (\psi(x) \rightarrow \psi(x+1))$, we see that a cannot be in M_{n1}. We thus see that $\forall\, x{\leq}a\ \psi(x)$ and again we obtain $M_{n1} \models \forall\, x\ \psi(x)$.

[19] Note that adding the $\{c_\varphi\}_{c_\varphi\in C}$ is not necessary as all the c_φ are already contained in the $\{m\}_{m\in M_n}$. It only provides us with a richer vocabulary.

[20] By coding techniques we could obtain here $M_{n1} \models \mathsf{B}\Sigma_1(\{\mathsf{Sup}_i\}_{i\in\omega})$.

Step 2.

Secondly we go from $M_{n1} \models \mathsf{B}\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})(+\mathrm{PRA})$[21] to a model $M_{n1} \prec_1 M_{n3} \models \mathrm{PRA} + \neg\mathsf{Tot}(\chi_n)$. If $M_{n1} \models \neg\mathsf{Tot}(\chi_n)$ nothing has to be done and we take $M_{n3} = M_{n1}$, so, we assume that $M_{n1} \models \mathsf{Tot}(\chi_n)$ here. We consider the set

$$\Gamma := \mathsf{Th}(M_{n1}, \{m\}_{m\in M_{n1}}) \cup \{g_n(c) > m + \mathsf{Sup}_k(c) \mid m \in M_{n1},\ k \in \omega\}$$

with c a fresh constant symbol.

As g_n is not majorizable in M_{n1} we see that any finite subset of Γ is satisfiable whence Γ is satisfiable. Let M_{n2} be a countable model of Γ in which we can naturally embed M_{n1}.

We will now see that $c > M_{n1}$. For suppose $c \leq m \in M_{n1}$. Then $M_{n1} \models \forall\, x{\leq}m\, \exists\, z\ \ g_n(x){=}z$.[22] By $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$-collection we get $M_{n1} \models \exists\, d_0\, \forall\, x{\leq}m\, \exists\, z{\leq}d_0\ g_n{=}z$. But then $M_{n1} \models g_n(c) \leq d_0$ whence $M_{n1} \models \neg(g_n(c) > d_0 + \mathsf{Sup}_k(c))$. A contradiction.

Define M_{n3} to be the (initial) submodel of M_{n2} with domain $\{m \in M_{m2} \mid \exists\, k \in \omega\ \ M_{n2} \models m < \mathsf{Sup}_k(c)\}$. As $c \geq M_{n1}$ we get $M_{n1} \subseteq M_{n3}$. We now see that $M_{n1} \prec_1 M_{n3}$. For suppose $M_{n1} \models \exists\, x\ \varphi(x)$ with $\varphi(x) \in \Pi_1$ then $M_{n1} \models \varphi(m_0)$ for some $m_0 \in M_{n1}$. Consequently $M_{n2} \models \varphi(m_0)$ and as $M_{n3} \subset_e M_{n2}$ and $\varphi(m_0) \in \Pi_1$, also $M_{n3} \models \varphi(m_0)$ whence $M_{n3} \models \exists\, x\ \varphi(x)$. Clearly $M_{n3} \models \neg\mathsf{Tot}(\chi_n)$ as $g_n(c)$ cannot have a value in M_{n3}. M_{n+1} will be the reduct of M_{n3} to the original language.

QED

References

[Avi02] J Avigad. Saturated models of universal theories. *Annals of Pure and Applied Logic*, 118(3):219–234, 2002.

[Bek96] L.D. Beklemishev. Induction rules, reflection principles, and provably recursive functions. Logic Group Preprint Series 168, University of Utrecht, 1996.

[Bek97] L.D. Beklemishev. Induction rules, reflection principles, and provably recursive functions. *Annals of Pure and Applied Logic*, 85:193–242, 1997.

[Bek99] L.D. Beklemishev. Proof-theoretic analysis by iterated reflection. Preprint, to appear in 2002 in Archive for Mathematical Logic, 1999.

[21] Or from the reduct of M_{n1} to the original language for that matter.
[22] We actually should substitute the $\Delta_0(\{\mathsf{Sup}_i\}_{i\in\omega})$ graph of g_n here.

[Bus98] S.R. Buss. First-order proof theory of arithmetic. In S.R. Buss, editor, *Handbook of Proof Theory*, pages 79–148, Amsterdam, 1998. Elsevier, North-Holland.

[End72] H.B. Enderton. *A mathematical introduction to logic.* Academic Press, New York, 1972.

[Fer02] F. Ferreira. Yet another proof of Parsons' theorem. Not yet published, 2002.

[GD82] H. Gaifman and C. Dimitracopoulos. Fragments of Peano's arithmetic and the MDRP theorem. In *Logic and algorithmic (Zurich, 1980), (Monograph. Enseign. Math., 30)*, pages 187–206. Genève, University of Genève, 1982.

[HB68] D. Hilbert and P. Bernays. *Grundlagen der Mathematik, Vols. I and II, 2d ed.* Springer-Verlag, Berlin, 1968.

[HP93] P. Hájek and P. Pudlák. *Metamathematics of First Order Arithmetic.* Springer-Verlag, Berlin, Heidelberg, New York, 1993.

[Ign90] A.D. Ignjatovic. *Fragments of first and Second Order Arithmetic and Length of Proofs.* PhD thesis, University of California, Berkeley, 1990.

[Joo01] J.J. Joosten. Arithmetics, a course by Lev Beklemishev, personal notes of Joost J. Joosten., 2001. Lecture notes online at `http://www.phil.uu.nl/~joosten/`.

[Kay91] R. Kaye. *Models of Peano Arithmetic.* Oxford University Press, Oxford, 1991.

[Min72] G Mints. Quantifier-free and one-quantifier systems. *Journal of Soviet Mathematics*, 1:71–84, 1972. First published in Russian in 1971.

[Par70] C. Parsons. On a number-theoretic choice schema and its relation to induction. In A. Kino, J. Myhill, and R.E. Vessley, editors, *Intuitionism and Proof Theory*, pages 459–473. North Holland, Amsterdam, 1970.

[Par72] C. Parsons. On n-quantifier induction. *Journal of Symbolic Logic*, 37(3):466–482, 1972.

[Sch77] H. Schwichtenberg. Some applications of cut-elimination. In J. Barwise, editor, *Handbook of Mathematical Logic*, pages 867–896. North Holland, Amsterdam, 1977.

[Sch87] D.G. Schwarz. On the equivalence between logic-free and logic-bearing systems of primitive recursive arithmetic. *Zeitschrift f. math. Logik und Grundlagen d. Math.*, 33:245–253, 1987.

[Sie91] W. Sieg. Herbrand analyses. *Archive for Mathematical Logic*, 30:409–441, 1991.

[Sim99] S. G. Simpson. *Subsystems of Second Order Arithmetic*. Springer-Verlag, 1999.

[Sko67] T. Skolem. The foundations of elementary arithmetic established by means of the recursive mode of thought, without the use of apparent variables ranging over infinite domains. In J. van Heijenoort, editor, *From Frege to Godel*, pages 302–333. Iuniverse, Harvard, 1967.

[Tai81] W. Tait. Finitism. *Journal of Philosophy*, 78:524–546, 1981.

[Tak75] G. Takeuti. *Proof Theory*. North–Holland, Amsterdam, 1975.

[Vis90] A. Visser. Notes on $\mathrm{I}\Sigma_1$. Unpublished manuscript, 1990?

[Zam96] D Zambella. Notes on polynomial bounded arithmetic. *Journal of Symbolic Logic*, 61:942–966, 1996.