<table>
<thead>
<tr>
<th>Title</th>
<th>Categorical characterizations of the natural numbers require primitive recursion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Kolodziejczyk, Leszek Aleksander; Yokoyama, Keita</td>
</tr>
<tr>
<td>Citation</td>
<td>Annals of Pure and Applied Logic, 166(2): 219-231</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2014-10-30</td>
</tr>
<tr>
<td>Type</td>
<td>Journal Article</td>
</tr>
<tr>
<td>Text version</td>
<td>author</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10119/12317">http://hdl.handle.net/10119/12317</a></td>
</tr>
</tbody>
</table>

**Description:**

NOTICE: This is the author's version of a work accepted for publication by Elsevier. Leszek Aleksander Kolodziejczyk and Keita Yokoyama, Annals of Pure and Applied Logic, 166(2), 2014, 219-231, http://dx.doi.org/10.1016/j.apal.2014.10.003
Categorical characterizations of the natural numbers require primitive recursion

Leszek Aleksander Kołodziejczyk* Keita Yokoyama†

October 14, 2014

Abstract

Simpson and Yokoyama [Ann. Pure Appl. Logic 164 (2013), 284–293] asked whether there exists a characterization of the natural numbers by a second-order sentence which is provably categorical in the theory RCA₀. We answer in the negative, showing that for any characterization of the natural numbers which is provably true in WKL₀, the categoricity theorem implies \( \Sigma^0_1 \) induction.

On the other hand, we show that RCA₀ does make it possible to characterize the natural numbers categorically by means of a set of second-order sentences. We also show that a certain \( \Pi^1_2 \)-conservative extension of RCA₀ admits a provably categorical single-sentence characterization of the naturals, but each such characterization has to be inconsistent with \( \text{WKL}_0 + \text{superexp} \).

Inspired by a question of Väänänen (see e.g. [Vään12] for some related work), Simpson and the second author [SY13] studied various second-order characterizations of \( \langle \mathbb{N}, S, 0 \rangle \), with the aim of determining the reverse-mathematical strength of their respective categoricity theorems. One of the general conclusions is that the strength of a categoricity theorem depends heavily on the characterization. Strikingly, however, each of the categoricity theorems considered in [SY13] implies RCA₀, even over the much weaker base theory RCA₀, that is, RCA₀ with \( \Sigma^0_1 \) induction replaced by \( \Delta^0_1 \) induction in the language with exponentiation. (For RCA₀*, see [SS86].)

This leads to the following question.

---

*Institute of Mathematics, University of Warsaw, Banacha 2, 02-097 Warszawa, Poland, lak@impan.pl. Supported in part by Polish National Science Centre grant no. 2013/09/B/ST1/04390.

†School of Information Science, Japan Advanced Institute of Science and Technology, Nomi, Ishikawa, Japan, y-keita@jaist.ac.jp. Supported in part by JSPS Grant-in-Aid for Research Activity Start-up grant no. 25887026.
Question 1. [SY13, Question 5.3, slightly rephrased] Does RCA₀ prove the existence of a second-order sentence or set of sentences $T$ such that $\langle \mathbb{N}, S, 0 \rangle$ is a model of $T$ and all models of $T$ are isomorphic to $\langle \mathbb{N}, S, 0 \rangle$? One may also consider the same question with RCA₀ replaced by $\Pi^0_2$-conservative extensions of RCA₀.

Naturally, to have any hope of characterizing infinite structures categorically, second-order logic has to be interpreted according to the standard semantics (sometimes also known as strong or Tarskian semantics), as opposed to the general (or Henkin) semantics. In other words, a second-order quantifier $\forall X$ really means “for all subsets of the universe” (or, as we would say in a set-theoretic context, “for all elements of the power set of the universe”).

Question 1 admits multiple versions depending on whether we focus on RCA₀ or consider other $\Pi^0_2$-equivalent theories and whether we want the characterizations of the natural numbers to be sentences or sets of sentences. The most basic version, restricted to RCA₀ and single-sentence characterizations, would read as follows:

**Question 2.** Does there exist a second-order sentence $\psi$ in the language with one unary function $f$ and one constant $c$ such that RCA₀ proves: (i) $\langle \mathbb{N}, S, 0 \rangle \models \psi$, and (ii) for every $\langle A, f, c \rangle$, if $\langle A, f, c \rangle \models \psi$, then there exists an isomorphism between $\langle \mathbb{N}, S, 0 \rangle$ and $\langle A, f, c \rangle$?

We answer Question 2 in the negative. In fact, characterizing $\langle \mathbb{N}, S, 0 \rangle$ not only up to isomorphism, but even just up to equicardinality of the universe, requires the full strength of RCA₀. More precisely:

**Theorem 1.** Let $\psi$ be a second-order sentence in the language with one unary function $f$ and one individual constant $c$. If WKL₀ proves that $\langle \mathbb{N}, S, 0 \rangle \models \psi$, then over RCA₀ the statement “for every $\langle A, f, c \rangle$, if $\langle A, f, c \rangle \models \psi$, then there exists a bijection between $\mathbb{N}$ and $A$” implies RCA₀.

Since RCA₀ is equivalent over RCA₀ to a statement expressing the correctness of defining functions by primitive recursion [SS86, Lemma 2.5], Theorem 1 may be intuitively understood as saying that, for provably true single-sentence characterizations at least, “categorical characterizations of the natural numbers require primitive recursion”.

Do less stringent versions of Question 1 give rise to “exceptions” to this general conclusion? As it turns out, they do. Firstly, characterizing the natural numbers by a set of sentences is already possible in RCA₀, in the following sense (for a precise statement of the theorem, see Section 4):

**Theorem 2.** There exists a $\Delta_0$-definable (and polynomial-time recognizable) set $\Xi$ of $\Sigma^1_1 \land \Pi^1_1$ sentences such that RCA₀ proves: for every $\langle A, f, c \rangle$, $\langle A, f, c \rangle$ satisfies all $\xi \in \Xi$ if and only if $\langle A, f, c \rangle$ is isomorphic to $\langle \mathbb{N}, S, 0 \rangle$. 

2
Secondly, even a single-sentence characterization is possible in a $\Pi^1_2$-conservative extension of RCA$_0$, at least if one is willing to consider rather peculiar theories:

**Theorem 3.** There is a $\Sigma^1_2$ sentence which is a categorical characterization of $\langle \mathbb{N}, S, 0 \rangle$ provably in RCA$_0 + \neg\text{WKL}$.

Theorem 3 is not quite satisfactory, as the theory and characterization it speaks of are false in $\langle \omega, \mathcal{P}(\omega) \rangle$. So, another natural question to ask is whether a single-sentence characterization of the natural numbers can be provably categorical in a true $\Pi^0_2$-conservative extension of RCA$_0^*$. We show that under an assumption just a little stronger than $\Pi^0_2$-conservativity, the characterization from Theorem 3 is actually “as true as possible”:

**Theorem 4.** Let $T$ be an extension of RCA$_0^*$ conservative for first-order $\forall \Delta_0(\Sigma_1)$ sentences. Let $\eta$ be a second-order sentence consistent with WKL$_0^* + \text{superexp}$. Then it is not the case that $\eta$ is a categorical characterization of $\langle \mathbb{N}, S, 0 \rangle$ provably in $T$.

The proofs of our theorems make use of a weaker notion of isomorphism to $\langle \mathbb{N}, S, 0 \rangle$ studied in [SY13], that of “almost isomorphism”. Intuitively speaking, a structure $\langle A, f, c \rangle$ satisfying some basic axioms is almost isomorphic to $\langle \mathbb{N}, S, 0 \rangle$ if it is “equal to or shorter than” the natural numbers. The two crucial facts we prove and exploit are that almost isomorphism to $\langle \mathbb{N}, S, 0 \rangle$ can be characterized by a single sentence provably in RCA$_0^*$, and that structures almost isomorphic to $\langle \mathbb{N}, S, 0 \rangle$ correspond to $\Sigma^0_1$-definable cuts.

The paper is structured as follows. After a preliminary Section 1, we conduct our study of almost isomorphism to $\langle \mathbb{N}, S, 0 \rangle$ in Section 2. We then prove Theorem 1 in Section 3, Theorems 2 and 3 in Section 4, and Theorem 4 in Section 5.

## 1 Preliminaries

We assume familiarity with subtheories of second-order arithmetic, as presented in [Sim09]. Of the “Big Five” theories featuring prominently in that book, we only need the two weakest: RCA$_0$, axiomatized by $\Delta^0_1$ comprehension and $\Sigma^0_1$ induction (and a finite list of simple basic axioms), and WKL$_0$, which extends RCA$_0$ by the axiom WKL stating that an infinite binary tree has an infinite branch.

We also make use of some well-known fragments of first-order arithmetic, principally $\Pi_0 + \text{exp}$, which extends induction for $\Delta_0$ formulas by an axiom exp stating the totality of exponentiation; $\mathcal{B} \Sigma_1$, which extends $\Pi_0$ by the $\Sigma_1$ collection (bounding) principle; and $\Sigma_1$. For a comprehensive treatment of these and other subtheories of first-order arithmetic, refer to [HP93].
The well-known hierarchies defined in terms of alternations of first-order quantifiers make sense both for purely first-order formulas and for formulas allowing second-order parameters, and we will need notation to distinguish between the two cases. For classes of formulas with first-order quantification but also arbitrary second-order parameters, we use the $\Sigma_n^0$ notation standard in second-order arithmetic. On the other hand, when discussing classes of first-order formulas, we adopt a convention often used in first-order arithmetic and omit the superscript "^0". Thus, for instance, a $\Sigma_1$ formula is a first-order formula (with no second-order variables at all) containing a single block of existential quantifiers followed by a bounded part. More generally, if we want to speak of a formula possibly containing second-order parameters $\bar{X}$ but no other second-order parameters, we use notation of the form $\Sigma_n(\bar{X})$ (to be understood as $\Sigma_n$ relativized to $\bar{X}$).

A formula is $\Delta_0(\Sigma_1)$ if it belongs to the closure of $\Sigma_1$ under boolean operations and bounded first-order quantifiers. $\forall \Delta_0(\Sigma_1)$ (respectively $\exists \Delta_0(\Sigma_1)$) is the class of first-order formulas which consist of a block of universal (respectively existential) quantifiers followed by a $\Delta_0(\Sigma_1)$ formula.

The theory $\text{RCA}_0^\sharp$ was introduced in [SS86]. It differs from $\text{RCA}_0$ in that the $\Sigma_1^0$ induction axiom is replaced by $\text{IA}_0^0 + \text{exp}$. $\text{WKL}_0^\sharp$ is $\text{RCA}_0^\sharp$ plus the WKL axiom. Both $\text{RCA}_0^\sharp$ and $\text{WKL}_0^\sharp$ have $\text{BS}_1 + \text{exp}$ as their first-order part, while the first-order part of $\text{RCA}_0$ and $\text{WKL}_0$ is $\text{I} \Sigma_1$.

We let superexp denote both the “tower of exponents” function defined by $\text{superexp}(x) = \text{exp}_x(2)$ (where $\text{exp}_0(2) = 1, \text{exp}_{x+1}(2) = 2^{\text{exp}_x(2)}$) and the axiom saying that for every $x$, superexp$(x)$ exists. $\Delta_0(\text{exp})$ stands for the class of bounded formulas in the language extending the language of Peano Arithmetic by a symbol for $x^3$. $\text{IA}_0(\text{exp})$ is a definitional extension of $\text{IA}_0^0 + \text{exp}$.

In any model $M$ of a first-order arithmetic theory (possibly the first-order part of a second-order structure), a cut is a nonempty subset of $M$ which is downwards closed and closed under successor. For a cut $J$, we sometimes abuse notation and also write $J$ to denote the structure $\langle J, S, 0 \rangle$, or even $\langle J, +, \cdot, \leq, 0, 1 \rangle$ if $J$ happens to be closed under multiplication.

If $\langle M, \mathcal{X} \rangle \models \text{RCA}_0^\sharp$ and $J$ is a cut in $M$, then $\mathcal{X}_J$ will denote the family of sets $\{X \cap J : X \in \mathcal{X} \}$. Throughout the paper, we frequently use the following simple but important result without further mention.

**Theorem** ([SS86], Theorem 4.8). If $\langle M, \mathcal{X} \rangle \models \text{RCA}_0^\sharp$ and $J$ is a proper cut in $M$ which is closed under exp, then $\langle J, \mathcal{X}_J \rangle \models \text{WKL}_0^\sharp$.

If $\langle M, \mathcal{X} \rangle \models \text{RCA}_0^\sharp$ and $A \in \mathcal{X}$, then $A$ is $M$-finite (or simply finite if we do not want to emphasize $M$) if there exists $a \in M$ such that all elements of $A$ are smaller than $a$. Otherwise, the set $A$ is (M)-infinite. For each $M$-finite set $A$ there is an element $a \in M$ coding $A$ in the sense that $A$ consists exactly of those $x \in M$ for which
the $x$-th bit in the binary notation for $a$ is 1. Moreover, RCA$_0^*$ has a well-behaved notion of cardinality of finite sets, which lets us define the internal cardinality $|A|_\mathcal{M}$ of any $A \in \mathcal{X}$ as $\sup\{x \in M : A \text{ contains a finite subset with at least } x \text{ elements}\}$. $|A|_\mathcal{M}$ is an element of $M$ if $A$ is $M$-finite, and a cut in $M$ otherwise.

$\mathbb{N}$ stands for the set of numbers defined by the formula $x = x$; in other words, $\mathbb{N}_M = M$. To refer to the set of standard natural numbers, we use the symbol $\omega$. The general notational conventions regarding cuts apply also to $\mathbb{N}$: for instance, if there is no danger of confusion, we sometimes write that some structure is “isomorphic to $\langle \mathbb{N}, S, 0 \rangle$” rather than “isomorphic to $\mathbb{N}$”.

We will be interested mostly in structures of the form $\langle A, f, c \rangle$, where $f$ is a unary function and $c$ an individual constant. The letter $A$ will always stand for some structure of this form. $A$ is a Peano system if $f$ is one-to-one, $c \notin \text{rng}(f)$, and $A$ satisfies the second-order induction axiom:

$$\forall X [X(c) \land \forall a [X(a) \rightarrow X(f(a))] \rightarrow \forall a X(a)].$$

Second-order logic is considered here in its full version — that is, non-unary second-order quantifiers are allowed — and interpreted according to the so-called standard semantics (cf. e.g. [End09]). Thus, the quantifier $\forall X$ with $X$ unary means “for all subsets of $A$”, $\forall X$ with $X$ binary means “for all binary relations on $A$”, etc. For instance, $A$ satisfies (1) exactly if there is no proper subset of $A$ containing $c$ and closed under $f$. Of course, from the perspective of a model $\mathcal{M} = \langle M, \mathcal{X} \rangle$ of RCA$_0^*$ or some other fragment of second-order arithmetic, “for all subsets of $A$” means “for all $X \in \mathcal{X}$ such that $X \subseteq A$”. After all, according to $\mathcal{M}$ there are no other subsets of $A$!

## 2 Almost isomorphism

A Peano system is said to be almost isomorphic to $\langle \mathbb{N}, S, 0 \rangle$ if for every $a \in A$ there is some $x \in \mathbb{N}$ such that $f^x(c) = a$. Here we take $f^x(c) = a$ to mean that there exists a sequence $\langle a_0, a_1, a_2, \ldots, a_x \rangle$ such that $a_0 = c$, $a_{z+1} = f(a_z)$ for $z < x$, and $a_x = a$. Note that we need to explicitly assert the existence of this sequence, which we often refer to as $\langle c, f(c), f^2(c), \ldots, f^x(c) \rangle$, because RCA$_0^*$ is too weak to prove that any function can be iterated an arbitrary number of times.

Being almost isomorphic to $\mathbb{N}$ is a definable property:

**Lemma 5.** There exists a $\Sigma^1_1 \land \Pi^1_1$ sentence $\xi$ in the language with one unary function $f$ and one individual constant $c$ such that RCA$_0^*$ proves: for every $A$, $A \models \xi$ if and only if $A$ is a Peano system almost isomorphic to $\langle \mathbb{N}, S, 0 \rangle$. 
By definition, $A$ is a Peano system precisely if it satisfies the $\Pi_1^1$ sentence $\xi_{\text{peano}}$:

$$f \text{ is } 1-1 \land c \not\in \text{rng}(f) \land \forall X \left[ X(c) \land \forall a [X(a) \to X(f(a))] \to \forall a X(a) \right].$$

The sentence $\xi$ will be the conjunction of $\xi_{\text{peano}}$, the $\Sigma_1^1$ sentence $\xi_{\varphi, \Sigma}$:

there exists a discrete linear ordering $\preceq$
for which $c$ is the least element and $f$ is the successor function,

and the $\Pi_1^1$ sentence $\xi_{\varphi, \Pi}$:

for every discrete linear ordering $\preceq$ with $c$ as least element and $f$ as successor
and for every $a$, the set of elements $\preceq$-below $a$ is Dedekind-finite.

We say that a set $X$ is Dedekind-finite if there is no bijection between $X$ and a proper subset of $X$. Note that $\xi$ involves quantification over non-unary relations: linear orderings and (graphs of) bijections.

In verifying that $\xi$ characterizes Peano systems almost isomorphic to $\mathbb{N}$, we will make use of the fact that provably in RCA$_0$, for any set $A$ and any $X \subseteq A$, $A \models "X \text{ is Dedekind-finite}"$ exactly if $X$ is finite. To see that this is true, note that if $X$ is infinite, then the map which takes $x \in X$ to the smallest $y \in X$ such that $x < y$ is a bijection between $X$ and its proper subset $X \setminus \{\min X\}$, and the graph of this bijection is a binary relation on $A$ witnessing $A \models "X \text{ is Dedekind-finite}"$. On the other hand, any witness for $A \models "X \text{ is Dedekind-finite}"$ must in fact be the graph of a bijection between $X$ and a proper subset of $X$, but such a bijection cannot exist for finite $X$ because all proper subsets of a finite set have strictly smaller cardinality than the set itself.

We first prove that Peano systems almost isomorphic to $\mathbb{N}$ satisfy $\xi_{\varphi, \Sigma}$ and $\xi_{\varphi, \Pi}$. Let $A$ be almost isomorphic to $\mathbb{N}$. Every $a \in A$ is of the form $f^x(c)$ for some $x \in \mathbb{N}$. Moreover, $x$ is unique. To see this, assume that $a = f^x(c) = f^{x+1}(c)$ and that $\langle c, f(c), \ldots, f^x(c) = a, f^{x+1}(c), \ldots, f^{x+y}(c) = a \rangle$ is the sequence witnessing that $f^{x+y}(c) = a$ (by $\Delta_0^0$-induction, this sequence is unique and its first $x+1$ elements comprise the unique sequence witnessing $f^x(c) = a$). If $y > 0$, then we have $c \neq f^y(c)$ and then $\Delta_0^0$-induction coupled with the injectivity of $f$ gives $f^w(c) \neq f^{w+y}(c)$ for all $w \leq x$. So, $y = 0$.

Because of the uniqueness of the $f^x(c)$ representation for $a \in A$, we can define $\preceq$ on $A$ by $\Delta_1^0$-comprehension in the following way:

$$a \preceq b := \exists x \exists y (a = f^x(c) \land b = f^y(c) \land x \leq y).$$

Clearly, $\preceq$ is a discrete linear ordering on $A$ with $c$ as the least element and $f$ as the successor function, so $A$ satisfies $\xi_{\varphi, \Sigma}$. 

6
For each \( a \in A \), the set of elements \( \preceq \)-below \( a \) is finite. Moreover, if \( \preceq \) is any ordering of \( A \) with \( c \) as least element and \( f \) as successor, then for each \( a \in A \) the set

\[
\{ b \in A : b \preceq a \iff b \preceq a \}
\]

contains \( c \) and is closed under \( f \). Since \( \mathbb{A} \) is a Peano system, \( \preceq \) has to coincide with \( \preceq \). Thus, \( \mathbb{A} \) satisfies \( \xi_{\preceq,\Pi} \).

For a proof in the other direction, let \( \mathbb{A} \) be a Peano system satisfying \( \xi_{\preceq,\Sigma} \) and \( \xi_{\preceq,\Pi} \). Let \( \preceq \) be an ordering on \( A \) witnessing \( \xi_{\preceq,\Sigma} \). Take some \( a \in A \). By \( \xi_{\preceq,\Pi} \), the set \( [c, a]_{\preceq} \) of elements \( \preceq \)-below \( a \) is finite. Let \( \ell \) be the cardinality of \( [c, a]_{\preceq} \) and let \( b \) be the \( \preceq \)-maximal element of \( [c, a]_{\preceq} \). By \( \Delta^0_1(\exp) \)-induction on \( x \) prove that there is an element below \( b^{x+1} \) coding a sequence \( (s_0, \ldots, s_\ell) \) such that \( s_0 = c \) and for all \( y < x \), either \( s_{y+1} = f(s_y) \preceq a \) or \( s_{y+1} = s_y = a \). Take such a sequence for \( x = \ell - 1 \). If \( a \) does not appear in the sequence, then by \( \Delta^0_1(\exp) \)-induction the sequence has the form \( (c, f(c), \ldots, f^{\ell-1}(c)) \) and all its entries are distinct elements of \( [c, a]_{\preceq} \setminus \{ a \} \); an impossibility, given that \( [c, a]_{\preceq} \setminus \{ a \} \) only has \( \ell - 1 \) elements. So, \( a \) must appear somewhere in the sequence. Taking \( w \) to be the least such that \( a = s_w \), we easily verify that \( a = f^w(c) \).

**Remark.** We do not know whether in \( \text{RCA}_0 \) it is possible to characterize \( \langle \mathbb{N}, S, 0 \rangle \) up to almost isomorphism by a \( \Pi^1_1 \) sentence. This does become possible in the case of \( \langle \mathbb{N}, \leq \rangle \) (given a suitable definition of almost isomorphism, cf. [SY13]), where there is no need for the \( \Sigma^1_1 \) part of the characterization which guarantees the existence of a suitable ordering.

An important fact about Peano systems almost isomorphic to \( \mathbb{N} \) is that their isomorphism types correspond to \( \Sigma^0_1 \)-definable cuts. This correspondence, which will play a major role in the proofs of our main theorems, is formalized in the following definition and lemma.

**Definition 6.** Let \( \mathcal{M} = \langle M, \mathcal{X} \rangle \) be a model of \( \text{RCA}_0 \). For a Peano system \( \mathbb{A} \) in \( \mathcal{M} \) which is almost isomorphic to \( \langle \mathbb{N}, S, 0 \rangle \), let \( J(\mathbb{A}) \) be the cut defined in \( \mathcal{M} \) by the \( \Sigma^0_1 \) formula \( \varphi(x) \):

\[
\exists a \in A \ f^x(c) = a.
\]

For a \( \Sigma^0_1 \)-definable cut \( J \) in \( \mathcal{M} \), let the structure \( \mathbb{A}(J) \) be \( \langle A_J, f_J, c_J \rangle \), where the set \( A_J \) consists of all the pairs \( \langle x, y_x \rangle \) such that \( y_x \) is the smallest witness for the formula \( x \in J \), the function \( f_J \) maps \( \langle x, y_x \rangle \) to \( \langle x+1, y_{x+1} \rangle \), and \( c_J \) equals \( \langle 0, y_0 \rangle \).

**Lemma 7.** Let \( \mathcal{M} = \langle M, \mathcal{X} \rangle \) be a model of \( \text{RCA}_0 \). The following holds:

(a) for a \( \Sigma^0_1 \)-definable cut \( J \) in \( \mathcal{M} \), the structure \( \mathbb{A}(J) \) is a Peano system almost isomorphic to \( \langle \mathbb{N}, S, 0 \rangle \), and \( J(\mathbb{A}(J)) = J \),

7
(b) if $\mathbb{A} \in \mathcal{X}$ is a Peano system almost isomorphic to $\langle \mathbb{N}, S, 0 \rangle$, then there is an isomorphism in $\mathcal{M}$ between $\mathbb{A}(J(\mathbb{A}))$ and $\mathbb{A}$.

c) if $\mathbb{A} \in \mathcal{X}$ is a Peano system almost isomorphic to $\langle \mathbb{N}, S, 0 \rangle$, then there is an isomorphism in $\mathcal{M}$ between $\mathbb{A}$ and $J(\mathbb{A})$, which also induces an isomorphism between the second-order structures $\langle \mathbb{A}, \mathcal{P}(\mathcal{A}) \rangle$ and $\langle J(\mathbb{A}), \mathcal{P}(J(\mathbb{A})) \rangle$.

Although all the isomorphisms between first-order structures mentioned in Lemma 7 are elements of $\mathcal{X}$, a cut is not itself an element of $\mathcal{X}$ unless it equals $M$ (because induction fails for the formula $x \in J$ whenever $J$ is a proper cut). Obviously, the isomorphism between second-order structures mentioned in part (c) is also outside $\mathcal{X}$.

**Proof.** For a $\Sigma_1^0$-definable cut $J \in \mathcal{M}$, it is clear that $A_J$ and $f_J$ are elements of $\mathcal{X}$, that $f_J$ is an injection from $A_J$ into $A_J$, and that $c_J$ is outside the range of $f_J$. Furthermore, for every $(x, y_x) \in A_J$, $\Sigma_1^0$ collection in $\mathcal{M}$ guarantees that there is a common upper bound on $y_0, \ldots, y_x$, so $\Delta_0^0$ induction is enough to show that the sequence $(c_J, f_J(c_J), \ldots, f_J^n(c_J) = \langle x, y_x \rangle)$ exists. If $X \subseteq A_J, X \in \mathcal{X}$, is such that $c_J \in X$ but $f_J(c_J) \notin X$, then $\Delta_0^0$ induction along the sequence $(c_J, f_J(c_J), \ldots, f_J^n(c_J))$ finds some $w < x$ such that $f_J^w(c_J) \notin X$ but $f_J^w(f_J^w(c_J)) \notin X$. Thus, $\mathbb{A}(J)$ is a Peano system almost isomorphic to $\mathbb{N}$, and clearly $J(\mathbb{A}(J))$ equals $J$, so part (a) is proved.

For part (b), if $\mathbb{A}$ is almost isomorphic to $\mathbb{N}$, then each $a \in A$ has the form $a = f^A(c)$ for some $x \in J(\mathbb{A})$, and we know from the proof of Lemma 6 that the element $x$ is unique. Thus, the mapping which takes $f^A(c) \in \mathbb{A}$ to $\langle x, y_x \rangle \in \mathbb{A}(J(\mathbb{A}))$ is guaranteed to exist in $\mathcal{M}$ by $\Delta_0^0$ comprehension. It follows easily from the definitions of $J(\mathbb{A})$ and $\mathbb{A}(J)$ that the mapping $f^A(c) \mapsto \langle x, y_x \rangle$ is an isomorphism between $\mathbb{A}$ and $\mathbb{A}(J(\mathbb{A}))$.

For part (c), we assume that $\mathbb{A}$ equals $\mathbb{A}(J(\mathbb{A}))$, which we may do w.l.o.g. by part (b). The isomorphism between $\mathbb{A}$ and $J(\mathbb{A})$ is given by $(x, y_x) \mapsto x$. To prove that this also induces an isomorphism between $\langle \mathbb{A}, \mathcal{P}(\mathcal{A}) \rangle$ and $\langle J(\mathbb{A}), \mathcal{P}(J(\mathbb{A})) \rangle$, we have to show that for any $X \subseteq A$, it holds that $X \in \mathcal{X}$ exactly if $\{x : (x, y_x) \in X\}$ has the form $Z \cap J(\mathbb{A})$ for some $Z \in \mathcal{X}$. This is easy if $J(\mathbb{A}) = M$, so below we assume $J(\mathbb{A}) \neq M$.

The “if” direction is immediate: given $Z \in \mathcal{X}$, the set $\{x, y_x : x \in Z\}$ is $\Delta_0(Z)$ and thus belongs to $\mathcal{X}$.

To deal with the other direction, we assume that $\mathcal{M}$ is countable. We can do this w.l.o.g. because $J(\mathbb{A})$ is a definable cut, so the existence of a counterexample in some model would imply the existence of a counterexample in a countable model by a downwards Skolem-Löwenheim argument.

By [SS86, Theorem 4.6], the countability of $\mathcal{M}$ means that we can extend $\mathcal{X}$ to a family $\mathcal{X}^+ \supseteq \mathcal{X}$ such that $\langle M, \mathcal{X}^+ \rangle \models \text{WKL}^\omega_0$. Note that there are no $M$-finite
Let it be shown in [SY13, Lemma 2.2] that in \( \mathcal{X}^+ \) there is some \( z \in M \) such that
\[
X = \{ x : \text{the } x\text{-th bit in the binary notation for } z \text{ is 1} \}.
\]
Therefore, \( X \) is \( \Delta_0 \)-definable with parameter \( z \) and so \( X \in \mathcal{X}^+ \).

Now consider some \( X \in \mathcal{X}^+ \). Write \( \langle \rangle \) addition resp. of this section, together with our Theorem 1, give precise meaning to the intuitive
\[
N \text{ which is almost isomorphic but not isomorphic to } \mathcal{X}^+ \text{ or that } \mathcal{X}^+ \text{ is actually isomorphic to } \mathcal{X}^+.
\]

Proof. \( \mathcal{M} = (M, \mathcal{X}) \) be a model of RCA\(^*_0\). Let \( \mathcal{A}_0 \in \mathcal{X}^+ \) be a Peano system almost isomorphic to \( (\mathbb{N}, S, 0) \). Assume that \( J(\mathcal{A}_0) \) is a proper cut closed under exp, that \( \triangleq \) is a linear ordering on \( A \) with least element \( c \) and successor function \( f \), and that \( \oplus, \otimes \) are operations on \( A \) which satisfy the usual recursive definitions of addition resp. multiplication with respect to least element \( c \) and successor \( f \). Then \( (A, \oplus, \otimes, \triangleq, c, f(c)) \text{, } \mathcal{X}^+ \cap \mathcal{P}(A) \models \text{WKL}^*_0 \).

Proof. Write \( \mathcal{A}_0 \) for \( (A, \oplus, \otimes, \leq, c, f(c)) \). By Lemma 7 part (b), we can assume w.l.o.g. that \( \mathcal{A}_0 = \mathcal{A}(J(\mathcal{A})) \). Using the fact that \( \mathcal{A}_0 \) is a Peano system, we can prove that for every \( x, z \in J(\mathcal{A}_0) \):
\[
\langle x, y_x \rangle \oplus \langle z, y_z \rangle = \langle x + z, y_{x+z} \rangle,
\]
\[
\langle x, y_x \rangle \otimes \langle z, y_z \rangle = \langle x \cdot z, y_{x\cdot z} \rangle,
\]
\[
\langle x, y_x \rangle \leq \langle z, y_z \rangle \text{ iff } x \leq z.
\]

By the obvious extension of Lemma 7 part (c) to structures with addition, multiplication and ordering, \( (\mathcal{A}, \mathcal{X}^+ \cap \mathcal{P}(A)) \) is isomorphic to \( (J(A), \mathcal{X}(J(A))) \). Since \( J(\mathcal{A}_0) \) is proper and closed under exp, this means that \( (\mathcal{A}_0, \mathcal{X}^+ \cap \mathcal{P}(A)) \models \text{WKL}^*_0 \).

Remark. It was shown in [SY13, Lemma 2.2] that in RCA\(^*_0\) a Peano system almost isomorphic to \( \mathbb{N} \) is actually isomorphic to \( \mathbb{N} \). In light of Lemma 7, this is a reflection of the fact that in RCA\(^*_0\) there are no proper \( \Sigma^0_1 \)-definable cuts.

Informally speaking, a Peano system which is not almost isomorphic to \( \mathbb{N} \) is “too long”, since it contains elements which cannot be obtained by starting at zero and iterating successor finitely many times. On the other hand, a Peano system which is almost isomorphic but not isomorphic to \( \mathbb{N} \) is “too short”. The results of this section, together with our Theorem 1, give precise meaning to the intuitive
idea strongly suggested by Table 2 of [SY13], that the problem with characterizing
the natural numbers in $\text{RCA}_0$ is ruling out structures that are “too short” rather than
“too long”.

3 Characterizations: basic case

In this section, we prove Theorem 1.

**Theorem 1.** Let $\psi$ be a second-order sentence in the language with one unary
function $f$ and one individual constant $c$. If $\text{WKL}_0$ proves that $(\mathbb{N}, S, 0) \models \psi$, then
over $\text{RCA}_0$ the statement “for every $A$, if $A \models \psi$, then there exists a bijection
between $\mathbb{N}$ and $A$” implies $\text{RCA}_0$.

We use a model-theoretic argument based on the work of Section 2 and a lemma
about cuts in models of $\text{I}_\Delta_0 + \text{exp} + -\text{I}_\Sigma_1$.

**Lemma 9.** Let $M \models \text{I}_\Delta_0 + \text{exp} + -\text{I}_\Sigma_1$. There exists a proper $\Sigma_1$-definable cut $J \subseteq M$
closed under exp.

**Proof.** We need to consider a few cases.

*Case 1.* $M \models \text{superexp}$. Since $M \not\models \Sigma_1$, there exists a $\Sigma_1$ formula $\varphi(x)$, possibly with parameters, which defines a proper subset of $M$ closed under successor.
Replacing $\varphi(x)$ by the formula $\hat{\varphi}(x)$: “there exists a sequence witnessing that for all $y \leq x$, $\varphi(y)$ holds”, we obtain a proper $\Sigma_1$-definable cut $K \subseteq M$. Define:

$$J := \{ y : \exists x \in K \ (y < \text{superexp}(x)) \}.$$ 

$J$ is a cut closed under exp because $K$ is a cut, and it is proper because it does not contain superexp($b$) for any $b \notin K$.

The remaining cases all assume that $M \not\models \text{superexp}$. Let $\text{Log}^*(M)$ denote the
domain of superexp in $M$. By the case assumption and the fact that $M \models \text{exp}$, $\text{Log}^*(M)$ is a proper $\Sigma_1$-definable cut in $M$.

*Case 2.* $\text{Log}^*(M)$ is closed under exp. Define $J := \text{Log}^*(M)$.

*Case 3.* $\text{Log}^*(M)$ is closed under addition but not under exp. Let $\text{Log}(\text{Log}^*(M))$ be the subset of $M$ defined as $\{ x : \text{exp}(x) \in \text{Log}^*(M) \}$. Since $\text{Log}^*(M)$ is closed under addition, $\text{Log}(\text{Log}^*(M))$ is a cut. Moreover, $\text{Log}(\text{Log}^*(M)) \subseteq \text{Log}^*(M)$, because $\text{Log}^*(M)$ is not closed under exp. Define:

$$J := \{ y : \exists x \in \text{Log}(\text{Log}^*(M)) \ (y < \text{superexp}(x)) \}.$$ 

$J$ is a cut closed under exp because $\text{Log}(\text{Log}^*(M))$ is a cut, and it is proper because it does not contain superexp($b$) for any $b \in \text{Log}^*(M) \setminus \text{Log}(\text{Log}^*(M))$. 

10
Case 4. $\text{Log}^*(M)$ is not closed under addition. Let $\frac{1}{2}\text{Log}^*(M)$ be the subset of $M$ defined as $\{x : 2x \in \text{Log}^*(M)\}$. Since $\text{Log}^*(M)$ is closed under successor, $\frac{1}{2}\text{Log}^*(M)$ is a cut. Moreover, $\frac{1}{2}\text{Log}^*(M) \subseteq \text{Log}^*(M)$, because $\text{Log}^*(M)$ is not closed under addition. Define:

$$J := \{y : \exists x \in \frac{1}{2}\text{Log}^*(M) \ (y < \text{superexp}(x))\}.$$

$J$ is a cut closed under exp because $\frac{1}{2}\text{Log}^*(M)$ is a cut, and it is proper because it does not contain $\text{superexp}(b)$ for any $b \in \text{Log}^*(M) \setminus \frac{1}{2}\text{Log}^*(M)$. 

Remark. Inspection of the proof reveals immediately that Lemma 9 relativizes, in the sense that in a model of $\text{I}\Delta_0(X) + \exp + -\text{I}\Sigma_1(X)$ there is a $\Sigma_1(X)$-definable proper cut closed under exp.

Remark. The method used to prove Lemma 9 shows the following result: for any $n \in \omega$, there is a definable cut in $\text{I}\Delta_0 + \exp$, provably closed under exp, which is proper in all models of $\text{I}\Delta_0 + \exp + -\text{I}\Sigma_n$. In contrast, there is no definable cut in $\text{I}\Delta_0 + \exp$ provably closed under $\text{superexp}$; otherwise, $\text{I}\Delta_0 + \exp$ would prove its consistency relativized to a definable cut, which would contradict Theorem 2.1 of [Pud85].

We can now complete the proof of Theorem 1. Assume that $\psi$ is a second-order sentence true of $\langle \mathbb{N}, S, 0 \rangle$ provably in WKL$_0^\ast$. Let $\mathcal{M} = \langle M, \mathcal{X} \rangle$ be a model of RCA$_0^\ast + -\text{I}\Sigma_0^0$. Assume for the sake of contradiction that according to $\mathcal{M}$, the universe of any structure satisfying $\psi$ can be bijectively mapped onto $\mathbb{N}$.

Let $J$ be the proper cut in $M$ guaranteed to exist by the relativized version of Lemma 9. Note that according to $\mathcal{M}$, there is no bijection between $A_J$ and $\mathbb{N}$. Otherwise, for every $y \in M$ the preimage of $\{0, \ldots, y - 1\}$ under the bijection would be a finite subset of $A_J$ of cardinality exactly $y$, which would imply $|A_J|_{\mathcal{M}} = M$. But it is easy to verify that $|A_J|_{\mathcal{M}} = J$.

From our assumption on $\psi$ it follows that $\mathcal{M}$ believes $\mathbb{A}(J) \models \neg \psi$.

By Lemma 7 and its proof, the mapping $\langle x, y \rangle \mapsto x$ induces an isomorphism between $\langle \mathbb{A}(J), \mathcal{X} \cap \mathcal{P}(A_J) \rangle$ and $\langle J, \mathcal{X} \rangle$. Since $J$ is closed under addition and multiplication, we can define the operation $\oplus$ on $A_J$ by $\langle x, y \rangle \oplus \langle z, y \rangle = \langle x + z, y + z \rangle$, and we can define $\otimes$ and $\preceq$ analogously. By $\Delta^0_0$ comprehension, $\oplus, \otimes, \preceq$ are all elements of $\mathcal{X}$. Write $\mathbb{A}(J)$ for $\langle \mathbb{A}(J), \oplus, \otimes, \preceq, \{0, y_0\}, \{1, y_1\} \rangle$.

Clearly, $A_J$ with the structure given by $\oplus, \otimes, \preceq$ satisfies the assumptions of Corollary 8, which means that $\langle \mathbb{A}(J), \mathcal{X} \cap \mathcal{P}(A_J) \rangle$ is a model of WKL$_0^\ast$. We also claim that $\langle \mathbb{A}(J), \mathcal{X} \cap \mathcal{P}(A_J) \rangle$ believes $\mathbb{N} \models -\psi$. This is essentially an immediate consequence of the fact that $\mathcal{M}$ thinks $\mathbb{A}(J) \models -\psi$, since the subsets of $A_J$ are exactly the same in $\langle \mathbb{A}(J), \mathcal{X} \cap \mathcal{P}(A_J) \rangle$ as in $\mathcal{M}$. There is one minor technical annoyance related to non-unary second-order quantifiers in $\psi$, as the integer pairing
function in $\hat{A}(J)$ does not coincide with that of $M$. The reason this matters is that
the language of second-order arithmetic officially contains only unary set variables,
so e.g. a binary relation is represented by a set of pairs, but a set of $M$-pairs of
elements of $A_J$ might not even be a subset of $A_J$. Clearly, however, since the graph
of the $\hat{A}(J)$-pairing function is $\Delta^0_0(\exp)$-definable in $\mathcal{M}$, a given set of $M$-pairs of
elements of $A_J$ belongs to $\mathcal{X}$ exactly if the corresponding set of $\hat{A}$-pairs belongs
to $\mathcal{X} \cap \mathcal{P}(A_J)$; and likewise for tuples of greater constant length.

Thus, our claim holds, and we have contradicted the assumption that $\psi$ is true
of $\mathbb{N}$ provably in $\text{WKL}_0^*$. □ (Theorem 1)

We point out the following corollary of the proof.

**Corollary 10.** The following are equivalent over $\text{RCA}_0^*$:

1. $\neg \text{RCA}_0$.
2. There exists $\mathcal{M} = \langle M, \mathcal{X} \rangle$ satisfying $\text{WKL}_0^*$ such that $|M| \neq |\mathbb{N}|$.

*Proof.* $\text{RCA}_0$ proves that all infinite sets have the same cardinality, which gives
(2) $\Rightarrow$ (1). To prove (1) $\Rightarrow$ (2), work in a model of $\text{RCA}_0 + \neg \text{RCA}_0$ and take the
inner model of $\text{WKL}_0^*$ provided by the proof of Theorem 1. □

*Remark.* The type of argument described above can be employed to strengthen
Theorem 1 in two ways.

Firstly, it is clear that $\langle \mathbb{N}, S, 0 \rangle$ could be replaced in the statement of Theorem
1 by, for instance, $\langle \mathbb{N}, \leq, +, - , 0, 1 \rangle$. In other words, the extra structure provided
by addition and multiplication does not help in characterizing the natural numbers
without $\Sigma^0_1$.

Secondly, for any fixed $n \in \omega$, the theories $\text{RCA}_0^* / \text{WKL}_0^*$ appearing in the statement
could be extended (both simultaneously) by an axiom expressing the totality of
$f_n$, the $n$-th function in the Grzegorczyk-Wainer hierarchy (e.g., the totality of
$f_2$ is $\exp$, the totality of $f_3$ is $\text{superexp}$). The proof remains essentially the same,
except that the argument used to show Lemma 9 now splits into $n + 2$ cases instead
of four.

By compactness, $\text{RCA}_0^* / \text{WKL}_0^*$ could also be replaced in the statement of the
theorem by $\text{RCA}_0^* + \text{PRA} / \text{WKL}_0^* + \text{PRA}$, where $\text{PRA}$ is primitive recursive arithmetic.

## 4 Characterizations: exceptions

In this section, we give a precise statement of Theorem 2, and prove Theorems 2
and 3.
Theorem 2 (restated). There exists a $\Delta_0$ formula $\Xi(x)$ defining a (polynomial-time recognizable) set of $\Sigma^1_1 \land \Pi^1_1$ sentences such that $\text{RCA}_0^*$ proves: “for every $A$, $A$ is isomorphic to $\langle \mathbb{N},S,0 \rangle$ if and only if $A \models \xi$ for all $\xi$ such that $\Xi(\xi)$”.

This is our formulation of “there exists a set of second-order sentences which provably in $\text{RCA}_0^*$ categorically characterizes the natural numbers”. Note that a characterization by a fixed set of standard sentences is ruled out by Theorem 1 (and a routine compactness argument).

Proof of Theorem 2. We will abuse notation and write $X$ for the set of sentences defined by the formula $X(x)$. Let $X$ consist of the sentence $x$ from Lemma 5 and the sentences

$$\exists a_0 \exists a_1 \ldots \exists a_{e-1} \exists a_x [a_0 = c \land a_1 = f(a_0) \land \ldots \land a_x = f(a_{e-1})],$$

for every $x \in \mathbb{N}$. (Note that in a nonstandard model of $\text{RCA}_0^*$, the set $X$ will contain sentences of nonstandard length.)

Provably in $\text{RCA}_0^*$, a structure $A$ satisfies all sentences in $X$ exactly if it is a Peano system almost isomorphic to $\mathbb{N}$ such that for every $x \in \mathbb{N}$, $f(x) = c$ exists. Clearly then, $\mathbb{N}$ satisfies all sentences in $X$. Conversely, if $A$ satisfies all sentences in $X$, then $J(A) = \mathbb{N}$ and so $A$ is isomorphic to $\mathbb{N}$.

Theorem 3. There is a $\Sigma^1_2$ sentence which is a categorical characterization of $\langle \mathbb{N},S,0 \rangle$ provably in $\text{RCA}_0^* + \neg \text{WKL}$.

Before proving the theorem, we verify that the theory it mentions is a $\Pi^1_1$-conservative extension of $\text{RCA}_0^*$.

Proposition 11. The theory $\text{RCA}_0^* + \neg \text{WKL}$ is a $\Pi^1_1$-conservative extension of $\text{RCA}_0^*$.

Proof. Let $X \forall Y \varphi(X,Y)$ be a $\Sigma^1_2$ sentence consistent with $\text{RCA}_0^*$. Take $\langle M, \mathcal{X} \rangle$ and $A \in \mathcal{X}$ such that $\langle M, \mathcal{X} \rangle \models \text{RCA}_0^* + \forall Y \varphi(A,Y)$. Let $\Delta_1(A) \cdot \text{Def}$ stand for the collection of the $\Delta_1(A)$-definable subsets of $M$. $\Delta_1(A) \cdot \text{Def} \subseteq \mathcal{X}$, so obviously $\langle M, \Delta_1(A) \cdot \text{Def} \rangle \models \text{RCA}_0^* + \forall Y \varphi(A,Y)$. Moreover, by a standard argument, there is a $\Delta_1(A)$-definable finite binary tree without a $\Delta_1(A)$-definable branch, so $\langle M, \Delta_1(A) \cdot \text{Def} \rangle \models \neg \text{WKL}$. 

Proof of Theorem 3. Work in $\text{RCA}_0^* + \neg \text{WKL}$. The sentence $\psi$, our categorical characterization of $\mathbb{N}$, is very much like the sentence $\xi$ described in the proof of Lemma 5, which expressed almost isomorphism to $\mathbb{N}$. The one difference is that the $\Sigma^1_2$ conjunct of $\xi$:
there exists a discrete linear ordering $\preceq$ for which $c$ is the least element and $f$ is the successor function.

is strengthened in $\varphi$ to the $\Sigma^1_2$ sentence:

there exist binary operations $\oplus, \otimes$ and a discrete linear ordering $\preceq$ such that

$\preceq$ has $c$ as the least element and $f$ as the successor function,

$\oplus$ and $\otimes$ satisfy the usual recursive definition of addition and multiplication,

and such that $\varDelta_0 + \exp + \neg \text{WKL}$ holds.

$I\Delta_0 + \exp$ is finitely axiomatizable [GD82], so there is no problem with expressing this as a single sentence. Note that $\varphi$ is $\Sigma^1_1$.

Since $\neg \text{WKL}$ holds, the usual $+ \cdot \leq$ and ordering on $\mathbb{N}$ witness that $\mathbb{N}$ satisfies the new $\Sigma^1_2$ conjunct of $\varphi$. Of course, $\mathbb{N}$ is a Peano system almost isomorphic to $\mathbb{N}$, and thus it satisfies $\varphi$.

Now let $\mathcal{A}$ be a structure satisfying $\varphi$. Then $\mathcal{A}$ is a Peano system almost isomorphic to $\mathbb{N}$, so we may consider $J(\mathcal{A})$. As in the proof of Corollary 8, we can show that the canonical isomorphism between $\mathcal{A}$ and $J(\mathcal{A})$ has to map $\oplus, \otimes, \preceq$ witnessing the $\Sigma^1_2$ conjunct of $\varphi$ to the usual $+, \cdot, \leq$ restricted to $J$. This guarantees that $J(\mathcal{A})$ is closed under $\exp$, because the $\Sigma^1_2$ conjunct of $\varphi$ explicitly contains $I\Delta_0 + \exp$. Moreover, Corollary 8 implies that $J(\mathcal{A})$ cannot be a proper cut, because otherwise $\mathcal{A}$ with the additional structure given by $\oplus, \otimes, \preceq$ would have to satisfy $\text{WKL}$. So, $J(\mathcal{A}) = \mathbb{N}$ and thus $\mathcal{A}$ is isomorphic to $\mathbb{N}$.  

5 Characterizations: exceptions are exotic

To conclude the paper, we prove Theorem 4 and some corollaries.

**Theorem 4.** Let $T$ be an extension of $\text{RCA}_0$ conservative for first-order $\forall \Delta_0(\Sigma^1_1)$ sentences. Let $\eta$ be a second-order sentence consistent with $\text{WKL}_0^\ast + \text{superexp}$. Then it is not the case that $\eta$ is a categorical characterization of $\langle \mathbb{N}, S, 0 \rangle$ provably in $T$.

**Proof.** Let $\mathcal{M} = \langle M, \mathcal{X} \rangle$ be a countable recursively saturated model of $\text{WKL}_0^\ast + \text{superexp} + \eta$.

Tanaka’s self-embedding theorem [Tan97] is stated for countable models of $\text{WKL}_0$. However, a variant of the theorem is known to hold for $\text{WKL}_0^\ast$ as well:

**Tanaka’s self-embedding theorem for $\text{WKL}_0^\ast$** (Wong-Yokoyama, unpublished). If $\mathcal{M} = \langle M, \mathcal{X} \rangle$ is a countable recursively saturated model of $\text{WKL}_0^\ast$ and $q \in M$, then there exists a proper cut $I$ in $M$ and an isomorphism $f : \langle M, \mathcal{X} \rangle \to \langle I, \mathcal{X}_I \rangle$ such that $f(q) = q$.

14
This can be proved by going through the original proof in [Tan97] and verifying that all arguments involving $\Sigma^0_1$ induction can be replaced either by $\Delta^0_0(\exp)$ induction plus $\Sigma^0_1$ collection or by saturation arguments\(^1\). A refined version of the result was recently proved by a different method in [EW14].

Thus, there is a proper cut $I$ in $M$ such that $\langle M, \mathcal{P} \rangle$ and $\langle I, \mathcal{P}_I \rangle$ are isomorphic. In particular, $\langle I, \mathcal{P}_I \rangle \models \eta$.

Let $a \in M \setminus I$. Define the cut $K$ in $M$ to be
\[
\{ y : \exists x \in I (y < \exp_{a+x}(2)) \}.
\]
Since $\exp_{a+x}(2) \in M \setminus K$, the cut $K$ is proper and hence $\langle K, \mathcal{P}_K \rangle \models \text{WKL}_0$. The set $I$ is still a proper cut in $K$, because $a \in K \setminus I$. Furthermore, $I$ is $\Sigma_1$-definable in $K$ by the formula $\exists x \exists y (y = \exp_{a+x}(2))$.

$T$ is conservative over $\text{RCA}_0$ for first-order $\forall \Delta_0(\Sigma_1)$ sentences, so there is a model $\langle L, \mathcal{V} \rangle \models T$ such that $K \equiv_{\Delta_0(\Sigma_1)} L$. We claim that in $\langle L, \mathcal{V} \rangle$ there is a Peano system $\mathcal{A}$ satisfying $\eta$ but not isomorphic to $\mathbb{N}$. This will imply that $T$ does not prove $\eta$ to be a categorical characterization of $\mathbb{N}$. It remains to prove the claim.

We can assume that $\eta$ does not contain a second-order quantifier in the scope of a first-order quantifier. This is because we can always replace first-order quantification by quantification over singleton sets, at the cost of adding some new first-order quantifiers with none of the original quantifiers of $\eta$ in their scope.

Note that $(K, \mathcal{P}_K)$ contains a proper $\Sigma_1$ definable cut, namely $I$, which satisfies $\eta$. Using the universal $\Sigma_1$ formula, we can express this fact by a first-order $\exists \Delta_0(\Sigma_1)$ sentence $\eta^{FO}$. The sentence $\eta^{FO}$ says the following:

there exists a triple “$\Sigma_1$ formula $\varphi(x, w)$, parameter $p$, bound $b$” such that

- $b$ does not satisfy $\varphi(x, p)$, the set defined by $\varphi(x, p)$ below $b$ is a cut,
- and this cut satisfies $\eta$.

To state the last part, replace the second-order quantifiers of $\eta$ by quantifiers over subsets of $\{0, \ldots, b-1\}$ (these are bounded first-order quantifiers) and replace the first-order quantifiers by first-order quantifiers relativized to elements below $b$ satisfying $\varphi(x, p)$. By our assumptions about the syntactical form of $\eta$, this ensures that $\eta^{FO}$ is $\exists \Delta_0(\Sigma_1)$.

$L$ is a $\Delta_0(\Sigma_1)$-elementary extension of $K$, so $L$ also satisfies $\eta^{FO}$. Therefore, $\langle L, \mathcal{V} \rangle$ also contains a proper $\Sigma_1$-definable cut satisfying $\eta$. The Peano system corresponding to this cut via Lemma 7 also satisfies $\eta$, but it cannot be isomorphic to $\mathbb{N}$ in $\langle L, \mathcal{V} \rangle$, because its internal cardinality is a proper cut in $L$. The claim, and the theorem, is thus proved.

---

\(^1\)The one part of Tanaka’s proof that does require $\Sigma^0_1$ induction is making $f$ fix (pointwise) an entire initial segment rather than just the single element $q$. See [Ena13].
Remark. The assumption that \( \eta \) is consistent with \( \text{WKL}_0^+ + \text{superexp} \) rather than just \( \text{WKL}_0 \) is only needed to ensure that there is a model of \( \text{RCA}_0^* \) with a proper \( \Sigma_1 \)-definable cut satisfying \( \eta \). The assumption can be replaced by consistency with \( \text{WKL}_0 \) extended by a much weaker first-order statement, but we were not able to make the proof work assuming only consistency with \( \text{WKL}_0^+ \).

One idea used in the proof of Theorem 4 seems worth stating as a separate corollary.

**Corollary 12.** Let \( \eta \) be a second-order sentence. The statement “there exists a Peano system \( A \) almost isomorphic but not isomorphic to \( \langle \mathbb{N}, S, 0 \rangle \) such that \( A \models \eta \)” is \( \Sigma_1 \) over \( \text{RCA}_0^* \).

**Proof.** By Lemma 7, a Peano system satisfying \( \eta \) and almost isomorphic but not isomorphic to \( \mathbb{N} \) exists exactly if there is a proper \( \Sigma_1^0 \)-definable cut satisfying \( \eta \). This can be expressed by a sentence identical to the first-order sentence \( \eta^{\text{FO}} \) from the proof of Theorem 4 except for an additional existential second-order quantifier to account for the possible set parameters in the formula defining the cut.

Theorem 4 also has the consequence that if we restrict our attention to \( \Pi_1^1 \)-conservative extensions of \( \text{RCA}_0^* \), then the characterization from Theorem 3 is not only the “truest possible”, but also the “simplest possible” provably categorical characterization of \( \mathbb{N} \).

**Corollary 13.** Let \( T \) be a \( \Pi_1^1 \)-conservative extension of \( \text{RCA}_0^* \). Assume that the second-order sentence \( \eta \) is a categorical characterization of \( \langle \mathbb{N}, S, 0 \rangle \) provably in \( T \). Then

(a) \( \eta \) is not \( \Pi_2^1 \).

(b) \( T \) is not \( \Pi_2^1 \)-axiomatizable.

**Proof.** We first prove (b). Assume that \( T \) is \( \Pi_2^1 \)-axiomatizable and \( \Pi_1^1 \)-conservative over \( \text{RCA}_0^* \). As observed in [Yok09], this means that \( T + \text{WKL}_0^+ \) is \( \Pi_1^1 \)-conservative over \( \text{RCA}_0^* \), so \( T \) is consistent with \( \text{WKL}_0^+ + \text{superexp} \). Hence, Theorem 4 implies that there can be no provably categorical characterization of \( \mathbb{N} \) in \( T \).

Turning now to part (a), assume that \( \eta \) is \( \Pi_2^1 \). Since \( T \) is \( \Pi_1^1 \)-conservative over \( \text{RCA}_0^* \) and proves that \( \mathbb{N} \models \eta \), then \( \text{RCA}_0^* + \eta \) must also be \( \Pi_1^1 \)-conservative over \( \text{RCA}_0^* \). But then, by a similar argument as above, \( \eta \) is consistent with \( \text{WKL}_0^+ + \text{superexp} \), which contradicts Theorem 4.

**Acknowledgement.** We are grateful to Stephen G. Simpson for useful remarks and to an anonymous referee for comments which helped improve the presentation.
References


