Metamathematical Investigations on the Theory of Grossone

Gabriele Lolli
Scuola Normale Superiore di Pisa, Italy
gabriele.lolli@sns.it

Abstract: We propose an axiomatization of Sergeyev's theory of Grossone, trying to comply with his methodological principles. We find that a simplified form of his Divisibility axiom is sufficient. We use for easier readability a second order language and a predicative second order logic. Our theory is not finitely axiomatizable and is a conservative extension of Peano's arithmetic.

Keywords: Grossone, infinite, arithmetic, divisibility, conservativeness.

1. Introduction

In the last ten years Yaroslav D. Sergeyev has introduced a new methodology for computing with infinities and infinitesimals, and he has applied it to a variety of problems, from numerical computations to ODE, Riemann's Zeta function, fractals and Turing machines.[1] Sergeyev's treatment is based on the use of a symbol ① for a numeral, called *Grossone*, which is meant to denote the number of elements of the set $\mathbb{N}$ of natural numbers. His theory differs both from the classical Cantorian introduction of infinite numbers and from non-standard analysis, as well as from Benci and Di Nasso's numerosity investigations in [9]. A first appraisal of Sergeyev's proposal against the background of the contemporary renewed interest in infinitesimals has been sketched by the author in [10].

The new numeral system conceived by Sergeyev aims at obtaining more accurate results concerning the infinite. It allows e.g. to assign different numbers to the set of natural numbers and to the set of even numbers, and quite properly the first number turns out to be the double of the latter.

Sergeyev's approach is more similar to that of an empirical scientist than of a deductive mathematician. He credits to himself unassumingly only the introduction of "a new computational methodology", which he justifies with the metaphor of the instruments resolution power: "Physicists decide the level of the precision they need and obtain a result depending of the chosen level of

[1] See e.g. Sergeyev's [1], [2], [3], [4], [5], Sergeyev and Garro's [6], [7] and De Cosmis and De Leone's [8] for applications to Linear Programming.

the accuracy. In the moment when they put a lens in the microscope, they have decided the minimal (and the maximal) measure of objects that they will be able to observe. If they need a more precise or a more rough answer, they change the lens of their microscope [. . .]. In natural sciences there always exists the triad – the researcher, the object of investigation, and tools used to observe the object – and the instrument used to observe the object bounds and influences results of observations. The same happens in Mathematics studying numbers and objects that can be constructed by using numbers. Numeral systems used to express numbers are instruments of observations used by mathematicians" ([5, p. 131]).

Sergeyev is wary of the axiomatic method because he thinks that by adopting it we would be tied to the expressive power of a language in the description of mathematical obiects and concepts (see [3, §2]); on the contrary, "the choice of the mathematical language depends on the practical problem that is to be solved and on the accuracy required for such a solution. In dependence of this accuracy, a numeral system that would be able to express the numbers composing the answer should be chosen" ([5, p. 131]).

Actually no mathematician ever works in a fixed theory. This is why logic is not popular among them. However, to assess the strength of an instrument, we have to consider it so to say *in vitro*.

In this paper we investigate the possibility of a formal axiomatic presentation of the arithmetical theory of ①,[2] at the risk of being unfaithful to Sergeyev's spirit; our aim is to make the theory accessible and acceptable, hopefully palatable, to traditionally minded mathematicians and to discuss a few of its metamathematical properties. To be able to offer a clear axiomatization, we chose to use a predicative second order logic, instead of first order logic, for reasons that will become clear in the ensuing text. Details will be provided for those who are not familiar with logical matters.

2. An informal presentation of Grossone

Sergeyev's philosophy is expressed by three postulates he assumes before plunging ahead into the computational methodology:

"P1. We postulate the existence of infinite and infinitesimal objects but accept that human beings and machines are able to execute only a finite number of operations.

P2. We shall not tell what are the mathematical objects we deal with. Instead, we shall construct more powerful tools that will allow us to improve our capacities to observe and to describe properties of mathematical objects.

P3. We adopt the principle : 'The part is less than the whole', and apply it to

[2]Grossone as a natural number is also of course a real number, and as such treated in many of Sergeyev's investigations. But we consider it here only in the arithmetical context.

all numbers, be they finite, infinite, or infinitesimal, as well as to all sets and processes, be they finite or infinite". [3, §2]

These three postulates constitute an unusual introduction to a mathematical theory. Their meaning becomes clearer with the development of the mathematical content, but a few preliminary comments are not out of place.

The postulates are not to be conceived as the "axioms" of an axiomatic theory. They have a pragmatic or methodological character, P2 and P3 a more general, P1 a more binding one; they explain how users of the methodology work and produce their results.

The principle in P3 goes contrary to the ideas prevailing after the success of set theory; its statement prepares the readers to accept results which are at odds with the received wisdom and it probably reveals what was behind the original intuition of the new conception. P3 is actually Euclid's "common notion" n. 5 (see [11, vol. 1, p. 155]).

All three postulates could probably be likened to Euclid's common notions, at least to n. 5, not to the others, which were basically properties of equality. According to Proclus, Euclid's common notions were general axioms not specific to geometry but valid for all sciences, as contemplated by Aristotle's theory of science.[3]

P1 and P2 are the programmatic statements of the new methodology, and have a multiple function. P1 explains a constraint to be obeyed that affects the results that will be obtained. Probably it is also a warning that if not respected a contradiction could follow. So it is a restriction of the acceptable arguments.

P2 is a reference to the philosophy of the instruments accuracy mentioned earlier. Mathematicians normally do not say what their objects are, so its first sentence could be superfluous; but mathematicians usually tacitly assume that in case of need their objects could be defined, for example in set theory (unless they are full fledged formalists); no such eventuality can present itself in this case, since the focus is on tools. It would be useless to ask Sergeyev what are the objects he studies with his new tools, and if he is a realist or otherwise; very likely he wouldn't take the bait, insisting that objects being what representation systems allow us to discover, we can only say of them what the instruments we use allow us to say.

Sergeyev's position does not fall under any of the traditional philosophies of mathematics: he accepts the realistic talk of the working mathematician but qualifies it as we have seen above; he has a finitistic restriction on logic, but he is not a costructivist. One could possibly draw a parallel with P. W. Bridgman's operationalism for the physical sciences, see [12], but the issue deserves a deeper scrutiny.

After the postulates, the first step is the introduction of the so called "Infinite Unit Axiom" (IUA), which consists of three parts, namely (from [3, §3]):

[3]The fifth common notion does not seem to be genuine however, being a generalization of a geometric inference used only in Euclid I,6, as noted in [11, vol. 1, p. 232].

"*Infinity*. Any finite natural number n is less than grossone, i.e., $n < ①$.
Identity. The following relations link ① to identity elements 0 and 1

$$0 \cdot ① = ① \cdot 0 = 0, \quad ① - ① = 0, \quad ①/① = 1, \quad ①^0 = 1, \quad 1^{①} = 1, \quad 0^{①} = 0.$$

Divisibility. For every finite natural number k, sets $\mathbb{N}_{k;n}$, $1 \leq k \leq n$, being the n-th part of the set, $\mathbb{N}$, of natural numbers have the same number of elements indicated by the numeral $①/n$ where

$$\mathbb{N}_{k;n} = \{k, k+n, k+2n, k+3n, \ldots\}, \qquad (1 \leq k \leq n), \qquad \bigcup_{k=1}^{n} \mathbb{N}_{k;n} = \mathbb{N}."$$

There are no other assumptions on Grossone. We will have to dissect carefully *Divisibility*, but a few preliminary remarks can be useful.

The divisibility axiom makes ① an "infinite unit of measure": "By using Postulate 3, we construct the sets $\mathbb{N}_{k;n}, 1 \leq k \leq n$, by separating the whole, i.e., the set $\mathbb{N}$, in n parts [...]. Again due to Postulate 3, we affirm that the number of elements of the nth part of the set, i.e., $①/n$, is n times less than the number of elements of the whole set, i.e., than ①. [...] Note that, since the numbers $①/n$ have been introduced as numbers of elements of sets $\mathbb{N}_{k;n}$, they are integer" ([3, §3]).

In fact, the axiom says that the set of natural numbers, which is $\mathbb{N}_{1;1}$, has a number of elements given by ①; moreover, if you split the natural numbers into, say, two disjoint sets as $\mathbb{N}_{1;2}$, the set of odd numbers, and $\mathbb{N}_{2;2}$, the set of even numbers, these two sets have the same number of elements $①/2$. Sameness of cardinality appears to be reasonable, if $\mathbb{N}$ is partitioned into equal parts, but why is it so? Everybody would accept that the two parts are in some sense equivalent, because there is an obvious bijection between them, but the appeal to bijections is dangerous, since there is also one between them and the whole set of natural numbers. To bring unrestricted bijections into the determination of cardinality would contradict *Divisibility*, and P3.

Sergeyev could possibly object that there is no question of bijections, between $\mathbb{N}_{1;2}$ and $\mathbb{N}_{2;2}$ or between $\mathbb{N}_{1;1}$ and $\mathbb{N}_{1;2}$, since against P1 they would be infinite processes. But then the part of *Divisibility* concerning the sameness of the cardinalities of $\mathbb{N}_{1;2}$ and $\mathbb{N}_{2;2}$ has no justification. However at this initial stage it is quite appropriate on Sergeyev's part to assume *Divisibility* without justifications, as an axiom. Axioms are justified by theorems, not by prior reasons, though they can be suggested by them.

We will see that it is possible, even necessary, to refer to bijections in studying cardinality; it can be done avoiding paradoxes, respecting P3 and justifying *Divisibility*.

In what follows, we will be mainly concerned with P3; this is because P3 is the only postulate with an enunciative content, realized e.g. in *Divisibility*. P1 plays a fundamental role in the mathematical investigations; in a metamathematical study of the methodology of ① we are not bound to adopt the directions of the methodology we are studying. What is important is that the frame we build is able to encompass all the results that are compatible with the use of the methodology.

3. A second order language

Since we have to talk of numbers and of sets of numbers, we use a second order language L^2, with two types of variables:

$x, y, z, \dots$ individual variables (first order variables)
$X, Y, Z, \dots$ set variables (second order variables).

The second order variables could be intended as variables for predicates, but we prefer to let them vary over sets, more familiar to mathematicians. We could have other sorts of second order variables, for relations and functions, but for simplicity we avoid them with a suitable coding of pairs.

There exists a primitive recursive function $P(x, y)$, actually there are many of them, which is a bijection of $\mathbb{N}^2$ onto $\mathbb{N}$,

$$P(x,y) = \frac{1}{2}[(x+y)^2 + 3x + y]$$

and there exist two functions P_1 and P_2, also primitive recursive, such that:[4]

$$\begin{array}{lcl} P(P_1(z), P_2(z)) & = & z \\ P_1(P(x,y)) & = & x \\ P_2(P(x,y)) & = & y. \end{array} \qquad (1)$$

We call P a pairing function, with $P_1(z)$ and $P_2(z)$ respectively the first and second projection of z; we write $P(x, y) = \langle x, y\rangle$ with the usual notation for ordered pairs; notice however that this is an arithmetical term.

We anticipate that we will say accordingly, when necessary, that a set X is a binary relation, or a function, viewed as $\{\langle P_1(x), P_2(x)\rangle \mid x \in X\}$, and extend all the set theoretic terminology and the usual notations to this interpretation. Recall that the domain of X is $dom(X) = \{x \mid \exists y(\langle x, y\rangle \in X)\}$ and the image $im(X) = \{y \mid \exists x(\langle x, y\rangle \in X)\}$; we will write: $F : X \longrightarrow Y$ for "F is a function from X to Y", formally

$$dom(F) = X \wedge im(F) \subseteq Y \wedge \forall x \in X \exists_1 y \in Y(\langle x, y\rangle \in F),$$ [5]

and $F : X \xrightarrow{1-1} Y$ for "F is a bijection between X and Y".

The language L^2 has as non-logical alphabet:

- symbols for 0, the successor function, addition, product, less than, equality (which make up the non-logical alphabet of the first order reduct L), and the coding functions:[6]

---

[4]For a verification see [13, pp. 43-5].

[5]$\exists_1$ is a shorthand for "there exists exactly one".

[6]The symbol $\prec$ can be primitive, hence requiring specific axioms, or defined, usually by the formula $x \prec y \leftrightarrow \exists z(y \approx x + s(z))$; in any case the same properties hold. $x \preccurlyeq y$ means $x \prec y \vee x \approx y$; $x \succ y$ is $y \prec x$.

For the order relation and equality we use curled symbols to distinguish them from the symbols $<$ and $=$ belonging to the metalanguage; for perspicuity reasons we avoid the duplication of the operation symbols.

$\underline{0}, \quad s, \quad +, \quad \cdot, \quad \prec, \quad \approx, \quad P, \quad P_1, \quad P_2$

- and a relational symbol for membership and auxiliary symbols:

$\in$, {, }, |.

Formulae are built from atomic formulae

$$t \approx s,\ t \prec s,\ t \in T,$$

where t, s are first order terms and T a second order term (to be defined below), with propositional connectives, first order quantifiers $\forall x$ and $\exists x$, second order quantifiers $\forall X$ and $\exists X$.

A formula is *predicative* if does not contain second order quantifiers (possibly second order parameters).

A numeral is a symbol or a finite array of symbols denoting a number; in a language like L^2 numerals are the terms

$$\underline{n} = \underbrace{s(\ldots s}_{n \text{ times}}(\underline{0})\ldots)$$

for every natural number $n \in \mathbb{N}$, including $\underline{0}$. $\mathbb{N}$ is the set of natural numbers in the metatheory. Sergeyev denotes by $\mathbb{N}$ the set of counting numbers $\{1, 2, 3, \ldots\}$ without 0, but in logical analyses it is more custumary to start from zero. The required adjustements are easy.

First order terms are: $\quad x, y, \ldots, \quad \underline{n}, \quad t+s, \quad t\cdot s, \quad \langle t, s\rangle, \quad P_1(t), \quad P_2(t),$
second order terms are: $\quad X, Y, \ldots, \quad \{x \mid A(x)\}$ for predicative A.

The term $\{x \mid A(x)\}$ is called an *abstract*, or the abstract of A and its intended denotation is the set of numbers which satisfy A.

With the abstracts all set theoretical notations: $\cup$, $\cap$, $\setminus$, $\subseteq$, $\emptyset$, $\{x\}$, $\{x, y\}$ ... can be introduced by definition, by means of predicative formulae. They belong to a definitional expansion of L^2.

Some of these symbols belong also to the metalanguage, but there should be no danger of confusion.

Notice that "F is a function from X to Y", as well as "F is a bijection between X and Y", are predicative formulae, while the formula normally used in set theory to define equivalence of sets X and Y

$$\exists F(F : X \xrightarrow{1-1} Y)$$

is not predicative.[7]

4. 2nd order predicative logic

Logical axioms and rules are those of predicate logic with equality, modulated according to the language. The axioms for quantifiers duplicate, for example

[7] It is called Σ^1_1 in the hierarchy of second order formulae.

the particularization axiom appears both as

$$\forall x A(x) \to A(x/t),$$

t a first order term, and

$$\forall X A(X) \to A(X/T),$$

T a second order term, and similarly for the $\exists$-introduction axiom, with the usual syntactical restrictions on the free variables of t and T.

Axioms for $\approx$ describe equality as an equivalence relation with the substitutivity property, and include a condition for the equality of sets:

Extensionality axiom:

$$X \approx Y \leftrightarrow \forall x(x \in X \leftrightarrow x \in Y),$$

for second order variables.

Finally the logic must provide for the introduction of second order terms (abstracts) associated to predicative formulae: for every predicative formula $A(x)$ we have the

Abstraction axiom:

$$x \in \{x \mid A(x)\} \leftrightarrow A(x),$$

which says what we have anticipated on the interpretation of an abstract.

From abstraction it follows the predicative comprehension schema

$$\exists X \forall x(x \in X \leftrightarrow A(x)),$$

for every predicative formula A.

A theory T in L^2 is a set of sentences of L^2. It would be more precise to call T a set of axioms for the theorems, which are the logical consequences of T. Logical axioms and the equality axioms are always tacitly added to any theory. We include also among logical axioms, by abuse of language, the three equations (1) above concerning P, P_1 and P_2 .

The relation of logical consequence can be defined either semantically, by means of models and validity in the models, or deductively, by the existence of a derivation.

A derivation of A from T is a finite sequence of formulae ending with A such that each formula of the sequence is either a logical or equality axiom, or an element of T or is obtained by preceding ones by means of logical rules.

The two definitions of logical consequence are equivalent, thanks to the logical completeness theorem. We will write $\mathsf{T} \vdash_2 A$ to say that A is a theorem of T, in any of the equivalent senses.

Recall that a structure for L^2 is a pair $\langle \mathcal{M}, \mathcal{X} \rangle$, where $\mathcal{M}$ is a structure for the reduct L and $\mathcal{X}$ a family of subsets of the universe M of $\mathcal{M}$, with suitable closure conditions; it must contain at least those definable in $\mathcal{M}$ by predicative formulae. In $\mathcal{M}$ there is an element $\underline{0}^{\mathcal{M}}$, and there are functions $s^{\mathcal{M}}$, $+^{\mathcal{M}}$, $\cdot^{\mathcal{M}}$, $P^{\mathcal{M}}$, $P_1^{\mathcal{M}}$, $P_2^{\mathcal{M}}$ and a relation $\prec^{\mathcal{M}}$ which are the interpretations of the corresponding symbols. There is a subset N of M which is the set of the interpretations $\underline{n}^{\mathcal{M}}$ in $\mathcal{M}$ of the numerals $\underline{n}$ for $n \in \mathbb{N}$. When $\mathcal{M}$ is a model of arithmetic, to be introduced in the next section, the map $n \mapsto \underline{n}^{\mathcal{M}}$ will preserve the arithmetical operations and the order. Hence N will be an isomorphic copy of $\mathbb{N}$ contained in M. We can assume for simplicity that $\mathbb{N} \subseteq M$. Then we stipulate that P be interpreted as the above described bijection between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$.

5. A formal arithmetic of Grossone

We will build Grossone arithmetic by successive approximations. The basis will be classical arithmetic as given by Peano's axioms.

Peano's second order arithmetic PA^2 has as axioms (the universal closures of)

$$
\begin{array}{ll}
1 & s(x) \approx s(y) \to x \approx y \\
2 & \underline{0} \not\approx s(x) \\
3 & x \not\approx \underline{0} \to \exists y(x \approx s(y)) \\
4 & x + \underline{0} \approx x \\
5 & x + s(y) \approx s(x+y) \\
6 & x \cdot \underline{0} \approx \underline{0} \\
7 & x \cdot s(y) \approx x \cdot y + x, \\
8 & x \prec y \leftrightarrow \exists z(x + s(z) \approx y), \\
9 & \text{Induction}
\end{array}
$$

where Induction is:

$$\forall X(\underline{0} \in X \wedge \forall x(x \in X \to s(x) \in X) \to \forall x(x \in X)).[8]$$

First order PA is 1-8 plus Induction formulated as the schema:

$$A(\underline{0}) \wedge \forall x(A(x) \to A(s(x))) \to \forall x A(x)$$

for each first order formula $A(x)$ of L, possibly with parameters. The axioms of PA are theorems of PA^2.

Notice that Induction of PA^2 can be applied, particularizing $\forall X$ to an abstract, only to sets defined by predicative formulae.

Before proceeding, it is worth to ponder whether induction is a permissible principle in a theory of ①. Induction has been considered questionable on the ground that it could happen that $A(\underline{n+1})$ could not be obtained from $A(\underline{n})$

[8]Axioms 1-7 constitute Raphael Robinson's finitely axiomatized theory Q, from [14], a minimal theory that satisfies the hypotheses of the first Gödel's incompleteness theorem; axiom 8 replaces the usual axioms that characterize $\prec$ as the relation of total order generated by s, from which it could be proved (see [15, Ch. 3] for an example of such a theory); conversely the laws of $\prec$ as a total order are derivable from axiom 8. This is to be expected, given that it has the same form of the definition of $\prec$, when $\prec$ is not a primitive symbol of the language.

due to the impossibility of writing $\underline{n+1}$, owing to the restrictions imposed by P1.[9]

Induction is a method of proof that condenses in two lemmas (the basis and the inductive step) an infinity of proofs. It has been so extolled by Blaise Pascal when in 1654 he gave the first deliberate application of this method: "Although this proposition has an infinity of cases, I will give a pretty short demonstration assuming two lemmas".[10]

In constructivistic mathematics induction is widely accepted. David Hilbert considered it, restricted to decidable predicates, as a finitistic tool; for the intuitionists it is valid, even if their series of natural numbers is a mental production that is always finite and progressively extended.

We are not aware that Sergeyev has discussed this principle in any of his writings, but we think that induction is not averse to the spirit of P1, it is at least compatible. When $\forall x A(x)$ has been proved by induction, its finite proof is a template that can be iterated to get a concrete proof of $A(\underline{n})$ for every n reachable with the means available in the numeral system one works with. If the iteration is not feasible, also the writing of $A(\underline{n})$ is not feasible.

We assume that PA^2 is consistent, or that it has a model.

Elementary arithmetic is developed in PA^2 by relying decisively on recursive definitions, after proving the recursion theorem. In this way for example truncated difference and division with remainder are introduced in PA^2. We are not going to prove the recursion theorem, but we will have to resort to it; for a proof see [15, §1.4].

In order to build a Grossone theory $\mathsf{T}_{①}$ we enrich the alphabet with a new (first order) constant symbol ①, and in $\mathsf{L}^2 \cup \{①\}$ we add to PA^2 the following infinite list of axioms:

$$\underline{n} \prec ①$$

one for each $n \in \mathbb{N}$.

When a new symbol is added, here and later for μ, definitions and schemata referring to formulae, such as abstraction, first order induction, ... extend automatically to formulae containing the new symbol.

Since ① is meant to denote a number, we call it a numeral, but its denotation is different from that of the numerals $\underline{n}$ corresponding to the numbers of the metatheory.

The above list of axioms $\underline{n} \prec ①$ translates the Infinity axiom of IUA.

The Identity axiom of IUA is a theorem of $\mathsf{T}_{①}$ because those relations are proved in PA for all x, and hold when x is replaced by ①. Sergeyev himself is well aware of this fact: "In reality, we could even omit this part of the axiom because, due to Postulate 3, all numbers should be treated in the same

[9] The author is grateful to a referee who has rightly urged the necessity to discuss this point.

[10] "Quoique cette proposition ait une infinité des cas, j'en donnerai une démonstration bien courte, en supposant 2 lemmes", [16, p. 103].

way and,[11] therefore, at the moment we have told that Grossone is a number, we have attributed to it the usual properties of numbers, i.e., the properties described in Identity, associative and commutative properties of multiplication and addition, distributive property of multiplication over addition [...]" ([3, §3]). More formally, all these algebraic properties are universal theorems of PA, that particularize to terms containing ①.

The Divisibility axiom will be introduced later, although it is of course the most characteristic, as a first epiphany of P3, and for its original consequences.

$\mathsf{T}_{①}$ is only a first approximation of the Grossone arithmetic, but it requires an infinite schema of axioms.

There is no way to do it with a single or a finite number of formulae. Suppose the set $\{\underline{n} \prec ① \mid n \in \mathbb{N}\}$ were equivalent over PA^2 to a single sentence $A(①)$; then $A(①)$ would be equivalent over PA^2 to a sentence of the form $\underline{n}_1 \prec ① \wedge \ldots \wedge \underline{n}_r \prec ①$ for some $n_1, \ldots, n_r \in \mathbb{N}$; for $A(①)$ would imply over PA^2 all sentences $\underline{n} \prec ①$ and conversely it would be derivable over PA^2 from the set of sentences $\{\underline{n} \prec ① \mid n \in \mathbb{N}\}$, hence from a finite number of them since any derivation if finite.

So $\mathsf{T}_{①}$ would be equivalent to $\mathsf{PA}^2 \cup \{\underline{n}_1 \prec ① \wedge \ldots \wedge \underline{n}_r \prec ①\}$. If $\langle \mathcal{M}, \mathcal{X} \rangle$ is a model of PA^2, and k is a number greater than $n_1, \ldots, n_r$, $\mathsf{T}_{①}$ would have as model the same structure with $\mathcal{M}$ expanded to a structure for $\mathsf{L} \cup \{①\}$ where the new symbol ① is interpreted as the value in $\mathcal{M}$ of $\underline{k}$.[12] But then $s(\underline{k}) \prec \underline{k}$ would be true in $\mathcal{M}$, which entails a contradiction.

Hence $\mathsf{T}_{①}$ is not a finite extension of PA:

Theorem 5.1. *The theory $\mathsf{T}_{①}$ is not finitely axiomatizable over PA^2.* □

Let us note, for further reference, that the set of finite natural numbers $\mathbb{N}$ is not definable by any predicative formula $A(x)$ of $\mathsf{L}^2 \cup \{①\}$. We claim that for any model $\langle \mathcal{M}, \mathcal{X} \rangle$ of $\mathsf{T}_{①}$, with universe M, there is no predicative formula $A(x)$ of $\mathsf{L}^2 \cup \{①\}$ such that A in $\mathcal{M}$ is satisfied by all and only the elements of $\mathbb{N} \subset M$. In this sense,

Theorem 5.2. *The set $\mathbb{N}$ is not definable in $\mathsf{T}_{①}$.*

Proof. Suppose there is a predicative formula $A(x)$ such that A is satisfied in a model $\langle \mathcal{M}, \mathcal{X} \rangle$ of $\mathsf{T}_{①}$ by an element n if and only if $n \in \mathbb{N}$. Applying the minimum principle (which is equivalent to induction) to $\neg A$, we would have that there is an a such that $\neg A(x)$ and $\forall y \prec x A(x/y)$ are satisfied by a as value for x. Since a cannot be the interpretation of $\underline{0}$, there is a b such that a is the

[11] Actually, this is only implicit in P3, which apparently refers exclusively to the whole-part principle. According to this specification P3 should be read as combining two assumptions: the whole-part principle, and the fact of ① sharing all properties of numbers. This is taken care in the logical construction by having ① as a term, on the same footing as $\underline{0}$ or any $\underline{n}$ or $\underline{n} + \underline{m}$.

[12] "Expanded" does not mean "extended": the universe is the same, only the interpretation of new symbols is added to $\mathcal{M}$.

successor of b, b is less than a, so b satisfies A. But then b would belong to $\mathbb{N}$, and so would a, while $a \notin \mathbb{N}$. □

Up to now Grossone hasn't played any role, apart from sharing the algebraic properties of natural numbers; to get more, the burden rests on the divisibility axiom.

For this we need a further extension of the language. We want to associate a measure to sets of numbers; functions assigning cardinalities to sets of numbers can certainly be defined in set theory, but we want to keep at a minimum the logical strength of the tools sufficient to formalize Grossone theory. So the most convenient way is to add a function symbol μ and state suitable axioms for it.

This seems the appropriate place for a remark on the choice of second order predicative logic. As will become apparent in a while, the measure function need only be defined for finite sets, since we cannot assign a finite cardinality to infinite set. Finite sets of numbers can be coded by numbers in many ways. So one could argue that second order logic is entirely dispensable, and it is; but only at the prize of an awful Byzantine nesting of codings: one should extend the coding from pairs to n-tuples, uniformly for variable n, then one could define for example that a number x is a function from u into v if x codes a triple whose second and third projections are respectively u and v, and whose first projection codes a set whose elements are codes of ordered pairs, whose first projections make up the set coded by u and whose second projections belong to the set coded by v.

The use of two types of variables, for numbers and sets of numbers, has no rival for perspicuity and manageability. With two sorts of variables one is still in the domain of first order logic for many sorted languages. The name "second order" refers only to the natural interpretation of the new variables as sets.[13] There is a hierarchy of logics according to the complexity (number of second order quantifiers) of the formulae used in the abstraction or comprehension axioms, and second order predicative logic is the weakest possible, no second order quantifiers. Second order predicative logic is really a first order logic in disguise, sharing the same metalogical properties of first order logic, such as completeness, compactness, Löweheim-Skolem.

6. 2-nd order predicative theory of ①

The symbol μ is a function symbol that the syntactical rules (that we are not going to write) bind to have only second order terms as arguments and first order terms as values; so μ denotes a function assigning natural numbers to sets of natural numbers.

Among these sets a special role is played by initial segments: let us denote by $\prec_x$ the intial segment of numbers less that x, that is:

$$\prec_x = \{y \mid y \prec x\}.$$

[13] See [15, Ch. 4] for a detailed explanation.

Recall that in a set theoretic foundation a natural number, *à la* von Neumann, *is* the set of preceding numbers ($\in$ being the order relation); a number thus coincides with the initial segment it determines; a set X is finite if there is a bijection between X and a natural number n, considered as a segment, and the cardinality of X is n. There cannot be two such distinct n's, since there is no bijection between different initial segments.

In the arithmetical language we have to distinguish between x and $\prec_x$, one is a number, the other a set, but we need to recover the same picture of cardinality; we will start from an abstract concept of measure, axiomatically defined, without mentioning bijections at the beginning, but we will later show that $\mu(\prec_x) \approx x$, and that the measure of a set is that of a segment equivalent to the set, that is the cardinality function.

Our strategy is opposite to that followed by Margestern in [17] to define the number of elements of a set; he defines a set as measured if there is a bijection between the set and an initial segment of a numeral system (Definition 1). Among other results, Margestern proves e.g. that two measured sets have the same number of elements if and only if there is a bijection between them (Theorem 2), and that any non-empty subset of an initial segment, finite or infinite, is measured (Theorem 8). We will obtain formally similar results below in Corollary 2 and 4 (for finite and definable segments). But should we define the measure of a set by the existence of a bijection between the set and an initial segment, which is not a predicative condition, we could not use induction to prove basic properties of the measure. We could not even prove that there is no bijection between different segments. It is even surprising that it can be done, in this roundabout way.[14]

For a start, natural assumptions on μ, irrespective of ①, are the following, with the restriction that X and Y be bounded sets (bounded sets being the sets contained in an initial segment $\prec_x$ for some x):

$$
\begin{array}{llcll}
1_\mu & \mu(\{x\}) & \approx & \underline{1} & \\
2_\mu & \mu(X \cup Y) & \approx & \mu(X) + \mu(Y) & \quad \text{if } X \cap Y \approx \emptyset.
\end{array}
$$

These assumptions justify the name of "measure" given to μ. Conditions $1_\mu, 2_\mu$ are predicative formulae in the language $\mathsf{L}^2 \cup \{\mu\}$. Let PA^2_μ be $\mathsf{PA}^2 \cup \{1_\mu, 2_\mu\}$.

Notice that by 1_μ and 2_μ if $x \notin X$, $\mu(X \cup \{x\}) = \mu(X) + \underline{1}$; so if $X \not\approx \emptyset$ $\mu(X) \succ \underline{0}$. In general

$$X \subseteq Y \to \mu(X) \preccurlyeq \mu(Y)$$

is provable, in PA^2_μ, since if $X \subseteq Y$ then $Y \approx X \cup (Y \setminus X)$, and

$$X \subset Y \to \mu(X) \prec \mu(Y)$$

[14] Given our metamathematical interest, we pay attention to restrict results and proofs within the limits of the weak logical frame adopted. Margestern's paper is a mathematical contribution to the theory of ①, driven and constrained by postulates P1 and P3, avoiding axiomatization or the explicit statement of its mathematical assumptions; therefore it is not plain to compare its content with our theory; his results are more general, referring to generic numeral systems, possibly including ①.

since $\mu(Y \setminus X) \not\approx \underline{0}$. Moreover:

$$\mu(\emptyset) \approx \underline{0},$$

since $\mu(X) \approx \mu(X \cup \emptyset) \approx \mu(X) + \mu(\emptyset)$. This is one of the reasons for having 0 among the natural numbers.

From 1_μ and 2_μ we can prove a relation between x and $\prec_x$ that mirrors what in the set theoretic foundaton is a bonus of the definitions:

Lemma 6.1. $\mathsf{PA}^2_\mu \vdash_2 \ \forall x(\mu(\prec_x) \approx x)$.

Proof. By induction on x. For $x \approx \underline{0}$, $\mu(\prec_0) \approx \mu(\emptyset) \approx \underline{0}$. Assuming $\mu(\prec_x) \approx x$ as inductive hypothesis, since $\prec_{s(x)} \approx \prec_x \cup \{x\}$ and $\prec_x \cap \{x\} \approx \emptyset$, it follows that $\mu(\prec_{s(x)}) \approx \mu(\prec_x) + 1 \approx x + 1 \approx s(x)$. □

To proceed, we need a lemma, to the effect that given any x, any $X \subseteq \prec_x$ can be put in one-to-one correspondence with an initial segment $\prec_z$ for a $z \preccurlyeq x$; actually we need to be more precise, since we do not want an $X \subset \prec_x$ to be in one-to-one correspondence with $\prec_x$, for P3, so we state the stronger proposition:

Lemma 6.2. $\mathsf{PA}^2 \vdash_2 \forall x \forall X (X \subset \prec_x \to \exists z \prec x \exists F(F : \prec_z \xrightarrow{1-1} X))$.

Proof. Given x, consider $X \subset \prec_x$, $X \not\approx \emptyset$, hence $x \not\approx \underline{0}$ (if $X \approx \emptyset$ we would take as F the empty set; if $x \approx \underline{0}$, there is no $X \subset \prec_x$). We define by recursion a function from numbers into X. Call $min(X)$, the minimum of X, the element of X such that $\forall y \prec x(y \notin X)$. The restriction of a function F to k, $F \restriction_k$, is the set $\{\langle u, v\rangle \in F \mid u \prec k\}$.

Define

$$\begin{cases} F(\underline{0}) & \approx \quad min(X) \\ F(s(u)) & \approx \quad \begin{cases} min(X \setminus im(F \restriction_{s(u)})) & \text{if } X \setminus im(F \restriction_{s(u)}) \not\approx \emptyset \\ F(u) & \text{otherwise} \end{cases} \end{cases}$$

Notice that the conditions on the right are predicative, so the recursion theorem applies.

There must be a first y such that $F(s(y)) \approx F(y)$ and F becomes constant after y, because until $F(s(u))$ is defined by the first clause, its values are strictly increasing, hence $u \preccurlyeq F(u)$, and should this go on up to x, $F(u)$ could not belong to X. The first y such that $F(s(y)) \approx F(y)$ occurs when $X \approx im(F \restriction_{s(y)})$; take the restriction of F to $s(y)$ and call it F_X; this restricted F_X has domain $\prec_{s(y)}$.

F_X is strictly increasing below $s(y)$, hence $y \preccurlyeq F_X(y)$; since $X \subset \prec_x$, and $F_X(y) \in X$, we can say that $F_X(y) \prec x$, hence $y \prec x$ and $s(y) \preccurlyeq x$; but in fact $s(y) \prec x$. As $X \not\approx \prec_x$, either X is contained in a smaller segment, or it has a gap, say at u (the first element of X greater than u is greater than $s(u)$); then $s(u) \prec F_X(s(u))$ and the relation is mantained onward, so that $y \prec F_X(y)$; in any case $s(y) \prec x$ and $s(y)$ is the z of the statement of the lemma. □

The proof gives us more, in terms of μ; let us imagine a counter such that, whenever we define $F_X(s(u))$ by the first clause, choosing a new element of X, it adds 1. When the second clause applies, it stops. The counter parallels the computation of the measure of X according to 1_μ and 2_μ; we conclude that $\mu(X)$ equals $\sum_1^y \underline{1} = s(y)$, $s(y)$ as given from the proof, and adding Lemma 1:

Corollary 6.1. $\mathsf{PA}^2_\mu \vdash_2 \forall x(\forall X(X \subset \prec_x \rightarrow \exists z \prec x(\mu(X) \approx z)) \wedge \mu(\prec_x) \approx x)$. □

Corollary 1 shows that the simple axioms 1_μ and 2_μ suffice to ensure that μ is far from trivial; we know that every bounded set has a measure; however in order to compute the measure of specific sets we might be interested in, we must have more information, which is given by the existence of bijections.

To talk of bijections in the context of Grossone theory seems to jeopardize the intended consequences of P3, but it will not be so. The restriction to bounded sets is essential to preserve the principle that the part is less than the whole. Consider for example that there is a bijection between the set of even numbers less than $2x$ and the set of odd numbers less than $2x$, which we will use later, but no bijection between any of them and $\prec_{2x}$. As Sergeyev repeatedly stresses, Grossone theory does not want to contradict Cantor's theory, but to introduce finer distinctions.

The proof of Lemma 2 shows that the z given by the proof is the smallest number for which there is an enumeration in increasing order of X of length z. Let us call such z the *collapse* of X, $c(X)$, and F_X the collapsing function of X.

If X and Y have the same collapse, then there exist a bijection $G : X \xrightarrow{1-1} Y$ by composition of the two collapsing functions, and their inverses. Conversely if there is a bijection $G : X \xrightarrow{1-1} Y$, and if it preserves the order than it is clear that the collapsing function of X composed with G is the collapsing function of Y, so $c(X) \approx c(Y)$; this is the case we will have to consider later in Theorem 3 with respect to the sets $\mathbb{N}_{k;n}$ and $\{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\}$.

If G does not preserve the order, the conclusion is reached by a devious route:[15] in defining the collapsing function of Y as in Lemma 2, F_Y is defined at $s(u)$ by the first clause only if the set of $F_Y(v)$'s for $v \prec s(u)$ does not exhaust Y; this happens if and only if the set of the G^{-1} images of these elements does not exhaust X; the same in the other direction, when we define F_X, using G images. This means that F_X and F_Y are defined by the first clause for $s(u)$ for the same $s(u)$'s; so X and Y have the same collapse.

The upshot of this reasoning is that in PA^2_μ for bounded sets holds:

Corollary 6.2. *If* $X, Y \subseteq \prec_x$, $\mu(X) \approx \mu(Y)$ *if and only if there is a bijection between* X *and* Y. □

We can also make more precise the statement of Corollary 1 by indicating explicitely that the z declared to exist is the collapse of X:

[15] The reason is that we do not want to make use of the fact that there is no bijection between different initial segment, a property whose usual proof apparently requires induction applied to a formula that is not predicative. This fact however will be a consequence of the Corollary 2: if there were a bijection between $\prec_a$ and $\prec_b$ then $a \approx \mu(\prec_a) \approx \mu(\prec_b) \approx b$.

Corollary 6.3. $\mathsf{PA}^2_\mu \vdash_2 \forall x \forall X(X \subset \prec_x \to (c(X) \prec x \wedge \mu(X) \approx c(X)))$. $\square$

Now enters Grossone; add 1_μ and 2_μ to $\mathsf{T}_{①}$, and by particularizing $\forall x$ to ① in Lemma 1 and Corollary 1,

Corollary 6.4.

$$\mathsf{T}_{①} \cup \{1_\mu, 2_\mu\} \vdash_2 \forall X(X \subset \prec_{①} \to \exists y \prec ①(\mu(X) \approx y)) \ \wedge \ \mu(\prec_{①}) \approx ① \quad \square$$

that is, all proper subsets of $\prec_{①}$ have a measure which is a number $\prec$ ① and $\prec_{①}$ has measure ①.

Before turning to the measure of the sets $\mathbb{N}_{k;n}$, and the divisibility axiom, some clarifications of Sergeyev's presentation are necessary.

After introducing ① with the Infinity axiom Sergeyev warns us that to him the set $\mathbb{N}$ is now what we would define as the initial segment up to and including ①:

> The new numeral ① allows one to write down the set, $\mathbb{N}$, of natural numbers in the form
>
> $$\mathbb{N} = \{1, 2, 3, \ldots, ① - 3,\ ① - 2,\ ① - 1, ①\} \tag{5}$$
>
> where the numerals
>
> $$\ldots, ① - 3,\ ① - 2,\ ① - 1, ① \tag{6}$$
>
> indicate infinite natural numbers.
>
> It is important to emphasize that in the new approach the set (5) is the same set of natural numbers
>
> $$\mathbb{N} = \{1, 2, 3, \ldots\} \tag{7}$$
>
> we are used to deal with and infinite numbers (6) also take part of $\mathbb{N}$. Both records, (5) and (7), are correct and do not contradict each other. They just use two different numeral systems to express $\mathbb{N}$. Traditional numeral systems do not allow us to see infinite natural numbers that we can observe now thanks to ①. ([3, §3])

Since we want to reserve $\mathbb{N}$ exclusively to the notation for our metamathematical numbers, we have better distinguish clearly between (5) and (7), and the set $\{1, 2, 3, \ldots, ① - 3, ① - 2, ① - 1, ①\}$ will be denoted differently below.

We interpret however Sergeyev's remark in the sense that when he defines, as we have seen above, $\mathbb{N}_{k;n} = \{k, k+n, k+2n, k+3n, \ldots\}$ he means by this the set of the numbers of the form $k + yn$ which are less than or equal to ①. They cannot be only finite numbers, since by Theorem 5.2 the set of finite numbers is not definable.

Sergeyev's $\{1, 2, 3, \ldots, ① - 3, ① - 2, ① - 1, ①\}$ will be first replaced, to compensate for the addition of 0, by $\{0, 1, 2, \ldots, ① - 3, ① - 2, ① - 1\}$, then denoted for what this is, namely

$$\{x \mid x \prec ①\}.$$

Consequently our $\mathbb{N}_{k;n}$ will be represented as

$$\{x \mid \exists y\, (x \approx \underline{k} + y \cdot \underline{n}) \wedge x \prec ①\}.$$

Moreover, we will allow k to be 0; small adjustements are needed to recover the correspondence with the original definition; for example, Sergeyev's set of even numbers $\mathbb{N}_{2;2} = \{2, 4, \ldots\}$ corresponds to $\{x \mid \exists y(x \approx y \cdot \underline{2}) \wedge x \prec ①\} \setminus \{\underline{0}\}$. By the same observation, we can also assume that $k < n$ because for $k = n$ Sergeyev's set $\mathbb{N}_{n;n}$ is $\{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\} \setminus \{\underline{0}\}$.

But apart from these correspondences, what is required is that our sets $\mathbb{N}_{k;n}$ determine a partition of $\prec_{①}$ into n equivalent parts, as k assumes n values between 0 and $n - 1$.

Let us first concentrate on the sets $\mathbb{N}_{0;n}$, which are definable, for each $n \in \mathbb{N}$, as $\{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\}$.

Consider $n = 2$ and the following fact: in PA^2 it can be proved that the number of even numbers less than $2x$ is x, for any x; define the function $F(y) \approx 2y$ and restrict it to $\prec_x$: x is the collapse of the set of even numbers less than $2x$, and F its collapsing function.[16]

In the same way, one proves that the number of odd numbers less than $2x$ is x.

Now if ① were an even number, we would get from Corollary 3 by particularizing $\forall x$ to ①

Theorem 6.1.

$$\mathsf{T}_{①} \cup \{1_\mu, 2_\mu\} \vdash_2 \mu(\{x \mid x\, \text{even} \wedge x \prec ①\}) \approx ①/\underline{2}$$

and

$$\mathsf{T}_{①} \cup \{1_\mu, 2_\mu\} \vdash_2 \mu(\{x \mid x\, \text{odd} \wedge x \prec ①\}) \approx ①/\underline{2}.$$

But is ① an even number? In Sergeyev's approach the only way to get this seems to be by appeal to the divisibility axiom, where it is assumed in particular that $\mathbb{N}_{0;2}$ has ①/2 elements, and ①/2 is a number, hence that ① is divisible by 2. Actually the divisibility axiom of IUA implicitely states that ① is divisible by all finite integers (whence the name, we surmise, although Sergeyev seems to give more importance to the idea of the splitting of the set of natural numbers into equal parts).

The divisibility axiom has two aspects: one is the purely arithmetical fact of the divisibility of ① by any finite natural number; the other is the partition of $\prec_{①}$ into n intuitively equivalent and equinumerous parts.

[16] In case of $\mathbb{N} = \{1, 2, 3, \ldots\}$ one should use segments $\preccurlyeq_x = \{y \mid y \preccurlyeq x\}$ and observe that the number of even numbers less than or equal to $2x$ is x.

Is it necessary to formulate the divisibility axioms exactly as Sergeyev does, or is it possible to assume explicitely only the divisibility of ① and to try to calculate the μ measure of the $\mathbb{N}_{k;n}$ sets? We explore this way, and accordingly we complete the definition of Grossone theory by adding the axiom schema:

Divisibility axiom: for all finite $n > 0$

$$\exists x(x \approx ①/\underline{n}).$$

So in the end the Grossone theory we propose is, in $\mathsf{L}^2 \cup \{①, \mu\}$,

$$\mathsf{T}^2_{①} = \mathsf{PA}^2_\mu \cup \{\underline{n} \prec ① \mid n \in \mathbb{N}\} \cup \{\exists x(x \approx ①/\underline{n}) \mid n \in \mathbb{N},\, n \neq 0\}.$$

Theorem 1 above is then already proved, by the argument preceding it, if we substitute $\mathsf{T}^2_{①}$ to $\mathsf{T}_{①} \cup \{1_\mu, 2_\mu\}$. Notice that the reasoning we did on the number of even numbers can be generalized. For example, informally, the number of numbers of the form $3y$ less than $3x$ is x; once this is proved in PA^2 for all x, if ① is divisible by 3 one can particularize x to ① and obtain that the set of numbers of the form $y \cdot \underline{3}$ less than ① has measure $①/\underline{3}$. The same in general for $\mathbb{N}_{0;n}$. So

Theorem 6.2. *For every $n \in \mathbb{N}$, $n \neq 0$,*

$$\mathsf{T}^2_{①} \vdash_2 \mu(\{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\}) \approx ①/\underline{n}. \quad \square$$

As for $\mathbb{N}_{k;n}$ with $k \neq 0$, notice that $\mathbb{N}_{k;n}$ is the translation by k of $\mathbb{N}_{0;n} = \{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\}$.

Since ① is divisible by $\underline{n}$, the last element of this set is $① - \underline{n}$, and since $\underline{k} \prec \underline{n}$, it follows $\mathbb{N}_{k;n} \subset \prec_{①}$. Thus both $\mathbb{N}_{k;n}$ and $\{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\}$ are proper subsets of $\prec_{①}$ and clearly they have the same collapse.

So by Theorem 2

Theorem 6.3. *For every $n \in \mathbb{N}$, $n \neq 0$, and every k, $0 \leq k < n$,*

$$\mathsf{T}^2_{①} \vdash_2 \mu(\mathbb{N}_{k;n}) \approx ①/\underline{n}. \quad \square$$

With this axiomatization and Theorem 3 to complete Sergeyev's divisibility axiom, we probably diverge from his intuition, according to which his axiom is motivated by P3; we separate the divisibility of ① from the measure of subsets of $\prec_{①}$; we see the possibility of such a measure as a provable extension of properties of the finite to ①. The worth of an axiomatized theory is that it is compatible with multiple intuitions.

This does not mean that in $\mathsf{T}^2_{①}$ the postulate P3 can be violated, even when working with sets not contained in $\prec_{①}$. Sergeyev considers such sets, exploring various properties, when he deals with what he calls the extended natural numbers, denoted by $\hat{\mathbb{N}}$; this is the structure generated by applying arithmetical operations to ①, including $①+1, ①+2, \ldots, n \cdot ①, \ldots, ①^2, \ldots, ①^{①}, \ldots$ We will mention a result of Sergeyev in Section 8.

In our perspective, $\hat{\mathbb{N}}$ is a model of $\mathsf{T}^2_{①}$. We have focused our attention on $\prec_{①}$ in order to discuss *Divisibility*, but of course in $\mathsf{T}^2_{①}$ all uses of ① and all techiniques employed by Sergeyev can be recovered as long as they rely on arithmetical computations. $\mu(X) \prec \mu(Y)$ holds for $X \subset Y$ whenever X and Y are measurable. We can particularize Corollary 3 to every set contained in $\prec_t$ where t is an arithmetical term built from ①, denoting an element of $\hat{\mathbb{N}}$.

7. Metamathematical properties of $\mathsf{T}^2_{①}$

In our research we have been motivated by the aim of untangling the divisibility axiom, but at the same time of formalizing ①'s properties in a weak and natural theory. We can extend Theorems 5.1 and 5.2 to $\mathsf{T}^2_{①}$ and prove other metamathematical properties, for example that $\mathsf{T}^2_{①}$ is a conservative extension of PA^2_μ, meaning that every sentence A of $\mathsf{L}^2 \cup \{\mu\}$ (not containing ①) provable in $\mathsf{T}^2_{①}$ is already provable in PA^2_μ.

Theorem 7.1. $\mathsf{T}^2_{①}$ *is a conservative extension of* PA^2_μ.

Proof. Suppose $\mathsf{T}^2_{①} \vdash_2 A$, A non containing ①. Then there is a finite number of the axioms of Infinity $\underline{n}_1 \prec ①, \ldots, \underline{n}_r \prec ①$, and a finite number of the axioms of divisibility $\exists x(x \approx ①/\underline{m}_1), \ldots, \exists x(x \approx ①/\underline{m}_s)$ such that

$$\mathsf{PA}^2_\mu \cup \{\underline{n}_1 \prec ①, \ldots, \underline{n}_r \prec ①, \exists x(x \approx ①/\underline{m}_1), \ldots, \exists x(x \approx ①/\underline{m}_s)\} \vdash_2 A,$$

or

$$\mathsf{PA}^2_\mu \vdash_2 \underline{n}_1 \prec ① \wedge \ldots \wedge \underline{n}_r \prec ① \wedge \exists x(x \approx ①/\underline{m}_1) \wedge \ldots \wedge \exists x(x \approx ①/\underline{m}_s) \to A.$$

Since ① does not occur neither in A nor in PA^2_μ, the logical rules allow us to claim that

$$\mathsf{PA}^2_\mu \vdash_2 \underline{n}_1 \prec u \wedge \ldots \wedge \underline{n}_r \prec u \wedge \exists x(x \approx u/\underline{m}_1) \wedge \ldots \wedge \exists x(x \approx u/\underline{m}_s) \to A,$$

where u is a new variable not occurring in the above last derivation.

It follows that

$$\mathsf{PA}^2_\mu \vdash_2 \exists u(\underline{n}_1 \prec u \wedge \ldots \wedge \underline{n}_r \prec u \wedge \exists x(x \approx u/\underline{m}_1) \wedge \ldots \wedge \exists x(x \approx u/\underline{m}_s) \to A,$$

but the antecedent of the implications is provable, already in PA, hence the conclusion. □

By a similar argument, $\mathsf{T}_{①}$ is a conservative extension of PA^2.

Notice that conservativeness is not a weakness of the theory, rather it seems consistent with Sergeyev's philosophy: the new tool allows us to see new things, it does not strengthen the old ones, nor falsifies their observations. At the same time, a traditional mathematician can be assured that all that is most cherished is preserved, as long as it is provable in the theory of which $T^2_{①}$ is a conservative extension.

Conservativity implies consistency of the extension, if the extended theory is consistent: if $\mathsf{T}^2_{①}$ were inconsistent, every formula could be proved, in particular $\underline{0} \approx \underline{1}$, but then PA^2_μ would also prove it, hence it would be inconsistent.

However, we outline also a direct proof, which is a straightforward application of the logical compactness theorem, and has a more mathematical flavour.

The compactness theorem says that a theory T is consistent if and only if every finite subset $S \subseteq T$ is consistent (see [15, §4.3] for a proof for many-sorted logic that applies also to second order predicative logic).

Theorem 7.2. *The theory* $\mathsf{T}^2_{①}$ *is consistent.*

Proof. We assume that PA^2_μ has a model.[17]

A finite subset S of $\mathsf{T}^2_{①}$ has as elements some axioms of PA^2, some sentences $\underline{n} \prec ①$, some sentences $\exists x(x \approx ①/\underline{n})$, and possibly some μ axioms.

Choose as the interpretation of ① an element of M greater than all the n's such that $\underline{n} \prec ①$ occur in S, and which is divisible by all the n occurring in the $\exists x(x \approx ①/\underline{n})$ in S. These sentences become true in the expanded $\mathcal{M}$, hence S has a model. $\square$

The statement of the theorem is of course conditional, as apparent from the proof, upon the consistency of PA^2_μ. It would be more satisfactory to assume only the consistency of PA^2, an improvement that would be attained if the consistency of PA^2_μ were implied by the consistency of PA^2. Now this is true but its model theoretic proof (transform any model of PA^2 in a model of PA^2_μ) is technically rather demanding. Alternatively, we sketch an informal syntactic argument, to convince the reader of the plausibility of the conclusion that $\mathsf{T}^2_{①}$ is consistent if PA^2 is consistent.

The μ axioms are properly one sentence, the conjunction of 1_μ, with a prefixed quantifier $\forall x$, and of 2_μ, with prefixed quantifiers $\forall X$ and $\forall Y$ restricted to bounded sets. Let us call it $A(\mu)$. Assuming PA^2 consistent, if $\mathsf{PA}^2 \cup \{A(\mu)\}$ were inconsistent then $\mathsf{PA}^2 \vdash_2 \neg A(\mu)$. Given such a derivation, a proof theoretic argument, already used in the proof of Theorem 1, shows that, since μ does not occur in PA^2, then $\mathsf{PA}^2 \vdash_2 \neg A(F)$ where $A(F)$ is obtained by substituting μ in A by F, F a new variable not occurring in the given derivation. There would follow $\mathsf{PA}^2 \vdash_2 \neg\exists F A(F)$. But this is preposterous, because such an F exists, it is definable in set theory; since PA^2 is included, or interpretable in set theory, it would mean that set theory is inconsistent.

[17] With respect to models $\langle \mathcal{M}, \mathcal{X} \rangle$ for PA^2, we should have here also a function from bounded subsets of $\mathbb{N}$ to $\mathbb{N}$, call it $\mu^{\mathcal{M}}$, as interpretation of μ, in an expansion of $\mathcal{M}$, but the details are immaterial.

8. Further investigations

Corollary 6.3 guarantees that every subset of $\prec_{①}$ has a measure; the result appears to be strong, but not immediately useful; an interesting question would be that of characterizing the sets to which we are able to assign a μ-value, those X for which there is a closed term t such that $\mathsf{T}^2_{①} \vdash_2 \mu(X) \approx t$. We know the answer for the finite sets, for the sets $\{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\}$ and for sets obtainable from these with simple set theoretic operations.

A plausible conjecture could be that the answer is the algebra generated by the finite initial segments and the $\{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\}$.

But the computations become soon complicated: for example the intersection of $\{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\}$ and $\{x \mid \exists y(x \approx y \cdot \underline{m}) \wedge x \prec ①\}$ is still of the same form: $\{x \mid \exists y(x \approx y \cdot (\underline{n} \cdot \underline{m})) \wedge x \prec ①\}$, but in order to compute the μ of the union, this must be substituted by the disjoint union and expressed as the union of $\{x \mid \exists y(x \approx y \cdot \underline{n}) \wedge x \prec ①\}$ and $\{x \mid \exists y(x \approx y \cdot \underline{m}) \wedge x \prec ①\}$ minus $\{x \mid \exists y(x \approx y \cdot (\underline{n} \cdot \underline{m})) \wedge x \prec ①\}$.

Notice however that the cardinality of many sets can be evaluated from that of the $\mathbb{N}_{k;n}$'s, and this fact shows Sergeyev's ingenuity in formulating his axiom of divisibility. As an example, consider the set of square numbers. We will use the same argument employed for the $\mathbb{N}_{k;n}$, looking for the number of square numbers $\leq x$. We leave out 0.

Define $\lfloor\sqrt{x}\rfloor$ as the maximum n such that $n^2 \leq x$. Remember the formula

$$1 + 3 + \ldots + (2n-1) = n^2, \text{ for } n \geq 1.$$

For a natural number $x > 0$ the maximum square $\leq x$ is $\lfloor\sqrt{x}\rfloor^2$, which is equal to the sum of odd numbers up to $2\lfloor\sqrt{x}\rfloor - 1$. Moreover the positive squares $\leq x$ are as many as the odd numbers $\leq 2\lfloor\sqrt{x}\rfloor - 1$. We already know there are $\lfloor\sqrt{x}\rfloor$ of them.

In $\mathsf{T}^2_{①}$ we can define the floor function and prove such facts for every x, hence prove that the subset of $\preccurlyeq_{①}$ made of the positive square numbers has measure $\lfloor\sqrt{①}\rfloor$. To include 0, add 1.

Sergeyev obtained this result in a different way as a special case of a more general formula in [3, §5.4, Example 5.7]. He considers sets of the form

$$G = \{g(i) \mid i \geq 1, 0 < g(i) \leq b\}$$

for a strictly increasing g and $b \in \hat{\mathbb{N}}$ (in Example 5.7 $g(i) = k + ni^j$ and $b = ①$), and observes that the number of elements of G can be determined as $J = \max\{i \mid g(i) \leq b\}$ if we are able to determine the inverse function g^{-1}; then $J = \lfloor g^{-1}(b)\rfloor$.

In our terminology, this amounts to note that $g^{-1}(G)$ is the collapse of $im(g)$. We could prove the result by appeal to Corollary 6.3 and the remarks preceding it on strictly increasing functions. This shows that the techniques we have employed in our axiomatized theory are not foreign to the mathematical practice.

A different pressing task is obviously that of considering ① as a full fledged real number; then one would have to take the infinitesimal 1/① into consideration, and to build its theory comparing it with current non-Archimedean fields theories.

[1] Ya. D. Sergeyev, Arithmetic of infinity, Edizioni Orizzonti Meridionali, CS, Italy, 2003.

[2] Ya. D. Sergeyev, Blinking fractals and their quantitative analysis using infinite and infinitesimal numbers, Chaos Solitons Fractals 33 (1) (2007) 50–75.

[3] Ya. D. Sergeyev, A new applied approach for executing computations with infinite and infinitesimal quantities, Inform. 19 (4) (2008) 567-96.

[4] Ya. D. Sergeyev, Evaluating the exact infinitesimal values of area of Sierpinski's carpet and volume of Menger's sponge, Chaos Solitons Fractals 42 (5) (2009) 3042–3046.

[5] Ya. D. Sergeyev, On Accuracy of Mathematical Languages Used to deal with the Riemann Zeta Function and the Dirichlet Eta Function, p-Adic Numbers, Ultrametr. Anal. Appl. 3 (2) (2011) 129-48.

[6] Ya. D. Sergeyev, A. Garro, Observability of Turing Machines: a refinement of the theory of computation, Inform. 21 (3) (2010) 425–454.

[7] YA. D. Sergeyev, A. Garro, Single-tape and multi-tape Turing machines through the lens of the Grossone methodology, Journal of Supercomputing 65 (2) (2013) 645-663.

[8] S. De Cosmis, R. De Leone, The Use of Grossone in Mathematical Programming and Operations Research, Appl. Math. Comput. 218 (16) (2012) 8029-8038.

[9] V. Benci, M. Di Nasso, Numerosities of labeled sets: a new way of counting, Adv. Math. 173 (2003) 50-67.

[10] G. Lolli, Infinitesimals and Infinites in the History of Mathematics: a Brief Survey, Appl. Math. Comput. 218 (16) (2012) 7977-8216.

[11] T. L. Heath (Ed.), The Thirteen Books of Euclid's Elements, 3 voll., Dover, New York, 1956.

[12] H. Chang, Operationalism, in : The Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/archives/fall2009/entries/ operationalism/.

[13] M. Davis, Computability & Unsolvability, McGraw-Hill, New York, 1958.

[14] R. Robinson, An essentially undecidable axiom system, Proceed. International Congr. Mat. 1950, Cambridge MA, 1952, vol. 1, pp. 729-30.

[15] H. B. Enderton, A Mathematical Introduction to Logic, second ed., Academic Press, New York, 2001.

[16] B. Pascal, Œuvres complètes, Gallimard, Paris, 1954.

[17] M. Margenstern, Using Grossone to count the number of elements of infinite sets and the connection with bijections, p-Adic Numbers, Ultrametr. Anal. Appl. 3 (3) (2011) 196-204.