A Defense of Strict Finitism

Jean Paul Van Bendegem • Vrije Universiteit Brussel, Belgium • jpvbende/at/vub.ac.be

> Context • Strict finitism is usually not taken seriously as a possible view on what mathematics is and how it functions. This is due mainly to unfamiliarity with the topic. **> Problem** • First, it is necessary to present a "decent" history of strict finitism (which is now lacking) and, secondly, to show that common counterarguments against strict finitism can be properly addressed and refuted. **> Method** • For the historical part, the historical material is situated in a broader context, and for the argumentative part, an evaluation of arguments and counterarguments is presented. **> Results** • The main result is that strict finitism is indeed a viable option, next to other constructive approaches, in (the foundations of) mathematics. **> Implications** • Arguing for strict finitism is more complex than is usually thought. For future research, strict finitist mathematics itself needs to be written out in more detail to increase its credibility. In as far as strict finitism is a viable option, it will change our views on such "classics" as the platonist-constructivist discussion, the discovery-construction debate and the mysterious applicability problem (why is mathematics so successful in its applications?). **> Constructivist content** • Strict finitism starts from the idea that counting is an act of labeling, hence the mathematician is an active subject right from the start. It differs from other constructivist views in that the finite limitations of the human subject are taken into account.
> Key words • **Mathematics, finite, largest number, infinite, limits, budget constraints.**

1. Introduction

When discussing with either mathematicians, philosophers acquainted with mathematics, or philosophers that are not, invariably, the idea of a mathematics that is without infinity being a conceivable or even workable one is met with a number of spontaneous counterarguments. Often these objections are of an informal nature, not requiring a formal rejoinder. In this paper a number of those (the most important ones, at least to my mind) are looked into and rejected. Should this succeed then of course the "real" work, to formally build a full-fledged strictly finite mathematics, has not even begun. But at least a first hurdle would have been taken. This, in a nutshell, is the modest goal of this contribution. In the next section, I will begin by clearing away a number of historical misconceptions about strict finitism, before dealing with a number of different objections in the sections to follow.

2. Some historical remarks

Strict finitism, over the course of the years, has mostly received a *bad press*. The reasons are many, and often involve a "natural" disgust for the subject (to be dealt with in the sections to follow). Also of importance is the lack of consensus about the label itself: strict finitism, ultrafinitism, and ultra-intuitionism are often used without distinction, which adds to the confusion.

Very often, the name Alexander Yessenin-Volpin (or Essenine-Volpin or Ésenine-Volpine) is mentioned as that of its founder. This is to be lamented for a couple of reasons. First of all, his papers, such as e.g.,Yessenin-Volpin (1970), the most famous among them, are extremely cryptical, not to say nonsensical. Moreover, he had rather specific purposes in mind, which are not necessarily shared by strict finitists (including the present author). Incidently, Yessenin-Volpin himself described his approach as ultra-intuitionism not strict finitism. Very briefly, and doing him injustice, the central idea was the following: consider classical set theory, *ZFC* (Zermelo-Fraenkel with the choice axiom). We know, thanks to Kurt Gödel, that the consistency of this system cannot be shown by *ZFC*'s own means. One would normally think that finitary means are surely excluded because these are easily formulated in set theory. But, as Gödel himself remarked in his famous 1931 paper, it is not excluded that there exist finitary proof methods that are themselves *not* expressible in *ZFC*.[1] This goal in mind, Yessenin-Volpin introduced so-called "Zenonian" sets. Characteristic of this kind of set is that: first, if n belongs to it, then so does $n+1$; and second, nevertheless, in its entirity, the set is finite. To give substance to this curious idea, he gives the following example: consider the set of all the heartbeats in one's youth. For each and every heartbeat there has been a next one. Still, there have only been but a finite number of heartbeats in one's youth. What has just been said probably already makes clear that Yessenin-Volpin's perspective can hardly be qualified as strict finitism, as in the latter the infinite is squarely banned from mathematics. This being said, recent authors such as David Isles (1992, 1994) have succeeded in translating (some of) Yessenin-Volpin's ideas into a formal and accessible presentation, allowing for some nuance in this debate.

A second source often mentioned in connection with strict finitism, and also causing a lot of confusion is Ludwig Wittgenstein, more particularly his *Bemerkungen über die Grundlagen der Mathematik* (Wittgenstein 1984). The view that this rather enigmatic work betrays strict finitist

1 | Here is what the Gödel paper says: "For this viewpoint presupposes only the existence of a consistency proof in which nothing but finitary means of proof is used, and it is conceivable that there exist finitary proofs that *cannot* be expressed in the formalism of P (or of M or A)" (Gödel 1971: 615).

sympathies in the author can be defended. Many paragraphs deal with *feasibility*, a concept referring to what is mathematically executable by real humans (situated in space and time), and some paragraphs contain very critical remarks with respect to infinity. It is however a very intricate affair to sort out whether or not Wittgenstein was indeed a strict finitist. The *Bemerkungen* in no way contain a systematic proposal to this effect, and it is even near to impossible to distil one out of them for oneself. Add to this that many mathematicians and philosophers do not in general have an awful amount of sympathy for Wittgenstein, and it becomes clear that, like Yessenin-Volpin, he is not likely to become the best promoter of the strict finitist's cause. Nevertheless, in the case of Wittgenstein also, a lot of authors have tried to deduce something useful (intelligible, coherent) from his writings, and this material is actually very interesting for the strict finitist, although it is of course not necessarily developed from within this spirit. Some examples are Wright (1980), Marion (1998), and Rodych (2000).

Should one be urged to point to an original source for strict finitism, then one of the best candidates would probably be an article from 1956 by David Van Dantzig[2] with the title "Is $10^{10^{10}}$ finite number?" Although no full-fledged, explicit proposal is formulated in it, it does contain a very clear defense of strict finitism as well as an impetus to get to work with various, finite number systems. Roughly, the idea is to start with n_1 natural numbers forming a row S_1, and then to look at the sums of these, obtaining another row S_2, and so on. At no one particular moment does one have to deal with the entire row of naturals, and supposing one is able to add and multiply, this does not imply the possibility of exponentiation. Interestingly, van Dalen (1978) and Epstein & Carnielli (2000) not only mention Yessenin-Volpin but also van Dantzig as pioneers. Whoever one wants to point to as its founder(s), it is clear that strict finitism as a school of thought is quite young, hence material is not abundant (yet). In addition to this, I want to draw the attention to three directions that, though not seeing themselves as strict finitist, in a number of aspects come very close to it and can thus, at the very least, be inspiring.

A first trend starts from classical mathematics, e.g., Peano Arithmetic (*PA*), and adds to this a certain operator expressing a limitation. A well-known approach is that of Parikh (1971), where a *feasibility* operator is added to *PA*, having the properties that, first, if natural numbers *n* and *m* are *feasible* then so is the sum $n+m$, and second, if natural numbers *n* and *m* are *feasible*, then so is the multiplication $n \times m$, but, third, 2 to the 10000th power is not. Briefly, adding and multiplying always works, exponentiation does not. Obviously, this is no strict finitist approach, because for the strict finitist summation itself is also not closed. Nevertheless, as said, this material is very inspiring for the strict finitist. Along the same lines, Shaughan Lavine (1994) has devised a beautiful model for transfinite set theory that assumes the potential infinite.

A second trend has to do with the study of fragments of *PA*. This seldom pertains to strict finitism but other philosophical and foundational concerns are involved. Presburger arithmetic (Presburger 1929), for example, retaining addition but omitting multiplication, has the nice property of being decidable, as opposed to regular arithmetic, e.g., *PA*, which is incomplete and undecidable. A lot of research is concerned with the question of how far the lower limit can be raised: what can one add to Presburger arithmetic while maintaining decidability and/or other pleasant properties (vis-à-vis *PA*) of the extension. Similarly, in Robinson arithmetic or system *Q* (see Nelson (1986)), the full induction axiom is dropped, bringing up questions as to the means of expression without and with several limited versions of the axiom. Needless to say, this kind of material is also very interesting to the strict finitist.

Finally, a third trend also takes *PA* as a starting point, and via (sometimes minimal) changes to the axioms tries to bring into play finite models, next to the classical infinite ones. One of the earliest proposals to this effect is Jan Mycielski (1981). This will moreover be relevant for our discussion of the argument from poverty below. In *PA*, the so-called successor function is of central importance. This function *s* maps numbers to numbers, such that, for each *n*, $s(n)$ is the next number in line. Two axioms determine the structure of this function, to wit:

(s1) $\sim(\exists n)(s(n)=0)$, i.e., 0 is the first number, and

(s2) $(\forall n)(\forall m)((s(n) = s(m) \supset (n = m))$, i.e., if the successors are equal, then so are the originals.

It suffices to replace (s2) by the following axiom (s2)*

(s2)* $(\forall n)(\forall m)((n \neq s(n)\ \&\ m \neq s(m)) \supset (s(n)=s(m) \supset (n=m))$.

For the classical model, it is obviously the case that for all *n*, the successors are different from the original, but at the same time (s2)* is perfectly compatible with an additional axiom stating:

(s3) $(\exists n)(n=s(n))$

Note that this number can serve as the greatest number. It has therefore been particularly this approach that inspired the present author to develop proposals for a finitary arithmetic himself: Van Bendegem (1994, 1999, 2003). In parallel, from within paraconsistent logic, Graham Priest has also become interested in strict finitism. Not being convinced of the correctness of this position, his main motivation is related to the possibility of a finite model, for this opens ways to clarify certain metamathematical properties; see Priest (1994a, 1994b, 1997, 2000).

The preceeding may give the impression that among strict finitists, the focus is predominantly or even exclusively on arithmetic and numbers. Although in the present paper, attention will indeed be limited to numbers, this is certainly not the case in general: see, e.g., Van Bendegem (1995, 2000, 2002a). Let it moreover be noted that there are several beautifully argued papers against strict finitism, such as Dummett (1978) or Wright (1982), that a number of very good surveys and reviews are available, such as Welti (1987), Groenink (1993), Mawby (2005) or Styrman (2009), and also that interest has even been shown from semiotics: Rotman (1988, 1993). It is not the intention to impress with or draw conclusions from this list, although it should be pointed out that the widespread belief that this theme has only attracted a limited number of (marginal) philosophers and mathematicians is pertinently wrong.

2| For an excellent introduction to the life and work of this author, see Alberts (2000) and Alberts & Blauwendraat (2000).

3. The argument from continuous counting

This first argument against strict finitism is without doubt the most popular one. It runs as follows. It is totally senseless to claim that one can only count up to a certain number n and then suddenly stop. Put otherwise: if one can write down a number (sign) n, then one can also write down number (sign) $n+1$. This argument is analogous to the one saying that surely the universe must be infinite, for supposing that there is a border, what is to be found beyond that border? Admittedly, this is a very strong argument. Therefore my defense shall proceed in two steps. In a first step, I want to show that something must be wrong with this reasoning, without exactly pinpointing what. In a second step, I want to clarify what is the origin of the misunderstanding, thereby refuting the argument.

The first step of the counter-argument uses the Sorites paradox, already implicitly evoked in the previous section, when dealing with Yessenin-Volpin. Most philosophers are acquainted with the paradox as dealing with a heap, a beard, a bald man, etc. What it all boils down to is also familiar to most: we are dealing with *vague* predicates. Basically, this is the reasoning:

(P1) $A(1)$

(P2) $(\forall n)(A(n) \supset A(n+1))$

hence

(C) $(\forall n)A(n)$.

(P1) says that 1 has the property A, (P2) that if n has property A, then the next number in line also has it, and (C) says that all numbers are A. A classic example: $A(n) \overset{\text{def}}{=}$ "the number n is small."[3] If (P1) and (P2) are acceptable – which on first sight seems to be the case – then (C) follows, or in other words, all numbers are. Which is obviously wrong. Ways to deal with the problem are also well known: (a) reject the reasoning, (b) reject (P1), or (c) reject (P2).

Strategy (a) is mostly avoided, for the argument in question is none other than mathematical induction (see the previous section). Few want to rob mathematics of one of its favorite reasoning schemes. Strategy (b), in its turn, is of little avail: one chooses property A in order to be able to start with the number 1. Put otherwise, the fact that there exist properties for which, no matter what, $A(1)$ is the case, is enough. Strategy (c) remains. However, to reject (P2) has a peculiar consequence. In terms of models, it means that there is at least one model where (P2) is false, i.e., the negation that (P2) is the case, so:

$$\sim(\forall n)(A(n) \supset A(n+1)),$$

or, according to the rules of classical logic,

$$(\exists n)(A(n) \mathbin{\&} \sim A(n+1)).$$

This is a strange result, because we have hereby transformed the vague A to a *precise* one. And who would be prepared to claim that 1000 is small, whereas 1001 is not? Some solutions proposed, such as Sorensen (2006), are based on the idea of a "grey" area in between A and not-A. The details need not be discussed here, but usually solutions like these are considered hardly elegant. In the present context, however, I would like to present this as an advantage.

But let me first show the connection. Very simply: for property $A(n)$, take "I can write down the number (sign) n." Let us carefully spell out the various steps of the reasoning involved:

(P1) I can write down the number 1.

No comment.

(P2) If I can write down the number n, then I can also write down the number $n+1$.

This claim also seems perfectly acceptable. I am even inclined to add to this: for anyone. Whether one is a Platonist, logicist, formalist, intuitionist, constructivist, strict finitist, ultra-intuitionist, or just a *working mathematician*, one has to agree that once n is available, $n+1$ can be derived via a simple manipulation. Supposing no issue is made of mathematical induction, the conclusion then follows with absolute force, again: for anyone. But what does (C) state?

(C) For any number n it is the case that I can write down the number n.

In contrast, this claim seems unacceptable, and again, one is inclined to add: for all. Or is it? Perhaps room should be left to interpret (C) conditionally. When one is asked to write down any specific number n, then there are circumstances imaginable in which this is perfectly possible. If necessary, we devise a handy and economical notational system, such that, however finite our universe might be, one can write down the number in question anyhow. But to be clear, we are talking here of a physical activity: one is asked to write down the number n. That is, on a carrier substance. And using a particular notation. For anyone appealing to this "way out," I will confront them with this version:

(P1) I can write down the number n on paper in the decimal system.

(P2) If I can write down the number n on paper in the decimal system, then I can also write down the number $n+1$ on paper in the decimal system.

(C) For each number n in the decimal system, it is the case that I can write down the number n on paper in the decimal system.

Of course, your average classical mathematician can reply that this precisely shows that there is no room in mathematics for properties such as "to write down the number n on paper in the decimal system." In all honesty, I have nothing against this. But what is the strict finitist's reaction to the reasoning just proposed? Whenever (s)he refuses to accept the conclusion (C), we find ourselves in a typical Sorites situation. One can discuss whether the transition from (P1) and (P2) to (C) is justified, but to doubt (P2) is surely the most interesting strategy. In order to avoid entering a situation in which to maintain that for a certain number n, we can write down n, but on the other hand are unable to write down $n+1$, one can make use of the "grey" area mentioned above. For a specific initial fragment of the naturals, one can hold that $A(n)$ is the case, from a certain number onwards that $A(n)$ is definitely not the case, and for the region in between those, that it is, e.g., indetermined whether $A(n)$ is the case or not. A situation like this is easily understandable. One sees a mathematician at work on a piece of paper. Once the paper is full, (s)he takes another sheet, and then another and another. Upon the announcement that one has run out of paper, then the walls are still available, but that too must come to an end. One can try and write as small as possible, but again: there comes a moment when the available blank space is used up.

3| This version, known as Wang's paradox, has been the object of a brilliant analysis in Dummett (1978).

A possible objection to this approach is that vagueness remains a problematic concept, and that thus, considering the fact that strict finitism is in need of vagueness, this is rather an argument against instead of for it. The answer to this objection is straightforward: the impact will be minimal. If statements about numbers in the grey zone are indetermined, then nothing much can be said about them in mathematical terms anyway. In other words: these numbers are not useful. I shall return to this point in the section about the argument from poverty. There arises another danger, however. In order to be able to speak about the above reasoning, I had to allow myself to write down expressions such as "from a certain number n onwards, $A(n)$ is definitely not the case." But does this not sound terribly paradoxical? After all, we are talking here about the statement "I can write down the number n on paper in the decimal system." How then, can I know n? If I do know n, then I need a description of n and that description can be written, which is contradictory. But if I do not know n, then how can I ever be able to make the statement that I am unable to write down n under certain conditions? This argument can be applied to the greatest number itself in particular.

4. The argument from the greatest number itself

The essence of this objection is that the closer the strict finitist gets to a precise characterization of the greatest number, say L, the easier it becomes for adversaries to imagine numbers that are greater than L, or to come up with questions about L that make it seem quite arbitrary. Classical examples of the latter are: Is L prime? If not, what or how many are its prime factors? Can L be divided by a specific number, say 3? Examples of the former are: Does $L+1$ exist? Does L^2? Is L different from $L+1$? Or, if all expressions of the $L+n$ have a meaning, does one not obtain a potentially infinite number of names that can serve as number signs? And so on.

The key to dealing with objections like these, it seems to me, is given by wondering what are the presuppositions of one of the above questions, viz. "Is L different from $L+1$?" One of these presuppositions is that I can rightfully speak of L and $L+1$. But surely that is the problem to start with! To the extent that one can speak about a number $L+1$, I should be able to make a representation of it, which immediately implies that L is no longer the greatest number. As a consequence, to reply to the above question by "I am sorry, but I cannot answer this question" is perfectly defensible. This might sound awfully silly, but consider the analogy with the matter of discrete vs continuous time. One who holds time to be discrete is unavoidably confronted with the question of what is to be found in between two subsequent moments in time. Here also, avoiding this question is legitimate because its presupposition is precisely that something *can* be found in between these two moments, which is not the case (or so is the claim). Note that few will object to the analogous question in the continuous case, viz. what moment in time immediately follows a given one. Here also, the evident answer will be that this question does not make sense, because such a moment is simply inexistent (out of the nature of the continuum). In essence, this is the same answer, although I admit that this is a pretty weak argument. To condone one's own mistakes (if they are mistakes, of course) by pointing to similar mistakes by others is indeed an intellectually not very satisfactory option. Nevertheless, with respect to the greatest number L, one can very well answer: it is that number about which no question whatsoever can be answered. Strictly speaking, I am thus even prohibited from claiming that L actually is the greatest number, unless by using the phrase "the number about which no question whatsoever can be answered" as a definition.[4] This has a very interesting consequence, for I shall *not* consider L as the greatest number because I can write down $L-1$ as opposed to L. So the objection is disarmed. A legitimate question remains: what can one do with L, mathematically speaking? I shall return to this point below.

Is that it? Unfortunately not, because the critic can still remark that, although it might be wrong to demand an answer to the question "Is L different from $L+1$?," the fact remains that two different expression L and $L+1$ have been used to formulate the question. These two expressions are very well determined: one consists of a capital letter L, the other of that same capital letter L followed by a "+" sign, followed by a 1. So the question about the expressions can be answered, and the answer has to be: yes, we are confronted here with two different expressions, one consisting of one sign, the other of three. No strict finitist will claim that one and three are actually the same. But is this not destructive for the strict finitist? For the question above can be reiterated: "If all expressions of the form $L+n$ have a meaning, do we not in this way obtain a potentially infinite amount of names that can serve as number signs?" The answer, again, will have to be "yes." Which is problematic.

5. The argument from object and language

In order to be in a position to answer the above problem without bringing the potentially infinite in again through the back door, it is necessary to clarify the relation between the objects we talk about, number signs for instance, and our actual talk about them. The latter comprises the production of mathematical proofs about a certain domain. This relation is not at all straightforward. Typical objections directed at the strict finitist are the following.

A. Suppose you have a domain consisting of numbers from 1 to 100. When written down in the decimal system, we need 9 signs for the first 9 digits, then 90 times 2 signs, and 3 signs for the last number in line, which adds up to 192 signs in total.

4| It has been pointed out to me, among others by Leon Horsten and Diderik Batens, that there is something paradoxical about this statement (so perhaps Graham Priest would like this). The paradox is this: can you answer the question "Is L a number?" Given the definition, I would be forced to answer "yes," but then there is at least one question one can ask about L and answer. If, nevertheless, I insist that the answer must be negative, then I cannot even claim that L is a number, and that is an extremely weak position. This problem has also motivated the idea of a "dummy" largest number, presented further on.

Suppose I want to write down all these numerals, only having 100 numbers at my disposal, how will I ever be able to count them (which I will need to do to keep track)? Let us not even discuss the possible sums of these numbers for, even if I only wanted to write down all these sums of the form $n+m$, then I would have 100 times 100, altogether 10,000 possibilities. That means that I cannot even write down all these sums, let alone count them afterwards.

B. Suppose you have a domain consisting of numbers from 1 to 100. Suppose you have a language in which expressions dealing with these numbers can be written down. Suppose that such expressions can take the form of a system of axioms and inference rules, so that one can speak of proofs in that language. Is it then not completely arbitrary to claim that proofs are limited in length, namely not exceeding 100 signs? For should the opposite be the case, then it suffices to take the number of signs used as a measure for a greater number. Allowing the latter, the argument can be repeated until the potentially infinite is brought in.

For the time being, I will pass around the problem of finite proofs that, *qua text*, are perfectly acceptable for the strict finitist but nevertheless deal with infinite domains. That problem will be (briefly) touched upon in the next section.

A first and crucial observation, also of importance for what follows in the next section, is that to work with all numbers up to the greatest one is an impossible affair. As the greatest number itself is indeterminate, there is no sense in making statements about *all* numbers. A universal quantifier such as "For all n up to L" must be meaningless, for, to briefly recapitulate, from the moment we make a representation of all numbers up to L, L ceases to be the greatest number.

A supportive argument for this is given by considering this case at scale. Imagine a world with N objects, where people in this world are asked to count objects by associating each object with another as its label. It is straightforward in this world: people using this particular, elementary counting act will only be able to count half of the objects present. In other words, what is being shown here is that there need not be a contradiction between claiming that, one the one hand, the world is finite and, on the other, no man or woman is able to count all objects. *Finitism does not imply countability.* As a consequence, to create a workable situation, one has to look at a fragment of the whole. Or, we determinate an *arbitrary* upper limit by choosing a number L as the greatest number, this number being a *fictitious* one. Perhaps it is better to speak of a *dummy* as the greatest number. From this moment, L is a number about which statements can be made, e.g., asking what could be $L+1$, and so. One could put it this way: by determining L, one establishes a budget, within the margins of which not just all activities involving operations on numbers have to be accounted for, but also all reasoning about these operations, particularly theory building about the domain in question. The whole of this should be seen as limited in extent. Let me give a very simple example to illustrate this point.

Suppose you have at your disposal a budget B and you are able to identify a specific amount of numbers, say up to N. Let us, for the sake of argument, assume that to label each number, one sign is sufficient. Then, N elements of the budget have been used, and what remains, i.e., $B-N$, can be used to talk about these N numbers. Let us limit ourselves to sums of two numbers, $n+m$. If one wants to write down all possible sums, then N^2 signs are needed to this effect. If one wants to write down all equalities, i.e., expressions of the form $n+m=k+l$, then N^4 signs are needed, and so on. If one wants to write down even more complex expressions, e.g., of the form $((n+m)+k)+l$, then this will increasingly burden the budget. In general, the cost of theoretical reflection is considerably higher than the mere representation of the entities involved (as one might expect). On the other hand, one is perhaps not often invited or urged to this kind of undertaking. Who on earth would want to write down all expressions of the form $n+m$?[5]

5| Undertakings like these are not usually associated with mathematics, but rather with the *Guinness Book of World Records*. E.g., in 1998, Les Stewart from Australia, after 16 years, finished his task of typing all numbers from one to one million in words, resulting in a book of 19,890 pages.

Besides, in mathematics we have at our disposal a very economical way of dealing with signs, also applied above, i.e., the use of variables. It is crucial to see that with the expression $n+m$, *not* all concrete expressions or instantiations of that form have been written down. When looked at this way, the first thing that should be determined is the theoretical budget. In function of this, it can then be calculated what the "dummy" greatest, workable number is. Which implies that all arguments of the form "The greatest number should be equal to the amount of elementary particles in the universe" or "The greatest number is a function of Planck time and length" can be dismissed right away because they totally disregard the contribution of theoretical considerations. I have not listed this argument separately because the objections against it are so straightforward that it did not seem worth it. When only considering the question of whether all subsets of elementary particles exist, and if so, what numbers could possibly correspond with them, it becomes clear that something is not right.

From this perspective, it is peculiar and highly relevant to observe that most of mathematical literature rarely makes use of concrete number signs of a specific number system. To give but one example, in the first chapter of Baker (1984), a famous, concise introduction to number theory, the following *specific* numbers appear (in a mathematical context): 1, 2, 3, 4, 5, 7, 11, 12, 16, 32, 35, 40, 41, 55, 77, 95, 187, 432, 641, 13395, 44497.

There is an obvious reason behind the remarkable jump from 641 to two clearly larger numbers. Namely, the latter two are related to Mersenne primes – numbers of the form 2^n-1, with n prime – appearing in a statement about the then greatest *concrete* Mersenne number known, i.e., the 27th in line, where the smaller number refers to the number of digits and the larger number to the prime exponent. I'll get back to this. In the same chapter, however, it is also shown that there are infinitely many prime numbers. This is made possible via the subtle use of expressions such as "1, 2, …, n." Though it may seem that the complete row from 1 to n is represented here, this is actually not the case, that is, except for the mathematician who has no problem with the infinite, for whom such a problem is but a meaningless detail since *in principle* such an enumeration is always possible. For the strict finitist, the hidden subtleties are the witnesses of a mathematical practice suggesting that there are no limits to the budgets.

It may be remarked that a curious game is being played here. I am working with a finite budget, but what I have just done is to present a number of reflections about that budget. Does this not require an even bigger budget? In other words: does this kind of approach not imply the need, given any budget, for an even bigger budget? Which again brings in the potentially infinite. The answer is simple: these costs at the meta-level should also be taken into account when drafting the original budget.

There is, however, another consequence that needs to be addressed. Let me illustrate it with a concrete example. Suppose you have chosen a "dummy" greatest number L and you wish to examine a mathematical proposition, e.g., Goldbach's conjecture, or the claim that each even number larger than 2 is the sum of two primes. There is no reason whatsoever to assume that your budget will allow you to check individually whether all even numbers smaller than or equal to L can indeed be written as the sum of two primes. This can be generalized to other, similar cases. It is certainly not the case that, for the strict finitist, mathematical activity, primarily consisting of the search for and formulation of proofs, is a pointless and superfluous affair just because one can always reduce such a proof to a finite number of cases that can be checked one by one. Most certainly not! For the strict finitist, too, mathematical proof has a crucial role to play. On the contrary, for him or her the "economical" aspect will be extremely important. Think about the question of how to calculate the sum of the first n natural numbers, viz. $1+2+3+\ldots+n$ (yes, I am now myself using the mysterious three dots). The rudimentary way of doing this is to perform $n-1$ sums. Assuming the execution of one such sum has a price K, this costs $(n-1)\cdot K$. However, if we can use the formula $n\cdot(n+1)/2$, and assuming that multiplication and division cost the same as a single sum, then the total cost is but $3K$ (one sum of n and 1, one multiplication of n and $n+1$ and one division of $n\cdot(n+1)$ by 2, neglecting the cost of showing that both expressions introduced above are equivalent. Note that the total cost now has become independent of n itself![6]

This establishes a smooth transition to the last argument against strict finitism to be dealt with in this paper, viz. the argument that a strictly finite mathematics is not worthwhile because of the loss of virtually all interesting mathematical theorems.

6. The argument from poverty

It can hardly be the intention to show, in this section, how mathematics can be recast in strict finitist terms. This is a huge task, and moreover a technically very intricate one. What I would like to present instead is a rather simple example that might have the status of *exemplar*. Put otherwise, anyone convinced by the following argument should have no reason to believe that this approach might not work for the whole of mathematics. There is only a minimum of formal background needed. Take three propositions about variables x, y, z and a relation R specified for a domain of objects, for example the natural numbers. R can be interpreted as "follows" (in the sense of the successor function), so xRy means "after x comes y."

(a) $(\forall x) \sim xRx$
(b) $(\forall x)(\exists y)xRy$
(c) $(\forall x)(\forall y)(\forall z)((xRy \,\&\, yRz) \supset xRz)$.

Following the interpretation given, these propositions, in words, say:

(a) No object succeeds itself (or R is irreflexive).
(b) For each object, there is another object succeeding it.
(c) If y succeeds x and z succeeds y, then z succeeds x (or R is transitive).

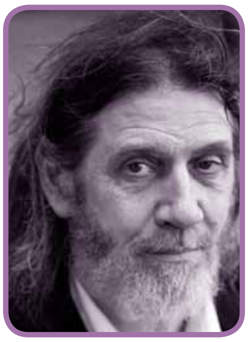
Informally speaking, it is easy to see that a finite set of objects can never meet conditions (a), (b), and (c) at the same time. Assume, for simplicity's sake, that there are but two objects, a and b. Consider statement (b). Both for a and for b an element has to exist that succeeds them. Take a. Can a succeed itself? No, that much is prohibited by (a). So aRb must be the case. But what about b? bRb is impossible, again because of (a), so all that remains is bRa. However, from aRb and (c) it follows that bRb, which is impossible. Summarizing, for each object there has to be a next object different from any previous object, in other words, you are in need of an infinite amount of elements. How can a strict finitist deal with this situation? That is the challenge. One thing stands out: whatever the budget available, (a), (b), and (c) are expressible, so the question of what are the corresponding models is perfectly legitimate. There is no escape there.

The way out is rather that, for a strict finitist, it is acceptable to ask the question as to how the universal quantifier be read, that is, what is the meaning of "∀" in expressions such as (a), (b), and (c)? Let us stick with the example of the two objects. A possible proposal could be to read the three statements in question as follows (the influence of Mycielski is clear here):

(a)* $(\forall x)(((x=a) \vee (x=b)) \supset \sim xRx)$
(b)* $(\forall x)(\exists y)((((x=a) \vee (x=b)) \,\&\, ((y=a) \vee (y=b))) \supset xRy)$
(c)* $(\forall x)(\forall y)(\forall z)((((x=a) \vee (x=b)) \,\&\, ((y=a) \vee (y=b)) \,\&\, ((z=a) \vee (z=b))) \supset ((xRy \,\&\, yRz) \supset xRz))$.

Let us not set out the general procedure, but focus on (a)*. What has been added? Part of the expression "$((x=a) \vee (x=b)) \supset$." Which just means that what immediately follows is limited to both elements a and b, or we relate the statement (a) to a domain with two *labelled* elements. In (b)* exactly the same happens, but as we have now two variables, x and y, both have to be related to a and b. Ditto for (c)* where this has to happen for x and y as well as z. One might think that this is a pretty trivial strategy, but nothing is further from the truth. In the expressions (a)*, (b)*, and (c)*, a and b are explicity mentioned, but that does not have to preclude that in the domain itself further elements are present, not mentioned in the propositions themselves. Now suppose the domain consists of three elements a, b, and c. The third element is thus not explicitly mentioned in the statements. For (a)* and

6| I fully realize that an important aside is in order. One can remark that this independence is presupposed, as the cost for a sum is considered as constant, regardless of the size of terms to be added. The point is, of course, that in the first calculation, n is a factor next to the cost K (which itself can be independent from n), and that in the second calculation, n has disappeared.



JEAN PAUL VAN BENDEGEM
is professor at the Vrije Universiteit Brussel and guest professor at Ghent University. He specializes in logic and philosophy of mathematics. He is also director of the Centre for Logic and Philosophy of Science and editor-in-chief of the logic journal *Logique et Analyse*

(c)* this poses no problems, only (b)* merits our attention. If we take a for x then (b)* is okay if we take b for y. But what if we take b for x? Then (b)* is turned into the following expression:

$$(\exists y)(((b=a) \vee (b=b)) \,\&\, (y=a) \vee (y=b)) \supset bRy)$$

As a small aside, as $b=b$ is certainly acceptable and therefore also $(b=a) \vee (b=b)$, this can be easily omitted, so the expression becomes:

$$(\exists y)(((y=a) \vee (y=b)) \supset bRy)$$

What shall we choose for y? a is no candidate, and neither is b, because of the previous argumentation. The only possibility left is to take the (or an) unnamed object c. This makes the statement trivially true, for if y stands for c, then neither $y=a$, nor $y=b$ is the case, so the implication is always true.

There is an alternative way to understand this procedure (and by explaining this, I clarify the connection with the work of Graham Priest). Start out with an infinite model with elements a, b, c, d, e, … and divide this model into a finite number of parts. An initial part is retained, so we allow a to be a, and b to be b, but all that follows is squeezed into one element. Doing this, one obtains a finite number of parts that let themselves be – informally – described as a, b, and "the rest," for which we write c. By the fact that a and b are still sharply delimited, statements can be made about these elements that perfectly coincide with classical ones. And from the moment one ends up in "the rest" of the elements, it does not matter what is being said, which has the very pleasant consequence that it cannot possibly contradict what is expressible in precise terms.

There is, however, more. Let us return to the original formulations (a), (b), and (c). It is easily seen that claim (d) can be shown (provided classical predicate logic can be used):

(d) $(\forall x)(\forall y)(xRy \supset {\sim}yRx)$

The proof could look as follows: suppose that xRy is the case, and so is yRx, then because of transitivity from (c) it is the case that xRx, which is, however, contradicted by (a). *QED*

Let us now rewrite this little proof, replacing the original claims by their starred variants. One shall see that mostly it only requires the addition of "$(x=a) \vee (x=b) \supset$" or "$(y=a) \vee (y=b) \supset$" in the appropriate place, nothing more. When dropping these extra passages, the classical proof appears. A nice graphical way to express what is going on here is to work with a main and margin text. The margin then contains the restrictions, while the main text only holds the claims without the restrictions. Forgetting the margin text, one only considers the classical proof in the main text, *as if* the proofs were dealing with infinite domains. I would describe this as the "book keeper's model." Although it seems that what is put in the margin is not very important from a classical point of view, once the available means are restricted, it has to be taken into account. If one forgets the margin altogether, then one runs a considerable risk of assuming that one's means are indeed infinite. This, however, should be considered a fiction in mathematics, too. This opens up a quite interesting perspective: are there relations between strict finitism and a new development in the philosophy of mathematics, where mathematical entities are considered to be fictions; see Leng (2010) for a fervent defense of this position.

For strict finitism, the immediately above is of central importance. Here is an alternative formulation: apparently there are mathematical proofs that allow the application of the book keeper's model. This means that, in the main text, there is no mention of limits and borders, or in other words, the main text can be considered as *budget independent*. This is exactly what makes mathematical activity so interesting for the strict finitist. To him or her, too, a proof independent of any budget will be much richer than one that is not and where, e.g., the "dummy" greatest number comes into play (just because it is a dummy, it is not worth much). So, to be clear, the strict finitist will, as actively as the classical mathematician, search for the most interesting and richest type of proof. So much for the alleged poverty!

To give an additional, mathematically interesting example, let us look at a proof showing that the number of primes is infinite. In essence, this proof is very simple. Assume that the number of primes is finite, i.e., $p_1, p_2, \ldots, p_n$. Define the number $P = p_1 \cdot p_2 \cdot \ldots p_n + 1$ and show that this number has prime factors, including P itself, that must differ from all p_i. The question, from a strict finitist point of view, is whether P can always be constructed. Since any budget is finite, there will indeed come a moment when P can no longer be calculated, and in this sense there is a greatest prime number. But take a closer look at the proof. In its margin, the budgetary constraints will appear, while the main text is nothing but the classical proof just introduced. This also means that, should a larger budget be available, the main text would not undergo any changes because it is independent of a budget.[7] Whoever is so sloppy as to forget about the margin text can thus be deceived into believing that indeed there are an infinite amount of primes. At the same time, it is perfectly sound to hold that the classical proofs can be retained, even within the frame of strict finitism. I hope the above shows that strict finitism, in practice, does not need to be a trivial affair.

Of course, I do realize that even a reader sympathetic towards this approach will still have many questions that cannot possibly be answered in this paper. For example: (a) How does one deal with the real and transcendental numbers, such as π in particular? π can be defined, e.g., as the sum of particular infinite series, but how does one capture that in the book keeper's model? (b) What about the possibility of analysis? What are differentials and integrals like? How does one define continuity? (c) Is transfinite number theory possible at all, or should this part of mathematics disappear? To be fair, some of these questions have already received a lot of study, and partial answers have been given (see the first section). Some of these have, moreover, revealed unexpected connections. In Van Bendegem (2002b), for example, I initiated a strict finitist approach to infinitesimals, thereby putting at stake the property of compactness, which has opened up a very interesting discussion. And this, it should be said, is already an achievement in itself, even if the position, in the end, might appear to be untenable.

7. Conclusion

In this paper, I have concentrated on a defense of strict finitism. In the second section, I have tried to show that a considerable number of researchers have, either directly or indirectly, been involved with this subject, and in the sections following it, that possible answers to the most famous and common counter arguments are available. Summarizing these, I would use the word "naive." Which means that, as a strict finitist, one often has to conclude that so-called critical remarks have not been correctly formulated, cutting short any discussion about the substance of the matter. This could also be observed in the course of the present paper. Nevertheless, especially towards the end, I hope to have been able to show in a positive way that the undertaking has indeed a chance of success. The philosophical benefits of a plausible strict finitism would, moreover, be considerable, I think. To end, I give a few examples. (a) The discussion about the mysterious applicability of mathematics can be handled in a radically different way. In a slogan: finite world, finite mathematics, end of problem. (b) As infinity has been the main reason for its collapse, the project of nominalizing physics by Field (1980) could be reassessed. (c) To the extent that the book keeper's model can be extended, there are also opportunities to enhance our understanding of mathematical practice. On the one hand, one accepts the theorem that there are infinitely many primes, while at the same time there is an impressive quest to find the (next) greatest one. Both seem to be in conflict but are not. From the viewpoint of strict finitism, both undertakings make perfect sense.

7| Note that one has to be very careful not to introduce the potential infinite through ever larger budgets. Then again, the advantage of budgets is that one is more likely to accept that there is an ultimate, finite budget rather than the fact that there is only a finite number of numbers.

Acknowledgement

This paper, in a slightly modified form, was first published in the peer-reviewed Dutch journal *Algemeen Nederlands Tijdschrift voor Wijsbegeerte* (Van Bendegem 2010). I wish to thank the publishers for their permission to present an English translation here, which has been modified in several places to take into account the remarks of the commentators, which were also included in the journal as was a final response by myself. In addition, I wish most profoundly to thank Bart Van Kerkhove for his unconditional willingness to do the actual translation. My sole task was to make some cosmetics here and there.

References

Alberts G. (2000) Twee geesten van de wiskunde: Biografie van David van Dantzig. CWI, Amsterdam.

Alberts G. & Blauwendraat H. (eds.) (2000) Uitbeelden in de wiskunde. CWI, Amsterdam.

Baker A. (1984) A concise introduction to the theory of numbers. Cambridge University Press, Cambridge.

Dummett M. (1978) Wang's paradox. In: Dummett M., Truth and other enigmas. Duckworth, London: 248–268.

Epstein R. L. & Carnielli W. A. (2000) Computability. Computable functions, logic, and the foundations of mathematics with computability: A timeline. Second Edition. Wadsworth, London.

Gödel K. (1971) On formally undecidable propositions of Principia mathematica and related systems I. In: van Heijenoort J. (ed.) From Frege to Gödel. Harvard University Press, Cambridge, MA: 592–617. Originally published in German in 1931.

Field H. (1980) Science without numbers: A defence of nominalism. Princeton University Press, Princeton.

Groenink A. (1993) Uism and short sighted models. Ph.D. Thesis, University of Utrecht.

Isles D. (1992) What evidence is there that 2^65536 is a natural number? Notre Dame Journal of Formal Logic 33: 465–480.

Isles D. (1994) A finite analog to the Löwenheim–Skolem theorem. Studia Logica 53: 503–532.

Lavine S. (1994) Understanding the infinite. Harvard University Press, Cambridge, MA.

Leng M. (2010) Mathematics and reality. Oxford University Press, Oxford.

Marion M. (1998) Wittgenstein, finitism, and the foundations of mathematics. Clarendon Press, Oxford.

Mawby J. (2005) Strict finitism as a foundation of mathematics. Ph.D. Thesis, Glasgow University.

Mycielski J. (1981) Analysis without actual infinity. Journal of Symbolic Logic 4: 625–633.

Nelson E. (1986) Predicative arithmetic. Princeton University Press, Princeton.

Parikh R. (1971) Existence and feasibility in arithmetic. Journal of Symbolic Logic 36: 494–508.

Presburger M. (1929) Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition

als einzige Operation hervortritt. Comptes Rendus du I congrés de Mathématiciens des Pays Slaves: 92–101.
Priest G. (1994a) Is arithmetic consistent? Mind 103: 337–349.
Priest G. (1994b) What could the least inconsistent number be? Logique et Analyse 37: 3–12.
Priest G. (1997) Inconsistent models of arithmetic I: Finite models. Journal of Philosophical Logic 26: 223–235.
Priest G. (2000) Inconsistent models of arithmetic II: The general case. Journal of Symbolic Logic 65: 1519–1529.
Rodych V. (2000) Wittgenstein's anti-modal finitism. Logique et Analyse 43: 301–333.
Rotman B. (1988) Toward a semiotics of mathematics. Semiotica 72: 1–35.
Rotman B. (1993) Ad infinitum. The ghost in Turing's machine. Taking god out of mathematics and putting the body back in. Stanford University Press, Stanford.
Sorensen R. (2006) Vagueness. In: Zalta E. N. (ed.) The Stanford encyclopedia of philosophy. Fall 2008 Edition. Available at http://plato.stanford.edu/archives/fall2008/entries/vagueness/
Styrman A. (2009) Finitist critique on transfinity: An investigation of infinity, collection theory and continuum. Master's Thesis, University of Helsinki.
Van Bendegem J. P. (1994) Strict finitism as a viable alternative in the foundations of mathematics. Logique et Analyse 37: 23–40.
Van Bendegem J. P. (1995) In defence of discrete space and time. Logique et Analyse 38: 127–150.
Van Bendegem J. P. (1999) Why the largest number imaginable is still a finite number. Logique et Analyse 42: 107–126.
Van Bendegem J. P. (2000) How to tell the continuous from the discrete. In: Beets F. & Gillet E. (eds.) Logique en perspective. Mélanges offerts à Paul Gochet. Ousia, Brussels: 501–511.
Van Bendegem J. P. (2002a) Finitism in geometry. In: Zalta E. N. (ed.) The Stanford encyclopedia of philosophy. Spring 2002 Edition. Available at http://plato.stanford.edu/entries/geometry-finitism/
Van Bendegem J. P. (2002b) Inconsistencies in the history of mathematics: The case of infinitesimals. In: Meheus J. (ed.) Inconsistency in science. Kluwer, Dordrecht: 43–57.
Van Bendegem J. P. (2003a) Classical arithmetic is quite unnatural. Logic and Logical Philosophy 11: 231–249.
Van Bendegem J. P. (2010) Een verdediging van het strikt finitisme. Algemeen Nederlands Tijdschrift voor Wijsbegeerte 102(3): 164–183.
Van Dalen D. (1978) Filosofische grondslagen van de wiskunde. Van Gorcum, Assen.
Van Dantzig D. (1956) Is $10^{10^{10}}$ a finite number? Dialectica 9: 273–278.
Yessenin-Volpin A. (1970) The ultra-intuitionistic criticism and the antitraditional program for foundations of mathematics. In: Kino A., Myhill J. & Vesley R. (eds.) Intuitionism and proof theory. North-Holland, Amsterdam: 3–45.
Welti E. (1987) Die Philosophie des strikten Finitismus. Entwicklungstheoretische und mathematische Untersuchungen über Unendlichkeitsbegriffe in Ideengeschichte und heutiger Mathematik. Peter Lang, Bern.
Wittgenstein L. (1984) Bemerkungen über die Grundlagen der Mathematik. Edited by Anscombe G. E. M., Rhees R. & von Wright G. H. Suhrkamp, Frankfurt. Originally published in 1956. English translation: Wittgenstein L. (1956) Remarks on the foundations of mathematics. Edited by von Wright G. H., Rhees R. & Anscombe G. E.M., translated by Anscombe G. E.M. Basil Blackwell, Oxford.
Wright C. (1980) Wittgenstein on the foundations of mathematics. Harvard University Press, Cambridge MA.
Wright C. (1982) Strict finitism. Synthese 51: 203–282.

Received: 2 December 2011
Accepted: 26 February 2012