Numbers, Infinities and Infinitesimals

N. J. Wildberger
School of Mathematics UNSW
Sydney 2052 Australia
webpages: http://web.maths.unsw.edu.au/~norman/

October 17, 2006

Problems with infinities and infinitesimals

For the last hundred years or so, mathematicians have finally 'understood the infinite'—or *so they think.* Despite several thousand years of previous uncertainty, and very stiff opposition from notable mathematicians (Kronecker, Poincaré, Weyl, Brouwer and numerous others), the theory of infinite point sets and hierarchies of cardinal and ordinal numbers introduced by Cantor is now the established orthodoxy, and dominates modern logic, topology and analysis. Most mathematicians believe that there are infinite sets, but can't demonstrate this claim, and have difficulty resolving the paradoxes that beset the subject a hundred years ago, except to assert that there are dubious 'axioms' that supposedly extract us from the quagmire. If you have read my recent diatribe *Set Theory: Should You Believe?* you will know that I no longer share this religion, and I tried to convince you that you shouldn't either.

Several readers of that paper have asked me: so what should we believe? What are the alternatives to the modern set theoretical paradigm? To begin the slow redevelopment of mathematics as a subject that actually makes complete sense, I now propose tell you what a *number is*, what an *infinity is*, and what an *infinitesimal is*. You'll see how to understand these notions without unnecessary philosophy. I am *not* going to define 'infinite sets'—for the simple reason that they don't really exist. But I will explain how some of the modern theory of 'ordinals' may be recast as a *purely finite theory* which makes complete and precise sense, requires no assumptions or 'axioms'—and reveals interesting natural connections with computer science. Then you'll see how the inverses of these infinities are naturally infinitesimals, giving a concrete, specifiable approach to nonstandard analysis.

After an initial (skeptical) look at the usual 'set-theoretical' approach to 'ordinals', we will start by defining *natural numbers* in a simplified way, without use of set theory. The fact is that *sets* are just not the all-important data structures that many believe—in computer science, for example, the notions of *list* and *multiset* are arguably more important. James Franklin told me about *heaps*, which are multisets without a 'containing box', and which philosophers studied half a century ago under the subject of *mereology,* see for example [LG].

So our approach is the following. *Zero* is defined as nothing, or more precisely as the empty set. *One* is defined as the set consisting of the single element zero. So far this agrees with the usual framework. But now:

A natural number is a heap of ones.

Once we know what a natural number is, the correct approach to the sequence $1, 2, 3, \cdots$ of natural numbers is as an inductive progression that trails off beyond our gaze—*not* as a completed infinite set. This is the ancient and more modest view. As we go further down this sequence it becomes increasingly complicated, unpredictable and inaccessible, and we can never get to the end. But the beautiful fact is that *we don't need to*, and we *don't need to pretend to be able to.*

Infinity is already right in front of us, and it turns out that we have been working with the concept from our high school days. So what is infinity, really? It is this:

An infinity is a growth rate.

We define inductively certain increasing functions from natural numbers to natural numbers called *natural functions* or just *infinities*, and we recover most of the useful aspects of ordinal arithmetic—*without 'infinite sets'.* With this approach, the contemplation of infinity reverts from a philosophical pastime to a powerful tool for analysis.

Because once you have understood infinity in the right way, *understanding infinitesimals becomes straightforward.* Infinitesimals are much more important historically than infinities, playing a key role in calculus since the early days of that subject, admittedly with rather shaky justifications. After a period of being ignored as imprecise by pure mathematicians, while nevertheless being used extensively by applied mathematicians, physicists and engineers, infinitesimals made a modest recovery with the advent of the *nonstandard analysis* of Robinson [Ro] and their use in calculus as implemented in H. J. Keisler's text [Ke]. However this has not caught on, basically because the modern nonstandard definition of an infinitesimal is based on the 'hyperreals', which are conjured up from thin air by the 'axiom of choice' and 'ultrafilters'—and so are incomprehensible to students beginning the subject.

In this paper we show how to harness our understanding of infinities to give a concrete approach to infinitesimals, thus contributing to a modern constructive rebirth of nonstandard analysis. The basic idea is this:

An infinitesimal is a decay rate.

By using explicit decay sequences—the basic one is simply $1/n$—to form augmented numbers such as $3+2\frac{1}{n}$, we may study the local behavior of functions, which is the start of analysis. Surprisingly, a good amount of this can be done even over the rational numbers.

The ideas in this paper are not really new. The notion of a number as a collection of basic objects (be they pebbles, twigs, or heaps) surely goes back tens or even hundreds of thousands of years, to the very first counting done by our tribal forefathers. The idea of infinity as a sequence was developed in the 1950's by C. Schmieden with the introduction of Ω as the sequence of natural numbers, playing essentially the same role as what we call n. Infinitesimals as sequences going to zero were introduced and studied by Chwistek [Ch], Schmieden and Laugwizt [SL], Laugwitz [L1], [L2], and more recently by J. M. Henle [He1], [He2], and used in a recent calculus text by Cohen and Henle [CH].

The current paper differs from most of these approaches by rejecting beforehand modern set theory, especially 'infinite sets', 'axioms' and 'equivalence classes of sequences'. It insists on building everything up by computational notions that are completely specifiable in a finite amount of time and that computers can utilize. This is done by elevating particular explicit computational sequences, both for infinities and infinitesimals, to the foreground.

This paper should be viewed as an introductory account of this material, as many statements are given with no proof or only a brief indication. I hope that some logicians, analysts and/or computer scientists will step up to the plate and provide a more formal structure for these ideas, and extend them to a coherent framework for *real analysis*. Let's think creatively and concretely about foundational issues, and develop approaches to the subject that our computers can implement, and that have real applications.

Ordinals of old

Ordinal numbers, we are told, measure the size of well-ordered sets. The usual theory of ordinals goes something like this. [Remember that I suffer from never having seen an infinite set, and don't believe there are any, so for most of this section I will just be repeating the usual mantras.]

First there is the ordinal number 0, which can be defined as the empty set $\{\ \}$. The ordinal number one is defined to be the set $1 \equiv \{0\} = \{\{\ \}\}$. Following von Neumann, one defines $2 \equiv \{0,1\}$, $3 \equiv \{0,1,2\}$ and so on, where each ordinal is a set, consisting of all the ordinals up to, but not including itself. So in particular the number 4 is the set

$$4 \equiv \{\{\ \},\{\{\ \}\},\{\{\ \},\{\{\ \}\}\},\{\{\ \},\{\{\ \}\},\{\{\ \},\{\{\ \}\}\}\}\}.$$

Even die-hard fans must admit that this is not entirely compelling. If we are going to start out with such complexity built in to our basic notions, how will we build a good coherent theory? Shortly I will offer an alternative view of the natural numbers that you may find more sensible and closer to your intuition.

In any case, Cantor believed that you could group 'all' the natural numbers into an 'infinite set' as follows: $\{0,1,2,3,\cdots\}$. For those with a metaphysical leaning, this is a proper 'construction', and defines a new mathematical object— an ordinal called ω.

The successor of any ordinal is obtained by taking the union with the ordinal itself, so that for example $\omega+1 = \{0,1,2,3,\cdots,\omega\} = \{0,1,2,3,\cdots,\{0,1,2,3,\cdots\}\}$. The limit of a sequence of ordinals is obtained by taking the union of all elements of the sequence. Thus for example the limit of the sequence

$$\omega, \omega+1, \omega+2, \cdots$$

is the ordinal

$$\left\{ \begin{array}{c} 0,1,2,3,\cdots,\{0,1,2,3,\cdots\},\{0,1,2,3,\cdots,\{0,1,2,3,\cdots\}\}, \\ \{0,1,2,3,\cdots,\{0,1,2,3,\cdots\},\{0,1,2,3,\cdots,\{0,1,2,3,\cdots\}\}\},\cdots \end{array} \right\}$$

which is then known as 2ω. The limit of $\omega, 2\omega, 3\omega, \cdots$ is ω^2, and the limit of $\omega, \omega^2, \omega^3, \cdots$ is ω^ω. Continuing on, one can then create ordinals such as

$$2\omega^{4\omega^{\omega^2+3\omega}+\omega^3} + \omega^{3\omega+1} + 7.$$

There is an arithmetic with such ordinals, and one can do induction with them. Up to this point they can pleasantly be *encoded by rooted trees*. Briefly, the reason is that we can recursively associate to an ordinal

$$\omega^a + \omega^b + \cdots \omega^d$$

the rooted tree with branches corresponding to $a, b, \cdots, d$. This association is a clue that the definition of ordinals might be an optical illusion, in other words that there might be a completely finite approach to the subject which yields essentially the same arithmetic, but dispenses with the need to conjure up 'infinite sets'. We'll shortly see just how to develop such an alternative theory.

The ordinals do not stop here—we are more or less at the third bus stop on an overland trip to Tibet. Ordinals go on and on in an increasingly bewildering fashion, and charting them becomes more a case of identifying some recognizable landmarks than taking a complete survey. The further and further you go down this path, the less relevance there is to ordinary mathematics and real life.

Natural numbers

When we begin to discuss numbers, we should not assume too much sophistication. Hundreds of thousands of years ago, one of our African ancestors would have discovered the idea of representing some animals at a waterhole, or some children under a tree, by using the fingers of his/her hand, or perhaps some pebbles or twigs. Thus was established the idea of counting with numbers.

With the advent of writing, the first numbers would have been represented simply by little sticks or circles to represent fingers, twigs or pebbles, such as

$$|||||||$$

to represent what we now call 'seven'.

The fundamental concepts upon which the notion of number is built are the dual ideas of *nothing* and *something*, and the idea of *adding* something to what we already have. Let's represent *nothing* by the symbol 0 and call it **zero**. Let's represent *something—the simplest possible something*—by the symbol | and call it **one**. You may consider *zero* to be the *empty set*, and *one* to be the *set consisting of the element zero*.

Now that we have one, we can consider one together with one, and then one together with one together with one. And so on. This gives the basic sequence

$$|, ||, |||, ||||, |||||, ||||||, |||||||, \cdots$$

where each term is defined to be the **successor** of the term before it, obtained by adding a single |. By definition a **natural number** is an element in this sequence.

Note that a number such as |||| is not exactly a set, or even a multiset. Why not? It's not a set because it has repetitions, and it's not a multiset because there are no brackets. It is rather a collection without any bounding box—a *heap*.

Now that we have the natural numbers, let's give them the usual names one, two, three, four, five, six, seven and so on, and use the usual Hindu-Arabic decimal notation to represent them. So the sequence will now be denoted

$$1, 2, 3, 4, 5, 6, 7, \cdots .$$

The familiarity of this notation masks a subtlety—when we write down 42, we are really only referring to a particular *representation* of the natural number

$$||||||||||||||||||||||||||||||||||||||||||.$$

A typical natural number will be denoted N, and its successor by N'. The natural numbers will definitely *not* be considered as a completed 'infinite set', but rather as a sequence that unfolds beyond our view. Perhaps the sequence keeps on going 'forever and ever', whatever that might mean. Perhaps the terms gets so overwhelmingly large that the world can no longer hold them past a certain point, or time itself runs out. Or perhaps there is a three-headed dragon somewhere down the sequence that devours anyone who gets there. Which of such alternatives might be the 'truth' seems a metaphysical issue, not a mathematical one, so let's leave the question to philosophers, cosmologists and poets.

In the mathematical practice of our actual experience, there is always a successor N' to any natural number N. The sequence of natural numbers has a natural order $<$ which makes the sequence **well ordered**—for any two distinct natural numbers N and M, either $N < M$ or $M < N$, and for any natural number N any sequence $N > N_1 > N_2 > \cdots$ must eventually stop.

Counting is easy with natural numbers viewed as heaps. If you have a multiset, or set, or list A that you want to count, simply dump all the contents out, and replace each of them by the simplest possible object, namely one. The result is a heap of ones which we call the **size** of A.

Natural functions, operators and sequences

A **function** is *a rule that takes an input of a specified kind to an output of a specified kind*. If the function f takes an object $\heartsuit$ to an object $\Diamond$ then we write $f : \heartsuit \to \Diamond$ or $f(\heartsuit) = \Diamond$. So to specify a function we must say what type of object it inputs, what type of object it outputs, and what the rule is that determines an output from an input.

We are going to build up inductively functions f which input and output natural numbers, and which are **increasing** in the sense that $f(N) \leq f(M)$ whenever $N \leq M$. The functions we will construct will be called **natural functions**. Not every increasing function is a natural function!

The first natural functions are the **constant functions**. For each natural number M, we can define a function $c_M : N \to M$. Thus for example

$$c_7 : 5 \to 7 \qquad \text{or} \qquad c_7(5) = 7.$$

The sequence $c_1, c_2, c_3, \cdots$ of constant functions corresponds to the sequence $1, 2, 3, \cdots$ of natural numbers, and occasionally it is useful to identify them.

The next natural function is the **identity function** $n : N \to N$. Note that N denotes an arbitrary natural number, while n is a specific function on natural numbers. We write for example

$$n : 5 \to 5 \qquad \text{or} \qquad n(5) = 5.$$

An **infinity** is by definition a *non-constant natural function.* Thus n is an infinity—in fact it is the first and most important infinity.

For any natural function f, define the **successor** f' by the rule

$$f' : N \to f(N)'.$$

We could also write $f'(N) = f(N) + 1$. Thus $c_2 = c_1'$, $c_3 = c_2'$ and so on. The successor of n is the function $n' : N \to N'$. We write this function in the more familiar way as

$$n + 1 : N \to N + 1.$$

Starting with n, by taking successors repeatedly we get the following sequence of natural functions:

$$n + 1, n + 2, n + 3, \cdots.$$

This sequence is itself a strictly increasing sequence, in the following sense. For two natural functions f and g, we say that $f < g$ precisely when there exists a natural number M such that if $N > M$ then $f(N) < g(N)$. All the sequences $f_1, f_2, f_3, \cdots$ of natural functions we will consider will be **strictly increasing**, meaning that $f_1 < f_2 < f_3 < \cdots$.

For an increasing sequence $f_1, f_2, f_3, \cdots$ of natural functions, we define the **limit function** f by the rule

$$f : N \to f_N(N)$$

and declare f to also be a natural function provided it is increasing. In all the examples we consider, this will be the case. Note that the identity function n can be viewed as the limit of the sequence $c_1, c_2, c_3, \cdots$ of constant functions.

Polynomial natural functions

So the limit of the sequence

$$n + 1, n + 2, n + 3, \cdots$$

is the function $2n : N \to N + N = 2N$. Then we can create the sequence of successors

$$2n + 1, 2n + 2, 2n + 3, \cdots$$

with limit $3n$. This gives a new type of sequence

$$n, 2n, 3n, \cdots$$

with limit n^2. After that we get

$$n^2 + 1, n^2 + 2, n^2 + 3, \cdots$$

with limit $n^2 + n$, then similarly $n^2 + 2n$, $n^2 + 3n$ and then $2n^2$, and eventually $3n^2$. The sequence

$$n^2, 2n^2, 3n^2, \cdots$$

has limit n^3, and continuing allows us to define any **polynomial natural function**, such as

$$4n^5 + 3n^2 + 7n + 5.$$

Exponential and pyramidal natural functions

The sequence of polynomial functions

$$n, n^2, n^3, \cdots$$

has limit n^n. This is a new type of natural function, which we will call an **exponential natural function**. Taking successors gives

$$n^n + 1, n^n + 2, n^n + 3, \cdots$$

with limit $n^n + n$. Then we get $n^n + 2n$, then $n^n + n^2$, then $n^n + n^3$ and so the sequence

$$n^n + n, n^n + n^2, n^n + n^3, \cdots$$

with limit $2n^n$.

One plays the same kind of game as with ordinals, except that at all times you are working with concrete, finitely specifiable functions that in principle you could program on your computer, and instead of taking *infinite unions* of '*infinite sets*' you are taking *limits of sequences* which are themselves finitely specifiable.

The sequence

$$n^n, 2n^n, 3n^n, \cdots$$

has limit $n \times n^n = n^{n+1}$. Then we get n^n,$n^{n+1}, n^{n+2}, \cdots$ with limit n^{2n}. The sequence $n, n^n, n^{2n}, \cdots$ has limit n^{n^2}, and then the sequence $n, n^n, n^{n^2}, n^{n^3}, \cdots$ has limit n^{n^n}.

You can then create anything of the form n^p where p is a polynomial function, and then also anything of the form n^{n^p}, and then anything of the form $n^{n^{n^p}}$ and so on. So if you can create q, then you can create n^q. By continuing on with this type of limit, we can get things that look like

$$2n^{3n^{n^{2n^n+n^2}+5n^3}+n^{4n+1}+n^2} + 3n^{n^{n^5+6n}+7n^2} + n^{13} + 7.$$

Note that we have not given a way of getting, for example, $(n+1)^n$. This is not a natural function in this framework.

Now we consider the sequence $n, n^n, n^{n^n}, \cdots$ with limit which we denote $n^{\triangle n}$ and call "n to the pyramid n". So for example

$$\begin{aligned} n^{\triangle n} &: 1 \rightarrow 1 \\ n^{\triangle n} &: 2 \rightarrow 2^2 = 4 \\ n^{\triangle n} &: 3 \rightarrow 3^{\triangle 3} = 3^{3^3} = 7625\,597\,484\,987 \end{aligned}$$

So starting with $n^{\triangle n}$, you can probably figure out how to get the sequence

$$0, n^{\triangle n}, 2n^{\triangle n}, \cdots$$

with limit $n \times n^{\triangle n} = nn^{\triangle n}$, then $n^2 n^{\triangle n}$, then $n^n n^{\triangle n}$. *And so on* (famous not last words)!

Sequences

Although we can keep going and generate more and more functions, let us pause and clarify *which type of function sequences* we are going to allow when forming limits. The sequence $1, 2, 3, \cdots$ of natural numbers and the sequence $c_1, c_2, c_3, \cdots$ of constant functions are considered **natural sequences**. Additional **natural sequences** will be particular types of function from natural numbers to natural functions which we are going to define inductively.

The function n can be viewed as the limit of the natural sequence $c_1, c_2, c_3, \cdots$ of constant functions. We define n to be the first **1-limit function**. For any limit function f, the **1-sequence** of f is the sequence

$$f', f'', f''', \cdots.$$

We say also that this 1-sequence **follows** f. More precisely such a sequence is a function, call it s, inputting natural numbers and outputting natural functions, defined inductively by the rules

$$\begin{aligned} s(1) &\equiv f' \\ s(N') &\equiv s(N)' \end{aligned}$$

for any natural number N, so that $s(N)(M) = f(M) + N$. Any 1-sequence is by definition a natural sequence. The limit of a 1-sequence will be called a **1-limit**. Starting with n, our first limit function, the following 1-sequence is

$$n+1, n+2, n+3, \cdots$$

and this has limit $2n$. So $2n$ is also a 1-limit. It has a following 1-sequence

$$2n+1, 2n+2, 2n+3, \cdots$$

which has the limit $3n$, also a 1-limit.

Now we are going to define a new type of natural sequence, called a **2-sequence**. The sequence of 1-limits we have constructed starts with n, and is

$$n, 2n, 3n, \cdots,$$

and it will be the first 2-sequence. Its limit is n^2, and this is called a **2-limit**, since it is the limit of a 2-sequence. Now since n^2 is a limit, it is followed by a 1-sequence

$$n^2+1, n^2+2, n^2+3, \cdots$$

which has the 1-limit n^2+n. This 1-limit has following 1-sequence

$$n^2+n+1, n^2+n+2, n^2+n+3, \cdots$$

with 1-limit n^2+2n. We now have a new 2-sequence which follows the 2-limit n^2, namely

$$n^2+n, n^2+2n, n^2+3n, \cdots$$

with 2-limit $2n^2$. Then we get another 1-sequence with 1-limit $2n^2+n$, another 1-sequence with 1-limit $2n^2+2n$, and so the 2-sequence

$$2n^2+n, 2n^2+2n, 2n^2+3n, \cdots$$

with 2-limit $3n^2$. The sequence of 2-limits we have constructed starts with n^2, and is

$$n^2, 2n^2, 3n^2, \cdots,$$

and it will be the first **3-sequence**. Its limit is n^3, and this is called a **3-limit**, since it is the limit of a 3-sequence.

Rules for limits

Now we may continue to define for any natural number N the notion of N-sequence and the notion of N-limit. Here is how to do this more formally.

- the limit n of the constant function sequence $c_1, c_2, c_3, \cdots$ is a **1-limit**
- for any limit f, the sequence $f', f'', f''', \cdots$ is a **1-sequence**

- the limit of an N-sequence is an N**-limit**
- a sequence

$$f_1, f_2, f_3, \cdots$$

  of consecutive N-limits is an N'**-sequence** provided that f_1 is either the first N-limit or the first N-limit following an M-limit for some $M > N$.

Let's clarify this last point. If f is an M-limit then it is a limit, so it has a following 1-sequence, then a following 2-sequence starting with the limit of the previous 1-sequence, then a following 3-sequence starting with the limit of the previous 2-sequence, and so on till we get to a following N-sequence. It is to be noted that there are no functions between f and the limit of this following N-sequence which are K-limits for any $K > N$.

This inductive setup allows us to construct all polynomial natural functions, such as

$$3n^4 + 2n^2 + n + 7.$$

The next step is to define a new type of natural sequence, an n**-sequence**. The first n-sequence is

$$n, n^2, n^3, \cdots$$

consisting of the first 1-limit, the first 2-limit, the first 3-limit and so on. The limit of this n-sequence is n^n. It is an n**-limit**. In the definition of n-sequence, why did we use the sequence 1-limit, 2-limit, 3-limit and so on instead of say 1-limit, 4-limit, 7-limit and so on? The reason is that the sequence $1, 2, 3, \cdots$ is *already a natural sequence*, while $1, 4, 7, \cdots$ is not.

After n^n we get a 1-sequence of successors with limit $n^n + n$, which is a 1-limit. Then another 1-limit $n^n + 2n$, then another $n^n + 3n$, and then the 2-limit $n^n + n^2$. Eventually after n^n there will be a first 1-limit, a first 2-limit, a first 3-limit and so on, which gives another n-sequence with n-limit $2n^n$.

The sequence $n^n, 2n^n, 3n^n, \cdots$ of n-limits is an $(n+1)$**-sequence**, with $(n+1)$-limit n^{n+1}. After n^{n+1} we get a 1-sequence of successors with limit $n^{n+1} + n$, which is a 1-limit. Then another 1-limit $n^{n+1} + 2n$, then another $n^{n+1} + 3n$ and so the 2-limit $n^{n+1} + n^2$. Eventually after n^{n+1} there will be a first 1-limit, a first 2-limit, a first 3-limit and so on, which gives another n-sequence with n-limit $n^{n+1} + n^n$.

The sequence of n-limits $n^{n+1}+n^n, n^{n+1}+2n^n, n^{n+1}+3n^n, \cdots$ is an $(n+1)$-sequence, with $(n+1)$-limit $2n^{n+1}$. The sequence $n^{n+1}, 2n^{n+1}, 3n^{n+1}, \cdots$ of $(n+1)$-limits is an $(n+2)$**-sequence**, with $(n+2)$**-limit** n^{n+2}. The first $(n+3)$**-sequence** will be $n^{n+2}, 2n^{n+2}, 3n^{n+2}, \cdots$ with $(n+3)$**-limit** n^{n+3}.

Now recall that $n+1, n+2, n+3, \cdots$ is a 1-sequence, with 1-limit $2n$. So the sequence $n^{n+1}, n^{n+2}, n^{n+3}, \cdots$ consisting of the first $(n+1)$-limit, the first $(n+2)$-limit, the first $(n+3)$-limit and so on will be a $2n$**-sequence**, with $2n$**-limit** n^{2n}.

So the rules of the previous section may be augmented by the following:

- If f_1, f_2, f_3 is a natural sequence with limit f, then the sequence of the first f_1-limit, the first f_2-limit, the first f_3-limit and so on is an f-sequence, and its limit is an f-limit.
- More generally if $g \geq f$ then following a g-limit h, the sequence consisting of the first f_1-limit, the first f_2-limit, the first f_3-limit and so on is by definition an f-sequence, and its limit an f-limit.

With these rules, we see that whenever f is a natural function, so is n^f, and it is an f-limit. For example n^{n^n} is an n^n-limit, and furthermore $3n^{2n+1}+5n^{n+3}$ is an $(n+3)$-limit, and $n^{n^{n+4}}+2n^{n^2}$ is an n^2-limit. These are all exponential natural functions, and they can be labelled by rooted trees. [Just like all the ordinals up to ε_0.]

The previous conventions do not apply to get us up to pyramidal functions. If we try to specify the sequence $n, n^n, n^{n^n}, \cdots$ we see they are the first 1-limit, the first n-limit, the first n^n-limit and so on. In this case the indexing sequence has essentially caught up with the sequence itself. This is analogous to what happens with the ordinal ε_0. So we need to extend our rules to include this as a new type of natural sequence.

Instead, let's have a quick look at higher forms of arithmetic.

Operations

An **operation** $\otimes$ is a function which takes an ordered pair of natural numbers $[M, N]$ and outputs another number $M \otimes N$. The successor function $N \to N' = N+1$ is not an operation in this sense, but it plays a useful role in building up operations, which we now do in an inductive fashion.

If we didn't know addition already, we could define it in terms of the successor operation by the inductive rules

$$\begin{aligned} M+1 &= M' \\ M+N' &= (M+N)'. \end{aligned}$$

Multiplication can be defined in terms of successor and addition by the rules

$$\begin{aligned} M \times 1 &= M \\ M \times N' &= M + (M \times N). \end{aligned}$$

Exponentiation can be defined using successor and multiplication by the rules

$$\begin{aligned} M \wedge 1 &= M \\ M \wedge N' &= M \times (M \wedge N). \end{aligned}$$

Pyramidation can be defined using successor and exponentiation by the rules

$$\begin{aligned} M \bigtriangleup 1 &= M \\ M \bigtriangleup N' &= M \wedge (M \bigtriangleup N). \end{aligned}$$

Clearly there is a pattern here. Let's introduce the idea of $\wedge$ and $\triangle$ as the **second** and **third multiplications** respectively. So we will write $\times = \times_1$, $\wedge = \times_2$ and $\triangle = \times_3$, and then define the k'**th multiplication** $\times_{k'}$ in terms of the kth $\times_k$ by the rule

$$\begin{aligned} M \times_{k'} 1 &= M \\ M \times_{k'} N' &= M \times_k (M \times_{k'} N) . \end{aligned}$$

These are all **natural operations** by definition. Note that it is easier here to not include zero as a natural number, as defining these operations when one or both of the inputs is zero is more problematic.

More generally, for any operation $\otimes$ define the **successor** $\otimes'$ to be the operation defined by

$$\begin{aligned} M \otimes' 1 &= 1 \\ M \otimes' N' &= M \otimes (M \otimes' N) . \end{aligned}$$

There is a notion of the **limit** of a natural sequence $\otimes_1, \otimes_2, \otimes_3, \cdots$ of natural operations. This is the natural operation $\otimes$ defined by

$$M \otimes N = M \otimes_N N.$$

For example the sequence $\times_1, \times_2, \times_3, \cdots$ has limit $\times_n$ given by

$$M \times_n N = M \times_N N.$$

But now we are in the position of having an addition operation $+$, a sequence of multiplication operations $\times_1, \times_2, \times_3, \cdots$, and a way of taking limits of sequences of operations and successors of operations. But this means that we can create natural operations in the same way we created natural functions. So we have operations $\times_{n+1}$, $\times_{n^2}$, $\times_{n^n}$ and so on. In other words, for a natural function f, we can create a well defined operation $\times_f$.

With such higher operations, we can create a range of new 'levels' of functions. Recall that the polynomial, exponential and pyramidal functions were closely connected with the multiplication operators $\times_1, \times_2$ and $\times_3$. So for any function f we can investigate the level of functions brought into existence by our new operator $\times_f$. This is a pretty nifty way of writing down biggish numbers fast. Consider the natural number $4 \times_{n^n} 7$—can you express it (approximately) in any more conventional manner?

At this point, I will leave the theory for logicians and computer scientists to develop further. I believe it is potentially a rewarding subject—perhaps the 'purest' of pure mathematics, as it soon becomes completely devoid of practical applications. As Leonard Cohen put it in *Death of a Lady's Man:*

> It's like our visit to the moon or to that other star
>
> I guess you go for nothing if you really want to go that far.

Infinitesimals

Infinitesimals augment the rational number system, that is fractions. Once we have rational numbers, we can extend our notion of natural function to allow natural number inputs and positive rational number outputs. If f is such a function then $1/f$ denotes the **reciprocal** function

$$1/f : N \to \frac{1}{f(N)}$$

provided that $f(N)$ is non-zero. All the natural functions we have defined are everywhere nonzero, so they all have reciprocals.

An **infinitesimal** is by definition the *reciprocal of an infinity.* The natural functions $n, n+1, 2n, n^2, n^n, n^{\triangle n}$ and so on have reciprocals which we respectively denote by

$$\frac{1}{n}, \frac{1}{n+1}, \frac{1}{2n}, \frac{1}{n^2}, \frac{1}{n^n}, \frac{1}{n^{\triangle n}}$$

and so on. Each *corresponds to a decay rate*, as we let n increase. The most important infinitesimals form the sequence

$$\frac{1}{n}, \frac{1}{n^2}, \frac{1}{n^3}, \cdots$$

and these are the ones we may use for most routine applications in the calculus.

Before we do that, let's augment the rational numbers with infinitesimals in a visual way. An **augmented number** is the sum of a rational number and some rational multiples of infinitesimals. For example $3 + 2\frac{1}{n}$ is an augmented number, and can be interpreted as the function

$$3 + 2\frac{1}{n} : N \to 3 + \frac{2}{N}.$$

If you think of the rational numbers as lying on the number line, then just imagine a given rational number (say $2/3$) as a regular sequence of lighted blips, one for every time period $n = 1, 2, 3, \cdots$, and each at the point $2/3$. In other words, instead of thinking of a point as a constant lit up point, think of it as a regular sequence of flashing lights, all at the same point. We visualize an augmented number as a *non-constant sequence* of lighted blips. For $3 + 2\frac{1}{n}$ the first blip is at $3 + 2$, the second at $3 + 1$, the third at $3 + 2/3$ and so on. As you watch this sequence of blips, they approach the number 3 at a particular rate. This sequence of blips, indexed by the natural numbers $1, 2, 3, \cdots$, **is** the augmented number. There is nothing more to it than that!

The **standard part** of this augmented number is 3, and the **infinitesimal part** is $2\frac{1}{n}$. The infinitesimal part of an augmented number is another augmented number which corresponds to a sequence of blips which approach 0.

Now we may do additive and multiplicative arithmetic with these numbers, for example

$$\left(3+2\frac{1}{n}\right)+\left(4-3\frac{1}{n}+5\frac{1}{n^2}\right) = 7-\frac{1}{n}+5\frac{1}{n^2}$$
$$\left(3+2\frac{1}{n}\right)\left(4-3\frac{1}{n}+5\frac{1}{n^2}\right) = 12-\frac{1}{n}+9\frac{1}{n^2}+10\frac{1}{n^3}.$$

We might also allow reciprocals, rational functions and selected power series such as

$$\frac{1}{1+2\frac{1}{n}} = 1-2\frac{1}{n}+4\frac{1}{n^2}-8\frac{1}{n^3}+\cdots.$$

Some care will have to be taken with the latter concept!

Derivatives with infinitesimals

Without developing a complete theory of nonstandard analysis, let's give some initial indications of how these concepts might be used. Note that in this section we work entirely in the rational number framework. Suppose we wish to study the function $f(x) = x^2$ near the rational number $x = a$. We can do this by setting $x = a+\frac{1}{n}$. The n variable allows you a motion-picture view as you zoom in towards a, since as n increases, $a+\frac{1}{n}$ approaches a. We get

$$f\left(a+\frac{1}{n}\right) = \left(a+\frac{1}{n}\right)^2 = a^2+2a\frac{1}{n}+\frac{1}{n^2}.$$

Such an equation describes the behavior of f infinitesimally close to a. An increase in a by the infinitesimal $\frac{1}{n}$ yields an increase in $f(a)$ by the infinitesimal $2a\frac{1}{n}+\frac{1}{n^2}$. The latter increase divided by the former increase is by definition the **full derivative** of f at a, which we denote by $f^{[1]}(a)$. More precisely

$$f^{[1]}(a) = \frac{f\left(a+\frac{1}{n}\right)-f(a)}{\frac{1}{n}} = 2a+\frac{1}{n}.$$

The standard part of the full derivative $f^{[1]}(a)$ is by definition the **standard derivative** of f at a, denoted by either $f'(a)$ or $f^{(1)}(a)$. Thus in the above example

$$f'(a) = f^{(1)}(a) = 2a.$$

Similarly the **full second derivative** of f at a is

$$f^{[2]}(a) = \frac{f^{[1]}\left(a+\frac{1}{n}\right)-f^{[1]}(a)}{\frac{1}{n}} = \frac{2\left(a+\frac{1}{n}\right)-2a}{\frac{1}{n}} = 2.$$

The **standard second derivative** $f''(a) = f^{(2)}(a)$ is defined to be the standard part of the full second derivative. In the example, it happens to be equal to the full second derivative $f^{[2]}(a)$. In general the **full** $(k+1)$**st derivative** of a function f at a point a may be defined in terms of the full kth derivative by the rule

$$f^{[k+1]}(a) = \frac{f^{[k]}\left(a+\frac{1}{n}\right) - f^{[k]}(a)}{\frac{1}{n}},$$

and the **standard kth derivative** is the standard part of the full kth derivative. Thus

$$f^{(k)}(a) = \text{standard part of } f^{[k]}(a).$$

Let's give a few more examples. Consider the function $g(x) = \frac{x-1}{x+1}$ near the point $x = a$. Then the full derivative is

$$\begin{aligned} g^{[1]}(a) &= \frac{g\left(a+\frac{1}{n}\right) - g(a)}{\frac{1}{n}} \\ &= \frac{2}{\left(a+1+\frac{1}{n}\right)(a+1)} \end{aligned}$$

and since the standard part is obtained by setting any infinitesimal terms to zero, we get

$$g^{(1)}(a) = \frac{2}{(a+1)^2}.$$

The full second derivative is

$$\begin{aligned} g^{[2]}(a) &= \frac{g^{[1]}\left(a+\frac{1}{n}\right) - g^{[1]}(a)}{\frac{1}{n}} \\ &= \frac{-4}{\left(1+a+2\frac{1}{n}\right)\left(1+a+\frac{1}{n}\right)(a+1)} \end{aligned}$$

and taking standard parts,

$$g^{(2)}(a) = \frac{-4}{(a+1)^3}.$$

Note that

$$f^{[2]}(a) = \frac{f^{[1]}\left(a+\frac{1}{n}\right) - f^{[1]}(a)}{\frac{1}{n}} = \frac{f\left(a+\frac{2}{n}\right) - 2f\left(a+\frac{1}{n}\right) + f(a)}{n^2}.$$

More generally

$$\begin{aligned} f^{[k]}(a) &= \frac{1}{n^k}\left\{ \begin{array}{c} f\left(a+\frac{k}{n}\right) - \binom{k}{1} f\left(a+\frac{k-1}{n}\right) \\ +\binom{k}{2} f\left(a+\frac{k-2}{n}\right) - \cdots + (-1)^k f(a) \end{array} \right\} \\ &= \frac{1}{n^k}\sum_{i=0}^{k} (-1)^k \binom{k}{i} f\left(a+\frac{k-i}{n}\right). \end{aligned}$$

Product, quotient and chain rules

For two functions f and g,

$$\begin{aligned}(fg)^{[1]}(a) &= \frac{(fg)\left(a+\frac{1}{n}\right)-(fg)(a)}{\frac{1}{n}}\\ &= \frac{(fg)\left(a+\frac{1}{n}\right)-f\left(a+\frac{1}{n}\right)g(a)+f\left(a+\frac{1}{n}\right)g(a)-(fg)(a)}{\frac{1}{n}}\\ &= f^{[1]}(a)\,g(a)+f\left(a+\frac{1}{n}\right)g^{[1]}(a).\end{aligned}$$

Taking standard parts gives

$$(fg)^{(1)}(a)=f^{(1)}(a)\,g(a)+f(a)\,g^{(1)}(a).$$

For two functions f and g,

$$\begin{aligned}\left(\frac{f}{g}\right)^{[1]}(a) &= \frac{\left(\frac{f}{g}\right)\left(a+\frac{1}{n}\right)-\left(\frac{f}{g}\right)(a)}{\frac{1}{n}}=\frac{\frac{f\left(a+\frac{1}{n}\right)}{g\left(a+\frac{1}{n}\right)}-\frac{f(a)}{g(a)}}{\frac{1}{n}}\\ &= \frac{f\left(a+\frac{1}{n}\right)g(a)-f(a)\,g\left(a+\frac{1}{n}\right)}{g\left(a+\frac{1}{n}\right)g(a)\frac{1}{n}}\\ &= \frac{f\left(a+\frac{1}{n}\right)g(a)-f(a)\,g(a)+f(a)\,g(a)-f(a)\,g\left(a+\frac{1}{n}\right)}{g\left(a+\frac{1}{n}\right)g(a)\frac{1}{n}}\\ &= \frac{f^{[1]}(a)\,g(a)-f(a)\,g^{[1]}(a)}{g\left(a+\frac{1}{n}\right)g(a)}.\end{aligned}$$

Taking standard parts gives

$$\left(\frac{f}{g}\right)^{(1)}(a)=\frac{f^{(1)}(a)\,g(a)-f(a)\,g^{(1)}(a)}{g(a)^2}.$$

For two functions f and g set $(f\circ g)(x)=f(g(x))$. Then

$$\begin{aligned}(f\circ g)^{[1]}(a) &= \frac{(f\circ g)\left(a+\frac{1}{n}\right)-(f\circ g)(a)}{\frac{1}{n}}\\ &= \frac{f\left(g\left(a+\frac{1}{n}\right)\right)-f(g(a))}{\frac{1}{n}}\end{aligned}$$

but

$$g\left(a+\frac{1}{n}\right)=g(a)+g^{[1]}(a)\,\frac{1}{n}$$

so

$$(f\circ g)^{[1]}(a)=\frac{f\left(g(a)+g^{[1]}(a)\frac{1}{n}\right)-f(g(a))}{g^{[1]}(a)\frac{1}{n}}\times\frac{g\left(a+\frac{1}{n}\right)-g(a)}{\frac{1}{n}}.$$

Take standard parts, and use the fact that for any decimal number b

$$f^{(1)}(a) = \text{standard part of } \frac{f\left(a + b\frac{1}{n}\right) - f(a)}{b\frac{1}{n}}$$

to obtain

$$\begin{aligned}(f \circ g)^{(1)}(a) &= \frac{f\left(g(a) + g^{(1)}(a)\frac{1}{n}\right) - f(g(a))}{g^{(1)}(a)\frac{1}{n}} g^{(1)}(a) \\ &= f^{(1)}(g(a))\, g^{(1)}(a).\end{aligned}$$

Directions for education

Of course many more details need to be worked out, definitions need be made more precise etc. In addition, analysts will want to develop this theory over the 'real numbers'. May I humbly suggest that they first provide a coherent account of what 'real numbers' might really be?

What is necessary is the following—to come to grips with the *computational aspects* of the *algorithms* that specify Cauchy sequences (or decimal expansions, or Dedekind cuts, or continued fraction expansions, or whatever tool you use—the problems are essentially the same in all cases, as discussed in [Wi]). Much work needs to be done, completely in the computable framework, divorced from metaphysics. *No more waffle about 'non-computable numbers', please*!

In practice, calculus can be approached as largely an algebraic subject. Most functions that arise in real life can be piecewise expanded in power series, and most numbers that come up can be dealt with by rational approximations. Ultimately knowing how to deal with power series involving rational numbers solves the great majority of real life calculus problems (in the spirit of Newton). The idea of an infinitesimal as a quantity gradually getting smaller and smaller at a certain fixed rate fits with this approach, and is simpler than epsilons and deltas for students.

Versions of nonstandard analysis using infinitesimals as sequences have already been studied. The first to have done so might have been the Polish philosopher and mathematician L. Chwistek in *The Limits of Science, Outline of Logic and Methodology of Science* published in 1935 (in Polish). D. Laugwitz [L1], [L2] developed a system of sequences called Ω-*Zahlen* to study distributions and operators, and this in turn was based on earlier work by Schmieden and Laugwitz [SL]. More recent developments in this direction include J. M. Henle's articles [He1] and [He2] as well as P. Schuster's work [Sc], and the recent calculus text by Cohen and Henle [CH].

In conclusion, an infinity is a particular kind of function on natural numbers, defined inductively and including n, n^2, n^n etc. An infinitesimal is the reciprocal of an infinity. In this way ordinal arithmetic and the basics of calculus can be recast using only natural numbers—as Kronecker urged us to do. Such an approach justifies some of the intuitions of greats such as Newton, Euler, Gauss and many other prominent mathematicians prior to the twentieth century in their understanding of analysis. It also leads to a simpler form of nonstandard analysis accessible to students.

References

[Ch] L. Chwistek, *The Limits of Science, Outline of Logic and Methodology of Science*, Lwów-Warszawa: Ksiażnica-Atlas, 1935.

[CH] D. Cohen and J. M. Henle, *Calculus: The Language of Change*, Jones and Bartlett, 2004.

[He1] J. M. Henle, 'Non-nonstandard Analysis: Real infinitesimals', *Mathematical Intelligencer*, **21**, Number 1 (1999), 67–73.

[He2] J. M. Henle, 'Second-order Non-nonstandard Analysis', *Studia Logica* **74** No. 3, (2003), 399-426.

[Ke] H. J. Keisler, *Elementary Calculus*, Weber and Schmidt, Boston, Mass., 1976.

[L1] D. Laugwitz, 'Anwendungen unendlich kleiner Zahlen I', *J. Reine Angew. Math.* **207** (1960), 53-60.

[L2] D. Laugwitz, 'Anwendungen unendlich kleiner Zahlen II', *J. Reine Angew. Math.* **208** (1961), 22-34.

[LG] H. S. Leonard and N. Goodman, 'The Calculus of Individuals and Its Uses', *Journal of Symbolic Logic* **5** (1940), 45-55.

[Ro] A. Robinson, *Nonstandard Analysis*, Amsterdan: North-Holland, 1966.

[SL] C. Schmieden and D. Laugwitz, 'Eine Erweiterung der infinitesimalrechnung', *Math. Zeitsch.* **69** (1958), 1-39.

[Sc] P. Schuster, 'A Constructive Look at Generalised Cauchy Reals', *Mathematical Logic Quarterly* **46** (2000) 1, 125–134.

[Wi] N. J. Wildberger, 'Real Fish, Real Numbers, Real Jobs', *Mathematical Intelligencer*, 21 No. 2, (1999), 4-7.