The way of the Infinitesimal:

Fabrizio Genovese

Quantum Group
University of Oxford

`fabrizio.genovese@hertford.ox.ac.uk`

In this document we provide a brief introduction to Non-Standard Analysis, with Categorical Quantum Mechanics in mind as application. Nevertheless, we are convinced this document will be helpful for everyone wanting to learn how to practically use Non-Standard Analysis without having to go too deep into the logical and model-theoretical complications of the subject. Our goal is to give the reader the necessary confidence to manipulate infinities, infinitesimals and the standard constructions, to employ Non-Standard Analysis as an everyday tool in scientific research.

This document is heavily based on "Lectures on the Hyperreals" by Goldblatt, that has to be credited for having given one of the simplest and friendliest explanations of the subject so far.

1 Introduction

When we think about limits, derivatives and other constructions commonly employed in calculus we often do it in intuitive terms. We manipulate "dx", "dy" and the like as they were algebraic objects, with ease. Nevertheless, when it comes to formalize these entities mathematically the definition usually employed is the one *à la Cauchy*: derivatives and limits become analytic constructions relying on the mantra "the more you go further, the more this quantity will get near to this other one". As an instance, take the usual undergrad textbook definition of limit for a succession, originally given by Weierstrass:

$$\lim_{n \to \infty} s_n = l \stackrel{def}{\iff} \forall \epsilon > 0, \exists n_0 : \forall n \geq n_0, |l - s_n| \leq \epsilon$$

We all know what this definition tries to capture: It does not matter how small ϵ is, for every ϵ you pick at some point the distance between s_n and l will constantly be less than ϵ, and hence we can say that s_n gets "closer and closer" to l. We would be tempted to say, then, that s_n becomes *infinitesimally close to l* but we can't, since we have no notion of what "infinitesimal" really means: This concept has to be relegated to our intuition. In the practice, this means that one has to deal with a lot of ugly inequalities, moving out from the comfortable and reassuring world of pure algebra. (I'm pretty sure this is the main reason why people that hate calculus hate calculus.)

Definitions such the one above work nicely for a plethora of applications, and this is clearly proved by the fact that calculus is probably the most popular tool in the history of Science. Nevertheless, not having any way to explicitly talk about infinities and infinitesimals becomes a problem when, say, one has to work with *compositional features* of a model.

To better explain this, we consider an example from Categorical Quantum Mechanics (abbreviated CQM). In this framework there are some very important tools that are called *Frobenius algebras*. It is not important at this point to know what they do or why they are useful, what is important is that, in a Hilbert space $\mathbb{H}$, to have Frobenius algebras the sum $\sum_1^n \langle e_i |$ has to be part of $\mathcal{H}$, where n is its and $|e_i\rangle$ are vectors of an orthonormal basis. Obviously using the standard definition of sum and limits this object does not belong to $\mathcal{H}$ unless its dimension is finite (that sum is divergent), and this is a problem that

essentially relegated CQM to the study of finite-dimensional quantum mechanics for long time. Similarly, in physics one manipulates all the time objects such as Dirac deltas and plain waves on spaces like $L^2[\mathbb{R}]$ or $L^2[\mathbb{C}]$. Again, it doesn't matter to know what Dirac deltas and plain waves are: The problem here is that they are not elements of $L^2[\mathbb{R}]$ or $L^2[\mathbb{C}]$, and this poses serious limitations when one wants to capture the compositional behaviour of physical systems using category theory.

The solution to this problem comes from Non-Standard Analysis, and our plan is more or less this: We want to extend the real numbers (or the complex numbers, or $L^2[\mathbb{R}]$, whatever) including infinities and infinitesimals. In this way entities like the diverging sum above become well-defined algebraic objects with their own dignity, and we can formalize our constructions (as in the CQM case) on these extended versions of our spaces. This construction will be explained in detail in the next section.

This is clearly not enough: We want to be able to control our extension. This means that we want clear and sharp answers to questions like "If T is a theorem in some environment, do I know when it is also a theorem in the environment extension, and vice-versa?" Satisfaction of this requirement allows us to go back and forth from the extension to the original environment, and is ultimately what characterizes our extension construction as meaningful, making us able to regard it as "my original environment plus algebraic entities that represent infinities and infinitesimals, entities that I can manipulate exactly as my intuition would suggest". The piece of machinery that puts us in control, attaining exactly what we wanted to do here, is called *transfer theorem*, and it will be explained in detail in section [].

In the last section we will put all this machinery into good use. We will take a lot of textbook examples à la Cauchy and we will re-express them "à la Robinson", that is, using non-standard formalism. This will hopefully give the practicing scientist in need of an algebraic formalization of calculus enough confidence to wander in the beautiful world of Non-Standard Analysis as he pleases.

As already mentioned in the abstract, we owe a lot to Robinson[], the father of Non-Standard Analysis itself, and to Goldblatt[], that was able to explain the subject in a very simple and intuitive way. In particular, this document can be considered a recap of his book "Lectures on the Hyperreals", to which we redirect the reader that wants to know more.

The original book by Robinson[] is, perhaps, still the Bible of Non-Standard Analysis: Countless non-standard formalisations of interesting mathematical structures can be found there, from Hilbert Spaces to Differential Geometry, and we redirect the reader to that beautiful world of wonders in case more in-dept information is needed.

2   The construction of $^*\mathbb{R}$

In logic there is a very famous theorem called *compactness theorem*. What it says goes more or less like this: If Σ is a set of statements such that every finite subset Σ' of Σ has a model (that is, some mathematical structure in which all the statements in Σ' are true), then Σ has a model itself. Now, suppose that Σ is the set of all statements that are true for real numbers, plus the statements

$$0 < \epsilon, \quad \epsilon < 1, \quad \epsilon < \frac{1}{2}, \quad \epsilon < \frac{1}{3}, \quad ,\ldots$$

Using the compactness theorem we can infer that Σ has a model, call it $^*\mathbb{R}$, that will be an ordered field in which the element ϵ is an infinitesimal, that is, something greater than zero but smaller than any real number. We also get that statements in $^*\mathbb{R}$ "formulated in the right way" will hold if and only if they hold in $\mathbb{R}$. This last sentence has been intentionally expressed in a murky way, and one of the objectives of this document is to understand what it precisely means.

Albeit the compactness theorem is what makes non-standard analysis ultimately work, extending the real numbers using this approach has a great disadvantage: We get a model of the reals containing infinities and infinitesimals, but in a non-constructive way. What we will do, instead, is to *constructively build* the extended real numbers $^*\mathbb{R}$, using mathematical tools called *ultraproducts*. The experienced logician will be aware that, in practice, what we will do amounts to re-proving the compactness theorem in a more specialized setting that suits our need. But in order to proceed with this construction, we have to figure out first what real numbers really are.

2.1 Structure of $\mathbb{R}$

There are many definitions of real numbers, all equivalent. We here state some of them, with which the reader may be familiar:

- A real number is something that can be expressed as an infinite decimal expression. As an instance,

$$\pi = 3.141592\dots$$

  The set of all such numbers is called $\mathbb{R}$.

- A real number is a set of a field $\mathbb{R}$ that is:

  - Ordered, meaning that there is an order relation $\leq$ compatible with the field operations of $\mathbb{R}$;
  - Complete, meaning that every non-empty set $S \subseteq \mathbb{R}$ with an upper bound[1] has a supremum[2].

  This notion identifies $\mathbb{R}$ uniquely since it can be shown that two ordered complete fields are always isomorphic.

- Given a sequence $p_1, p_2, \dots, p_n, \dots$, we say that it is a *Cauchy sequence* if $\lim_{i,j\to\infty} |p_i - p_j| = 0$. This is nothing more than the usual definition of Cauchy sequence seen in undergrad calculus. Now consider the set of all the Cauchy sequences of rational numbers (call it $\overline{\mathbb{Q}}$), that is, all the Cauchy sequences in which $p_i \in \mathbb{Q}$ for every i.

  Given two sequences $p_1, p_2, \dots, p_n, \dots, s_1, s_2, \dots, s_n, \dots$, we say that they are *equivalent* if

$$\lim_{n\to\infty} |p_n - s_n| = 0$$

  meaning that two sequences are equivalent if they approach the same limit. It can be shown that the definition above is an equivalence relation on $\overline{\mathbb{Q}}$. The equivalence classes of this relation can be endowed with an ordered field structure that is complete, and this again characterizes $\mathbb{R}$ since we already said that two complete ordered fields are isomorphic.

  Note that in this approach two equivalent Cauchy sequences define the same real number.

As we already mentioned, it can be proved that all the definitions above are equivalent, but the reader can already appreciate some differences: The third one is constructive, while the second one is not. We will extend $\mathbb{R}$ with infinities and infinitesimals following steps similar to the ones in third construction, starting with sequences of real numbers and quotienting them in an opportune way. The next section is about finding what this opportune way is.

---

[1] An upper bound for $S \subseteq \mathbb{R}$ is an $x \in \mathbb{R}$ such that for all $s \in S$ it is $s \leq x$.
[2] This means that the set of the upper bounds of S has a minimum.

2.2 Comparing Sequences

In the last subsection we said that we begin our construction considering sequences of real numbers. If r is a sequence, we will indicate with r_n the n-th element of the sequence, that is of course just a real number.

Given two sequences r, s, we want to find some criterion to compare them. Intuitively, one way to go is to say that r, s are equivalent if they agree on a *large number* of elements, that is, if the set

$$E_{r,s} := \{n \in \mathbb{N} : r_n = s_n\}$$

is *large*, whatever this means. Since we want this notion of equivalence to define an equivalence relation, we can already state some properties we want this notion of bigness to have:

- $\mathbb{N}$ has to be large, this is because our equivalence relation has to be reflexive, and clearly $E_{r,r} = \mathbb{N}$;

- Equivalence relations are transitive. This means that if $E_{r,s}$ and $E_{s,k}$ are large, also $E_{r,k}$ has to be large. But $E_{r,k}$ is contained in $E_{r,s} \cap E_{s,k}$, and hence the property we are looking for is

  If A, B are large and $A \cap B \subseteq C$, then C is large.

  In particular this means that the set of large sets is closed for intersection and supersets, meaning that

  If A, B are large and $A \subseteq C$, then $A \cap B$ and C are large.

- $\emptyset$ cannot be large: Every other set $K \subset \mathbb{N}$ has the property $\emptyset \subseteq K$, so if $\emptyset$ were to be large every subset of $\mathbb{N}$ would be large, making our equivalence relation trivial.

Remark 2.1. The first notion of largeness that comes to mind is to say that a set L is large if it is *cofinite*, meaning that its complement $\mathbb{N} - L$ is a finite set. Unfortunately this notion does not work well: What we want ultimately do is to take our extension of the real numbers to be the sequences of reals quotiented by the equivalence relation we are defining. Moreover we want this extension to be *totally ordered*, meaning that given two equivalence classes of sequences r, s it is always $[r] < [s]$, $[l] = [s]$ or $[l] > [s]$. The natural way to say that $[r]$ is less than a sequence $[s]$ is to require that the set[3]

$$L_{r,s} := \{n \in \mathbb{N} : r_n < s_n\}$$

is large. Now consider the sequences $r = (1, 0, 1, 0, \dots), s = (0, 1, 0, 1, \dots)$. Clearly $E_{r,s} = \emptyset$, so these two sequences are not identified by our relation, and we should be able to say $[r] < [s]$ or $[r] > [s]$. But both $L_{r,s}, L_{s,r}$ are not cofinite, and so we are not able to compare $[r], [s]$ if we take cofiniteness as a definition of large.

From this remark we deduce that we have to impose an extra requirement, namely that

Given a set A, either A or $\mathbb{N} - A$ is large.

We can immediately infer that if we require this A and $\mathbb{N} - A$ cannot be both large, otherwise $A \cap \mathbb{N} - A = \emptyset$ would be large as well.

The experienced reader will already have note that our requirements to define what largeness is match the definition of *ultrafilter*. This is indeed the tool we are going to use.

---

[3]Obviously we have to check that this definition is independent of the representatives of $[r], [s]$, but this is an easy exercise.

2.3 Ultrafilters

After the discussion in the last section we are able to formally state what we want.

Definition 2.2. Let I be a nonempty set, and denote with $\mathcal{P}(I)$ its powerset (the set of all subsets of I). A *filter* on I is a nonempty collection $\mathcal{F}$ of subsets of I (that is, an $\mathcal{F} \subseteq \mathcal{P}(I)$) such that

- If $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$ ($\mathcal{F}$ is closed with respect to intersections);

- If $A \in \mathcal{F}$ and $A \subseteq B \subseteq I$, then $B \in \mathcal{F}$ ($\mathcal{F}$ is closed with respect to supersets).

Moreover, a filter $\mathcal{F}$ is called *ultrafilter* if it satisfies the additional condition

$$\text{For any } A \subseteq I, \text{ it is } A \in \mathcal{F} \text{ or } I - A \in \mathcal{F}.$$

Note that a filter containing $\emptyset$ coincides with $\mathcal{P}(I)$. Thus, we call $\mathcal{F}$ *proper* if $\emptyset \notin \mathcal{F}$.

Example 2.3. The powerset $\mathcal{P}(I)$ of any nonempty set I is a filter (the trivial one, if you want) but not an ultrafilter.

Example 2.4. The *cofinite* (or *Fréchet*) filter briefly discussed in the previous subsection in the case of natural numbers, defined as

$$\mathcal{F} := \{N \subseteq I : \ I - N \text{ is finite}\}$$

is a filter, but not in general an ultrafilter.

Example 2.5. The set $\{I\}$ is an ultrafilter. Note that I is an element of every filter because of the superset closure, hence $\{I\}$ is the smallest filter on I.

Example 2.6. The *principal ultrafilter generated by an element $i \in I$*, defined as

$$\mathcal{F} := \{S \subseteq I : \ i \in S\}$$

is an ultrafilter (as the name suggests, I'd say...)

A very important fact, the proof of which can be found on [], is that

Theorem 2.7. *Every filter on a finite set I is a principal ultrafilter. If I is infinite, then it admits a NON principal ultrafilter on it.*

We will use this theorem later, so keep it in mind.

2.4 Construction of $^*\mathbb{R}$

Finally we developed enough stuff to really begin our journey. As we said, we start from sequences of real numbers. Since a sequence $(r_1, r_2, \ldots)$ of real numbers can be seen as a function $\mathbb{N} \to \mathbb{R}$ and vice-versa, we can denote with $\mathbb{R}^\mathbb{N}$ the set of all sequences of real numbers. We will usually denote with r the sequence $(r_1, r_2, \ldots)$ (mnemonic rule: unless otherwise specified, such as in "$r \in \mathbb{R}$", if r has a subscript it's a real, otherwise it's a sequence of reals!), and with $\mathbf{r}$ the constant sequence $(r, r, \ldots)$ for $r \in \mathbb{R}$.

The first thing we want to do is to endow $\mathbb{R}^\mathbb{N}$ with a ring structure. Given sequences r, s we can define

$$r \oplus s := (r_1 + s_1, r_2 + s_2, \ldots)$$
$$r \odot s := (r_1 \cdot s_1, r_2 \cdot s_2, \ldots)$$

$(\mathbb{R}^\mathbb{N}, \oplus, \otimes)$ is a commutative ring, with zero $\mathbf{0} = (0, 0, \ldots)$, unity $\mathbf{1} = (1, 1, \ldots)$ and addictive inverses $-r = (-r_1, -r_2, \ldots)$ for every r. It is not a field since $(0, 1, 0, 1, \ldots) \odot (1, 0, 1, 0, \ldots) = \mathbf{0}$ and so the sequences on the right do not possess multiplicative inverses.

Our aim is to turn this ring into an ordered field. As expected, we will reach our goal quotienting $(\mathbb{R}^\mathbb{N}, \oplus, \otimes)$ by an ultrafilter on the natural numbers. We will show in a bit that we need this ultrafilter to be non principal to get infinities and infinitesimals from our extension.

Lemma 2.8. *Let $\mathcal{F}$ be a non principal ultrafilter on $\mathbb{N}$, the existence of which is guaranteed by theorem 2.7. The relation $\equiv$ on $\mathbb{R}^{\mathbb{N}}$ defined by*

$$r \equiv s \overset{def}{\iff} \{n \in \mathbb{N} : \ r_n = s_n\} \in \mathcal{F}$$

is an equivalence relation, that is moreover compatible with the operations $\oplus, \odot$, meaning that if $r \equiv r'$, $s \equiv s'$, then $r \oplus s \equiv r' \oplus s'$ and $r \odot s \equiv r' \odot s'$.

When $r \equiv s$ we say that r and s *agree almost everywhere modulo $\mathcal{F}$*. Borrowing some logical notation the set $\{n \in \mathbb{N} : \ r_n = s_n\}$ can be denoted as $[[r = s]]$, and can be thought as a measure of the extent to which "$r = s$" is true. Quotienting $\mathbb{R}^{\mathbb{N}}$ by $\mathcal{F}$ can be tought as collapsing this measure into a binary one: r, s agree "enough to be equal" or they do not. The notation $[[r = s]]$ is utterly convenient to generalise this procedure to statements other than equality, as an instance setting

$$[[r \leq s]] := \{n \in \mathbb{N} : \ r_n \leq s_n\}$$

Clearly we can now say that $r \leq s$ iff $[[r \leq s]] \in \mathcal{F}$. Moreover, it is easy to prove that if $r \equiv r'$, $s \equiv s'$ then $[[r \leq s]] \in \mathcal{F}$ if and only if $[[r' \leq s']] \in \mathcal{F}$, meaning that also $\leq$ is somehow well-behaved with respect to $\equiv$. Generalization to $[[r < s]], [[r \geq s]], [[r > s]]$ is obvious.

Since $\oplus, \odot, <$ are well-behaved with respect to $\equiv$, we can extend these operations to the quotient $\mathbb{R}^{\mathbb{N}}/\mathcal{F}$. We denote with $[r]$ an element of $\mathbb{R}^{\mathbb{N}}/\mathcal{F}$, that, being an equivalence class modulo $\mathcal{F}$, is the set of all $r' \in \mathbb{R}^{\mathbb{N}}$ such that $r \equiv r'$. We set

$$[r] + [s] := [r \oplus s]$$
$$[r] \cdot [s] := [r \odot s]$$
$$[r] < [s] \overset{def}{\iff} r \leq s$$

Theorem 2.9. $(\mathbb{R}^{\mathbb{N}}/\mathcal{F}, +, \cdot, <)$ *is an ordered field.*

Proof. Showing that $\mathbb{R}^{\mathbb{N}}/\mathcal{F}$ is a ring with zero $[\mathbf{0}]$ and unit $[\mathbf{1}]$ is an easy exercise. To define multiplicative inverses, take an element $[r] \neq [\mathbf{0}]$. This means that r is not zero almost everywhere modulo $\mathcal{F}$, and so that $[[r = \mathbf{0}]] \notin \mathcal{F}$. Being $\mathcal{F}$ an ultrafilter, this implies that its complement $[[r \neq \mathbf{0}]] \in \mathcal{F}$. Setting

$$s_n := \begin{cases} \frac{1}{r_n} & \text{if } n \in [[r \neq \mathbf{0}]] \\ 0 & \text{otherwise} \end{cases}$$

then defines a sequence s with the property that $[[r \odot s = \mathbf{1}]] = [[r \neq \mathbf{0}]] \in \mathcal{F}$, and hence $[r] \cdot [s] = [\mathbf{1}]$, meaning that $[s]$ is a multiplicative inverse of $[r]$. We now have to check that the order relation $<$ on $\mathbb{R}^{\mathbb{N}}/\mathcal{F}$ is total, meaning that for $[r], [s]$ it is always $[r] > [s]$, $[r] = [s]$ or $[r] < [s]$. Noting that $[[r > s]], [[r = s]], [[r < s]]$ are disjoint sets and that their union is the whole $\mathbb{N}$, the conclusion follows from a general property about ultrafilters (easy to prove!):

If $A_1, A_2, \dots$ are pairwise disjoint sets such that $\bigcup_i A_i \in \mathcal{F}$, then $A_i \in \mathcal{F}$ *for exactly one i.*

There are still some trifles to prove (like that the set of $[\mathbf{r}]$ such that $[\mathbf{0}] \leq [\mathbf{r}]$ is closed under addition), that are left to the reader. From now on, the ordered field $\mathbb{R}^{\mathbb{N}}/\mathcal{F}$ will be denoted with $^*\mathbb{R}$ and called the set of *Hyperreals* (equivalently, *Non-standard reals*). $\qquad \square$

We can send every real number $r \in \mathbb{R}$ to the equivalence class $[\mathbf{r}] \in {}^*\mathbb{R}$, where $\mathbf{r}$ is the constant sequence $(r, r, \dots)$. This correspondence is a ordered field homomorphism, and so we can regard $\mathbb{R}$ as a subfield of ${}^*\mathbb{R}$: This tells us that the latter is an extension of the former in the "intuitive" sense. The common notation to define an element of ${}^*\mathbb{R}$ that corresponds to some $r \in \mathbb{R}$ is ${}^*r$. That is, ${}^*r = [\mathbf{r}]$.

2.5   What about $\mathcal{F}$?

We built the hyperreal numbers as a quotient by some non principal ultrafilter $\mathcal{F}$. But which one? What if we have multiple choices of $\mathcal{F}$? There is a right one to pick?

The answer to this question is *it doesn't matter*: If we assume the *continuum hypothesis* it can be proved that all the quotients of $\mathbb{R}^{\mathbb{N}}$ by a non principal ultrafilter are isomorphic as ordered fields, and this is why we left the choice of $\mathcal{F}$ completely undetermined.

The continuum hypothesis is an important conjecture about cardinality of sets, and it has been proved by Paul Cohen that this conjecture is indecidable in the usual Zermelo-Fraenkel-Choice formalization of set theory. In particular, assuming it is true or not does not produce any contradiction, so it is up to us to decide if we want it or not in our axiomatization of set theory. Clearly we do, and in this way the problem of picking the right ultrafilter for our construction vanishes completely.

2.6   Infinities and infinitesimals

Let ϵ be the sequence $(1, \frac{1}{2}, \frac{1}{3}, \dots)$. Clearly $[[\mathbf{0} \leq \epsilon]] = \mathbb{N}$ and hence is in $\mathcal{F}$. On the other hand, for every *positive* $r \in \mathbb{R}$ it is $[[\mathbf{0} \leq \mathbf{r}]]$; moreover the set $[[\epsilon < \mathbf{r}]]$ is cofinite. We now use another general property of ultrafilters (again, easy to prove!):

> If $\mathcal{F}$ is non principal, it contains all the cofinite sets.

From this we infer that $[[\epsilon < \mathbf{r}]] \in \mathcal{F}$ for every positive real r, and hence we conclude that $[\epsilon]$ is a *positive infinitesimal*. We can generalize this to define *negative infinitesimals* in the obvious way, and we will refer to *infinitesimals* in general when we do not specify a sign. If we set $[\omega] = [\epsilon]^{-1} = (1, 2, 3, \dots)$ then for every positive $r \in \mathbb{R}$ the set $[[\mathbf{r} < \omega]]$ is again cofinite, and hence we can regard $[\omega]$ as a positive unlimited quantity. The existence of such $[\epsilon], [\omega]$ shows that ${}^*\mathbb{R}$ is a *proper extension* of $\mathbb{R}$, that is, it contains elements that cannot be put in correspondence with any real number.

Remark 2.10. Note how having such $[\epsilon], [\omega]$ crucially depends on the fact that $\mathcal{F}$ is non principal. Otherwise, we would have $\mathcal{F} := \{S \subseteq \mathbb{N} : n \in S\}$ for a fixed natural number n. But then any sequence $s = (s_1, s_2, \dots, \dots)$ would agree with $(s_n, \dots, s_n)$ on the n-th number and hence $[[s = \mathbf{s_n}]] \in \mathcal{F}$, meaning that $s = [\mathbf{s_n}] = {}^*s_n$. In this case, then, every element of $\mathbb{R}/\mathcal{F}$ comes from a real number and we do not obtain anything new from our construction.

If r is a sequence that converges to zero, using the strategy above we can easily check that $[r]$ and hence that it defines an infinitesimal on ${}^*\mathbb{R}$ (note that, however, it is not guaranteed that $[r] = [\epsilon]$: there are many different infinities and infinitesimals in ${}^*\mathbb{R}$!). Similarly, if $r \to \infty$ then $[r]$ is unlimited, so the notion of infinities and infinitesimals in our extension ${}^*\mathbb{R}$ captures the intuitive concepts of infinity and infinitesimal that we know from undergrad calculus.

Moreover, infinities (that from now on we will prefer to call "unlimited quantities") and infinitesimals are elements of ${}^*\mathbb{R}$, and hence we can multiply, sum them etc. Their arithmetic will be explored in detail later.

2.7   Extending sets, functions, relations

Remark 2.11. For all the definitions given below one has to check that they are well defined, in the sense that they do not depend on the choice of representatives in an equivalence class $[r] \in {}^*\mathbb{R}$. This is indeed the case and verifying it can be a useful exercise.

If A is a subset of $\mathbb{R}$, we can consider the set ${}^*A$ defined by

$$[r] \in {}^*A \overset{def}{\iff} \{n \in \mathbb{N} : r_n \in A\} \in \mathcal{F}$$

That is, $[r] \in {}^*A$ if $r \in A$ almost everywhere. We call the elements of ${}^*A - A$ *non-standard*. As an example, using the definition it is easy to see that $[\omega] \in {}^*\mathbb{N}$, but clearly $[\omega] \neq [\mathbf{n}]$ for any natural n. Hence $[\omega]$ is a non-standard natural number.

Similarly, we can extend functions. If $f : \mathbb{R} \to \mathbb{R}$ then ${}^*f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ is defined as

$$ {}^*f([(r_1, r_2, \dots)]) = [(f(r_1), f(r_2), \dots]$$

We can also extend functions $f : A \to \mathbb{R}$. To do this, for each $[r] \in {}^*A$ define

$$s_n := \begin{cases} f(r_n) \text{ if } r_n \in A \\ 0 \text{ otherwise} \end{cases}$$

Then we set ${}^*f([r]) = [s]$. We need to do this because $[r] \in {}^*A$ means that $r \in A$ *almost everywhere*, and hence there could be some r_n of the sequence r that are not in A, on which f is not well defined. With the definition above we formally set f to 0 on these members. Notice moreover that if r is a real number in A, then ${}^*f({}^*r) = {}^*(f(r))$: f and ${}^*f$ agree on the real numbers. For this reason it makes sense to drop the * symbol to denote the extension, and this is what we will do for most functions. As an instance, we will use x^2 both to denote the usual function $\mathbb{R} \to \mathbb{R}$ and its extension ${}^*\mathbb{R} \to {}^*\mathbb{R}$.

Note in particular that a sequence is a function $s : \mathbb{N} \to \mathbb{R}$, hence we can extend it to an *hypersequence* $s : {}^*\mathbb{N} \to {}^*\mathbb{R}$: We can now look which values it assumes at infinite since this corresponds to just unlimited numbers in ${}^*\mathbb{N}$. This is exactly what we will use to calculate limits.

Finally, we can extend k-ary relations: These are nothing but subsets of $\mathbb{R}^k$, and hence we can set

$$ {}^*P([r^1], [r^k]) \overset{def}{\iff} \left\{n \in \mathbb{N} : P(r_n^1, \dots r_n^k)\right\} \in \mathcal{F}$$

It is easy to see that the previous extensions are just particular cases of this one: A set $A \subseteq R$ is just an unary relation and every function $f : A \to R$ can be identified with the relation $\{(a, r) : f(a) = r\}$. You can check that extending sets and function seeing them as relations gives us the same results we got performing the extension directly.

3   The transfer theorem

The *transfer theorem* (also called *transfer principle*, according to your taste) is what tells us which properties get preserved from statements on standard sets, like $\mathbb{N}$, $\mathbb{R}$ etc. to their extensions ${}^*\mathbb{N}$, ${}^*\mathbb{R}$, ... Up to now we built our stuff "manually", having to verify every time some things: Namely, we said that some property held for a $[r]$ if it held almost everywhere on its components, that is, if the set of the r_n for which the property held in the "standard" universe was in the ultrafilter $\mathcal{F}$. Clearly this ain't an easy

way to do mathematics since at every step lots of conditions have to be checked and verified. The transfer theorem can be seen as a huge shortcut to avoid all this meticulous work.

To fully understand how the transfer theorem work, quite a bit of a background in logic is needed. In here, we will provide a practical recipe with examples and counterexample that should be more than sufficient to use it with confidence.

First of all we have to specify what do we mean when we say "statement on a set". Given a set S, we consider the triplet $(S, \mathfrak{R}_S, \mathfrak{F}_S)$, where $\mathfrak{R}_S$ and $\mathfrak{F}_S$ are the set of all the possible relations and functions on S, respectively. Note that since a k-ary relation on S is just a subset of S^k, so S^k is a relation itself $\mathfrak{R}_S$ includes also all the subsets of S (seen as unary relations) and all the finite products of subsets of S.

This triplets are called *relational structures*, and each of them comes with a *language*: This language is the set of all the well-formed formulas in which we allow the use of variables (ex. the x in $x \in \mathbb{N}$), constants, that are just elements of S (ex. the 1 in $1 \in \mathbb{N}$), functional symbols in $\mathfrak{F}_S$ (ex. the f in $f(x) \in \mathbb{N}$), relational symbols in $\mathfrak{R}_S$ (ex. the $\in$ in $x \in \mathbb{N}$) and the usual logical connectives $\wedge, \vee, \neg, \rightarrow, \leftrightarrow$, parentheses and ",". We moreover allow universal and existential quantification of variables as long as the quantifiers range on a relation in $\mathfrak{R}_S$. This means that we allow *only* quantification over elements of subsets of finite products of S, hence statements like $\forall K \in \mathcal{P}(S)$ are in general not allowed since $\mathcal{P}(S)$ cannot be expressed as a subset of S^k for some finite k if S is infinite.

A *statement on a set S* is a formula in the language of the relational structure on S in which every variable is bounded by a quantifier, that is, every variable can "range" over a defined set.

Example 3.1. Here some examples:

- $\forall K \subset \mathbb{N}, \emptyset \in K$ is not a formula in the language of $\mathbb{N}$, since quantification over K is not allowed;

- $x \in \mathbb{N}$ is a formula in the language of $\mathbb{N}$, but not a statement since x is not bounded;

- $1 \in \mathbb{N}$ is a statement in the language of $\mathbb{N}$ since there are no variables around, and hence everything is trivially bounded;

- $\forall x \in \mathbb{N}, x + 1 \in \mathbb{N}$ is a statement in the language of $\mathbb{N}$ since every variable is bound;

- $\forall n \in \mathbb{N}, \exists r \in \mathbb{R} : n < r$ is a statement in the language of $\mathbb{R}$: both variables are bounded, one on a subset of $\mathbb{R}$ and one on $\mathbb{R}$ itself;

- $\forall c \in \mathbb{C}, c + 1 \in \mathbb{R}$ is a statement (clearly false, but still a statement) in the language of $\mathbb{R}$: $\mathbb{C}$ can be seen as $\mathbb{R}^2$, and hence our quantification is legit!

3.1 *-transformations

We can transform a statement on S to a statement on $^*S$ using the tools developed in section 2.7 as follows:

- We replace every set A in the statement with its extension $^*A$;

- We replace every function f with its extension $^*f$;

- We replace every constant s with its interpretation in the extension $^*s$.

As an instance, the statement $\forall x \in \mathbb{N}, x + 1 \in \mathbb{N}$ becomes $\forall x(^*\in)^*\mathbb{N}, x(^*+)^*1(^*\in)^*\mathbb{N}$. For some special relation and function symbols though, such as "+", "=", "$\in$", "$\leq$" and the like, we avoid to write explicitly the * symbol to avoid clutter. In the previous example then we will usually write $\forall x \in {}^*\mathbb{N}, x + {}^*1 \in {}^*\mathbb{N}$. Note that this is not ambiguous: The $\in$ symbol, having $^*\mathbb{N}$ on the right, must necessarily be the extended version of $\in$, otherwise the formula would not make sense. Similarly, the $+$ is acting on elements $x, {}^*1$ that are both in $^*\mathbb{N}$ and hence must denote the extension of the function $+$.

Finally, the relation between a statement and its *-transformation is expressed by the transfer theorem:

Theorem 3.2 (Transfer theorem)**.** *A statement ϕ is true if and only if $^*\phi$ is true.*

This in particular means that:

Theorem 3.3 (Universal transfer)**.** *If a property ϕ holds for* all *real numbers, then $^*\phi$ holds for* all *hyperreal numbers.*

If you think about it, this corollary is obvious: A property that holds for all real numbers is a statement on the language of the reals, hence its extension must hold, by transfer theorem, on all the hyperreals, exactly as the corollary says. Similarly,

Theorem 3.4 (Existential transfer)**.** *If there is an hyperreal satisfying some property $^*\phi$, then there is some real satisfying ϕ.*

Example 3.5. The existential transfer theorem can be very useful: As an example, take Rolle's theorem from undergrad calculus:

If a function f is continuous and differentiable on $[a,b]$ and $f(a) = f(b)$, then $f'(x) = 0$ for some $x \in [a,b]$.

This can be expressed saying that if f is "such and such", then it exists a real number x such that "blablabla". The idea is that if we can express the "such and such" and the "blablabla" part of the statement (that is, the bits involving derivatives, continuity and the like) as statements on the reals (and we obviously can, as we will see later), then we can consider the following

If a function $^*f$ is *-continuous and *-differentiable on $^*([a,b])$ and $^*f(^*a) = {}^*f(^*b)$
then $^*f'(x) = {}^*0$ for some $x \in {}^*([a,b])$.

And the two statements will be completely entangled: if one is true so is the other. The difference is that in the second case we have much more choice for our x, since it can be an hyperreal. The idea is that even if we find an x that satisfies the condition turns out to be in $^*\mathbb{R} - {}^*\mathbb{R}$, the transfer theorem ensures there will be some $x' \in \mathbb{R}$ for which the first statement holds.

The core of non-standard analysis, then, is the following: Every time we have a statement on $^*S$ that is the *-transform of some statement ϕ on S, we can "drag" the result back to S. Clearly the statements on $^*S$ that are not in the form $^*\phi$ for some ϕ on S are *NOT* transferable, and this is exactly where one has to pay attention.

Example 3.6. Consider a sequence $s : \mathbb{N} \to \mathbb{R}$ and extend it to $^*s : {}^*\mathbb{N} \to {}^*\mathbb{R}$. Suppose we can show that the sequence $^*s$ never takes infinite values. We then want to prove that s is a bounded sequence. Since $^*s$ never takes infinite values the following statement must be true for some unlimited ω:

$$(\forall n \in {}^*\mathbb{N})(|{}^*s(n)| \le \omega)$$

Since ω is unlimited, though, it must be $\omega \in {}^*\mathbb{R} - \mathbb{R}$ and hence the statement above is *not* in a transferrable form, because ω cannot be expressed as $^*r$ for some real r. We then try to massage the statement in a transferrable form: In fact, since $\omega \in {}^*\mathbb{R}$, the formula above implies that the following is true:

$$(\exists r \in {}^*\mathbb{R})(\forall n \in {}^*\mathbb{N})(|{}^*s(n)| \le r)$$

but now we can transfer this back to our "standard" world! In fact, from here we can apply the transfer theorem and deduce that the following must be true too:

$$(\exists r \in \mathbb{R})(\forall n \in \mathbb{N})(|s(n)| \le r)$$

That obviously means that s is bounded on the reals.

Remark 3.7 (A bit of jargon). Consider now the triplet $(^*S, \mathfrak{R}_{*S}, \mathfrak{F}_{*S})$, that is, the relationa structure on $^*S$. Relations in $\mathfrak{R}_{*S}$ that are of the form $^*R$ for some R in $\mathfrak{R}_S$ are called *internal relations*. Similarly we get a definition of *internal function*. What the transfer theorem says is that a statement can be transferred from $^*S$ to S if and only if it comprises only internal relations and functions.

<div align="center">

- Mantra of non-standard analysis -

</div>

Pay attention! Transfer theorem from S to $^*S$ is always ok. When you go in the other direction make sure, before applying tit, that 100% of your stuff is in transferrable form, that is, it is made of *statements* (and not just formulas) in which only internal relations and functions are present.

4   Examples

5   Conlusion

"I know kung fu."