NUI Galway
OÉ Gaillimh

Infinitesimals, Nonstandard Analysis and Applications to Finance

Author:
Eoghan STAUNTON

ID Number:
09370803

FINAL YEAR PROJECT

NATIONAL UNIVERSITY OF IRELAND, GALWAY

Supervisor: Dr. Ray RYAN

February 8, 2013

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of degree is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Author:_____ **ID No:**_____

Eoghan Staunton   *09370803*

Contents

1 Introduction

"There are good reasons to believe that non-standard analysis, in some version or other, will be the analysis of the future" [33].
Kurt Gödel, 1974.

An infinitesimal is a number that is smaller in magnitude than every positive real number. The word infinitesimal comes from the Latin word *infinitesimus* and was coined by the German mathematician Gottfried Wilhelm Leibniz around 1710 [1]. We learn early on in our study of standard analysis that nonzero infinitesimals cannot exist. It is also true however that many people use the intuitive notion when trying to understand basic concepts in analysis and calculus such as derivatives or integrals. For example a student may think of the derivative of a function f at a point x as the slope of the secant line between the point $(x, f(x))$ and a point an infinitesimal distance away. Informal notions of infinitesimals have been used throughout history, indeed the concept was used by Isaac Newton and Leibniz in their formulation of calculus [9].

The use of infinitesimals in the formative years of calculus however lacked rigour and their use was criticised by the Irish philosopher George Berkeley among others [44]. Efforts were made to come up with a system that would admit the existence of infinitesimals in a consistent manner. Little progress was made however, and most efforts were abandoned when in the 1870's Karl Weierstrass came up with the formal 'epsilon-delta' theory of limits which became the rigorous foundation needed for calculus. Modern mathematicians however again made efforts in the 20th century to formalise a theory of nonstandard numbers and in 1961 Abraham Robinson succeeded in producing a consistent nonstandard analysis.

I was introduced to the area by my supervisor who directed me to read a piece by 2006 Fields Medal winner Terence Tao on the subject in his book *Structure and Randomness: Pages from Year One of a Mathematical Blog* [39]. In the piece Tao gives a nice introduction to the area explaining many of the ideas in a nice intuitive way which piqued my interest in the subject. In his book Tao attributes the reluctance of many mathematicians to use non-standard methods to the tendency to *"gloss over the actual* construction *of non-standard number systems"*. Perhaps this is one of the reasons why although nonstandard analysis may still be the *"analysis of the future"*, as predicted by Gödel, in mainstream mathematics, and certainly in undergraduate mathematics, it has yet to become the analysis of the present.

The main aim of my project therefore is to give a clear introduction to the construction of the hyperreal numbers and the transfer principle of nonstandard analysis, suitable for any undergraduate mathematics student without any background in the area. A simple introduction to nonstandard analysis is given by Jerome Keisler in his book *Elementary Calculus: An Infinitesimal Approach* [23]. His construction of the hyperreals is based on introducing infinitesimals in an axiomatic way. My introduction will be based on the so-called ultraproduct construction of nonstandard analysis and uses some of the ideas of Tao and Jaap Ponstein [39], [30]. I will give an overview of the fascinating history of infinitesimals. I will also present some of the interesting applications of nonstandard analysis with a focus

on the applications to economics and finance.

2 Construction of the Hyperreals

2.1 Our aim

In the first article I read on the subject of non-standard analysis Tao attributes the reluctance of many mathematicians to use non-standard methods to the tendency to *"gloss over the actual* construction *of non-standard number systems"*. This causes the transfer principle and the construction behind it to be viewed as *"some sort of* "black box" *which mysteriously bestows some certificate of rigour on non-standard arguments"* [39]. In this section I will attempt to clearly explain just how the hyperreals are constructed and to begin to demystify this *"black box"*.

First we must be clear on exactly what we are attempting to do. We wish to introduce non-zero infinitesimals to our set of real numbers. It is clear that 0 is the only element of the real numbers that is an infinitesimal. Recall, an infinitesimal is a number that is smaller in magnitude than any non-zero real number and so no non-zero real number ε is an infinitesimal since $|\frac{\varepsilon}{2}| < |\varepsilon|$ for every $\varepsilon \neq 0$. Our aim therefore must be to introduce non-zero infinitesimals to the set of real numbers and come up with an extension of the real numbers. We also wish to treat these infinitesimals as we would classical numbers and so each non-zero infinitesimal δ can be inverted to give $\gamma = \frac{1}{\delta}$. Now $\forall n \in \mathbb{N} : |\gamma| > n$, we will call such numbers hyperlarge. These hyperlarge numbers are also clearly greater than any real number x since $\forall x \in \mathbb{R} : \exists n \in \mathbb{N} : n > x$.

Our ultimate goal therefore is to introduce these numbers to the reals and to come up with a consistent system of hyperreals $^*\mathbb{R}$. So how will we do this? To motivate our method of constructing the hyperreals we'll first look at how we can introduce $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ from our starting point of the natural numbers $\mathbb{N}$.

2.2 $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ from $\mathbb{N}$

Just as when we are attempting to construct the system of hyperreals we are trying to construct is an extension of the set of real numbers, the set of integers is an extension of the set of natural numbers. The set of rationals is in turn an extension of the integers and the set of reals is an extension of the rationals. The elements of each of these sets can be generated by the elements of the set we are trying to extend, for example the elements of $\mathbb{Z}$ can be generated by elements of $\mathbb{N}$. Leopold Kronecker is famously quoted as saying *"God created the natural numbers; all else is the work of man"* [42]. When we are generating these new elements we must always follow two rules:

1. We must define the equality of elements by using an equivalence relation.

2. If the element we have generated is not new we must identify it explicitly with the element we already knew.

2.2.1　The Integers

We can generate each integer using an ordered pair of natural numbers. If (x, y) is a pair of natural numbers then let $Z(x, y)$ be the integer generated by it. So the set of all integers $\mathbb{Z} = \{Z(x, y) : x \in \mathbb{N}, y \in \mathbb{N}\}$.

We now must define when two integers $Z(x, y)$ and $Z(w, z)$ are equal. To do this we first define an equivalence relation $\sim_z$ on pairs of natural numbers:

$$(x, y) \sim_z (w, z) \Leftrightarrow x + z = y + w$$

and then set $Z(x, y) = Z(w, z)$ precisely when $(x, y) \sim_z (w, z)$.

Finally we must identify when a pair of natural numbers generates one of our original set of natural numbers. We use the following rule:

$$\forall n \in \mathbb{N} : Z(n, 0) = n.$$

2.2.2　The Rationals

We can generate each rational number by using an ordered pair of integers. If (p, q) is a pair of integers with $q \neq 0$ then let $Q(p, q)$ be the rational generated by it. So the set of all rationals $\mathbb{Q} = \{Q(p, q) : p \in \mathbb{Z}, q \in \mathbb{Z}, q \neq 0\}$.

We now must define when two rationals $Q(p, q)$ and $Q(r, s)$ are equal. To do this we first define an equivalence relation $\sim_q$ on pairs of integers:

$$(p, q) \sim_q (r, s) \Leftrightarrow ps = qr$$

and then set $Q(p, q) = Q(r, s)$ precisely when $(p, q) \sim_q (r, s)$.

Finally we must identify when a pair of integers generates one of our original set of integers. We use the following rule:

$$\forall x \in \mathbb{Z} : Q(z, 1) = z.$$

2.2.3　The Reals

We can generate each real number by using a Cauchy sequence of rationals. If $(q_1, q_2, \ldots)$ is a Cauchy sequence of rationals then let $R(q_1, q_2, \ldots)$ be the real generated by it. So the set of all reals $\mathbb{R} = \{R(q_1, q_2, \ldots) : \forall i : q_i \in \mathbb{Q}\}$.

As in the two cases above we now must define when two reals $R(q_1, q_2, \ldots)$ and $R(r_1, r_2, \ldots)$ are equal. To do this we first define an equivalence relation $\sim_r$ on Cauchy sequences of rationals:

$$(q_1, q_2, \ldots) \sim_r (r_1, r_2, \ldots) \Leftrightarrow \forall m \in \mathbb{N} : \exists k \in \mathbb{N} : \forall n \in \mathbb{N}, n > k : |q_n - r_n| < \frac{1}{m}$$

and then set $R(q_1, q_2, \ldots) = R(r_1, r_2, \ldots)$ precisely when $(q_1, q_2, \ldots) \sim_r (r_1, r_2, \ldots)$.

Finally, we must identify when a Cauchy sequence of rationals generates one of our original set of rationals. We use the following rule:

$$\forall q \in \mathbb{Q} : R(q, q \ldots) = q.$$

We have seen that we can extend the natural numbers $\mathbb{N}$ to the integers $\mathbb{Z}$ by using two natural numbers to generate each integer. The integers $\mathbb{Z}$ can again be

extended to the rationals $\mathbb{Q}$ by using two integers to generate each rational. Finally, the rationals $\mathbb{Q}$ can be extended to the reals $\mathbb{R}$ by using infinite Cauchy sequences of rationals to generate each real number. Now we wish to extend the reals $\mathbb{R}$ to the hyperreals *$\mathbb{R}$ by using elements of the reals to generate each hyperreal. If man can construct $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ by simply using the natural numbers given to us by God why not go one step further and construct *$\mathbb{R}$?

2.3 Free Ultrafilters

To extend the real numbers to the hyperreals *$\mathbb{R}$ we are going to use infinite sequences of real numbers to generate each hyperreal number. To do this we also need to have rules for equality and identification, just as we had for $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ so we can come up with a mathematically consistent and sensible system of hyperreals. To help us do this we are going to use a *free utrafilter*.

Ultrafilters were not originally introduced for the purpose of constructing the hyperreal numbers and have other applications outside non-standard analysis. The notion of an ultrafilter was first introduced by the French mathematician Henry Cartan in two short notes in The Proceedings of The Academy of Sciences Paris in 1937, for use in the area of general topology [12]. Ultrafilters are particularly important when dealing with Hausdorff spaces. They are used in the construction of the Stone-Čech and Wallman Compatifications.

First we will explore the more general idea of a filter.

Definition A non-empty collection F of subsets of $\mathbb{N}$ is called a filter (over $\mathbb{N}$) if:

- $\varnothing \notin F$

- if $A \in F$ and $\mathbb{N} \supseteq B \supseteq A$ then $B \in U$,

- if $A \in U$ and $B \in U$, then $A \cap B \in U$,

A filter U is called an ultrafilter if for any $A \subseteq \mathbb{N}$, either A or A^c is an element of U, but not both. (Here A^c is the complement of A. $A^c = \mathbb{N} \setminus A$.)

A filter is called free (or non-principal) if all of its elements are infinite sets. Combining our above definition of a filter and the two conditions above we come up with the following definition of a free ultrafilter:

Definition A non-empty collection U of subsets of $\mathbb{N}$ is called a free ultrafilter (over $\mathbb{N}$) if:

- if $A \in U$ and $\mathbb{N} \supseteq B \supseteq A$ then $B \in U$,

- if $A \in U$ and $B \in U$, then $A \cap B \in U$,

- if $A \in U$, then A is infinite, and,

- if $A \subseteq \mathbb{N}$, then either $A \in U$ or $A^c \in U$, but not both.

Now we have given the definition of a free ultrafilter we must ask ourselves if such a filter exists. The answer is that free ultrafilters over any infinite set do exist. A proof of their existence, relying on the axiom of choice, was given by

Tarski in 1930. In our proof below we will invoke the axiom of choice through the use of Zorn's Lemma. This reliance on the axiom of choice leads to the area of non-standard analysis being criticised by constructivist mathematicians due to the extremely nonconstructive nature of the axiom. We will discuss this criticism further in the final section of this paper.

Theorem 2.3.1. *Free ultrafilters over $\mathbb{N}$ exist.*

Proof. Let F_0 be the filter consisting of all cofinite subsets of $\mathbb{N}$, sometimes known as the cofinite or Fréchet filter, ($Q \subseteq \mathbb{N}$ is cofinite if and only if Q^c is finite). Let E be the set of all filters F s.t. $F \supseteq F_0$. E is nonempty (since $F_0 \in E$) and can be partially ordered by set inclusion.

Let G be any totally ordered subset of E. Then $B = \bigcup\{F : F \in G\} \in E$ and B is an upper bound for G. $B \supseteq F_0$ since $\forall F \in G : F \supseteq F_0$. B is also a filter:

1. $\varnothing \notin B$ since B is the union of filters.

2. If $Q \in B$ and $\mathbb{N} \supseteq R \supseteq Q$ then $Q \in F$ for some $F \in G$. Since F is a filter $R \in F$ also $\Rightarrow R \in B$.

3. Let $Q, R \in B$, then $Q \in F_1$ and $R \in F_2$ for some $F_1, F_2 \in G$. G is totally ordered so $F_1 \subseteq F_2$ or $F_2 \subseteq F_1$. Suppose $F_1 \subseteq F_2$, then $Q, R \in F_2$. But F_2 is a filter so $Q \cap R \in F_2 \subseteq B$.

Now Zorn's lemma tells us that E contains a maximal element U. We wish to show that U is a free ultrafilter. $U \supseteq F_0$ and so U is a free filter. We must now show that U is also an ultrafilter.

Let $Q' \subseteq \mathbb{N}$ and consider the following cases:

Case 1: Suppose $\forall Q \in U : Q' \cap Q$ is infinite. Let,

$$V = \{T : \mathbb{N} \supseteq T \supseteq Q' \cap Q : Q \in U\}.$$

where Q is some arbitrary element of U. Then V is a filter. Also $V \in E$, we can see this by taking $Q = \{n, n+1, \ldots\}$[1]. We also have that $U \subseteq V$ but since $V \in E$ and U is the maximal element of E $U \supseteq V$. So $V = U$ and since $\mathbb{N} \supseteq Q' \supseteq Q' \cap Q$ we have that $Q' \in V \Rightarrow Q' \in U$.

Case 2: Suppose on the other hand that $\exists Q_0 \in U : Q' \cap Q_0$ is finite then $Q' \notin U$. Now take any $Q \in U$. $Q_0 \in U$ also $\Rightarrow Q_0 \cap Q \in U$ and $Q_0 \cap Q$ is infinite. Since $Q' \cap Q_0 \cap Q$ is finite $(Q_0 \cap Q) \setminus (Q' \cap Q_0 \cap Q) = Q'^c \cap Q_0 \cap Q$ is infinite. So $Q'^c \cap Q$ must be infinite. Applying Case 1, replacing Q' with Q'^c, gives $Q'^c \in U$. □

Theorem 2.3.2. *Suppose that $\mathbb{N} = A_1 \cup A_2 \cup \ldots \cup A_n$ for mutually disjoint A_i and $n \in \mathbb{N}$. Then $A_i \in U$ for exactly one i.*

Proof. Let $B_i = A_i^c$. Suppose that there is no $i \in \{1, 2, \ldots, n\}$ such that $A_i \in U$. Then $\forall i \in \{1, 2, \ldots, n\} : B_i \in U$. So we now have that $\varnothing = B_1 \cap B_2 \cap \ldots \cap B_{n-1} \cap B_n \in U$ which is a contradiction. So $\exists i \in \{1, 2, \ldots, n\} : A_i \in U$. But we also have that the A_i are mutually disjoint and so there can only be one i such that $A_i \in U$ since if $A_i \in U$ and $A_j \in U$ for $i \neq j$ we would have that $\varnothing = A_i \cap A_j \in U$. □

---

[1] $U \in E \Rightarrow U \supseteq F_0$. Now $Q \in F \Leftrightarrow \exists n \in \mathbb{N} : \{n, n+1, \ldots\} \subseteq Q$ (since Q must be cofinite). So $Q' \cap Q$ is infinite iff $\exists m \in \mathbb{N} : \{m, m+1, \ldots\} \subseteq Q' \cap Q$.

2.4 Generating elements of $^*\mathbb{R}$

We now fix an ultrafilter U on $\mathbb{N}$. We will use infinite sequences of real numbers, in conjunction with U, to generate the hyperreals. Just as in the case of $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ we will need to establish rules for the equality and identification of these infinite sequences of real numbers. This is where we will use our free ultrafilter.

A nice way to think about how this works is to think of the infinite sequence real numbers as the votes of an infinite electorate. The ultrafilter helps us decide which votes matter when deciding the winner of the election. The fact that the ultrafilter is free means that if one number gets a cofinite number of "votes" it will always win against a number that gets only a finite number of "votes".

2.4.1 Rules for Equality and Identification

We can generate each hyperreal number by using an infinite sequence of real numbers. If $(x_1, x_2, \ldots)$ is an infinite sequence of real numbers let $H(x_1, x_2, \ldots)$ be the hyperreal number generated by it. So the set of all hyperreals $^*\mathbb{R} = \{H(x_1, x_2, \ldots) : \forall i : x_i \in \mathbb{R}\}$.

- **Equality**

  Again we must define when two hyperreals $H(x_1, x_2, \ldots)$ and $H(y_1, y_2, \ldots)$ are equal. To do this we first define an equivalence relation $\sim_u$ on infinite sequences of reals:

  $$(x_1, x_2, \ldots) \sim_u (y_1, y_2, \ldots) \Leftrightarrow \{i : x_i = y_i\} \in U$$

  and then set $H(x_1, x_2, \ldots) = H(y_1, y_2, \ldots)$ precisely when $(x_1, x_2, \ldots) \sim_u (y_1, y_2, \ldots)$.

  Theorem 2.4.1. $\sim_u$ *is an equivalence relation.*

  Proof. We must show that $\sim_u$ is reflexive, symmetric and transitive.

  1. $\sim_u$ is reflexive. $\forall (x_1, x_2, \ldots) : (x_1, x_2, \ldots) \sim_u (x_1, x_2, \ldots)$ since $\{i : x_i = x_i\} = \mathbb{N} \in U$ for every free ultrafilter U.
  2. $\sim_u$ is symmetric. $(x_1, x_2, \ldots) \sim_u (y_1, y_2, \ldots) \Leftrightarrow \{i : x_i = y_i\} = \{i : y_i = x_i\} \in U \Leftrightarrow (y_1, y_2, \ldots) \sim_u (x_1, x_2, \ldots)$
  3. $\sim_u$ is transitive. $(x_1, x_2, \ldots) \sim_u (y_1, y_2, \ldots)$ and $(y_1, y_2, \ldots) \sim_u (z_1, z_2, \ldots) \Leftrightarrow A = \{i : x_i = y_i\} \in U$ and $B = \{i : y_i = z_i\} \in U$. Since U is a free ultrafilter we have that if $A \in U$ and $B \in U$, then $A \cap B = \{i : x_i = z_i\} \in U$. $\{i : x_i = z_i\} \in U \Leftrightarrow (x_1, x_2, \ldots) \sim_u (z_1, z_2, \ldots)$

  So $\sim_u$ satisfies all three properties and is therefore an equivalence relation.

  $\square$

- **Identification**

  We now must identify when an infinite sequence of reals generates one of our original sets of reals. We use the following rule:

  $$\forall x \in \mathbb{R} : H(x_1, x_2, \ldots) = x \Leftrightarrow \{i : x_i = x\} \in U.$$

2.4.2 Some Examples

1. $H(\sqrt{2}, \sqrt{2}, \sqrt{2}, \ldots) = \sqrt{2}$.
   This is true since $\{i : x_i = x\} = \mathbb{N}$ and $\mathbb{N} \in U$ for all free ultrafilters U.

2. What about the sequence $(\sqrt{2}, \pi, \sqrt{2}, \pi, \sqrt{2}, \ldots)$?
   What number does it generate? Is it possible that our definition could lead us to the obvious contradiction that $H(\sqrt{2}, \pi, \sqrt{2}, \pi, \sqrt{2}, \ldots) = \sqrt{2}$ and $H(\sqrt{2}, \pi, \sqrt{2}, \pi, \sqrt{2}, \ldots) = \pi$?

   To answer this we must recall the definition of an ultrafilter given above. If U is an ultrafilter and $A \subseteq \mathbb{N}$, then either $A \in U$ or $A^c \in U$, but not both. So either the set of odd numbers $\{1, 3, 5, \ldots\}$ or the set of even numbers $\{2, 4, 6, \ldots\}$ is in our free ultrafilter U but not both. In this case if $\{1, 3, 5, \ldots\} \in U$ (we can think of this as the case where only the odd voters votes are taken into consideration) then $H(\sqrt{2}, \pi, \sqrt{2}, \pi, \sqrt{2}, \ldots) = \sqrt{2}$ but if $\{2, 4, 6, \ldots\} \in U$ (we can think of this as the case where only the even voters votes are taken into consideration) then $H(\sqrt{2}, \pi, \sqrt{2}, \pi, \sqrt{2}, \ldots) = \pi$.

3. What number does the sequence $(1, \frac{1}{2}, \frac{1}{3}, \ldots)$ generate?
   Recall that if U is a free ultrafilter and $A \in U$, then A is infinite. Since no real number appears more than once the set $A_x = \{i : x_i = x\}$ is finite $\forall x \in \mathbb{R}$. So $\forall x : A_x \notin U \Rightarrow H(1, \frac{1}{2}, \frac{1}{3}, \ldots)$ is an entirely new number that is not equal to any of our classical real numbers. Later we will show that $H(1, \frac{1}{2}, \frac{1}{3}, \ldots)$ is a non-zero infinitesimal.

   Similarly the sequence $(1, 2, 3, \ldots)$ does not generate any of our classical real numbers. Later we will show that $H(1, 2, 3, \ldots)$ is a positive hyperlarge.

2.5 Arithmetic operations and inequalities in $^*\mathbb{R}$

Now that we have constructed the hyperreals we must be able to carry out simple operations and define when we consider when one hyperreal number to be larger than another. For example for two hyperreal numbes $x = H(x_1, x_2, \ldots)$ and $y = H(y_1, y_2, \ldots)$ how do we define $x + y$ or x/y? When can we say that $x < y$?

The definitions and rules we use follow very intuitively from our construction of the hyperreals. If & is an operation such as addition, subtraction, taking the absolute value, multiplication or division we introduce our version of & for the hyperreal numbers $^*\& = H(\&_1, \&_2, \ldots)$ by simply taking $\&_i = \&$ for all i. So for example $H(x_1, x_2, \ldots) \;^* + H(y_1, y_2, \ldots) = H(x_1 + y_1, x_2 + y_2, \ldots)$ for any $H(x_i)$ and $H(y_i) \in \mathbb{R}$. Since the context makes it clear whether we mean the classical version of the operation or our version for the hyperreals we will drop the $*$. This is a list of simple definitions and operations in $^*\mathbb{R}$:

(i) **Definition of $^*\mathbb{R}$**
   $^*\mathbb{R} = \{H(x_1, x_2, \ldots) : \forall i : x_i \in \mathbb{R}\}$.

(ii) **Equality**
   $H(x_1, x_2, \ldots) = H(y_1, y_2, \ldots)$ precisely when $(x_1, x_2, \ldots) \sim_u (y_1, y_2, \ldots)$.
   (Recall that: $(x_1, x_2, \ldots) \sim_u (y_1, y_2, \ldots) \Leftrightarrow \{i : x_i = y_i\} \in U$)

(iii) **Identification**
$\forall x \in \mathbb{R} : H(x_1, x_2, \ldots) = x \Leftrightarrow \{i : x_i = x\} \in U.$

(iv) **Addition**
$H(x_1, x_2, \ldots) + H(y_1, y_2, \ldots) = H(x_1 + y_1, x_2 + y_2, \ldots),\ x_i, y_i \in \mathbb{R}.$

(v) **Subtraction**
$H(x_1, x_2, \ldots) - H(y_1, y_2, \ldots) = H(x_1 - y_1, x_2 - y_2, \ldots),\ x_i, y_i \in \mathbb{R}.$

(vi) **Multiplication**
$H(x_1, x_2, \ldots) \times H(y_1, y_2, \ldots) = H(x_1 \times y_1, x_2 \times y_2, \ldots),\ x_i, y_i \in \mathbb{R}.$

(vii) **Absolute Value**
$|H(x_1, x_2, \ldots)| = H(|x_1|, |x_2|, \ldots),\ x_i \in \mathbb{R}.$

(viii) **Division**
When we are defining division we use the same method as usual however we must be careful since just like in the case of classical mathematics our divisor cannot be zero. We can deal with this extra condition easily. For $H(x_1, x_2, \ldots) \neq 0$ we define:
$1/H(x_1, x_2, \ldots) = H(r_1, r_2, \ldots),\ x_i \in \mathbb{R}$
with $r_i = 1/x_i$ if $x \neq 0$ and r_i arbitrary if $x_i = 0$. Note that our arbitrary choice of r_i when $x_i = 0$ has no effect on the value of $1/H(x_1, x_2, \ldots)$ since we can see from part (iii) that $H(x_1, x_2, \ldots) \neq 0 \Leftrightarrow \{i : x_i \neq 0\} \in U.$

(viii) **Inequalities**
$H(x_1, x_2, \ldots) < H(y_1, y_2, \ldots) \Leftrightarrow \{i : x_i < y_i\} \in U,\ x_i, y_i \in \mathbb{R}.$
(We have similar definitions for $\leq, >, \geq$.)

Theorem 2.5.1. *Let $x, y \in {}^*\mathbb{R}$ then exactly one of $x < y$, $x = y$ and $x > y$ is true.*

Proof. Let $x = H(x_1, x_2, \ldots)$ and $y = H(y_1, y_2, \ldots)$. Then $A_1 = \{i \in \mathbb{N} : x_i < y_i\}$, $A_2 = \{i \in \mathbb{N} : x_i = y_i\}$ and $A_1 = \{i \in \mathbb{N} : x_i > y_i\}$ are mutually disjoint sets with $A_1 \cup A_2 \cup A_3 = \mathbb{N}$. Now by Theorem 2.3.2 exactly one of A_1, A_2 and A_3 is in U. $\qquad\square$

2.5.1 Some examples

1. From (iii) above we have that $H(1, 1, \ldots) = 1$ and $H(2, 2, \ldots) = 2$ and so $H(1, 1, \ldots) + H(2, 2, \ldots) = 1 + 2 = 3$.
   This is consistent with the definition of addition for ${}^*\mathbb{R}$ given in (iv):
   $H(1, 1, \ldots) + H(2, 2, \ldots) = H(1 + 2, 1 + 2, \ldots) = H(3, 3, \ldots) = 3$

2. From (ii) and (iii) above we have that $H(0, 2, 2, \ldots) = H(2, 2, 2, \ldots) = 2$ and so $1/H(0, 2, 2, \ldots) = 1/2$.
   This is consistent with our definition of division for ${}^*\mathbb{R}$ given in (viii):
   $1/H(0, 2, 2, \ldots) = H(r, 1/2, 1/2, 1/2, \ldots)$, for arbitrary $r \in \mathbb{R}$. Now from (ii) and (iii) we have that:
   $H(r, 1/2, 1/2, 1/2, \ldots) = H(1/2, 1/2, 1/2, \ldots) = 1/2.$

3. From (iii) above we have that $H(1, 1, \ldots) = 1$ and $H(2, 2, \ldots) = 2$. We have that $1 < 2$ and so $H(1, 1, \ldots) < H(2, 2, \ldots)$.

   This is consistent with the definition of $<$ for $^*\mathbb{R}$ given in (ix):

   $H(1, 1, \ldots) < H(2, 2, \ldots) \Leftrightarrow \{i : 1 < 2\} = \mathbb{N} \in U$. But from the definition of a free ultrafilter $\mathbb{N} \in U$ for all ultrafilters U. So $H(1, 1, \ldots) < H(2, 2, \ldots)$.

2.6 Some Notation & Definitions

Before we go much further we will quickly introduce some notation that we will use to deal with the new elements we can now introduce.

Hyperlarge Numbers

Let $x \in {}^*\mathbb{R}$ be a positive hyperlarge i.e. $\forall n \in \mathbb{N} : x > n$ then we write $x \sim \infty$. For $x \in {}^*\mathbb{R}$, x a negative hyperlarge i.e. $\forall n \in \mathbb{N} : x < -n$ we write $x \sim -\infty$.

Infinitesimals

Let $\delta \in {}^*\mathbb{R}$ be an infinitesimal i.e. $\forall n \in \mathbb{N} : x < 1/n$ then we write $\delta \simeq 0$.

If δ is non-zero we write $\delta \sim 0$.

Limited Numbers

Let $x \in {}^*\mathbb{R}$ be a number that is not hyperlarge. Then we call x a limited number.

Appreciable Numbers

Let $x \in {}^*\mathbb{R}$ be a limited number that is not an infinitesimal. Then we call x an appreciable number.

Standard Part of a Limited Number

Let $x \in {}^*\mathbb{R}$ be a limited number, then the unique real number r that is infinitesimally close to x is called the standard part of x and we write $st(x) = r$.

2.7 Other Ultrapower Constructions

We have shown that it is possible to use an ultrafilter to generate hyperreal numbers, but is there anything special about real numbers or can we generate new hyperconstants from other types of mathematical constants such as functions, sequences, sets and n-tuples in a similar manner? For example, is it possible to generate hyper*sets* or hyper*functions*? Unsurprisingly perhaps, the answer is yes. Although our focus in this paper will be on hyperreals we will briefly deal with hyperfunctions and hypersets. Our first task is to give a generalised definition of when two hyperconstants are equal, a definition that will hold for any type of mathematical constant. These constants could be sets, functions, n-tuples or sequences for example.

2.7.1 Equality

We have already used the equivalence relation $\sim_u$ to define when two hyperreals are equal. Recall that for two infinite sequences, this relation is given by:

$$(x_1, x_2, \ldots) \sim_u (y_1, y_2, \ldots) \Leftrightarrow \{i : x_i = y_i\} \in U.$$

Now, to remain consistent, and to allow us to have a generalised definition of when any two hyperconstants are equal we shall use it again, setting
$H(x_1, x_2, \ldots) = H(y_1, y_2, \ldots)$ precisely when $(x_1, x_2, \ldots) \sim_u (y_1, y_2, \ldots)$.

2.7.2 Identification of Hypersets

We now wish to give a definition of a hyperset that will be consistent with this definition of equality. To motivate this we consider the following theorem.

Theorem 2.7.1. *Let $(X_1, X_2, \ldots)$ and $(Y_1, Y_2, \ldots)$ be infinite sequences of sets. Then*

$$\{H(x_i) : x_i \in X_i\} = \{H(y_i) : y_i \in Y_i\} \Leftrightarrow H(X_1, X_2, \ldots) = H(Y_1, Y_2, \ldots).$$

Proof. Suppose $H(X_1, X_2, \ldots) = H(Y_1, Y_2, \ldots)$ and let $Q = \{i : X_i = Y_i\}$ then $Q \in U$. Let $X = \{H(x_i) : x_i \in X_i\}$ and let $Y = \{H(y_i) : y_i \in Y_i\}$. Now for any $H(x_i) \in X$ for all i we have that $x_i \in X_i$ and for $i \in Q$, $x_i \in Y_i$. Now by the definition of equality we have that $H(x_i) \in Y$. So $X \subseteq Y$, similarly we have that $Y \subseteq X \Rightarrow X = Y$.

Conversely suppose $H(X_1, X_2, \ldots) \neq H(Y_1, Y_2, \ldots)$, then $Q \notin U$ and so by the definition of a free ultrafilter $Q^c = \{i : X_i \neq Y_i\} \in U$. Now we have that either $\{i : x_i \notin Y_i \text{ for some } x_i \in X_i\} \in U$ or, $\{i : y_i \notin X_i \text{ for some } y_i \in Y_i\} \in U$, or both.

Suppose $R = \{i : x_i \notin Y_i \text{ for some } x_i \in X_i\} \in U$. If $i \in R$ take $x_i \in X_i$ s.t $x_i \notin Y_i$, and if $i \notin R$ take any $x_i \in X_i$. Now $\{i : x_i \notin Y_i\} \supseteq R$ so $\{i : x_i \notin Y_i\} \in U$. Suppose now that $H(x_i) \in Y$, then $\{i : x_i \in Y_i\} \in U$. But since U is a free ultrafilter we have that it is closed under intersection and so:

$$\{i : x_i \notin Y_i\} \cap \{i : x_i \in Y_i\} = \varnothing \in U.$$

which is a contradiction. So $H(x_i) \notin Y$ and $X \neq Y$. The proof in the other case is similar. □

Our definition for identification of hypersets follows intuitively from the result of the previous theorem. We use the following definition: If all S_i are sets then

$$H(S_1, S_2, \ldots) = \{H(s_i) : s_i \in S_i\}.$$

2.7.3 Identification of Hyperfunctions

Again when giving the definition for a hyperfunction we wish for it to be consistent with the definition of equality that we have given above. Although we have omitted the proof, the following definition of a hyperfunction is consistent with that definition of equality.
Given an infinite sequence of functions $(f_1, f_2, \ldots)$ with $f_i \colon X \to Y_i$, we let $H(f_1, f_2, \ldots) \colon H(X_1, X_2, \ldots) \to H(Y_1, Y_2, \ldots)$ be the function defined by

$$H(f_i)(H(x_i)) = H(f_i(x_i)).$$

Our focus will be on hyperfunctions of the form $^*f = H(f, f, f, \ldots)$ where $f \colon \mathbb{R} \to \mathbb{R}$. This is known as the $*$-transform of f. We note that since $^*f \colon {}^*\mathbb{R} \to {}^*\mathbb{R}$ we have that $^*f \neq f$. However, $^*f$ is an extension of f. We should also note that since a sequence (x_n) of real numbers is simply a special type of function $x \colon \mathbb{N} \to \mathbb{R}$ where $x(n) = x_n$ we use the same rules as we use for functions to generate hypersequences. Again we will be most interested in the $*$-transform

of a sequence which is a function $^*x\colon {}^*\mathbb{N} \to {}^*\mathbb{R}$ which we will denote $^*(x_n)$. When referring to the mth term of the $*$-transform of a sequence (x_n) we will simply write x_m. Our meaning will be clear from the context but in any case of possible ambiguity a note will be included for clarity.

2.8 The $*$-transform

For any classical constant w the $*$-transform of w is the hyperconstant $H(w, w, w, \ldots)$ generated by the infinite sequence $(w, w, w, \ldots)$. We denote this hyperconstant $^*w$. Note that as shown above in the case of a function from $\mathbb{R}$ to $\mathbb{R}$, $^*w$ is not necessarily equal to w. Another example is the $*$-transform of the set of natural numbers $^*\mathbb{N}$, since the hyperlarge number $H(1, 2, 3, \ldots) \in {}^*\mathbb{N}$ we have that $^*\mathbb{N} \neq \mathbb{N}$. However it is also true that the two can be equal. Consider for example a real number x or a finite set of real numbers A, then $^*x = x$ and $^*A = A$.

2.9 Internal vs. External constants

Later when dealing with the transfer principle it will be important to distinguish between two types of constants in our nonstandard system. We define an internal constant to be any constant that is an element of a some set $^*X$ where $^*X$ is the $*$-transform of the classical set X. We will see that these internal constants will behave "well" i.e. in a manner very similar to classical constants. We will find on the other hand that external constants i.e. constants which are not internal, can behave very unpredictably. Our focus will be on results for internal constants and we will be careful to avoid the problems introduced by external constants.

2.9.1 Some Important Examples

1. **Every hyperreal number is internal.**
   This is immediately clear since $^*\mathbb{R}$ is the $*$-transform of $\mathbb{R}$.

2. $\mathcal{P}(^*A) \setminus {}^*\mathcal{P}(A)$ **are the external subsets of** $^*A$.
   Clearly every element of $^*\mathcal{P}(A)$ is internal since $\mathcal{P}(A)$ is a classical set. Now it remains to show that $\mathcal{P}(^*A) \supseteq {}^*\mathcal{P}(A)$. Let $X \in {}^*\mathcal{P}(A)$ then $X = H(S_i)$ for $S_i \subseteq A$, so
   $$X = H(S_i) = \{H(s_i) : s_i \in S_i\} \subseteq {}^*A,$$
   and so $X \in \mathcal{P}(^*A)$ as required.

   This inclusion is strict if and only if the set A is infinite [30] and so $^*A$ has external subsets if and only if A is infinite. In fact if A is an infinite set then A is an external subset of $^*A$.

2.10 Infinitesimals and Hyperlarge numbers in $^*\mathbb{R}$

Theorem 2.10.1. *Hyperlarge numbers and non-zero infinitesimals exist in $^*\mathbb{R}$.*

Proof. First we will prove the existence of non-zero infinitesimals. Consider the hyperreal number $H(1, \frac{1}{2}, \frac{1}{3}, \ldots)$. Clearly $0 < H(1, \frac{1}{2}, \frac{1}{3}, \ldots)$ since $\{i : 0 < \frac{1}{i}\} = \mathbb{N}$. Now let $x \in \mathbb{R}$ be positive. $\exists M \in \mathbb{N} : \forall n > M : \frac{1}{n} < x$. Since $\{1, 2, 3, \ldots, M\}$ is

finite it is not an element of U but its complement $\{M+1, M+2, \ldots\}$ is. This implies that $\{i : \frac{1}{i} < x\} \in U \Rightarrow H(1, \frac{1}{2}, \frac{1}{3}, \ldots) < H(x, x, \ldots) = x$. So $H(1, \frac{1}{2}, \frac{1}{3}, \ldots)$ is a non-zero infinitesimal. Recall we denote this $H(1, \frac{1}{2}, \frac{1}{3}, \ldots) \sim 0$.

Similarly, $H(1, 2, 3, \ldots)$ is hyperlarge since $\forall x \in \mathbb{R} : \exists M \in \mathbb{N} : \forall n > M : x < n \Rightarrow \{i : x < i\} = \{M+1, M+2, \ldots\} \in U$. Recall we denote this $H(1, 2, 3, \ldots) \sim \infty$. $\square$

Theorem 2.10.2. *Every infinite sequence of positive real numbers converging to zero generates a positive infinitesimal.*

Proof. Let (x_i) be a sequence of positive real numbers converging to zero. For any $\varepsilon > 0$, $\varepsilon \in \mathbb{R}$ there exists $M \in \mathbb{N}$ s.t. $\forall n > M : 0 < x_n < \varepsilon$. Now since $\{1, 2, 3, \ldots, M\}$ is finite it is not an element of U but its complement $\{M+1, M+2, \ldots\}$ is. This implies that $\{i : x_i < \varepsilon\} \in U \Rightarrow 0 < H(x_1, x_2, \ldots) < H(\varepsilon, \varepsilon, \ldots) = \varepsilon$. So $H(x_1, x_2, \ldots)$ is a positive infinitesimal. $\square$

Corollary 2.10.3. *Every infinite sequence of negative real numbers converging to zero generates a negative infinitesimal.*

Corollary 2.10.4. *Every infinite sequence of real numbers with a finite number of nonpositive terms which converges to zero generates a positive infinitesimal.*

Corollary 2.10.5. *Every infinite sequence of real numbers with a finite number of nonnegative terms which converges to zero generates a negative infinitesimal.*

Theorem 2.10.6. *An infinite sequence of real numbers generates an infinitesimal for some ultrafilter U_0 if and only if it has a infinite subsequence converging to zero.*

Proof. Let (x_i) be a sequence of real numbers with an infinite subsequence (x_{i_n}) converging to zero. Now since $A = \{i_1, 1_2, \ldots\}$ is an infinite set, there is some ultrafilter U_0 such that $A \in U_0$. (For proof that such an ultrafilter exists we refer to Theorem 1.16.1 of [30].) Now for any $\varepsilon \in \mathbb{R}, \varepsilon > 0$, we have $|x_{i_n}| < \varepsilon$ for all but finitely many $i_n \in A$, and so

$$B = \{i_n \in \mathbb{N} : |x_n| < \varepsilon\} \in U_0.$$

This implies that $H(x_1, x_2, \ldots) < H(\varepsilon, \varepsilon, \ldots) = \varepsilon$, for our chosen ultrafilter U_0 and since ε was an arbitrary positive real number we have that $H(x_1, x_2, \ldots) \simeq 0$.

Conversely suppose (x_i) be a sequence of real numbers with no infinite subsequence (x_{i_n}) converging to zero. Then $\exists \varepsilon \in \mathbb{R}, \varepsilon > 0$ such that the set

$$C = \{i \in \mathbb{N} : x_i < \varepsilon\}$$

is finite. Since C is finite it cannot be in any ultrafilter U. This implies that $H(x_1, x_2, \ldots) \geq H(\varepsilon, \varepsilon, \ldots) = \varepsilon$ for any ultrafilter U. $\square$

2.10.1 Some examples

1. **Addition and hyperlarge numbers**
   What happens when we add two positive hyperlarge numbers? Can we say that the resulting number is also hyperlarge? What about when we add a hyperlarge number to a positive classical real number?

   Consider first an example of the second case:
   $H(1, 2, 3, \ldots) + 1 = H(1, 2, 3, \ldots) + H(1, 1, 1, \ldots) = H(2, 3, 4, \ldots)$. Which is a hyperlarge number by the same argument as used in our theorem above. In fact it is true in general for any $H(x_i) \sim \infty$ and positive real number y since $\forall i : x_i + y > x_i$.

   Now consider an example of the first case:
   $H(1, 2, 3, \ldots) + H(2, 4, 6, \ldots) = H(3, 6, 9, \ldots)$. Which is a hyperlarge number by the same argument as used in our theorem above. Clearly it is also true in general for any $H(x_i), H(y_i) \sim \infty$ since if $A, B \in U$ then $A \cap B \in U$.

2. **Subtraction and hyperlarge numbers**
   What happens if we subtract one positive hyperlarge number from another? What can we say about subtracting a positive classical real number from a positive hyperlarge number?

   Again we will look at an example of the second case first:
   $H(1, 2, 3, \ldots) - 1 = H(1, 2, 3, \ldots) - H(1, 1, 1, \ldots) = H(0, 1, 2, \ldots)$. Which is a hyperlarge number by the same argument as used in our theorem above. In fact it is true in general for any $H(x_i) \sim \infty$ and positive real number y since if $\forall x \{i : x < x_i\} \in U$ then $\{i : x + y < x_i\} \in U$.

   Now consider an example of the first case:
   $H(1, 2, 3, \ldots) - H(0, 1, 2, \ldots) = H(1, 1, 1, \ldots) = 1$. So by subtracting one hyperlarge number from another we get a finite real number, but we can't say this is true in general since: $H(2, 4, 6, \ldots) - H(1, 2, 3, \ldots) = H(1, 2, 3, \ldots)$ which is again a hyperlarge number and $H(2, 5/2, 10/3, \ldots) - H(1, 2, 3, \ldots) = H(1, 1/2, 1/3, \ldots)$ which is an infinitesimal.

3. **Addition and Subtraction of non-zero infinitesimals**
   What happens when we add two positive non-zero infinitesimals? What happens when we subtract one non-zero infinitesimal from another?

   First, consider an example of the first case:
   $H(1, 1/2, 1/3, \ldots) + H(2, 1/4, 1/6, \ldots) = H(3, 3/4, 3/6, \ldots)$ which is again an infinitesimal. This is again true in general.

   Now consider an example of the second case:
   $H(2, 1/4, 1/6, \ldots) - H(1, 1/2, 1/3, \ldots) = H(1, 1/2, 1/3, \ldots)$ which is again an infinitesimal. This is again true in general since if $x_i, y_i > 0$ then $|x_i - y_i| < \max\{x_i, y_i\}$.

4. **The $*$-transform of a function evaluated at an infinitesimal**
   What happens when we evaluate the $*$-transform of a function at a non-zero infinitesimal?

Let's take for example the $*$-transform of $\sin(x)$, $^*\sin(x)$, evaluated at an infinitesimal. Since $\sin(x)$ is continuous and is positive on $(0, \pi)$ with $\sin(0) = 0$ we would expect $^*\sin(\delta)$ where δ is a positive infinitesimal to also be a positive infinitesimal.

Let $\delta = H(1, 1/2, 1/4, \ldots)$ then $^*\sin(\delta) = H(\sin(1), \sin(1/2), \ldots)$. Now $\forall x \in (0, 1) : \sin(x) > 0$, so $(\sin(1), \sin(1/2), \ldots)$ is a sequence of positive real numbers. We also have that $\sin(x)$ is continuous with $\sin(0) = 0$ and so $(\sin(1), \sin(1/2), \ldots)$ converges to zero. Now by Theorem 1.8.2 above we have that $^*\sin(\delta)$ is a positive infinitesimal.

In general we can apply Theorem 2.10.6 to prove that $^*\sin(\delta) \simeq 0$ for any $\delta \simeq 0$. We have only to notice that if any infinite subsequence (x_{i_n}) converges to zero so too does the sequence $(\sin[x_{i_n}])$.

3   The Transfer Principle

3.1   History and Importance

This section brings us finally to the transfer principle. The transfer principle is the powerful *"black box"* that allows us to use the methods of non-standard analysis to prove results in standard analysis. Essentially it provides a 'bridge' between nonstandard analysis and classical mathematics. We will interpret the term classical to mean 'not involving any nonstandard mathematical ideas'. It is therefore, in my opinion, the most important result pertaining to nonstandard analysis. It gives us one of our greatest motivations to study the area. It is as a result of the transfer principle that non-standard methods are powerful tools, tools that we can use to help our understanding of other areas of mathematics.

Being able to transfer reasoning from a system of numbers that included infinitely large and small numbers to a system which does not such as the real numbers was naturally of great interest to the founders of calculus. Since Leibniz and Newton both used such infinitely small and large numbers when developing calculus the validity of these results depended on such a principle. The idea was described by Leibniz and given the name the *"Law of Continuity"*.

In a 1702 letter to the French Mathematician Pierre Varignon, Leibniz formulated the Law of Continuity as follows:

"...et il se trouve que les règles du fini réussissent dans l'infini comme sil y avait des atomes (c'est à dire des éléments assignables de la nature) quoiquil ny en ait point la matière étant actuellement sousdivisée sans fin; et que vice versa les régles de l'infini réussissent dans le fini, comme s'il y'avait des infiniment petits métaphysiques, quoiqu'on n'en n'ait point besoin; et que la division de la matière ne parvienne jamais à des parcelles infiniment petites: c'est parce que tout se gouverne par raison, et qu'autrement il n'aurait point de science ni règle, ce qui ne serait point conforme avec la nature du souverain principe" [24].

Many academics including Robinson identify this passage as a formulation of the law of continuity, which can be summarized as follows: *"the rules of the finite succeed in the infinite, and conversely"* [21]. This principle was a forerunner to the transfer principle that we will discuss in this section. It is a consequence of

a theorem proved by the Polish mathematician Jerzy Łoś in 1955 [29]. Before we tackle Łoś' Theorem we will first give a quick outline of some ideas in mathematical logic.

3.2 Mathematical Logic

3.2.1 Atomic Statements

Atomic relations are simple mathematical relations that don't contain either logical connectives or quantifiers such as $=, <, >$ and $\in$. A relation on n arguments is called n-ary. An n-ary relation can be thought of as a function $R : X_1 \times X_2 \times \ldots \times X_n \to B$. Where B is the set of Boolean constants $B = \{\mathbf{TRUE}, \mathbf{FALSE}\}$. For example $(-1 \in \mathbb{N}) \equiv \mathbf{FALSE}$ and $(1 < 2) \equiv \mathbf{TRUE}$. Since $=$ is one of our atomic relations for clarity we will always use $\equiv$ to denote equivalence. For example $(0 = 1) \equiv \mathbf{FALSE}$ and $(1 = 1) \equiv \mathbf{TRUE}$. An atomic statement is a statement given by applying atomic relations to suitable arguments.

We define $^*\mathbf{TRUE} \equiv \mathbf{TRUE}$ and $^*\mathbf{FALSE} \equiv \mathbf{FALSE}$. Note that since B is finite $^*B \equiv \{^*\mathbf{TRUE}, {}^*\mathbf{FALSE}\} \equiv \{\mathbf{TRUE}, \mathbf{FALSE}\} \equiv B$. Now since we have a correspondence between these relations R and functions it follows intuitively that our corresponding relation $^*R$ in non-standard analysis is given by:

$$H(x_i)^*RH(y_i) \equiv H(x_i R y_i)$$

where by definition $H(x_i R y_i) \equiv (\{i : x_i R y_i\} \in U)$, so

$$H(x_i)^*RH(y_i) \equiv \mathbf{TRUE} \Leftrightarrow \{i : x_i R y_i \equiv \mathbf{TRUE}\} \in U.$$

Lemma 3.2.1. *(A Special Case of Łoś' Theorem)*
Let R be a binary relation $R : X \times Y \to B$. Then $^*[xRy] \equiv xRy$ for $x \in X$, $y \in Y$.

Proof. $^*[xRy] \equiv {}^*x^*R^*y \equiv (\{i : xRy\} \in U)$ Note that xRy does not depend on i so either $\{i : xRy\} = \mathbb{N} \in U$ if $xRy \equiv \mathbf{TRUE}$ or, $\{i : xRy\} = \varnothing \notin U$ if $xRy \equiv \mathbf{FALSE}$. So we have that $^*[xRy] \equiv {}^*x^*R^*y \equiv xRy$ as required. $\square$

This is an example of a transfer principle for simple atomic statements. Although not very powerful it is an interesting result that lets us know that if the atomic statement $^*x^*R^*y$ is equivalent to the classical statement xRy. For example the statement

$$(^*f(^*x^*y) < {}^*x^*y^*z) \equiv \mathbf{TRUE} \Leftrightarrow (f(xy) < xyz) \equiv \mathbf{TRUE}.$$

3.2.2 Arbitrary Statements

Building on the notion of atomic statements, an arbitrary statement is one which is made up of a finite number of atomic relations, logical connectives, quantifiers, constants, free variables and bound variables. using these we can construct more complex mathematic statements.

Logical connectives
Given two basic statements P and Q we can combine them with logical connectives to construct a more complex statement. The basic logical connectives are:

1. **Negation** ("not"), denoted $\neg$.
   Not P has the opposite Boolean value to P.

2. **Conjunction** ("and"), denoted $\wedge$.
   P and Q is true only when both P and Q are both true.

3. **Disjunction** ("or"), denoted $\vee$.
   P or Q is true only when either one or both of P and Q are true.

4. **Conditional** ("if-then" or "implication"), denoted $\Rightarrow$.
   P implies Q is true unless P is true and Q is false.

5. **Biconditional** ("if and only if" or "double implication"), denoted $\Leftrightarrow$.
   $P \Leftrightarrow Q$ is true when P and Q are both true or both false but is false otherwise.

Quantifiers

Quantifiers are used in statements containing variables. There are two quantifiers:

1. **The universal quantifier** ("for all"), denoted $\forall$.

2. **The existential quantifier** ("there exists"), denoted $\exists$.

Constants, free variables and bound variables

Apart from relations, logical connectives and quantifiers, each statement also contains a number of variables and constants.

1. **Constants**
   Specified or unspecified fixed numbers, n-tuples, sets and functions.

2. **Free variables**
   If replacing any variable occurring in a statement by some constant leads to another meaningful statement that variable is called a free variable with respect to the statement.

3. **Bounded variables**
   A variable that is not free is called a bounded or dummy variable.

Notation and conventions for arbitrary statements

While it was useful to write an atomic statement derived from the binary relation R as xRy above we now write this statement as $R(x, y)$ and a general atomic statement as $(P(x, x', \ldots)$ where P is an atomic relation and $x, x', \ldots$ is an expression of constants and free variables. We assume that arbitrary statements are in their prenex normal form[2] are free variables with all logical connectives to the right of the quantifiers. Every statement of first-order logic can be converted to an equivalent statement in prenex normal form [34]. It is also assumed that each bound variable occurs to the left of the $\in$ relation, helping us to ensure that each bound variable is internal.

Now instead of regarding a statement R as a function of substatements $P(s, s', \ldots)$, $Q(t, t', \ldots), S(u, u', \ldots), \ldots$, and of sets $X, X', \ldots$, required in the quantifications we

---

[2]For example the following statement where $P(a, b, \ldots, q, r, \ldots, xyz)$ is a statement containing no quantifiers, $a, b, \ldots$ are constants and $q, r, \ldots$ are free variables is in its prenex normal form: $\forall x : \exists y : \exists z : P(a, b, \ldots, q, r, \ldots, xyz)$.

can regard it simply as a function of the constants and free variables $X, X', X'', \ldots$; $s, s', s'', \ldots,$. We will write a general arbitrary statement with a finite number of constants or free variables $X, X', X'', \ldots; s, s', s'', \ldots$, and a finite number of logical connectives and quantifiers as,

$$R(X, X', X'', \ldots; s, s', s'', \ldots),.$$

Here $X, X', X'', \ldots$ are the sets required to formulate the quantifications properly; that is to say that X must occur in $\exists x \in X$ or in $\forall x \in X$, for some suitable bound variable x, and similarly for $X', X'', \ldots$ and that conversely each quantification is taken care of this way.

3.3 Łoś' Theorem

In this section we will present Łoś' Theorem which is also sometimes known as The Fundamental Theorem of Ultraproducts. This name reflects its significance in our study of non-standard analysis. In essence the Theorem tells us that a first-order statement is true in the ultraproduct if and only if the set of indices for which the formula is true is an element of our ultrafilter U. A proof consistent with our approach can be found in [30]. We give its formal statement below.

Theorem 3.3.1. *(Łoś' Theorem)*
Let any classical statement,

$$R(X, X', X'', \ldots; s, s', s'', \ldots),$$

with a finite number of constants or free variables $X, X', X'', \ldots; s, s', s'', \ldots$, and a finite number of logical connectives and quantifiers be given. $X, X', X'', \ldots$ are the sets required to formulate the quantifications properly; that is to say that X must occur in $\exists x \in X$ or in $\forall x \in X$, for some suitable bound variable x, and similarly for $X', X'', \ldots$ and that conversely each quantification is taken care of this way. Then,

$$H[R(X_i, X_i', X_i'', \ldots; s_i, s_i', s_i'', \ldots)] \equiv$$
$$R(H(X_i), H(X_i'), H(X_i''), \ldots; H(s_i), H(s_i'), H(s_i''), \ldots)$$

3.4 The Transfer Principle

The transfer principle comes as a direct consequence of Łoś' Theorem.

Theorem 3.4.1. *(Transfer Principle)*
Let any classical statement,

$$R(X, X', X'', \ldots; s, s', s'', \ldots),$$

with a finite number of constants or free variables $X, X', X'', \ldots; s, s', s'', \ldots$, and a finite number of logical connectives and quantifiers be given. $X, X', X'', \ldots$ are the sets required to formulate the quantifications properly; that is to say that X must occur in $\exists x \in X$ or in $\forall x \in X$, for some suitable bound variable x, and similarly for $X', X'', \ldots$ and that conversely each quantification is taken care of this way. Then,

$$R(X, X', X'', \ldots; s, s', s'', \ldots) \equiv R(^*X, ^*X', ^*X'', \ldots; ^*s, ^*s', ^*s'', \ldots)$$

Proof. Taking $X_i = X$ and $s_i = s$ for every i and similarly for $X_i', s_i', \ldots$ etc. and applying Łoś' Theorem we get that

$$^*[R(X, X', X'', \ldots; s, s', s'', \ldots)] \equiv H[R(X, X', X'', \ldots; s, s', s'', \ldots)] \equiv$$

$$R(H(X), H(X'), H(X''), \ldots; H(s), H(s'), H(s''), \ldots) \equiv$$

$$R(^*X, ^*X', ^*X'', \ldots; ^*s, ^*s', ^*s'', \ldots).$$

But, we also have that

$$^*[R(X, X', X'', \ldots; s, s', s'', \ldots)] \equiv R(X, X', X'', \ldots; s, s', s'', \ldots).$$

And so

$$R(X, X', X'', \ldots; s, s', s'', \ldots) \equiv R(^*X, ^*X', ^*X'', \ldots; ^*s, ^*s', ^*s'', \ldots).$$

$\square$

The transfer principle in this formulation tells us that any classical statement is equivalent to the non-standard statement we get by replacing everything by its $*$-transform except the bound variables in the statement. This is so important, we do not think of real numbers as infinite Cauchy sequences and now we no longer need to think of hyperreal numbers as infinite sequences of real numbers. Instead we can treat them in a similar way as we treat the real numbers. It is this transfer principle, that acts like a *"bridge"* between analysis in $\mathbb{R}$ and analysis in $^*\mathbb{R}$, that makes our study of nonstandard analysis so useful. Consider the following examples.

Theorem 3.4.2. *(The Archimedean Law)*
Let x be a real number. Then there exists a natural number n that is greater than x.

We can write this statement using the tools of mathematical logic as follows:

$$\forall x \in \mathbb{R} : \exists n \in \mathbb{N} : n > x.$$

Now applying the transfer principle given in Theorem 1.4.1 above we get that this is equivalent to:

$$\forall x \in ^*\mathbb{R} : \exists n \in ^*\mathbb{N} : n > x.$$

So for any hyperreal number x there exists a hypernatural number n that is greater than x. Obviously this is not true if we replace $^*\mathbb{N}$ with $\mathbb{N}$ or the word hypernatural with the word natural in the statement above.

Theorem 3.4.3. *Let n be a natural number that is greater than 1. Then n has at least one prime factor.*

Let $\mathbb{P} = \{P \in \mathbb{N} : p \text{ is prime}\}$. We can now write this statement using the tools of mathematical logic as follows:

$$\forall n \in \mathbb{N} : n > 1 : \exists p \in \mathbb{P} : n/p \in \mathbb{N}.$$

18

Now applying the transfer principle given in Theorem 1.4.1 above we get that this is equivalent to:
$$\forall n \in {}^*\mathbb{N} : n > 1 : \exists p \in {}^*\mathbb{P} : n/p \in {}^*\mathbb{N}.$$

So every hypernatural number greater than 1 has at least one hyperprime factor. We will use this result in section 3.6 to give an elegant proof that $\mathbb{P}$ is an infinite set.

Theorem 3.4.4. *Let p and q be real numbers and let q be greater than p. Then there is a real number r that is greater than p but less than q. (i.e. The real numbers are dense.)*

We can write this statement using the tools of mathematical logic as follows:

$$\forall p, q \in \mathbb{R} : p < q : \exists r \in \mathbb{R} : p < r < q.$$

Now applying the transfer principle given in Theorem 1.4.1 above we get that this is equivalent to:
$$\forall p, q \in {}^*\mathbb{R} : p < q : \exists r \in {}^*\mathbb{R} : p < r < q.$$

In other words the hyperreal numbers are also dense.

Theorem 3.4.5. *(Dedekind Completeness of the Real Numbers)*
Let X be a non-empty subset of $\mathbb{R}$ that has an upper bound $b \in \mathbb{R}$, then X has a least upper bound $\beta \in \mathbb{R}$.

We can write this statement using the tools of mathematical logic as follows:

$$\forall X \in \mathcal{P}(\mathbb{R}) : \{x \neq \varnothing \wedge [\exists b \in \mathbb{R} : \forall x \in X : x \leq b]\} \Rightarrow$$

$$\exists \beta \in \mathbb{R} : [\forall x \in X : x \leq \beta] \wedge [\forall \varepsilon \in \mathbb{R}, \varepsilon > 0 : \exists x \in X : x > \beta - \varepsilon],$$

Now applying the transfer principle given in Theorem 1.4.1 above we get that this is equivalent to:

$$\forall X \in {}^*\mathcal{P}(\mathbb{R}) : \{x \neq \varnothing \wedge [\exists b \in {}^*\mathbb{R} : \forall x \in X : x \leq b]\} \Rightarrow$$

$$\exists \beta \in {}^*\mathbb{R} : [\forall x \in X : x \leq \beta] \wedge [\forall \varepsilon \in {}^*\mathbb{R}, \varepsilon > 0 : \exists x \in X : x > \beta - \varepsilon].$$

In other words if X is an internal subset of ${}^*\mathbb{R}$ that is bounded above by some hypperreal number b, which could be hyperlarge or indeed an infinitesimal, then there is a hyperreal number β that is a least upper bound for X. (Again this could be hyperlarge or indeed an infinitesimal).

Note that it is important that X is an internal set, for example the statement above is not true for the set of real numbers $\mathbb{R}$ since $\mathbb{R}$ is external. Suppose β was a least upper bound for $\mathbb{R}$ in ${}^*\mathbb{R}$, then the β is a hyperlarge number but $\beta - 1$ is also hyperlarge and so is also an upper bound for $\mathbb{R}$ which is a contradiction since β was our least upper bound for $\mathbb{R}$. So ${}^*\mathbb{R}$ is not Dedekind complete[3].

---

[3]This it turns out is a major relief since every Dedekind complete ordered field is isomorphic to $\mathbb{R}$.

3.5 Definitions

Nonstandard analysis can be used to give simplified, elegant definitions of many concepts in classical mathematics. It is especially useful to give intuitive alternative definitions of things that are defined using ε's and δ's in classical mathematics. Such definitions are often found to be very difficult and unintuitive for many undergraduate mathematicians. I have tried to include some nonstandard definitions that are not easily found in the literature.

Since Weierstrass developed the concept of a limit to eliminate the need to use infinitesimals in calculus before a valid model of nonstandard analysis was developed, it makes sense to first give a nonstandard version of the $\varepsilon - \delta$ definition of a limit. As an added bonus the definition we will give is a very intuitive definition of a concept so many undergraduates find difficult to grasp when starting their studies.

Definition *(Nonstandard Definition of a Limit)*
Let $f : \mathbb{R} \to \mathbb{R}$ and let $a, l \in \mathbb{R}$ then we say that "the limit of f as x tends to a is l" and write $\lim\limits_{x \to a} f(x) = l$ if and only if

$$\forall \delta \in {}^{*}\mathbb{R}, \delta \sim 0 : {}^{*}f(a + \delta) - l \simeq 0.$$

In other words if and only if $\forall \delta \sim 0 : l = st({}^{*}f(a + \delta))$.

This can be read as *"The limit of the function f at a is l if and only if the value of the function when we are infinitesimally close to a is infinitesimally close to l."*, which is the intuitive way that many people think of a limit.

Analogously we can give the following definition of one-sided limits.

Definition *(Nonstandard Definition of One-sided Limits)*
Let $f : \mathbb{R} \to \mathbb{R}$ and let $a, l^{+}, l^{-} \in \mathbb{R}$. Then $\lim\limits_{x \to a^{+}} f(x) = l^{+}$ if and only if

$$\forall \delta, \delta > 0, \delta \sim 0 : {}^{*}f(a + \delta) - l^{+} \simeq 0,$$

and $\lim\limits_{x \to a^{-}} f(x) = l^{-}$ if and only if

$$\forall \delta, \delta > 0, \delta \sim 0 : {}^{*}f(a - \delta) - l^{-} \simeq 0.$$

Theorem 3.5.1. *The nonstandard definition of a limit given above is equivalent to the classic "$\varepsilon - \delta$" definition of a limit:*

$$\forall \varepsilon \in \mathbb{R}, \varepsilon > 0 : \exists \delta \in \mathbb{R}, \delta > 0 : \forall x \in \mathbb{R}, 0 < |x - a| < \delta : |f(x) - l| < \varepsilon.$$

Proof. By the transfer principle the statement above is equivalent to

$$\forall \varepsilon \in {}^{*}\mathbb{R}, \varepsilon > 0 : \exists \delta \in {}^{*}\mathbb{R}, \delta > 0 : \forall x \in {}^{*}\mathbb{R}, 0 < |x - a| < \delta : |{}^{*}f(x) - l| < \varepsilon,$$

and this can be simplified to

$$\forall \delta \in {}^{*}\mathbb{R}, \delta \sim 0 : {}^{*}f(a + \delta) - l \simeq 0.$$

$\square$

From this nonstandard definition of a limit the following two intuitive nonstandard definitions of continuity and differentiability quickly follow.

Definition *(Nonstandard Definition of Continuity)*
Let $f : \mathbb{R} \to \mathbb{R}$ and let $a \in \mathbb{R}$. Then f is continuous at a if and only if,

$$\forall \delta \sim 0 : {}^*f(a + \delta) - f(a) \simeq 0.$$

"A function f is continuous at a if and only if the value of the function infinitesimally close to a is infinitesimally close to $f(a)$."

Theorem 3.5.2. *The nonstandard definition of continuity given above is equivalent to the classic definition of continuity:*
$f : \mathbb{R} \to \mathbb{R}$ is continuous at $a \in \mathbb{R}$ if and only if

$$\lim_{x \to a} f(x) = f(a).$$

Proof. By our nonstandard definition of a limit

$$\lim_{x \to a} f(x) = f(a) \Leftrightarrow \forall \delta \sim 0 : {}^*f(a + \delta) - f(a) \simeq 0.$$

$\square$

Definition *(Nonstandard Definition of Differentiability)*
Let $f : \mathbb{R} \to \mathbb{R}$ and let $a, d \in \mathbb{R}$. Then f is differentiable at a if and only if,

$$\forall \delta \sim 0 : \frac{{}^*f(a + \delta) - f(a)}{\delta} \simeq d = f'(a).$$

And so $f'(a) = st[\frac{{}^*f(a+\delta) - f(a)}{\delta}]$.

"The derivative of the function f at a is the slope of the line between $f(a)$ and f evaluated at a point infinitesimally close to a."

The fact that this is equivalent to the classic definition of differentiability again follows directly from our nonstandard definition of a limit.

Definition *(Nonstandard Definition of a Convergent Sequence)*
Let (x_n) be an infinite sequence of real numbers then the sequence converges to $l \in \mathbb{R}$ $(x_n \to l)$ if and only if

$$\forall N \in {}^*\mathbb{N}, N \sim \infty : x_N - l \simeq 0.$$

In other words $st(x_N) = l$. (Here x_N is the Nth element of ${}^*(x_n)$.)

"The sequence (x_n) converges to l if and only if for infinitely large values of n, x_n is infinitesimally close to l."

Definition *(Nonstandard Definition of a Cauchy Sequence)*
Let (x_n) be an infinite sequence of real numbers then the sequence is a Cauchy sequence if and only if

$$\forall N, M \in {}^*\mathbb{N}, N, M \sim \infty : x_N - x_M \simeq 0.$$

(Here x_N is the Nth and x_M is the Mth element of ${}^*(x_n)$.)

"The sequence (x_n) is Cauchy if and only if for infinitely large values of n the terms of the sequence are infinitesimally close."

Theorem 3.5.3. *The nonstandard definition of a Cauchy sequence given above is equivalent to the standard definition of a Cauchy sequence:*

$$\forall n \in \mathbb{N} : \exists k \in \mathbb{N} : \forall N, M \in \mathbb{N}, N, M > k : |x_N - x_M| < 1/n. \ (*)$$

Proof. Fixing $n \in \mathbb{N}$ and $k \in \mathbb{N}$ the statement

$$\forall N, M \in \mathbb{N}, N, M > k : |x_N - x_M| < 1/n$$

by the transfer principle is equivalent to the statement

$$\forall N, M \in {}^*\mathbb{N}, N, M > k : |x_N - x_M| < 1/n$$

Now letting $N, M \sim \infty$, then $\forall k \in \mathbb{N} : N, M > k$ so the statement

$$\forall n \in \mathbb{N} : \exists k \in \mathbb{N} : \forall N, M \in {}^*\mathbb{N}, N, M > k : |x_N - x_M| < 1/n$$

can be simplified to the statement

$$\forall N, M \in {}^*\mathbb{N}, N, M \sim \infty : \forall n \in \mathbb{N} : |x_N - x_M| < 1/n$$

which is equivalent to

$$\forall N, M \in {}^*\mathbb{N}, N, M \sim \infty : x_N - x_M \simeq 0. \ (**)$$

Conversely if we consider the negation of $(*)$

$$\exists n \in \mathbb{N} : \forall k \in \mathbb{N} : \exists N, M \in \mathbb{N}, N, M > k : |x_N - x_M| \geq 1/n$$

Fixing $n \in \mathbb{N}$ the statement

$$\forall k \in \mathbb{N} : \exists N, M \in \mathbb{N}, N, M > k : |x_N - x_M| \geq 1/n$$

by the transfer principle is equivalent to the statement

$$\forall k \in {}^*\mathbb{N} : \exists N, M \in {}^*\mathbb{N}, N, M > k : |x_N - x_M| \geq 1/n$$

now fixing $k \sim \infty$ we have that $N, M \sim \infty$ and so the negation of $(*)$ implies that

$$\exists N, M \in {}^*\mathbb{N}, N, M \sim \infty : \exists m \in \mathbb{N} : |x_N - x_M| \geq 1/n$$

which is equivalent to the negation of $(**)$;

$$\exists N, M \in {}^*\mathbb{N}, N, M \sim \infty : {}^\neg[x_N - x_M \simeq 0].$$

$\square$

Definition *(Nonstandard Definition of Uniform Convergence)*
Let (f_n) be a sequence of functions with $f_n : \mathbb{R} \to \mathbb{R}$. Then (f_n) converges uniformly to the function $f : \mathbb{R} \to \mathbb{R}$ on $\mathbb{R}$ if and only if

$$\forall x \in {}^*\mathbb{R} : \forall N \in {}^*\mathbb{N}, N \sim \infty : {}^*f(x) - {}^*f_N(x) \simeq 0.$$

"The sequence of functions (f_n) converges uniformly to f on $\mathbb{R}$ if and only if for an infinitely large N, ${}^*f_N$ is infinitesimally close to ${}^*f$ at all points of ${}^*\mathbb{R}$"

Theorem 3.5.4. *The nonstandard definition of uniform convergence given above is equivalent to the classical definition given by:*

$$\forall \varepsilon \in \mathbb{R}, \varepsilon > 0 : \exists N \in \mathbb{N} : \forall n \in \mathbb{N}, n > N : \forall x \in \mathbb{R} : |f_n(x) - f(x)| < \varepsilon.$$

Proof. Suppose (f_n) converges to the function f uniformly on $\mathbb{R}$ then by the transfer principle the statement

$$\forall \varepsilon \in \mathbb{R}, \varepsilon > 0 : \exists N \in \mathbb{N} : \forall n \in \mathbb{N}, n > N : \forall x \in \mathbb{R} : |f_n(x) - f(x)| < \varepsilon$$

is equivalent to

$$\forall \varepsilon \in {}^*\mathbb{R}, \varepsilon > 0 : \exists N \in {}^*\mathbb{N} : \forall n \in {}^*\mathbb{N}, n > N : \forall x \in {}^*\mathbb{R} : |{}^*f_n(x) - {}^*f(x)| < \varepsilon$$

Now letting N be hyperlarge we must have that ε is infinitesimal and hence

$$\forall n \in {}^*\mathbb{N}, n \sim \infty : \forall x \in {}^*\mathbb{R} : {}^*f(x) - {}^*f_n(x) \simeq 0.$$

Conversely suppose that

$$\forall n \in {}^*\mathbb{N}, n \sim \infty : \forall x \in {}^*\mathbb{R} : {}^*f(x) - {}^*f_n(x) \simeq 0.$$

Then for taking N such that $N \in {}^*\mathbb{N}, N \sim \infty$ fixing any $\varepsilon \in \mathbb{R}, \varepsilon > 0$ we have that

$$\forall n \in {}^*\mathbb{N}, n > N : \forall x \in {}^*\mathbb{R} : |{}^*f_n(x) - {}^*f(x)| < \varepsilon$$

in other words

$$\exists N \in {}^*\mathbb{N} : \forall n \in {}^*\mathbb{N}, n > N : \forall x \in {}^*\mathbb{R} : |{}^*f_n(x) - {}^*f(x)| < \varepsilon$$

which by the transfer principle is equivalent to

$$\exists N \in \mathbb{N} : \forall n \in \mathbb{N}, n > N : \forall x \in \mathbb{R} : |f_n(x) - f(x)| < \varepsilon$$

and since epsilon was an arbitrary positive real number we have that

$$\forall \varepsilon \in \mathbb{R}, \varepsilon > 0 : \exists N \in \mathbb{N} : \forall n \in \mathbb{N}, n > N : \forall x \in \mathbb{R} : |f_n(x) - f(x)| < \varepsilon.$$

$\square$

3.6 Nonstandard Analysis as a Tool in Classical Mathematics

Many classical theorems and many classical problems can also be proved in a very elegant way using nonstandard analysis. Again in this section I've tried to include some nice illustrative examples that are not easily found in the literature available on the subject. The first proof is simple and uses our new nonstandard definitions of both differentiability and continuity.

Theorem 3.6.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable at $a \in \mathbb{R}$. Then f is continuous at a.*

Proof. Suppose f is not continuous at a. Then

$$\exists \delta \sim 0 : {}^*f(a + \delta) - f(a) = k \not\simeq 0.$$

Where k is either appreciable or hyperlarge. Recall a hyperreal number x is appreciable if it is neither an infinitesimal nor hyperlarge. So

$$\exists \delta \sim 0 : \frac{{}^*f(a + \delta) - f(a)}{\delta} \simeq d \sim \pm\infty.$$

This is a contradiction and so f must be continuous at a. $\qquad\square$

An extremely elegant proof that a sequence of real numbers converges if and only if it is a Cauchy sequence comes as a result of our alternative nonstandard definitions of convergent sequences and Cauchy sequences given above. It is shorter and in my opinion more intuitive than the usual classical proof which can be widely found in introductory real analysis textbooks.

Theorem 3.6.2. *A sequence of real numbers (x_n) is convergent if and only if it is Cauchy.*

Proof. Suppose (x_n) converges to $l \in \mathbb{R}$ then $\forall N \in {}^*\mathbb{N}, N \sim \infty : x_N \simeq l$ and so $\forall N, M \in {}^*\mathbb{N}, N, M \sim \infty : x_N - x_M \simeq l - l = 0$.

Conversely suppose (x_n) is a Cauchy sequence. Suppose for some hyperlarge number N, x_N is hyperlarge, then $\forall m \in \mathbb{N} : x_N > x_m + 1$ and so $\exists M \in \{n \in {}^*\mathbb{N} : n < N\}$ that is the greatest number such that : $x_N > x_M + 1$. This number cannot be limited, but neither can it be hyperlarge since that would imply that $x_M \not\simeq x_N$.

So we have that for every hyperlarge N x_N is limited. Since the sequence is a Cauchy sequence $\forall N, M \sim \infty : st(x_N) = st(x_M) = l \in \mathbb{R}$ and our sequence converges to l. $\qquad\square$

The following proof exploits our intuitive nonstandard definitions of limits and continuity, allowing us to prove that the limit of a uniformly convergent sequence of continuous functions is again continuous without having to use any classical $\varepsilon - \delta$ arguments.

Theorem 3.6.3. *Let $f : X \to Y$ be the limit of a uniformly convergent sequence of continuous functions (f_n), where $f_n : X \to Y$, $X, Y \subseteq \mathbb{R}$. Then f is continuous.*

Proof. Let $N \in {}^*\mathbb{N}, N \sim \infty$. Then since (f_n) converges to f uniformly on X we have that $\forall x \in X : f_N(x) - f(x) \simeq 0$ $(*)$. We also have that $\forall \delta \sim 0 : \forall x \in X :$ ${}^*f_N(x + \delta) - {}^*f(x + \delta) \simeq 0$ $(**)$. Furthermore since f_n is continuous for every $n \in \mathbb{N}$ we have that $\forall \delta \sim 0 : {}^*f_N(x + \delta) - f_N(x) \simeq 0$ $(***)$. Putting $(*), (**)$ and $(***)$ together gives $\forall \delta \sim 0 : f(x) - {}^*f(x + \delta) \simeq 0$ and so f is continuous. $\qquad\square$

This final proof concerning the prime numbers is an interesting example of how useful and elegant using nonstandard analysis to prove classical theorems can be. The example is quite different to the others we have presented in this section in that we are not using infinitesimals or giving a nonstandard proof that involves the use of limits in its classical formulation. It is a slightly more surprising application of nonstandard analysis to give a neat proof of a result in number theory. This is perhaps an area of mathematics you may not expect to have much cause to apply nonstandard analysis to.

Theorem 3.6.4. *The set of all primes, $\mathbb{P}$, is infinite.*

Proof. Showing that $\mathbb{P}$ is infinite is equivalent to showing that $^*\mathbb{P}$ contains non-standard elements. Let $n \in {}^*\mathbb{N}$ be divisible[4] by every natural number. For example one such n could be $H(1!, 2!, 3!, \ldots)$. Next, consider the hyperprime number $p \in {}^*\mathbb{P}$ that divides $n + 1$. This number exists by a result of section 3.4. Then p must be nonstandard; if it was not it would divide n and then since p would then divide n and $n + 1$, it would also divide their difference, 1, which is not true for standard primes. $\qquad\square$

4 The History of Infinitesimals

The concept of an infinitesimal has a long and rich history. For centuries before the rigorous formulation of nonstandard analysis, presented by Robinson in 1961, the idea of an infinitesimal was used by leading mathematicians from Archimedes to Leibniz. The intuitive nature of the use of infinitesimals can explain their popularity, however their use was always controversial and had many outspoken critics due to its lack of a rigorous footing.

4.1 Use in Ancient Greek Mathematics

The use of the intuitive notion of an infinitesimal goes back at least as far as the Ancient Greek mathematicians. They were used to find the area and volume of curved surfaces such as the circle. Antiphon of Athens, who was born around 480 B.C.E., used the idea of infinitesimals in an attempt to square a circle by inscribing a polygon with sides so small that the polygon would be indistinguishable from the circle [26]. Ponstein cites this as perhaps the first time infinitesimals were contemplated by a mathematician [30]. The idea is also known to have used by the Greek atomist philosopher Democritus around 450 B.C.E., it was however dismissed by Eudoxus as a rigorous concept around 350 B.C.E. [3] when he presented the Theorem of Eudoxus[5]. The idea was essentially banished as a rigorous concept from Greek mathematics when the standard was set in Euclid's *Elements* around 250 B.C.E.. In *Elements* Euclid included a version of the Theorem of Eudoxus:
"Two quantities are said to have a ratio, the one to the other, when, if multiplied, they can override themselves." [26]

   Mathematical rigour was very important to ancient Greek mathematicians and they made extensive use of Eudoxus' method of exhaustion and proof by contradiction in order to ensure rigour. Later mathematicians were frustrated by ancient Greek proofs such as the proofs of Archimedes which utilised these methods. Although they cannot be faulted on the grounds of rigour, how the theorems had been conceived seemed to be almost deliberately omitted. However the discovery of the Archimedes palimpset in the 20th century gave some clues as to how he conceived his theorems [36]. It seems the intuitive notion of an infinitesimal was still used by Archimedes but he was well aware that the concept was not rigorous.

---

[4]By divisible here we mean that $\forall m \in \mathbb{N} : n/m \in {}^*\mathbb{N}$.

[5]This is also commonly known as the Archimedean property of the real numbers, which Archimedes credited to Eudoxus.

Indeed Archimedes is quoted of saying of one of his results proven using infinitesimals *"...this has not therefore been proved, rather a certain impression has been created that the conclusion is true"* [10]. It was not until after he had applied Eudoxus' method of exhaustion and proof by contradiction that Archimedes accepted the conclusion as fact.

4.2   Geometers of the 17th century and Indivisibles

In the 1620's the idea of infinitely small numbers was again explored and exploited by mathematicians to achieve some remarkable results. Rather than infinitesimals as we have introduced them they thought about indivisibles. These were numbers so small they could not conceivably be divided any further. The most influential theory of indivisibles was published by Bonaventura Cavalieri in his *Geometria indivisibilibus continuorum nova quadam ratione promota* in 1635 [36]. Again Cavalieri was aware of the problems that arose when dealing with infinitely small quantities, he made great efforts to deal with these problems by adding precise rules to his work.

Many of those who worked to add to Cavalieri's new results were far less careful. In the early 1650's John Wallis used the method of indivisibles and had no problem with using infinitely small and large numbers. He thought of areas as infinite sums of lines and volumes as infinite sums of planes. He also introduced the symbol ∞ to denote the number of them [36]. His work however was doubted and criticised by many due to the use of indivisible. Chief among these critics was Thomas Hobbes who had a problem with the fact that Wallis considered *"an infinitely little Altitude"* to be "both nothing and something and an aliquot part." [18]. He finished his stinging attack by stating that *"All this proceeds from not understanding the grounds of your Profession."* [18].

4.3   The Development of Calculus

One of the most often cited, intuitive and important historical uses of infinitesimals was their use in the development of calculus. Calculus was developed independently by Isaac Newton in the 1660s and Gottfried Wilhelm Leibniz in the 1670s. Both used the concept of an infinitesimal in their development of calculus. Newton used *"fluents"*, *"fluxions"* and *"vanishing increments"* in his developments and Leibniz considered dx and dy as infinitesimals.

Their work allowed them to produce extraordinary results in the areas of differentiation and integration that proved extremely useful in many areas including physics and engineering. The use of infinitesimals in the formative years of calculus lacked rigour however and Leibniz in particular remained uneasy about his own use of infinitesimals stating:

"In any supposed transition, which ends up in a final result, it is admissible to develop a general argument [concerning the transition] such that it comprises also the final result" [26].

Their use was also criticised by the Irish philosopher George Berkeley among others [44]. In 1734 Berkeley published *The Analyst*, the piece was addressed *To an Infidel Mathematician.* In it he provided a sophisticated and powerful critique of infinitesimals and their use in the calculus being developed. His attitude is summed

up by this 'axiom' from *Philosophical commentaries* (no. 354):
"Axiom. No reasoning about things whereof we have no idea. Therefore no reasoning about Infinitesimals." [21]. In fact in 1821 Augustin Cauchy provided damning evidence of the problems associated with the use of infinitesimals. He used infinitesimals to prove a result that Abel showed by counter-example in 1826 could not be logically correct [16].

Despite the fact that there were huge problems with the use of infinitesimals in calculus the amazing results and sheer usefulness helped meant that it continued to progress. However efforts were made by many, especially those who were highly critical about this use of infinitesimals, to establish a rigorous method. One such critic was Georg Cantor he described infinitesimals as *"castles in the air, or rather just nonsense"* his opposition to their use was such that he described them as the *"cholera-bacilli of mathematics"* [3]. Cantor's work to eliminate the infinitesimals was based on attempting to establish mathematical analysis on the basis of number alone, to 'arithmetize' it — in effect, to replace the continuous by the discrete. Instead of presupposing the existence of real numbers he based their definition on sequences of rational numbers. It was Cantor's work along with the work of mathematicians such as Augustin Cauchy, Bernard Bolzano, Richard Dedekind and Karl Weierstrass who formulated the '$\varepsilon - \delta$' definition of a limit that gave a rigorous fondation for calculus. This led to the idea of an infinitesimal in calculus to be abandoned for some time until our modern concept of nonstandard analysis began to form in the late 1950s.

4.3.1   An Example from Newton's Calculus

In this section we present a proof from Newton's *De analysi per aequationes numero terminorum infinitas (On Analysis by Infinite Series)* [28].The proof was provided by Newton almost as an afterthought for an *"attentive reader"*. Throughout Newton extensively uses his intuitive but imprecise notion of an infinitesimal. The proof is essentially a proof of the fundamental theorem of calculus. Below is the translation given in [28] of the original proof, omitting the numerical example first given by Newton.

Quadrature as the inverse of fluxions
Rule 1. The quadrature of simple curves: If $y = ax^{m/n}$ is the curve AD, where a is a constant and m and n are positive integers, then the area of region ABD is $z(x) = [n/(m + n)]ax^{(m+n)/n}$.
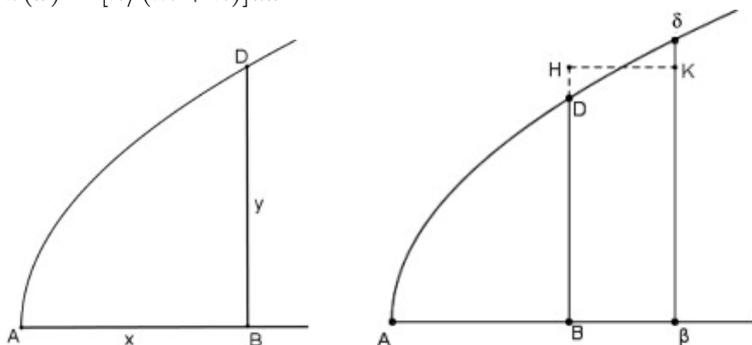
Taken from *Newton's attempt to construct a unitary view of mathematics* [11].

Proof. Let any curve $AD\delta$ have base $AB = x$, perpindicular ordinate $BD = y$ and area $ABD = z$. Take $B\beta = o$, $BK = v$ and the rectangle $B\beta HK(ov)$ equal to the space $B\beta\delta D$. It is therefore, $A\beta = x + o$ and $A\delta\beta = z + ov$. With these premisses, from any arbitrarily assumed relationship between x and z I seek y in the way you see following. ... (numerical example omitted) ...

Or in general if $[n/(m + n)]ax^{(m+n)/n} = z$, that is by setting $na/(m + n) = c$ and $m + n = p$, if $cx^{p/n} = z$ or $c^n x^p = z^n$, then when $x + o$ is substituted for x and $z + ov$ (or, what is its equivalent $z + oy$) for z there arises $c^n(x^p + pox^{p-1}\ldots) = z^n + noyz^{n-1}\ldots$, omitting the other terms, to be precise, which would ultimately vanish. Now, on taking the equal terms $c^n x^p$ and z^n and dividing the rest by o, there remains $c^n px^{p-1} = nyz^{n-1}(= nyz^n/z) = nyc^n x^p/cx^{p/n}$. That is, on dividing by $c^n x^p$, there will be $px^{-1} = ny/cx^{p/n}$ or $pcx^{(p-n)/n} = y$; in other words, by restoring $na/(m + n)$ for o and $m + n$ for p, that is, m for $p - n$ and na for pc , there will come $ax^{m/n} = y$. Conversely therefore if $ax^{m/n} = y$, then will $[n/(m + n)]ax^{(m+n)/n} = z$ as was to be proved. $\square$

Below is a version of the proof written in a rigorous way using modern nonstandard analysis, but remaining as true as possible to Newton's original proof. We have used some more modern language throughout and attempted to present the ideas Newton used in a more clear manner. We have also added to the proof notes to highlight the key differences between our proof and Newton's original, and the flaws in Newton's argument.

Theorem 4.3.1. *(A Modern Rewriting)*
Suppose the area under the continuous curve $y(x)$, is given by $z(x) = [n/(m + n)]ax^{(m+n)/n}$. Then $y(x) = z'(x) = ax^{m/n}$.

Proof. Let $y(x)$ be the curve as shown above, with $AB = x_0 \in \mathbb{R}$, $BD = y(x_0)$ and area under the curve $ABD = z(x_0)$. Let $B\beta = o$ where $o \in {}^*\mathbb{R}, o \sim 0$. Now the area $A\beta\delta = {}^*z(x_0 + o)$. Construct a rectangle $B\beta HK$ of height $BK = \beta H = v \in {}^*\mathbb{R}$ such that the area $B\beta HK$ is exactly equal to the area under the curve $B\beta\delta D$. Hence ${}^*z(x_0 + o) = z(x_0) + ov$.

For notational ease let $na/(m + n) = c$ and $m + n = p$ so that $z(x) = cx^p/n$ and $[z(x)]^n = c^n x^p$. Now we have that $[{}^*z(x_0 + o)]^n = [z(x_0) + ov]^n = c^n(x_0 + o)^p$. Expanding we have that

$$[z(x_0)]^n + \binom{n}{1}[z(x_0)]^{n-1}ov + \binom{n}{2}[z(x_0)]^{n-2}o^2v^2 + \ldots + o^nv^n$$

$$= c^n x_0{}^p + \binom{p}{1}c^n x_0{}^{p-1}o + \binom{p}{2}c^n x_0{}^{p-2}o^2 + \ldots + o^p.$$

Now since $[z(x_0)]^n = c^n x_0{}^p$ these two terms cancel, dividing by o we find that

$$n[z(x_0)]^{n-1}v + \binom{n}{2}[z(x_0)]^{n-2}ov^2 + \ldots + o^{n-1}v^n$$

$$= pc^n x_0{}^{p-1} + \binom{p}{2}c^n x_0{}^{p-2}o + \ldots + o^{p-1}. \; (*)$$

In Newton's proof he now simply disregards all of the terms containing o by setting $o = 0$, however this would mean that in hs previous step Newton could not have possibly divided by o. In this step he also sets $v = y$ without any justification. We will instead take a rigorous, modern nonstandard approach.

Since we have chosen v to be the height of the rectangle with equal area to the area underneath the curve $y(x)$ of the same base o we have that $v = {}^*y(x)$ for some $x \in [x_0, x_0 + o]$. And now since $y(x)$ is continuous we have that $\forall \delta \in {}^*\mathbb{R}, \delta \sim 0 : st[{}^*y(x_0 + \delta)] = y(x_0)$ and so $st(v) = y(x_0)$. Recall also that $st(\delta) = 0$ for any infinitesimal δ. Now taking the standard part of both sides of $(*)$ we have

$$n[z(x_0)]^{n-1}y(x_0) = pc^n x_0^{p-1},$$

substituting in our values for c and p and solving for $y(x_0)$ yields

$$y(x_0) = \frac{pc^n x_0^{p-1}}{n[z(x_0)]^{n-1}} = \frac{(m+n)\left[\dfrac{an}{m+n}\right]^n x_0^{m+n-1}}{n\left[\dfrac{an}{m+n}x_0^{(m+n)/n}\right]^{n-1}} = ax_0^{m/n}.$$

And since x_0 was arbitrary this implies that $y(x) = ax^{m/n}$.

It was without further justification that Newton stated that the converse was also true, that if the curve was given by the formula $y(x) = ax^{m/n}$ then the area under the curve is given by $z(x) = [n/(m+n)]ax^{(m+n)/n}$. $\qquad\qquad\square$

4.4 Modern Nonstandard Analysis

In the late 1950s and early 1960s the idea of infinitesimals and nonstandard analysis was revisited. In 1958 Curt Schmieden and Detlef Laugwitz used a cofinite filter on $\mathbb{N}$ rather than a free ultrafilter (while the cofinite filter is free it is not an ultrafilter) to produce a weak version of nonstandard analysis [35]. It was not until 1961 that a mathematically valid and rigorous system of nonstandard analysis was finally presented by Abraham Robinson.

In the preface to his book *Nonstandard Analysis* he wrote:
"In the fall of 1960 it occurred to me that the concepts and methods of contemporary Mathematical Logic are capable of providing a suitable framework for the development of the Differential and Integral Calculus by means of infinitely small and infinitely large numbers." [31]
His construction relied on advanced logic and was quite inaccessible. Since then many other constructions have been presented. These include Edward Nelson's Internal Set Theory and Jerome Keisler's axioms for hyperreals [27], [22].

5 Applications of Nonstandard Analysis

We have already seen that nonstandard analysis is an interesting area of study in its own right, as Arend Heyting put it, nonstandard analysis is *"a standard model of important mathematical research"* [20]. We have also briefly seen how it can be applied to approach and solve some of the problems of classical mathematics in an intuitive way. Despite the fact that the area is very new it has also been

applied to many areas of science and social science with great success. Nonstandard analysis helps us to achieve interesting and powerful results and the intuitive use of infinitesimals can often bring an interesting philosophical dimension to our work. In this section we will discuss some of these applications with a focus on the applications to economics and finance.

5.1 Economics and Finance

Our focus as mentioned above will be on applications to economics and finance, an area of particular interest due to our studies in mathematics thus far. Nonstandard analysis has been applied in many areas of economics and finance and there is certainly potential for the application of nonstandard analysis to economics to become more widespread. In his piece on the subject Kumaraswamy V. Velupillai claims that nonstandard analysis is one of the areas of mathematics that is *" more consistent with the intrinsic nature and ontology of economic concepts"* than the standard real analysis which currently dominates in the formalization of economic theory [41].

In the introduction to his book on the applications of nonstandard analysis to economics Robert Anderson attributes the fact that nonstandard analysis is not used to a greater extent in economics to the limited number of economists trained in nonstandard analysis. As a consequence he notes that the papers using the methodology of nonstandard analysis are *"necessarily restricted to a small audience"* and that *"the use of nonstandard methods in economics has been largely limited to certain problems in which the advantages of the methodology are greatest"* [2].

Some of the mathematics used in the applications we shall discuss is beyond the scope of the mathematics presented in our simple introduction to the area. This section will therefore be a discussion of the applications and the economic consequences and the ideas involved, rather than a detailed, formal, mathematical presentation. The first application that we will discuss is a model of a large economy based on the nonstandard idea of a so-called hyperfinite set.

5.1.1 Hyperfinite Exchange Economies

In economics it is possible to represent an exchange economy by a function γ that assigns to each agent in the economy a preference and an endowment vector.

$$\gamma : A \to P \times \mathbb{R}_+^k$$

where A is the set of agents P is a the set of preferences and $\mathbb{R}_+^k$ is the commodity space.

One of the most important areas in economics that has been explored using nonstandard analysis is the behaviour of large economies. When analysing large economies using standard analysis it is usual for economists to explore the properties of the limit economy μ, of the sequence of exchange economies (μ_n) where μ_n represents the nth exchange economy $\mu_n : A_n \to P \times \mathbb{R}_+^k$. Since our aim is to investigate large economies and so we consider such sequences where the number of agents in the economy $|A_n| \to \infty$. The limit economy is then formulated as $\mu : A \to P \times \mathbb{R}_+^k$ where A is a nonatomic measure space.

One of the problems with this formulation is that there are some conditions which are inherent in this measure theoretic formulation that can be looked at as strong endogenous assumptions about out limit exchange economy μ [2]. For this reason a nonstandard approach which allows us to construct a similar limit economy without having to impose these strong endogenous assumptions would be very useful. Our approach to the problem is to look at an exchange economy which has what is called a hyperfinite set of agents.

Definition (Hyperfinite Set)
Let $\mathcal{FP}(X)$ denote the set of finite subsets of a set X. A hyperfinite set in $^*\mathbb{R}$ is a set F such that $F \in {}^*\mathcal{FP}(\mathbb{R})$. Any hyperfinite set in $^*\mathbb{R}$ is therefore generated by a sequence of sets $(A_1, A_2, \ldots)$ where $\forall n \in \mathbb{N} : A_n \in \mathcal{FP}(\mathbb{R})$.
It is also true that an internal set F is a hyperfinite set if and only if there exists an internal bijection between F and $G = \{n \in {}^*\mathbb{N} : n \leq g\}$ for some $g \in {}^*\mathbb{N}$.

Hyperfinite sets share many of the properties of finite sets and so can be very useful to work with. We use this definition of a hyperfinite set to construct a hyperfinite exchange economy which does not have the same strong endogenous assumptions that are inherent in the measure theoretic case automatically imposed on it.

Definition (Hyperfinite Exchange Economy)
Let the set of agents in the economy A be a hyperfinite set. A hyperfinite economy is an internal function μ,
$$\mu : A \to {}^*(P \times \mathbb{R}_+^k).$$
Where P is a set of preferences and $\mathbb{R}_+^k$ is the commodity space.

Anderson notes in [2] that due to the endogenous assumptions inherent in the measure theoretic formulation some phenomena that can occur in our new formulation cannot occur in the measure theoretic formulation. These are usually phenomena that occur when a small number of agents are endowed with, or end up consuming most of the resources present in the economy.
One example where the use of the hyperfinite formulation is useful is when introducing an atom into the economy. In the standard formulation, the consumption set of an agent represented by an atom cannot be an element of $\mathbb{R}_+^k$. The limit economy must allow consumptions infinitely large compared to other agents in the economy. In the nonstandard formulation, we do not have this problem as our preferences over $^*\mathbb{R}_+^k$ are 'rich' enough to work with atoms. Anderson also comments that *"in the situations in which the behavior of the measure-theoretic and hyperfinite economies differ, it is the hyperfinite economy rather than the measure-theoretic economy which captures the behavior of large finite economies"* [2].

5.1.2   The van der Pol Equation

In his piece on the subject Velupillai presents the work done using nonstandard analysis by Mikhail Shubin and Alexander Zvonkin, to better understand the nature of the van der Pol equation, as an important contribution of nonstandard

analysis to economics and finance [41], [45]. The van der Pol equation is an ordinary differential equation named after the Dutch physicist Balthasar van der Pol who proposed the equation in the 1920s. The van der Pol equation, and its integrated form as the Rayleigh equation, played an important role in development of the nonlinear endogenous theory of the business cycle [41]. It first appeared in economic literature on the business cycle in a works by Hamburger [14], [15]. In [15] the equation is given in the form

$$\frac{d^2y}{dx^2} - \alpha(1 - y^2)\frac{dy}{dt} + \omega^2 y = 0.$$

In [45] the authors analysed the van der Pol equation rigorously using nonstandard techniques and discovered interesting phase portraits. They referred to what they discovered as *"ducks"*. In their own words *"Ducks are certain singular solutions of equations with a small parameter, which are studied in the theory of relaxation oscillations. These solutions were first found for the van der Pol equation, and their form resembled that of a flying duck"* [45].

In relaxation oscillations, which are also known as fast-slow systems, there is an interaction between slow and fast variables in the system. In finance we could take for example the set of financial markets, clearing infinitely quickly and the set of real markets, clearing relatively slowly. Because of the infinite speed of clearance of the financial market such a system is difficult to analyse using standard techniques. On the other hand, in the nonstandard world the use of hyperlarge and infinitesimal magnitudes is not a problem. Velupillai notes that the possibilities given to us by nonstandard analysis *"for exploring a dynamical system with parameters and variables taking infinitesimal and infinite values is indispensable"*. This is because it allows us to study the interaction of such markets and for example to study the results of infinitesimal changes to parameters in turbulent financial markets.

5.1.3 The Black-Scholes Model

The final application of nonstandard analysis to finance we will discuss is its use in the analysis of the famous Black-Scholes model. The Black-Scholes model is a mathematical model used to value European style options[6], first presented in 1973 [5]. The Black-Scholes PDE is

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0$$

where t is time, S is the price of the underlying stock, V is the price of a derivative (which is a function of time and stock price), r is the force of interest and σ is the volatility of the stocks returns.

Many economists have an intuition that the Black-Scholes model has a 'built-in' version of the Cox-Ross-Rubinstein (CRR) model for the valuation of options. The CRR model is a discrete time model based on a geometric random walk and was first introduced in 1979 [7]. In *Nonstandard Methods in Option Pricing* the

---

[6]In finance a European style option is a derivative instrument, a contract which offers the buyer the right, but not the obligation, to buy or sell an underlying asset at an agreed-upon price on a specific date.

authors use nonstandard analysis to show that the Black-Scholes model can be obtained as the standard part of a hyperfinite CRR model [8]. This is a formal justification of the idea that the Black-Scholes model has a 'built-in' version of the CRR model, allowing for better and clearer understanding about the nature of both models and the link between them.

5.2   Selected Other Applications

When studying infinitesimals one is immediately struck by how natural and intuitive an idea they represent. In contrast to the quite complicated '$\varepsilon - \delta$' definitions and proofs of standard real analysis, the proofs and definitions of nonstandard analysis, as we have seen earlier, often correspond exactly to our intuitive understanding. It is for this reason that using nonstandard analysis and infinitesimals in education has been explored by a number of mathematicians. Chief among these are Keisler who published an introduction to calculus based on the infinitesimals, *Elementary Calculus: An Infinitesimal Approach* [23]. The book was used by five high schools in an experiment by one of Keisler's PhD students with favourable results [37]. Another advocate of the use of infinitesimals in education is David Tall who found evidence for the existence of an intuitive notion of infinitesimals used by students in surveys he carried out [38].

Other applications include the elimination of the need for what Tao calls *"epsilon management"*. This is the use of nonstandard analysis to simplify and shorten proofs containing many small real quantities. He refers to one example in his own work where a paper which he produced using a nonstandard approach came to 28 pages. He contrasted this with a similar paper he wrote before he knew of the approach which stretched to 85 pages due to epsilon management [39]. Some theorems relating to classical mathematics have been originally proven using nonstandard methods. It is only after this that an alternative proof using a standard approach has been given. Without the use of nonstandard analysis these proofs may not have been discovered for some time. One example is a nonstandard proof by Abraham Robinson and Allen Bernstein that every polynomially compact linear operator on a Hilbert space has an invariant subspace [4]. A standard proof was given by Paul Halmos who reinterpreted their proof after being sent a preprint by Robinson. The proof was published in the very same issue of the same journal [13].

There has also been success in applying nonstandard analysis to physics. The combination of nonstandard analysis and physics is described by Sylvia Wenmackers as a *"natural"* one [43]. This is due to the fact that physicists have continued to speak in terms of infinitesimals despite the formal development of calculus. Nonstandard analysis has been applied to many problems in physics using differential equations, and to quantum mechanics. One interesting example is the application of nonstandard analysis to special and general relativity. Robert Herrmann provides some corrections to Einstein's original work in the area which used an intuitive concept of an infinitesimal in [16].

6 Appraisal and Conclusion

In this final section it remains to conclude whether or not the study of nonstandard analysis is a worthwhile activity. We have seen the intuitive appeal of infinitesimals, and a selection of the powerful applications of nonstandard analysis in our previous sections. We must ask ourselves however, how useful nonstandard analysis truly is if every nonstandard proof of a standard theorem in analysis also has a standard proof due to the transfer principle. We must also examine the critiques of nonstandard analysis that have emerged since Robinson's first presentation of the theory.

We believe that the first question is easily answered. The intuitive nature of nonstandard proofs is enough to make them extremely worthwhile. We also believe that it is true that mathematicians may not even discover standard proofs without first developing a nonstandard proof. A quote from Zvonkin and Shubin whose work we cited in our section on the applications to economics and finance: *"It was not by chance that ducks were discovered with the help of nonstandard analysis and in connection with it. We think that the language of non-standard analysis will make it easy for a wide circle of mathematicians to become acquainted with the theory of ducks and with the theory of relaxation oscillations in general."* [45] sums it up nicely. While we have also shown that in some cases nonstandard methods have at the very least allowed mathematicians to formulate proofs quicker.

The first criticisms of the use of infinitesimals we come across are of course the historical criticisms. Criticisms such as those by Berkeley, which were entirely valid at the time as there was no valid rigorous footing for the use of infinitesimals. This was until Robinson presented his version of nonstandard analysis. Once this system was presented by Robinson all such historical criticisms had been addressed. Now that a rigorous theory of infinitesimals had been presented we knew much about the nature of infinitesimals and so Berkeley's 'axiom' *"No reasoning about things whereof we have no idea. Therefore no reasoning about Infinitesimals."* [21] no longer held.

There are however modern criticisms of the theory as first presented by Robinson. One major figure associated with the criticism of Robinson's nonstandard analysis is the French mathematician Alain Connes. His criticisms of Robinson's system of infinitesimals began in 1995 and appear in his books, research articles, interviews and a blog, describing the hyperreals as a *"virtual theory"* and a *"chimera"* [19]. However not only does Connes' criticism undermine some of his own earlier work which used the ideas of Robinson's nonstandard analysis, but his argument also appears to be circular, relying on the transfer principle. In *Tools, objects and chimeras: Connes on the role of hyperreals in mathematics* [19] the authors provide a convincing defence of Robinson's nonstandard analysis and attempt to show the flaws in Connes' arguments. This article also addresses a critique due to Moshé Machover. Other challenges to the theory such as the those by Adam Elga and Erret Bishop have also been addressed [17], [20].

Other criticisms of nonstandard analysis come mainly from constructivist mathematicians. The constructivist school of mathematics asserts that to prove that any mathematical object exists one must first be able to 'construct' it. The axiom of choice, which is required to prove the existence of an ultrafilter, which in turn

underpins our entire theory of nonstandard analysis, is highly nonconstructive in nature. For this reason *"to some constructivists nonstandard analysis represents the worst extreme of nonconstructive mathematics"* [32]. Weak, restricted forms of nonstandard analysis have been presented that do not require the axiom of choice, mainly relying on the use of a cofinite filter instead of a free ultrafilter. Such models of a hyperreal system would be far more appealing to constructivist mathematicians but do not have the power of the full system of nonstandard analysis presented by Robinson. One example is the system presented by Schmieden and Laugwitz [35] and another is discussed by Tao in a blog post [40].

We started this project with a quote from the great Kurt Gödel, *"There are good reasons to believe that non-standard analysis, in some version or other, will be the analysis of the future"* [33]. After our study of nonstandard analysis and its applications as presented in this project, and in light of the criticisms discussed and analysed above, we agree with this statement. However we do not believe that nonstandard analysis will ever entirely replace standard analysis. Rather we feel that its use will become more widespread as more mathematicians, scientists and social scientists become familiar with its methods. We feel that more and more work will be originally proved and presented using nonstandard analysis as time goes on, although it may be some time before these are routinely given without an attempt to also provide a version using standard methods.

References

[1] Online etymology dictionary. Oct 2012. URL `http://dictionary.reference.com/browse/infinitesimal`.

[2] R.M. Anderson. Nonstandard analysis with applications to economics. In *Handbook of Mathematical Economics, Vol. 4*, pages 2145–2208. Elsevier, 1991.

[3] J. L. Ball. Continuity and infinitesimals. *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*, 2009. URL `http://plato.stanford.edu/archives/fall2009/entries/continuity/`.

[4] A.R. Bernstein and A. Robinson. Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos. *Pacific Journal of Mathematics*, 16(3): 421–431, 1966.

[5] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[6] A. Blass. Book review. *Bulletin of the American Mathematical Society*, 84(1): 34–41, 1978.

[7] J.C. Cox, S.A. Ross, and M. Rubinstein. Option pricing: A simplified approach. *Journal of Political Economy*, 7:229–263, 1979.

[8] N.J. Cutland, P.E. Kopp, and W. Willinger. Nonstandard methods in option pricing. In *Decision and Control, 1991., Proceedings of the 30th IEEE Conference on*, volume 2, pages 1293 –1298, Dec 1991.

[9] J. Fauvel and J. Gray. *The History of Mathematics: A Reader*. Macmillan Education in association with the Open University, 1987.

[10] J. L. Fera. Making the infinitesimal precise. 2008. URL `http://jfera.web.wesleyan.edu/docs/nonstandard.pdf`.

[11] M. Galuzzi. Newton's attempt to construct a unitary view of mathematics. *Historia Mathematica*, 37(3):535 – 562, 2010.

[12] L. Haddad. Un outil incomparable: l'ultrafiltre. *Tatra Mt. Math. Publ.*, 31: 131–176, 2005.

[13] P.R. Halmos. Invariant subspaces of polynomially compact operators. *Pacific Journal of Mathematics*, 16(3):433–437, 1966.

[14] L. Hamburger. Een nieuwe weg voor conjunctuur-onderzoek, een nieuwe richtlijn voor conjunctuur-politiek. *De Economist*, LXXIX:1–38, 1930.

[15] L. Hamburger. Analogie des fluctuations économiques et des oscillations de relaxation. *Institut de Statistique de l'Université de Paris, Supplement aux Indices du Mouvement des Affaires*, pages 1–35, Janvier 1931.

[16] R.A. Herrmann. Nonstandard analysis applied to special and general relativity - the theory of infinitesimal light-clocks. 1995. URL `http://http://www.raherrmann.com/ecpart1.pdf`.

[17] F. Herzberg. Intenal laws of probability, generalized likliehoods and Lewis' infinitesimal chances-a response to Adam Elga. *The British Journal for the Philosophy of Science*, 58(1):25–43, 2007.

[18] T. Hobbes. Hobbes' response to Wallis. In *Six lessons to the Professors of Mathematicks*, page 46. 1656.

[19] V. Kanovei, M. G. Katz, and T. Mormann. Tools, objects and chimeras: Connes on the role of hyperreals in mathematics. *Foundations of Science*, To appear 2013. URL `http://arxiv.org/pdf/1211.0244.pdf`.

[20] K.U. Katz and M.G. Katz. Meaning in classical mathematics: Is it at odds with intuitionism? *Intellectica*, 56(2):223–302, 2011.

[21] M. G. Katz and D. Sherry. Leibniz's Infinitesimals: Their fictionality, their modern implementations, and their other foes from Berkeley to Russell and beyond. *Erkenntnis*, 2012.

[22] H. J. Keisler. Limit ultrapowers. *Transactions of the American Mathematical Society*, 107:383–408, 1963.

[23] H.J. Keisler. *Elementary Calculus: An Infinitesimal Approach*. Prindle, Weber & Schmidt, second edition, 1986.

[24] G. Leibniz. Letter to Varignon. *In Leibniz's Infinitesimals: Their fictionality, their modern implementations, and their other foes from Berkeley to Russell and beyond, M. G. Katz And D. Sherry*, 1702.

[25] P.A. Loeb and M. Wolff. *Nonstandard Analysis for the Working Mathematician*. Kluwer, 2000.

[26] G. Lolli. Infinitesimals and infinites in the history of mathematics: A brief survey. *Applied Mathematics and Computation*, pages 7979–7988, 2012.

[27] E. Nelson. Internal set theory: a new approach to nonstandard analysis. *Bulletin of the American Mathematical Society*, 83:1165–1198, 1977.

[28] I. Newton. De analysi per aequationes numero terminorum infinitas (1669). In *The History of Mathematics: A Reader*, page 384. Macmillan Education in association with the Open University, 1987.

[29] J. Łoś. Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. In *Mathematical interpretation of formal systems*, pages 98–113. North-Holland Publishing Co., 1955.

[30] J. Ponstein. *Nonstandard Analysis*. Faculty of Economics at the University of Groningen, 2002.

[31] A. Robinson. *Nonstandard Analysis*. North-Holland Publishing Co., 1966.

[32] D. A. Ross. A nonstandard proof of a lemma from constructive measure theory. Technical report, 1999.

[33] T. Runge. Hyperfinite probability theory and stochastic analysis within Edward Nelsons internal set theory. 2011. URL http://www10.informatik.uni-erlangen.de/Publications/Theses/2010/Runge_DA10.pdf.

[34] A. Sakharov. Prenex normal form. *MathWorld - A Wolfram Web Resource, created by E.W. Weisstein*, 2012. URL http://mathworld.wolfram.com/PrenexNormalForm.html.

[35] C. Schmieden and D. Laugwitz. Eine erweiterung der infinitesimalrechnung. *Mathematisches Zeitschrift*, 69:1–39, 1958.

[36] J. Stedall. *Mathematics Emerging: A Sourcebook 1540-1900*. Oxford University Press, 2008.

[37] K. Sullivan. The teaching of elementary calculus using the nonstandard analysis approach. *The American Mathematical Monthly (Mathematical Association of America)*, 83(5):370–375, 1976.

[38] D. Tall. Intuitive infinitesimals in the calculus. 1980. URL http://homepages.warwick.ac.uk/staff/David.Tall/pdfs/dot1980c-intuitive-infls.pdf.

[39] T. Tao. *Structure and Randomness: Pages from Year One of a Mathematical Blog*. American Mathematical Society, 2008.

[40] T. Tao. A cheap version of nonstandard analysis. 2012. URL http://terrytao.wordpress.com/2012/04/02/a-cheap-version-of-nonstandard-analysis/.

[41] K. V. Velupillai. Varieties of mathematics in economics. 2007. URL http://aran.library.nuigalway.ie/xmlui/bitstream/handle/10379/328/paper_0126.pdf?sequence=1.

[42] H. Weber. Leopold Kronecker. *Mathematische Annalen*, 43:1–25, 1893.

[43] S. Wenmackers. Hyperreals and their applications. 2012. URL http://http://fitelson.org/few/wenmackers_notes.pdf.

[44] D. R. Wilkins. The 'analyst' controversy. 2012. URL http://www.maths.tcd.ie/pub/HistMath/People/Berkeley/AnalCont.html.

[45] A. K. Zvonkin and M. A. Shubin. Nonstandard analysis and singular peturbations of ordinary differential equations. *Russian Mathematical Surveys*, 39 (2):69–131, 1984.