Non-standard Analysis

Mahouclo Anicet Hounkanrin (anicet@aims.ac.za)
African Institute for Mathematical Sciences (AIMS)

Supervised by: Dr. Gareth Boxall
Stellenbosch University, South Africa
Co-supervised by: Georg Anegg
AIMS, South Africa

18 May 2017

Submitted in partial fulfillment of a structured masters degree at AIMS South Africa

Abstract

Most definitions in standard analysis are given using the ϵ-δ approach to limiting processes which is not always straightforward and makes proofs difficult. On the other hand, non-standard analysis uses the notion of infinitesimal and infinite numbers to make these definitions more direct and easy to understand. In this essay, we set up non-standard analysis rigorously and give an overview of how it can be used to make definitions and proofs of well-known results from Calculus simpler and easy to understand. Furthermore, we explore non-standard proofs of more advanced theorems which, in fact, reveal non-standard analysis as a powerful tool for mathematical proofs.

Keywords: Standard analysis, infinitesimal, infinite numbers, non-standard analysis.

Résumé

La plupart des définitions en mathématiques sont formulées utilisant l'approche ϵ-δ qui n'est pas toujours directe et rend les preuves difficiles. Par contre, l'analyse non-standard, introduisant la notion de nombres infinitésimaux et infinis, rend ces définitions plus directes et faciles à comprendre. Dans ce mémoire, après une définition rigoureuse de l'analyse non-standard, nous donnons un vaste apperçu de comment elle peut être utilisée pour formuler de façon plus claire et compréhensible des définitions et théorèmes bien connus. Nous explorons par ailleurs, l'usage de l'analyse non-standard à travers des preuves plus avancées qui révelent l'analyse non-standard comme un outil puissant de preuve en mathématiques.

Declaration

I, the undersigned, hereby declare that the work contained in this research project is my original work, and that any work done by others or by myself previously has been acknowledged and referenced accordingly.

Mahouclo Anicet Hounkanrin, 18 May 2017

Contents

1. Introduction

The concept of *infinitely small* and *infinitely large* numbers was introduced to Calculus in the seventeenth century by Gottfried Leibniz and Isaac Newton. Since then, a number of great mathematicians tried to justify the existence of such numbers to no avail. As a result, this concept felt out of favour in the nineteenth century and was replaced by the ϵ-δ approach of Karl Weierstrass. In 1960, using methods from model theory, Abraham Robinson created non-standard analysis in which he rigorously extended the set of real numbers to include *infinitely small* and *infinitely large* numbers (Davis, 2009). Non-standard analysis has then become a powerful mathematical tool which gives easier definitions for standard concepts and more direct proofs of well-known mathematical theorems (Goldbring, 2014).

Abraham Robinson and Allen Bernstein published a paper (Bernstein and Robinson, 1966) where they used non-standard analysis to solve the invariant subspace problem. This was a three hundred years old problem that had never been solved using standard analysis.

Although non-standard analysis is intuitively clear, more appealing than standard analysis and rigorously founded, it is still subject to controversies and not familiar to most mathematicians. The purpose of this essay is to explore how non-standard analysis can be used to give not only more comprehensible definitions of well-known Calculus concepts, but also nice and rigorous proofs of theorems.

In Chapter 2, we will see how the usual set of real numbers can be extended to include *infinitely small* and *infinitely large* numbers. Chapter 3 will introduce the main theorem of non-standard analysis, the Transfer Principle, which will allow us to carry over results from non-standard analysis to standard analysis and vice-versa. In Chapter 4, we will show some basic applications of non-standard analysis in Calculus. The last chapter will explore more advanced applications.

2. Construction of the set $^*\mathbb{R}$ of hyperreals

The first step towards the study of non-standard analysis is the construction of a set $^*\mathbb{R}$ called the non-standard reals or hyperreal numbers (Henson, 1997). This new set is an extension of the usual set $\mathbb{R}$ of real numbers.

The purpose of constructing the hyperreals $^*\mathbb{R}$ is to create a field that not only embeds every standard real number but also includes new types of numbers: infinitesimal and infinite numbers.

2.1 Real-valued sequences construction of $^*\mathbb{R}$

Consider the set $\mathbb{R}^{\mathbb{N}}$ of real-valued sequences. Let $x = (x_n)_{n\in\mathbb{N}}$ and $y = (y_n)_{n\in\mathbb{N}}$ be elements of $\mathbb{R}^{\mathbb{N}}$.

We define addition and multiplication on $\mathbb{R}^{\mathbb{N}}$ as follows:

- $x + y = (x_n + y_n)_{n\in\mathbb{N}}$
- $x \cdot y = (x_n \cdot y_n)_{n\in\mathbb{N}}$

From the properties of the field $(\mathbb{R}, +, \cdot)$, we can easily derive that $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$ is a commutative ring.

However, it turns out that $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$ cannot be a field for the simple reason that it has zero divisors.

Indeed, consider for example two non-zero real-valued sequences $x = (1, 0, 1, 0, \ldots)$ and $y = (0, 1, 0, 1, \ldots)$. The product $x \cdot y$ is the zero sequence $(0, 0, 0, \ldots)$ whereas none of the factors x and y is the zero sequence. Therefore, $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$ is not a field.

It follows that $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$ is not the extension of $\mathbb{R}$ we are looking for. Now the idea is to set up an equivalence relation on $\mathbb{R}^{\mathbb{N}}$ to avoid this problem. We can define such an equivalence relation by introducing the notion of ultrapowers.

2.2 Ultrapowers construction of $^*\mathbb{R}$

In this section, we are going to see how ultrapowers can be used to extend the usual set of real numbers $\mathbb{R}$. Toward this end, we first need to define the notion of *filter* and *ultrafilter*.

2.2.1 Definition (Filter)**.** Let J be a non-empty set. Denote by $\mathcal{P}(J)$ the power set of J. A filter $\mathcal{F}$ on J is a subset of $\mathcal{P}(J)$ that satisfies the following properties:

- $\emptyset \notin \mathcal{F}$ (Proper filter property)
- If $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$ (Finite intersection property)
- Let $A \in \mathcal{F}$ and $B \subseteq J$. If $A \subseteq B$, then $B \in \mathcal{F}$ (Superset property).

2.2.2 Definition (Ultrafilter)**.** Let J be a non-empty set. An ultrafilter on J is a filter $\mathcal{U}$ on J such that for all $A \subseteq J$, either $A \in \mathcal{U}$ or $J \setminus A \in \mathcal{U}$.

The ultrafilter $\mathcal{U}$ is said to be principal if it is of the form $\{A \subseteq J : j \in A\}$ for some $j \in J$.

If $\mathcal{U}$ is not principal, it is called a **non-principal ultrafilter**. Non-principal ultrafilters contain no finite subset of J.

2.2.3 Remark. Let J be a non-empty infinite set. There exists a non-principal ultrafilter $\mathcal{U}$ on J. This can be proved using Zorn's Lemma.

For the rest of this paper, we will be considering non-principal ultrafilters.

2.2.4 Lemma. Let $\mathcal{U}$ be an ultrafilter on $\mathbb{N}$. Let $\{A_1, \ldots, A_m\}$ be a finite collection of subsets of $\mathbb{N}$ such that $\bigcup_{i=1}^{m} A_i = \mathbb{N}$ and $A_i \cap A_j = \emptyset$ whenever $i \neq j$. Then there is exactly one $i \in \{1, \ldots, m\}$ such that $A_i \in \mathcal{U}$.

Proof. Suppose $A_i \notin \mathcal{U}$ for all $i \in \{1, \ldots, m\}$. Then $\mathbb{N} \setminus A_i \in \mathcal{U}$ for all $i \in \{1, \ldots, m\}$ (because $\mathcal{U}$ is an ultrafilter). So $\bigcap_{i=1}^{m} (\mathbb{N} \setminus A_i) \in \mathcal{U}$ (Finite intersection property). But, by De Morgan's law, we have $\bigcap_{i=1}^{m} (\mathbb{N} \setminus A_i) = \mathbb{N} \setminus \bigcup_{i=1}^{m} A_i$. So $\left(\mathbb{N} \setminus \bigcup_{i=1}^{m} A_i\right) \in \mathcal{U}$. But $\bigcup_{i=1}^{m} A_i = \mathbb{N}$. So $\left(\mathbb{N} \setminus \bigcup_{i=1}^{m} A_i\right) \in \mathcal{U}$ implies $\emptyset \in \mathcal{U}$. That violates the proper filter property of $\mathcal{U}$. Therefore, there must be some $i \in \{1, \ldots, m\}$ such that $A_i \in \mathcal{U}$.

Now, suppose $A_i, A_j \in \mathcal{U}$ for some $i, j \in \{1, \ldots, m\}$ such that $i \neq j$. Then $A_i \cap A_j \in \mathcal{U}$. But $A_i \cap A_j = \emptyset$. So $\emptyset \in \mathcal{U}$. Again, this violates the proper filter property of $\mathcal{U}$. Therefore, there is exactly one $i \in \{1, \ldots, m\}$ such that $A_i \in \mathcal{U}$. □

2.2.5 Definition (Relation on $\mathbb{R}^{\mathbb{N}}$)**.** Consider two sequences $x = (x_n)_{n\in\mathbb{N}}$ and $y = (y_n)_{n\in\mathbb{N}}$ in $\mathbb{R}^{\mathbb{N}}$. Let $\mathcal{U}$ be an ultrafilter on $\mathbb{N}$. We define a relation $\sim$ on $\mathbb{R}^{\mathbb{N}}$ as follows:

$$x \sim y \quad \text{if and only if} \quad \{n \in \mathbb{N} : x_n = y_n\} \in \mathcal{U}.$$

2.2.6 Proposition (Equivalence relation on $\mathbb{R}^{\mathbb{N}}$)**.** The relation $\sim$ is an equivalence relation on $\mathbb{R}^{\mathbb{N}}$.

Proof. Reflexivity: Let $x = (x_n)_{n\in\mathbb{N}}$ be an element of $\mathbb{R}^{\mathbb{N}}$. We have $\{n \in \mathbb{N} : x_n = x_n\} = \mathbb{N} \in \mathcal{U}$ because $\mathcal{U}$ is an ultrafilter on $\mathbb{N}$ and $\mathbb{N} \setminus \mathbb{N} = \emptyset \notin \mathcal{U}$. So $x \sim x$.

Symmetry: Let $x = (x_n)_{n\in\mathbb{N}}$ and $y = (y_n)_{n\in\mathbb{N}}$ be elements of $\mathbb{R}^{\mathbb{N}}$. Suppose $x \sim y$. Then $\{n \in \mathbb{N} : x_n = y_n\} \in \mathcal{U}$. So $\{n \in \mathbb{N} : y_n = x_n\} \in \mathcal{U}$. Thus $y \sim x$.

Transitivity: Let $x = (x_n)_{n\in\mathbb{N}}$, $y = (y_n)_{n\in\mathbb{N}}$ and $z = (z_n)_{n\in\mathbb{N}}$ be elements of $\mathbb{R}^{\mathbb{N}}$. Suppose $x \sim y$ and $y \sim z$. Then $\{n \in \mathbb{N} : x_n = y_n\} \in \mathcal{U}$ and $\{n \in \mathbb{N} : y_n = z_n\} \in \mathcal{U}$. So, $\{n \in \mathbb{N} : x_n = y_n\} \cap \{n \in \mathbb{N} : y_n = z_n\} \in \mathcal{U}$ (Finite intersection property). Then $\{n \in \mathbb{N} : x_n = y_n = z_n\} \in \mathcal{U}$. But $\{n \in \mathbb{N} : x_n = y_n = z_n\} \subseteq \{n \in \mathbb{N} : x_n = z_n\}$. Therefore, $\{n \in \mathbb{N} : x_n = z_n\} \in \mathcal{U}$ (Superset property of $\mathcal{U}$). Hence $x \sim z$.

It follows that $\sim$ is indeed an equivalence relation. □

The equivalence class of x, denoted by $[x]$, is defined as:

$$[x] = \{y \in \mathbb{R}^{\mathbb{N}} : x \sim y\}.$$

2.2.7 Proposition (Congruence)**.** The relation $\sim$ is a congruence on the ring $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$, i.e. $\sim$ is compatible with $+$ and $\cdot$ on $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$.

Proof. Let $x = (x_n)_{n\in\mathbb{N}}$, $y = (y_n)_{n\in\mathbb{N}}$, $x' = (x'_n)_{n\in\mathbb{N}}$ and $y' = (y'_n)_{n\in\mathbb{N}}$ be elements of $\mathbb{R}^{\mathbb{N}}$. Suppose $x \sim x'$ and $y \sim y'$. Then $\{n \in \mathbb{N} : x_n = x'_n\} \in \mathcal{U}$ and $\{n \in \mathbb{N} : y_n = y'_n\} \in \mathcal{U}$. So $\{n \in \mathbb{N} : x_n = x'_n\} \cap \{n \in \mathbb{N} : y_n = y'_n\} \in \mathcal{U}$ (Finite intersection property). Then $\{n \in \mathbb{N} : x_n = x'_n \text{ and } y_n = y'_n\} \in \mathcal{U}$. But $\{n \in \mathbb{N} : x_n = x'_n \text{ and } y_n = y'_n\} \subseteq \{n \in \mathbb{N} : x_n + y_n = x'_n + y'_n\}$. So $\{n \in \mathbb{N} : x_n + y_n = x'_n + y'_n\} \in \mathcal{U}$ (Superset property of $\mathcal{U}$). Then $x + y \sim x' + y'$. Hence $\sim$ is compatible with addition on $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$.

Likewise, $\{n \in \mathbb{N} : x_n = x'_n \text{ and } y_n = y'_n\} \subseteq \{n \in \mathbb{N} : x_n \cdot y_n = x'_n \cdot y'_n\}$. So $\{n \in \mathbb{N} : x_n \cdot y_n = x'_n \cdot y'_n\} \in \mathcal{U}$. Then $x \cdot y \sim x' \cdot y'$. Hence $\sim$ is compatible with multiplication on $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$.

Therefore, the relation $\sim$ is a congruence on the ring $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$. □

2.2.8 Definition (Ultrapower)**.** Consider an ultrafilter $\mathcal{U}$ on $\mathbb{N}$ and the equivalence relation $\sim$ induced by $\mathcal{U}$ on $\mathbb{R}^{\mathbb{N}}$. The quotient set defined by $\mathbb{R}^{\mathbb{N}}/\mathcal{U} = \mathbb{R}^{\mathbb{N}}/\sim = \left\{[a] : a \in \mathbb{R}^{\mathbb{N}}\right\}$, where $[a]$ is the equivalence class of a, is called the ultrapower of $\mathbb{R}$ with respect to $\mathcal{U}$.

2.2.9 Operations and relations on $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$.

Let $x = [(x_n)_{n\in\mathbb{N}}]$ and $y = [(y_n)_{n\in\mathbb{N}}]$ be elements of $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$. We define addition and multiplication on $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$ as follows:

- $x + y = [(x_n + y_n)_{n\in\mathbb{N}}]$
- $x \cdot y = [(x_n \cdot y_n)_{n\in\mathbb{N}}]$

The fact that $\sim$ is a congruence on the ring $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$ (Proposition 2.2.7) guarantees that these operations are well-defined.

Likewise, we can define the following relations on $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$:

- $x < y$ if and only if $\{n \in \mathbb{N} : x_n < y_n\} \in \mathcal{U}$.

  Claim: The relation $<$ is well-defined on $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$.

  Proof of claim: Let $x' = [(x'_n)_{n\in\mathbb{N}}]$ and $y' = [(y'_n)_{n\in\mathbb{N}}]$ be elements of $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$ such that $(x_n)_{n\in\mathbb{N}} \sim (x'_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}} \sim (y'_n)_{n\in\mathbb{N}}$. Then we have $I_1 = \{n \in \mathbb{N} : x_n = x'_n\} \in \mathcal{U}$ and $I_2 = \{n \in \mathbb{N} : y_n = y'_n\} \in \mathcal{U}$.

  Suppose $x < y$. Then $I_3 = \{n \in \mathbb{N} : x_n < y_n\} \in \mathcal{U}$. So $I_4 = I_1 \cap I_2 \cap I_3 \in \mathcal{U}$. That is $I_4 = \{n \in \mathbb{N} : x_n < y_n \text{ and } x_n = x'_n \text{ and } y_n = y'_n\} \in \mathcal{U}$. But $I_4 \subseteq \{n \in \mathbb{N} : x'_n < y'_n\}$. Then $\{n \in \mathbb{N} : x'_n < y'_n\} \in \mathcal{U}$. Thus $x' < y'$.

  Conversely, suppose $x' < y'$. Then $I'_3 = \{n \in \mathbb{N} : x'_n < y'_n\} \in \mathcal{U}$. So $I'_4 = I_1 \cap I_2 \cap I'_3 \in \mathcal{U}$. That is $I'_4 = \{n \in \mathbb{N} : x'_n < y'_n \text{ and } x_n = x'_n \text{ and } y_n = y'_n\} \in \mathcal{U}$. But $I'_4 \subseteq \{n \in \mathbb{N} : x_n < y_n\}$. Then $\{n \in \mathbb{N} : x_n < y_n\} \in \mathcal{U}$. Thus $x < y$.

  Therefore, $x < y$ if and only if $x' < y'$. Hence $<$ is well-defined on $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$.

- $x = y$ if and only if $\{n \in \mathbb{N} : x_n = y_n\} \in \mathcal{U}$

  The fact that the relation $=$ is well-defined follows immediately from the fact that $\sim$ is an equivalence relation.

2.2.10 Theorem (Field structure on $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$)**.** *The structure* $(\mathbb{R}^{\mathbb{N}}/\mathcal{U}, +, \cdot)$ *is an ordered field.*

Proof. Since the relation $\sim$ is a congruence on the ring $(\mathbb{R}^{\mathbb{N}}, +, \cdot)$, then $(\mathbb{R}^{\mathbb{N}}/\mathcal{U}, +, \cdot)$ is a commutative ring with zero $0 = [(0)_{n\in\mathbb{N}}]$.

Let $x = [(x_n)_{n\in\mathbb{N}}] \in \mathbb{R}^{\mathbb{N}}/\mathcal{U}$ such that $x \neq 0$. Let us define $y = [(y_n)_{n\in\mathbb{N}}] \in \mathbb{R}^{\mathbb{N}}/\mathcal{U}$ such that

$$\begin{cases} y_n = x_n & \text{if } x_n \neq 0 \\ y_n = 1 & \text{if } x_n = 0 \end{cases} \tag{2.2.1}$$

Claim: $x = y$

Proof of claim: We have $\{n \in \mathbb{N} : x_n = y_n\} = \mathbb{N} \setminus \{n \in \mathbb{N} : x_n = 0\}$. But $\{n \in \mathbb{N} : x_n = 0\} \notin \mathcal{U}$ since $x \neq 0$. So $\mathbb{N} \setminus \{n \in \mathbb{N} : x_n = 0\} \in \mathcal{U}$. Then $\{n \in \mathbb{N} : x_n = y_n\} \in \mathcal{U}$. Thus $x = y$. Hence this proves the claim.

Since $x = y$, we can define the multiplicative inverse of x as $x^{-1} = \left[\left(\frac{1}{y_n}\right)_{n\in\mathbb{N}}\right]$. It follows that $(\mathbb{R}^{\mathbb{N}}/\mathcal{U}, +, \cdot)$ is a field.

Moreover, let $x = [(x_n)_{n\in\mathbb{N}}]$ and $y = [(y_n)_{n\in\mathbb{N}}]$ be elements of $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$. The set $\mathbb{N}$ is the disjoint union of the following: $\{n \in \mathbb{N} : x_n < y_n\}$, $\{n \in \mathbb{N} : x_n = y_n\}$ and $\{n \in \mathbb{N} : y_n < x_n\}$. By Lemma 2.2.4, exactly one of these sets is in $\mathcal{U}$. Therefore, either $x < y$ or $x = y$ or $y < x$. Furthermore, one can derive that $(\mathbb{R}^{\mathbb{N}}/\mathcal{U}, +, \cdot)$ is an ordered field. □

2.2.11 Embedding $\mathbb{R}$ in $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$.

Let us define a function g from $\mathbb{R}$ to $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$ by $g(x) = [(x_n)_{n\in\mathbb{N}}]$, where $x_n = x$ for all $n \in \mathbb{N}$.

Claim 1: g is an injection.

Proof of Claim 1: Let $x, y \in \mathbb{R}$.

Suppose $g(x) = g(y)$. Then $[(x_n)_{n\in\mathbb{N}}] = [(y_n)_{n\in\mathbb{N}}]$, where $x_n = x$ and $y_n = y$ for all $n \in \mathbb{N}$. So $(x)_{n\in\mathbb{N}} \sim (y)_{n\in\mathbb{N}}$. Then $\{n \in \mathbb{N} : x = y\} \in \mathcal{U}$. But $\{n \in \mathbb{N} : x = y\}$ is either $\mathbb{N}$ (when $x = y$) or $\emptyset$ (when $x \neq y$). By the proper filter property of $\mathcal{U}$, we cannot have $x \neq y$. Therefore $x = y$. Hence g is an injection.

Claim 2: g is not surjective

Proof of Claim 2: Take $[(n)_{n\in\mathbb{N}}] \in \mathbb{R}^{\mathbb{N}}/\mathcal{U}$ and suppose $g(x) = [(n)_{n\in\mathbb{N}}]$ for some $x \in \mathbb{R}$. Then $[(n)_{n\in\mathbb{N}}] = [(x_n)_{n\in\mathbb{N}}]$, where $x_n = x$ for all $n \in \mathbb{N}$. So $\{n \in \mathbb{N} : n = x\} \in \mathcal{U}$. Then either $\{x\} \in \mathcal{U}$ or $\emptyset \in \mathcal{U}$. So $\{x\} \in \mathcal{U}$. That contradicts the fact that $\mathcal{U}$ is a non-principal ultrafilter (no finite subset of $\mathbb{N}$ is in $\mathcal{U}$). Therefore g is not surjective.

Hence we can identify $\mathbb{R}$ with $g(\mathbb{R}) \subsetneq \mathbb{R}^{\mathbb{N}}/\mathcal{U}$.

For ease of notation, we will write $^*\mathbb{R}$ instead of $\mathbb{R}^{\mathbb{N}}/\mathcal{U}$. We call $^*\mathbb{R}$ the set of non-standard real numbers or the set of hyperreal numbers. This set of hyperreal numbers includes not only (a copy of) the usual real numbers but also new kinds of numbers: infinitesimals and infinite numbers and finite hyperreals that are not real. The power of non-standard analysis leans on these new kinds of numbers.

Furthermore, it can be shown that $\mathbb{R}$ is a subfield of $^*\mathbb{R}$.

2.2.12 Remark. Since we can identify the embedded of the copy $g(\mathbb{R})$ with $\mathbb{R}$, if r is a real number, then its corresponding element $g(r) = [(r, r, r, \ldots)]$ in $^*\mathbb{R}$ is still denoted by r.

2.3 Infinitesimal and infinite numbers

In the previous section, we have mentioned that the set of hyperreals ${}^*\mathbb{R}$ includes new types of numbers namely, infinitesimal and infinite numbers. In this section, we will define clearly those types of numbers.

2.3.1 Definition.

- A hyperreal x is finite if and only if there exists some $r \in \mathbb{R}^+$ (where $\mathbb{R}^+$ is the set of positive real numbers) such that $|x| < r$. Otherwise, x is infinite.
- A hyperreal x is infinitesimal if $|x| < r$ for all $r \in \mathbb{R}^+$.
- Two hyperreals x and y are said to be infinitely close if $|x - y|$ is infinitesimal which we denote by $x \approx y$.

2.3.2 Remark.

- A hyperreal x is infinitesimal if x is infinitely close to zero, i.e. if $x \approx 0$.
- The only infinitesimal real number is zero.

2.3.3 Example.

- $H = [(n)_{n\in\mathbb{N}}]$ is an infinite hyperreal.
- $\epsilon = [(\frac{1}{n})_{n\in\mathbb{N}}]$ is infinitesimal.

2.3.4 Proposition (Equivalence relation on ${}^*\mathbb{R}$)**.** The relation $\approx$ defines an equivalence relation on ${}^*\mathbb{R}$.

Proof. Let $x, y, z \in {}^*\mathbb{R}$

Reflexivity: We have $x - x = 0$ which is infinitesimal. So $x \approx x$.

Symmetry: Suppose $x \approx y$ then $|x - y|$ is infinitesimal. But $|x - y| = |y - x|$. So $|y - x|$ is also infinitesimal. Then $y \approx x$.

Transitivity: Suppose $x \approx y$ and $y \approx x$. Then $|x - y|$ and $|y - x|$ are both infinitesimal. Let $r \in \mathbb{R}^+$. Then $|x - y| < \frac{r}{2}$ and $|y - x| < \frac{r}{2}$. But $|x - z| \leq |x - y| + |y - z|$ (Triangle inequality). So $|x - z| < r$. Thus $|x - z|$ must be infinitesimal because r was chosen arbitrarily. Then $x \approx z$.

Therefore $\approx$ is an equivalence relation. □

2.4 Arithmetic of hyperreals

Here we are using usual arithmetic operations on the set ${}^*\mathbb{R}$ of hyperreals.

2.4.1 Theorem. *(Keisler, 1986; Goldblatt, 2012) Consider hyperreals $\epsilon, \delta, b, c, H$ and K such that $\epsilon \approx 0$ and $\delta \approx 0$, b and c are finite but not infinitesimal, H and K are infinite. Then we have:*

- *Additive inverse: $-\epsilon \approx 0$, $-b$ is finite and $-H$ is infinite.*
- *Sum: $\epsilon + \delta \approx 0$, $b + \epsilon$ is finite but not infinitesimal, $b + c$ is finite; $H + \epsilon$ and $H + b$ are infinite.*
- *Product: $\delta \cdot \epsilon \approx 0$ and $b \cdot \epsilon \approx 0$, $b \cdot c$ is finite but not infinitesimal; $H \cdot b$ and $H \cdot K$ are infinite.*

- *Quotient:* $\epsilon/b \approx 0$, $\epsilon/H \approx 0$ *and* $b/H \approx 0$; b/c *is finite but not infinitesimal;* b/ϵ, H/ϵ *and* H/b *are infinite* $(\epsilon \neq 0)$.
- *Reciprocal:* $1/\epsilon$ *is infinite (*$\epsilon \neq 0$*),* $1/b$ *is finite but but not infinitesimal,* $1/H \approx 0$.
- *Indeterminate forms:* ϵ/δ, H/K, $\epsilon.H$ *and* $H + K$ *are indeterminate forms - they can be infinitesimal, finite or infinite.*

  Examples: $x = \frac{H}{H^2}$ *and* $y = \frac{H^2}{H}$ *are both ratios of infinite numbers but* x *is infinitesimal and* y *is infinite.*

The proofs of theses properties of hyperreals are straightforward. We will prove some of them, the rest can be proved following similar arguments.

Proof.

- *Sum*: Take $r \in \mathbb{R}^+$. Since ϵ and δ are infinitesimal, we have $|\epsilon| < \frac{r}{2}$ and $|\delta| < \frac{r}{2}$. But $|\epsilon + \delta| \leq |\epsilon| + |\delta|$. So $|\epsilon + \delta| < r$. Hence $\epsilon + \delta \approx 0$ since r was chosen arbitrarily.

  Furthermore, since b is finite, then there exists some $s \in \mathbb{R}$ such that $|b| < s$. But $|\epsilon + b| \leq |\epsilon| + |b|$. So $|\epsilon + b| < r + s$ for all $r \in \mathbb{R}^+$ and some $s \in \mathbb{R}$. Take $m = 1 + s$. Then $|\epsilon + b| < m$. Hence $\epsilon + b$ is finite. Now, we need to show that $\epsilon + b$ cannot be infinitesimal. Suppose $\epsilon + b \approx 0$. Then $b \approx -\epsilon \approx 0$. This contradicts the fact that $b \not\approx 0$. So $\epsilon + b \not\approx 0$.
- *Product*: Take $r \in \mathbb{R}^+$. We have $|b| < s$ for some $s \in \mathbb{R}^+$ and $|\epsilon| < \frac{r}{s}$. But $|\epsilon \cdot b| \leq |\epsilon| \cdot |b|$. So $|\epsilon \cdot b| < r$. Hence $\epsilon \cdot b \approx 0$ since r was chosen arbitrarily.

□

2.4.2 Examples.

Let $H = [(n)_{n\in\mathbb{N}}]$ and $\epsilon = [(\frac{1}{n})_{n\in\mathbb{N}}]$.

- Let $x = \dfrac{\epsilon}{H - 2\epsilon}$.

  Since ϵ is infinitesimal and $H - 2\epsilon$ is infinite, then x is infinitesimal.
- Let $y = \dfrac{H}{3 + \epsilon}$.

  Since H is infinite and $3 + \epsilon$ is finite (but not infinitesimal), then y is infinite.
- Let $z = H^2 + 3\epsilon + 1$. Then z is infinite.

2.5 Standard part

2.5.1 Theorem (Standard Part Principle)**.** *Every finite hyperreal* x *is infinitely close to a unique real number called the standard part of* x*, denoted by* $st(x)$ *(sometimes called the shadow* $sh(x)$*).*

Proof.

- *Existence:* Let x be a finite hyperreal. Consider the set $A = \{t \in \mathbb{R} : t < x\} \subseteq \mathbb{R}$.

Since x is finite, then there exists some $r \in \mathbb{R}$ such that $|x| < r$. So $-r < x < r$. Thus, A is not empty (because $-r \in A$). Moreover, for all $t \in A$, we have $t < x < r$. So A is bounded above by r. By the completeness of real numbers, A must have a least upper bound.

Let s be the least upper bound of A. Given $\epsilon \in \mathbb{R}^+$, we have $x < s + \epsilon$ (otherwise, will have $s + \frac{\epsilon}{2} \in A$ which contradicts the fact that s is an upper bound of A). Moreover, we have $s - \epsilon < t < x$ for some $t \in A$ (otherwise, we would have $t \leq s - \epsilon < s$ for all $t \in A$ which contradicts the fact that s is the least upper bound of A in $\mathbb{R}$).

It follows that, for all $\epsilon \in \mathbb{R}^+$, $s - \epsilon < x < s + \epsilon$. So $|x - s| < \epsilon$ for all ϵ in $\mathbb{R}^+$. Then $|x - s| \approx 0$. So $x \approx s$.

- *Uniqueness:* Suppose there exists another real s' such that $x \approx s'$. Then we have $x \approx s'$ and $x \approx s$. So $s \approx s'$ because $\approx$ is transitive. Then $s - s' \approx 0$. But s and s' are both real numbers. So $s - s'$ is also a real number. Then $s - s' = 0$ (since zero is the unique infinitesimal real number). Hence $s = s'$.

□

2.5.2 Corollary. Every finite hyperreal x can be uniquely written as $x = st(x) + \epsilon$, for some infinitesimal ϵ.

Proof. Let x be a finite hyperreal. From Theorem 2.5.1, $x \approx st(x)$. Then for $\epsilon = x - st(x)$, $\epsilon \approx 0$ and $x = \epsilon + st(x)$.

Suppose there is another infinitesimal ϵ' such that $x = st(x) + \epsilon'$. Then $x = st(x) + \epsilon$ and $x = st(x) + \epsilon'$. So $st(x) + \epsilon = st(x) + \epsilon'$. Hence $\epsilon = \epsilon'$. □

2.5.3 Theorem. *For all finite hyperreals x and y, we have:*

1. $x \approx y$ *if and only if* $st(x) = st(y)$
2. $st(x) = x$ *if and only if* $x \in \mathbb{R}$
3. *If* $x \leq y$ *then* $st(x) \leq st(y)$
4. $st(x + y) = st(x) + st(y)$
5. $st(x \cdot y) = st(x) \cdot st(y)$
6. *If* $st(y) \neq 0$ *then* $st(\frac{x}{y}) = \frac{st(x)}{st(y)}$

Proof.

1. Let x and y be finite hyperreals. Suppose $x \approx y$. We have $st(x) \approx x \approx y \approx st(y)$. So $st(x) \approx st(y)$ (by transitivity of $\approx$). Thus $st(x) - st(y) \approx 0$. But $st(x)$ and $st(y)$ are both real, hence $st(x) = st(y)$.

   Conversely, suppose $st(x) = st(y)$. Then $x \approx st(x) = st(y) \approx y$. So $x \approx y$ (by transitivity of $\approx$). Hence $x \approx y$ if and only if $st(x) = st(y)$.

2. Assume $x \in \mathbb{R}$. We have $st(x) \approx x$. So $x = st(x)$ since x and $st(x)$ are both real. Conversely, if $st(x) = x$ then $x \in \mathbb{R}$ since $st(x) \in \mathbb{R}$ (by definition). Hence $x \in \mathbb{R}$ if and only if $st(x) = x$.

3. Let x and y be finite hyperreals such that $x \leq y$.

Suppose $st(x) > st(y)$. Let $st(x) - st(y) = \delta \in \mathbb{R}^+$. Take $\epsilon \in \mathbb{R}^+$ such that $\epsilon < \delta$. Then $|st(y) - y| < \frac{\delta-\epsilon}{2}$ (since $|st(y) - y|$ is infinitesimal and $\frac{\delta-\epsilon}{2}$ is a positive real number). So $\frac{\epsilon-\delta}{2} < st(y) - y < \frac{\delta-\epsilon}{2}$.

Moreover, $st(x)-x = (st(x)-st(y))+(st(y)-y)+(y-x)$. But $st(x)-st(y) = \delta$, $st(y)-y > \frac{\epsilon-\delta}{2}$ and $y - x \geq 0$. So $st(x) - x > \delta + \frac{\epsilon-\delta}{2} = \frac{\epsilon+\delta}{2} > \epsilon$. This contradicts the fact that $st(x) - x$ is infinitesimal. Therefore, $st(x) \leq st(y)$.

4. We have $st(x + y) = (x + y) + \epsilon$ for some infinitesimal ϵ. But $x = st(x) + \epsilon_1$ and $y = st(y) + \epsilon_2$ for some infinitesimals ϵ_1 and ϵ_2. So $st(x + y) = st(x) + st(y) + \epsilon_1 + \epsilon_2 + \epsilon$. But the sum of infinitesimals is infinitesimal, hence $st(x + y) \approx st(x) + st(y)$. Hence $st(x + y) = st(x) + st(y)$.

5. We have $st(x \cdot y) = (x \cdot y) + \epsilon$ for some infinitesimal ϵ. But $x = st(x) + \epsilon_1$ and $y = st(y) + \epsilon_2$ for some infinitesimals ϵ_1 and ϵ_2. So $st(x\cdot y) = (st(x)+\epsilon_1)\cdot(st(y)+\epsilon_2)+\epsilon = \epsilon_1\epsilon_2+\epsilon_1 st(y)+\epsilon_2 st(x)+st(x)st(y) + \epsilon$. But $\epsilon, \epsilon_1\epsilon_2, \epsilon_1 st(y)$ and $\epsilon_2 st(x)$ are infinitesimal. So $st(x \cdot y) \approx st(x)st(y)$. Hence $st(x \cdot y) = st(x)st(y)$.

6. Let $X = st(\frac{x}{y}) - \frac{st(x)}{st(y)}$. We have:

$$\begin{aligned} X &= \left(\frac{x}{y} + \epsilon\right) - \frac{st(x)}{st(y)} \quad \text{for some infinitesimal } \epsilon \\ &= \left(\frac{st(x) + \epsilon_1}{st(y) + \epsilon_2} + \epsilon\right) - \frac{st(x)}{st(y)} \quad \text{for some infinitesimals } \epsilon_1 \text{ and } \epsilon_2 \\ &= \frac{st(x) + \epsilon_1 + \epsilon st(y) + \epsilon\epsilon_2}{st(y) + \epsilon_2} - \frac{st(x)}{st(y)} \\ &= \frac{\epsilon_1 st(y) + \epsilon(st(y))^2 + \epsilon\epsilon_2 st(y) - \epsilon_2 st(x)}{(st(y))^2 + \epsilon_2 st(y)} \end{aligned}$$

But $\epsilon_1 st(y) + \epsilon(st(y))^2 + \epsilon\epsilon_2 st(y) - \epsilon_2 st(x) \approx 0$ (sum of infinitesimals) and $(st(y))^2 + \epsilon_2 st(y)$ is finite and not infinitesimal. So $X \approx 0$. Thus $st(\frac{x}{y}) \approx \frac{st(x)}{st(y)}$. Hence $st(\frac{x}{y}) = \frac{st(x)}{st(y)}$.

□

2.6 Non-standard extensions

2.6.1 Definition (Extension of a set)**.** Let A be subset of $\mathbb{R}$. We can extend A to some ${}^*A \subseteq {}^*\mathbb{R}$ as follows: ${}^*A = [(A)_{n\in\mathbb{N}}] = \{x \in {}^*\mathbb{R} : \{n \in \mathbb{N} : x_n \in A\} \in \mathcal{U}\}$, where $x = [(x_n)_{n\in\mathbb{N}}] \in {}^*\mathbb{R}$. We call ${}^*A$ the non-standard extension (or $*$-transform) of A in ${}^*\mathbb{R}$.

2.6.2 Proposition. The set ${}^*A$ is well-defined.

Proof. Let $x = [(x_n)_{n\in\mathbb{N}}]$ and $x' = [(x'_n)_{n\in\mathbb{N}}]$ be elements of ${}^*\mathbb{R}$ such that $(x_n)_{n\in\mathbb{N}} \sim (x'_n)_{n\in\mathbb{N}}$. Suppose $x \in {}^*A$. Then $I_1 = \{n \in \mathbb{N} : x_n \in A\} \in \mathcal{U}$. But since $(x_n)_{n\in\mathbb{N}} \sim (x'_n)_{n\in\mathbb{N}}$, then $I_2 = \{n \in \mathbb{N} : x_n = x'_n\} \in \mathcal{U}$. So $I_3 = I_1 \cap I_2 \in \mathcal{U}$. That is, $I_3 = \{n \in \mathbb{N} : x_n \in A \text{ and } x_n = x'_n\} \in \mathcal{U}$. But $I_3 \subseteq \{n \in \mathbb{N} : x'_n \in A\}$. So $\{n \in \mathbb{N} : x'_n \in A\} \in \mathcal{U}$. Thus $x' \in {}^*A$. Similarly, assuming that $x' \in {}^*A$ leads to $x \in {}^*A$.

Therefore, ${}^*A$ is a well-defined set. □

2.6.3 Theorem. *For all $A, B \subseteq \mathbb{R}$, we have:*

1. $A \subseteq B$ *if and only if* ${}^*A \subseteq {}^*B$
2. $A = B$ *if and only if* ${}^*A = {}^*B$
3. ${}^*(A \cup B) = {}^*A \cup {}^*B$
4. ${}^*(A \cap B) = {}^*A \cap {}^*B$
5. ${}^*(A \setminus B) = {}^*A \setminus {}^*B$
6. ${}^*\emptyset = \emptyset$

Proof. We shall prove Property 1 and Property 4. The rest can be proved in a similar way.

Let $x = [(x_n)_{n\in\mathbb{N}}]$. We have ${}^*A = [(A)_{n\in\mathbb{N}}] = \{x \in {}^*\mathbb{R} : \{n \in \mathbb{N} : x_n \in A\} \in \mathcal{U}\}$ and ${}^*B = [(B)_{n\in\mathbb{N}}] = \{x \in {}^*\mathbb{R} : \{n \in \mathbb{N} : x_n \in B\} \in \mathcal{U}\}$.

1. $A \subseteq B$ if and only if ${}^*A \subseteq {}^*B$

   Suppose $A \subseteq B$. Assume $x \in {}^*A$. Then $\{n \in \mathbb{N} : x_n \in A\} \in \mathcal{U}$. But $A \subseteq B$. So $\{n \in \mathbb{N} : x_n \in A\} \subseteq \{n \in \mathbb{N} : x_n \in B\} \subseteq \mathbb{N}$. Then $\{n \in \mathbb{N} : x_n \in B\} \in \mathcal{U}$ (Superset property of $\mathcal{U}$). So $x \in {}^*B$. Thus ${}^*A \subseteq {}^*B$.

   Now let us suppose ${}^*A \subseteq {}^*B$. Assume $A \nsubseteq B$. Then there exists some $y \in A \setminus B$. So $y = [(x)_{n\in\mathbb{N}}] \in {}^*A$ and $y \notin {}^*B$. This contradicts the fact that ${}^*A \subseteq {}^*B$. Hence $A \subseteq B$. This completes the proof.

4. ${}^*(A \cap B) = {}^*A \cap {}^*B$

   Let us first prove ${}^*(A \cap B) \subseteq {}^*A \cap {}^*B$.

   Let $x \in {}^*(A \cap B)$. Then $\{n \in \mathbb{N} : x_n \in (A \cap B)\} \in \mathcal{U}$. But $\{n \in \mathbb{N} : x_n \in (A \cap B)\} \subseteq \{n \in \mathbb{N} : x_n \in A\}$. So $\{n \in \mathbb{N} : x_n \in A\} \in \mathcal{U}$ (superset property of $\mathcal{U}$). Thus $x \in {}^*A$. Similarly, we get $x \in {}^*B$. Therefore, $x \in {}^*A \cap {}^*B$. Hence ${}^*(A \cap B) \subseteq {}^*A \cap {}^*B$.

   Now let us prove ${}^*A \cap {}^*B \subseteq {}^*(A \cap B)$.

   Let $x \in {}^*A \cap {}^*B$. So $x \in {}^*A$ and $x \in {}^*B$. Then $\{n \in \mathbb{N} : x_n \in A\} \in \mathcal{U}$ and $\{n \in \mathbb{N} : x_n \in B\} \in \mathcal{U}$. So $\{n \in \mathbb{N} : x_n \in A\} \cap \{n \in \mathbb{N} : x_n \in B\} \in \mathcal{U}$. Then $\{n \in \mathbb{N} : x_n \in (A \cap B)\} \in \mathcal{U}$. Thus $x \in {}^*(A \cap B)$. Hence ${}^*A \cap {}^*B \subseteq {}^*(A \cap B)$.

□

2.6.4 Remark. The set ${}^*\mathbb{R}$ contains ${}^*\mathbb{N}$ (set of hypernatural numbers), ${}^*\mathbb{Z}$ (set of hyperinteger numbers) and ${}^*\mathbb{Q}$ (set of hyperrational numbers).

2.6.5 Extension of functions.

Consider a function $f : \mathbb{R} \to \mathbb{R}$. The non-standard extension (or $*$-transform) of f is a function ${}^*f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ defined as follows:

Let $y = [(y_n)_{n\in\mathbb{N}}] \in {}^*\mathbb{R}$. We have ${}^*f(x) = y$ if and only if $\{n \in \mathbb{N} : f(x_n) = y_n\} \in \mathcal{U}$.

Claim: ${}^*f$ is well-defined.

Proof of Claim: Take $x = [(x_n)_{n\in\mathbb{N}}]$, $x' = [(x'_n)_{n\in\mathbb{N}}]$, $y = [(y_n)_{n\in\mathbb{N}}]$ and $y' = [(y'_n)_{n\in\mathbb{N}}]$ in ${}^*\mathbb{R}$.

Suppose $x = x'$. Assume ${}^*f(x) = y$ and ${}^*f(x') = y'$.

Then we have: $I_1 = \{n \in \mathbb{N} : f(x_n) = y_n\} \in \mathcal{U}$ (since ${}^*f(x) = y$), $I_2 = \{n \in \mathbb{N} : f(x'_n) = y'_n\} \in \mathcal{U}$ (since ${}^*f(x') = y'$) and $I_3 = \{n \in \mathbb{N} : x_n = x'_n\} \in \mathcal{U}$ (since $x = x'$).

So $I_4 = I_1 \cap I_2 \cap I_3 \in \mathcal{U}$. That is $\{n \in \mathbb{N} : f(x_n) = y_n$ and $f(x'_n) = y'_n$ and $x_n = x'_n\} \in \mathcal{U}$. But $I_4 \subseteq \{n \in \mathbb{N} : y_n = y'_n\}$. So $\{n \in \mathbb{N} : y_n = y'_n\} \in \mathcal{U}$. Then $y = y'$. Hence ${}^*f$ is well-defined.

2.6.6 Remark. For all $x = [(x_n)_{n\in\mathbb{N}}] \in {}^*\mathbb{R}$, we have ${}^*f(x) = [(f(x_n))_{n\in\mathbb{N}}]$.

3. The Transfer Principle

In this chapter, we are going to introduce the main result of non-standard analysis: the Transfer Principle. This principle builds a bridge between non-standard and standard analysis. It shows how results from $^*\mathbb{R}$ can be rigorously carried over to $\mathbb{R}$ and vice-versa.

In order to state the Transfer Principle, we need to define some terminology.

For the definitions and results in this chapter, we mostly follow the presentation in Boxall (2016).

3.1 Structures

3.1.1 Definition (Language)**.** A language is a pair (L, p), where L is a set of relation symbols (also called predicates) and function symbols; p is a function from L to $\mathbb{N}$. Some examples of predicates are: $=, <$ and function symbols are: $+, -, \times, 1, 0$.

The function p assigns to each element R of L a natural number n called the arity of R. We say R is an n-ary predicate, or function symbol.

For ease of notation, we shall refer to a language by L instead of (L, p).

3.1.2 Definition (Interpretation)**.** Let L be a language. Let M be a non-empty set. We define an interpretation I of L in M as follows:

For each element of L, I associates a function $I(f) : M^n \to M$ if f is an n-ary function symbol and $I(R) \subseteq M^n$ if R is a n-ary predicate.

3.1.3 Definition (L-structure)**.** Let L be a language. An L-structure is a pair (M, I) where M is a non-empty set and I is an interpretation of L in M.

3.2 Ultrapower of structures

Let J be a non-empty set. Let $(M_j)_{j \in J}$ be a non-empty family of non-empty sets. Let $\mathcal{U}$ be an ultrafiter on J. Let L be a language. For each $j \in J$, let I_j be an interpretation of L in M_j such that $(M_j, I_j)_{j \in J}$ defines a family of L-structures.

We define an interpretation $I_\mathcal{U}$ of L in $\prod_{j \in J} M_j/\mathcal{U}$ as follows:

- for each n-ary predicate $R \in L$, $I_\mathcal{U}(R) = [(I_j(R))_{j \in J}]$ where $[(I_j(R))_{j \in J}] = \{([a^1], \dots, [a^n]) \in (\prod_{j \in J} M_j/\mathcal{U})^n : \{j \in J : (a_j^1, \dots, a_j^n) \in I_j(R)\} \in \mathcal{U}\}$ (Boxall, 2016).
- for each n-ary function symbol $f \in L$, $I_\mathcal{U}(f) : (\prod_{j \in J} M_j/\mathcal{U})^n \to \prod_{j \in J} M_j/\mathcal{U}$ such that $([a^1], \dots, [a^n]) \mapsto [(I_j(f)(a_j^1, \dots, a_j^n))_{j \in J}]$.

We call $\left(\prod_{j \in J} M_j/\mathcal{U}, I_\mathcal{U}\right)$ the ultraproduct structure of $(M_j, I_j)_{j \in J}$ with respect to $\mathcal{U}$.

3.3 Formulas

3.3.1 First-order L-formulas.

Let L be a language. A first-order L-formula is a finite sequence of symbols such that each of these symbols is one of the following:

- a bracket: $(,)$
- a comma : ,
- a variable: $x_1, x_2, x_3, \ldots$
- an element of L
- a connective: $\wedge, \vee, \neg, \rightarrow$
- a quantifier: $\exists, \forall$

But not all such sequences are first-order L-formulas.

The concept of first-order formula is inductively defined by the basic ones called atomic L-formulas and then giving the rules by which new first-order L-formulas may be formed (Boxall, 2016).

3.3.2 Definition (Terms)**.** We define terms as follows:

- Variables $x_1, \ldots, x_n$ are terms.
- If $x_1, \ldots, x_n$ are terms and f is an n-ary function symbols, then $f(x_1, \ldots, x_n)$ is a term.

3.3.3 Definition (Atomic L-formulas)**.** Let $n \in \mathbb{N}$. Let L be a language. An atomic L-formula is an expression of the form:

- $R(x_1, \ldots, x_n)$, where $x_1, \ldots, x_n$ are terms and R an n-ary predicate of L.
- $s = t$, where s and t are terms.

3.3.4 Definition (L-formulas)**.**

- Every atomic L-formula is a first-order L-formula.
- If φ and ψ are first-order L-formulas, then so are $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\neg\varphi)$, $(\varphi \longrightarrow \psi)$, $(\exists x\varphi)$ and $(\forall x\varphi)$.

3.3.5 Remark (Defining sets with formulas)**.** Let L be a language. Let (M, I) be an L-structure. An L-formula φ together with an n-tuple of variables containing every free variable (unquantified) of φ define a set $X \subseteq M^n$.

3.3.6 Definition (L-sentences)**.** Let L be a language. Let (M, I) be an L-structure. A first-order L-sentence is a first-order L-formula σ with no free variables (i.e. all variables of σ are bound by quantifiers).

A first-order L-sentence σ together with the empty tuple () defines a subset of M^0. So $\sigma()$ is either $\emptyset$ or M^0. If $\sigma()$ defines M^0, we say σ is true in (M, I) and we write $(M, I) \models \sigma$. Otherwise, we say σ is false in (M, I).

3.4 Łoś's Lemma

3.4.1 Lemma (Łoś's Lemma)**.** Let L be a language. Let J be a non-empty set. For each $j \in J$, let (M_j, I_j) be an L-structure. Let $\mathcal{U}$ be an ultrafilter on J. Let $\left(\prod_{j\in J} M_j/\mathcal{U}, I_{\mathcal{U}}\right)$ be the ultraproduct of $(M_j, I_j)_{j\in J}$ with respect to $\mathcal{U}$. Let φ be a first-order L-formula. Let $x_{i_1}, \ldots, x_{i_n}$ be such that all free variables of φ appear in $x_{i_1}, \ldots, x_{i_n}$.

For each $j \in J$, let $X_j \subseteq M_j^n$ be the set defined by $\varphi(x_{i_1}, \ldots, x_{i_n})$ in (M_j, I_j). Let $X \subseteq \left(\prod_{j\in J} M_j/\mathcal{U}\right)^n$ be the set defined by $\varphi(x_{i_1}, \ldots, x_{i_n})$ in $\left(\prod_{j\in J} M_j/\mathcal{U}, I_{\mathcal{U}}\right)$.

Then $X = [(X_j)_{j\in J}]$ where $[(X_j)_{j\in J}] = \{([a^1], \ldots, [a^n]) \in (\prod_{j\in J} M_j/\mathcal{U})^n : \{j \in J : (a_j^1, \ldots, a_j^n) \in X_j\} \in \mathcal{U}\}$.

Proof. This lemma is stated here without proof but a proof can be found in Boxall (2016). □

3.4.2 Corollary. Let $L, J, \mathcal{U}$ and $(M_j, I_j)_{j\in J}$ be defined as in Lemma 3.4.1. Let σ be a first-order L-sentence.

Then $\left(\prod_{j\in J} M_j/\mathcal{U}, I_{\mathcal{U}}\right) \models \sigma$ if and only if $\{j \in J : (M_j, I_j) \models \sigma\} \in \mathcal{U}$.

Proof. Consider the Łoś's Lemma (3.4.1) in the special case where $n = 0$.

Let σ be a first-order L-sentence. For each $j \in J$, let $X_j \subseteq M_j^0$ be the set defined by $\sigma()$ in (M_j, I_j). Let $X \subseteq \left(\prod_{j\in J} M_j/\mathcal{U}\right)^0$ be the set defined by $\sigma()$ in $\left(\prod_{j\in J} M_j/\mathcal{U}, I_{\mathcal{U}}\right)$.

Then by Łoś's Lemma, we have $X = [(X_j)_{j\in J}]$. So $X = \{() \in (\prod_{j\in J} M_j/\mathcal{U})^0 : \{j \in J : () \in X_j\} \in \mathcal{U}\}$. Then $X = (\prod_{j\in J} M_j/\mathcal{U})^0$ if and only if $\{j \in J : () \in X_j\} \in \mathcal{U}$ for all $() \in (\prod_{j\in J} M_j/\mathcal{U})^0$. That is the case if and only if $\{j \in J : X_j = M_j^0\} \in \mathcal{U}$. So $\sigma()$ defines $\left(\prod_{j\in J} M_j/\mathcal{U}\right)^0$ if and only if $\{j \in J : (M_j, I_j) \models \sigma\} \in \mathcal{U}$.

Hence, $\left(\prod_{j\in J} M_j/\mathcal{U}, I_{\mathcal{U}}\right) \models \sigma$ if and only if $\{j \in J : (M_j, I_j) \models \sigma\} \in \mathcal{U}$. □

3.5 Transfer Principle

Consider a language $L = \{=, <, +, -, \times, 1, 0, \ldots\}$ and an interpretation I of L in $\mathbb{R}$.

3.5.1 Theorem (Transfer Principle)**.** *Let σ be a first order L-sentence. Then σ is true in $\mathbb{R}$ if and only if it is true in ${}^*\mathbb{R}$.*

Proof. Consider the special case of Corollary 3.4.2 where $J = \mathbb{N}$, $M_j = \mathbb{R}$ and $I_j = I$ for all $j \in \mathbb{N}$.

Then we have $(^*\mathbb{R}, I_{\mathcal{U}}) \models \sigma$ if and only if $\{j \in \mathbb{N} : (\mathbb{R}, I) \models \sigma\} \in \mathcal{U}$. But $\{j \in \mathbb{N} : (\mathbb{R}, I) \models \sigma\}$ is either $\mathbb{N}$ (when $(\mathbb{R}, I) \models \sigma$) or $\emptyset$ (when $(\mathbb{R}, I) \not\models \sigma$). Then by the proper filter property of $\mathcal{U}$, we cannot have $(\mathbb{R}, I) \not\models \sigma$ when $(^*\mathbb{R}, I_{\mathcal{U}}) \models \sigma$.

So $(^*\mathbb{R}, I_{\mathcal{U}}) \models \sigma$ if and only if $(\mathbb{R}, I) \models \sigma$. Hence the Transfer Principle is proved. □

As a direct application of the Transfer Principle, we can show that the set $^*\mathbb{N} \setminus \mathbb{N}$ contains no finite element.

3.5.2 Proposition. $^*\mathbb{N} \setminus \mathbb{N}$ is the set of infinite hypernatural numbers.

Proof. Let $n \in \mathbb{N}$. We know that $(\forall x \in \mathbb{N})((n \leq x < n+1) \longrightarrow x = n)$. By transfer, we have $(\forall x \in {}^*\mathbb{N})((n \leq x < n+1) \longrightarrow x = n)$. So $(\forall x \in {}^*\mathbb{N})(x \neq n \longrightarrow (x < n \vee x \geq n+1))$. Thus for all $x \in {}^*\mathbb{N}$, if $x \notin \mathbb{N}$ then x is infinite since n was chosen arbitrary (and we cannot have $x < 0$ because 0 is the minimum of $\mathbb{N}$ and the minimum of $^*\mathbb{N}$ by transfer). Therefore, for all $x \in {}^*\mathbb{N}$, if $x \in {}^*\mathbb{N} \setminus \mathbb{N}$ then x is infinite. This finishes the proof. □

4. Basic applications in Calculus

One of the main uses of non-standard analysis is to give alternative definitions of basic concepts of standard analysis that are more appealing. In this chapter, we are going to explore how non-standard analysis can be applied to Calculus.

4.1 Limits

4.1.1 Definition (Standard definition)**.** Consider a function $f : \mathbb{R} \to \mathbb{R}$. Let $l, c \in \mathbb{R}$. We say $f(x)$ has limit l as x approaches c, and we write $\lim_{x \to c} f(x) = l$, if and only if $(\forall \epsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})(0 < |x - c| < \delta \longrightarrow |f(x) - l| < \epsilon)$.

4.1.2 Proposition (Equivalent definition to 4.1.1)**.** Consider a function $f : \mathbb{R} \to \mathbb{R}$. Let $l, c \in \mathbb{R}$. Then $\lim_{x \to c} f(x) = l$ if and only if for all $x \in {}^*\mathbb{R}$ such that $x \approx c$ (but $x \neq c$), we have ${}^*f(x) \approx l$.

Proof. Suppose $\lim_{x \to c} f(x) = l$. Let $\epsilon \in \mathbb{R}^+$. Take $\delta \in \mathbb{R}^+$ such that $(\forall x \in \mathbb{R})(0 < |x - c| < \delta \longrightarrow |f(x) - l| < \epsilon)$. By transfer, we have $(\forall x \in {}^*\mathbb{R})(0 < |x - c| < \delta \longrightarrow |{}^*f(x) - l| < \epsilon)$.

If $x \approx c$ (but $x \neq c$) then $|x - c| \approx 0$. In particular $|x - c| < \delta$. Then $|{}^*f(x) - l| < \epsilon$. So $|{}^*f(x) - l| \approx 0$ since ϵ was taken arbitrarily. Thus ${}^*f(x) \approx l$ whenever $x \approx c$ and $x \neq c$.

Conversely, suppose ${}^*f(x) \approx l$ whenever $x \approx c$ (but $x \neq c$). Let $\epsilon \in \mathbb{R}^+$. Take $\delta \in {}^*\mathbb{R}^+$ such that $\delta \approx 0$. Then for all $x \in {}^*\mathbb{R}$, we have:

If $|x - c| < \delta$ (but $x \neq c$), then $|x - c| \approx 0$. Hence ${}^*f(x) \approx l$. So $|{}^*f(x) - l| \approx 0$. In particular, $|{}^*f(x) - l| < \epsilon$. Hence the following sentence is true in ${}^*\mathbb{R}$: $(\exists \delta \in {}^*\mathbb{R}^+)(\forall x \in {}^*\mathbb{R})(|x - c| < \delta \longrightarrow |{}^*f(x) - l| < \epsilon)$.

By transfer, the following is also true in $\mathbb{R}$: $(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})(0 < |x - c| < \delta \longrightarrow |f(x) - l| < \epsilon)$. Since ϵ was chosen arbitrarily and δ accordingly, then we have: $(\forall \epsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})(0 < |x - c| < \delta \longrightarrow |f(x) - l| < \epsilon)$. Therefore, $\lim_{x \to c} f(x) = l$. □

4.2 Continuity

4.2.1 Definition (Standard definition)**.** Consider a function $f : \mathbb{R} \to \mathbb{R}$. We say f is continuous at $c \in \mathbb{R}$ if and only if for any $\epsilon \in \mathbb{R}^+$, there exists $\delta \in \mathbb{R}^+$ such that $(\forall x \in \mathbb{R})(|x - c| < \delta \to |f(x) - f(c)| < \epsilon)$.

4.2.2 Proposition (Equivalent definition to 4.2.1)**.** Consider a function $f : \mathbb{R} \to \mathbb{R}$. Then f is continuous at $c \in \mathbb{R}$ if and only if ${}^*f(c + \epsilon) \approx {}^*f(c)$ for any infinitesimal ϵ. Or equivalently, f is continuous at $c \in \mathbb{R}$ if and only if ${}^*f(x) \approx {}^*f(c)$ whenever $x \approx c$.

Proof. Suppose f is continuous at $c \in \mathbb{R}$. Let $\epsilon \in \mathbb{R}^+$. Take $\delta \in \mathbb{R}^+$ such that $(\forall x \in \mathbb{R})(|x - c| < \delta \to |f(x) - f(c)| < \epsilon)$.

By Transfer Principle, the following is true in ${}^*\mathbb{R}$: $(\forall x \in {}^*\mathbb{R})(|x - c| < \delta \to |{}^*f(x) - {}^*f(c)| < \epsilon)$.

For any $x \in {}^*\mathbb{R}$ such that $x \approx c$, we always have $0 < |x - c| < \delta$ (since δ is not infinitesimal). This implies that $|{}^*f(x) - {}^*f(c)| < \epsilon$. But ϵ was arbitrarily chosen. So $|{}^*f(x) - {}^*f(c)| \approx 0$. Thus

${}^*f(x) \approx {}^*f(c)$.

Conversely, suppose ${}^*f(x) \approx {}^*f(c)$ whenever $x \approx c$. Let $\epsilon \in \mathbb{R}^+$. Let δ be a positive infinitesimal. If $0 < |x - c| < \delta$ then $|x - c| \approx 0$. So $|{}^*f(x) - {}^*f(c)| \approx 0$. In particular $|{}^*f(x) - {}^*f(c)| < \epsilon$. So the sentence $(\exists\delta \in {}^*\mathbb{R}^+)(\forall x \in {}^*\mathbb{R})(|x - c| < \delta \to |{}^*f(x) - {}^*f(c)| < \epsilon)$ is true in ${}^*\mathbb{R}$. By transfer, $(\exists\delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})(|x - c| < \delta \to |f(x) - f(c)| < \epsilon)$ is true in $\mathbb{R}$. Hence f is continuous at $c \in \mathbb{R}$. This completes the proof. □

As a basic application of non-standard analysis, we can quickly prove the Composite Functions Theorem.

4.2.3 Theorem (Composite Functions Theorem)**.** *(Ponstein, 2001) Let f and g be real valued functions with $g(w)$ defined for w in a neighbourhood of $c \in \mathbb{R}$, and $f(x)$ defined for x in a neighbourhood of $g(c)$. Then $f \circ g$ is continuous at c if g is continuous at c and f is continuous at $g(c)$.*

Proof. Suppose g is continuous at c and f is continuous at $g(c)$ for some $c \in \mathbb{R}$.

For all $x \in {}^*\mathbb{R}$, if $x \approx c$ then ${}^*g(x) \approx {}^*g(c)$ (since g is continuous at c). Also ${}^*g(x) \approx {}^*g(c)$ implies ${}^*f({}^*g(x)) \approx {}^*f({}^*g(c))$ (since f is continuous at $g(c)$ and ${}^*g(c) = g(c)$). Therefore $f \circ g$ is continuous at c. □

4.3 Uniform continuity

4.3.1 Definition (Standard definition)**.** Let $X \subset \mathbb{R}$. A function $f : X \to \mathbb{R}$ is uniformly continuous if and only if $(\forall\epsilon \in \mathbb{R}^+)(\exists\delta \in \mathbb{R}^+)[(\forall x, y \in X)(|x - y| < \delta \longrightarrow |f(x) - f(y)| < \epsilon)]$.

4.3.2 Proposition (Equivalent definition to 4.3.1)**.** A function f is uniformly continuous on $X \subseteq \mathbb{R}$ if and only if ${}^*f(x) \approx {}^*f(y)$ whenever $x \approx y$ on ${}^*X \subseteq {}^*\mathbb{R}$.

Proof. The proof is quite similar to the one given for normal continuity (*Proof* 4.2.2). □

4.3.3 Proposition. The function $f : x \mapsto x^2$ is not uniformly continuous on $\mathbb{R}$.

Proof. Let H be an infinite hyperreal. We have $H + \frac{1}{H} \approx H$. But $f(H + \frac{1}{H}) - f(H) = (H^2 + \frac{1}{H^2} + 2) - (H^2) = 2 + \frac{1}{H^2} \approx 2$. So $f(H + \frac{1}{H}) - f(H) \not\approx 0$. Therefore f is not uniformly continuous on $\mathbb{R}$. □

4.3.4 Proposition. If f is uniformly continuous on $X \subseteq \mathbb{R}$, then f is continuous on X.

Proof. Let $x \in {}^*X$ and $a \in X$. Suppose $x \approx a$. Then $a \in {}^*X$. So ${}^*f(x) \approx {}^*f(a)$ (since f is uniformly continuous on X). Thus f is continuous on X. □

4.3.5 Proposition. Consider a function $f : [a, b] \to \mathbb{R}$. If f is continuous, then f is uniformly continuous.

Proof. Let $x, y \in {}^*[a, b]$ such that $x \approx y$. Then $st(x) \approx st(y) = c$ for some $c \in [a, b]$. So we have $x \approx c$ and $y \approx c$. Then ${}^*f(x) \approx {}^*f(c)$. and ${}^*f(y) \approx {}^*f(c)$ (since f is continuous at c). So ${}^*f(x) \approx {}^*f(y)$. Hence f is uniformly continuous. □

4.3.6 Theorem (Intermediate Value Theorem)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Suppose f is continuous on $[a, b]$ and $f(a) < f(b)$. Then for any $c \in [f(a), f(b)]$, there exists some $x \in [a, b]$ such that $f(x) = c$.*

Proof. Let $n \in \mathbb{N}$. Consider the interval $I = [a, b] \subseteq \mathbb{R}$. This interval can be subdivided into n intervals of the form $I_k = [a + \delta k, a + \delta(k + 1)]$, where $k \in J(n) = \{0, \ldots, n - 1\}$ and $\delta = \frac{b-a}{n}$.

Take $c \in \mathbb{R}$ such that $c \in [f(a), f(b)]$. So there exists some $k \in J(n)$ such that $f(a+\delta k) \leq c$. Let K be the largest value of k for which we still have $f(a+\delta k) \leq c$. Then, $f(a+\delta K) \leq c \leq f(a+\delta(K+1))$.

Transferring this to ${}^*\mathbb{R}$ gives: for any $c \in {}^*[f(a), f(b)]$ there is some $K \in {}^*J(n)$ such that ${}^*f(a+\delta K) \leq c \leq {}^*f(a+\delta(K+1))$. In particular, this is true when we choose $n = N \in {}^*\mathbb{N} \setminus \mathbb{N}$ (N is infinite).

But $(a+\delta(K+1)) - (a+\delta K) = \delta \approx 0$. So $(a+\delta(K+1)) \approx (a+\delta K)$. Then ${}^*f(a+\delta(K+1)) \approx {}^*f(a+\delta K)$ (since f is continuous on I and therefore uniformly continuous by Proposition 4.3.5).

Since c is bounded by two infinitely close hyperreals, then c must be infinitely close to each of them. That is: ${}^*f(a+\delta K) \approx c \approx {}^*f(a+\delta(K+1))$.

Since $a+\delta K$ is bounded by two finite hyperreals (a and b), then $a+\delta K$ is infinitely close to a unique real number $x \in [a, b]$ (Standard Part Principle). So $a+\delta K \approx x$. This implies ${}^*f(x) \approx {}^*f(a+\delta K) \approx c$. Thus ${}^*f(x) = c$ since ${}^*f(x)$ and c are both real numbers.

Therefore, $(\forall c \in [f(a), f(b)])(\exists x \in [a, b])(f(x) = c)$. Hence the theorem is proved. □

4.3.7 Theorem (Extreme Value Theorem)**.** *Consider a function* $f : [a, b] \to \mathbb{R}$*. If* f *is continuous on* $[a, b]$ *then* f *attains both a maximum and a minimum on* $[a, b]$*. That is, there exist* $c, d \in [a, b]$ *such that* $\forall x \in [a, b]$*,* $f(c) \leq f(x) \leq f(d)$*.*

Proof. Let $n \in \mathbb{N}$. Consider the interval $I = [a, b] \subseteq \mathbb{R}$. This interval can be subdivided into n intervals of the form $I_k = [a+\delta k, a+\delta(k+1)]$, where $k \in J(n) = \{0, \ldots, n-1\}$ and $\delta = \frac{b-a}{n}$.

Consider the set $P(n) = \{f(a+\delta k) : 0 \leq k \leq n-1\}$. Since P is discrete and finite, the following sentence is true in $\mathbb{R}$:

$\sigma : (\exists k_1, k_2 \in J(n))(\forall k \in J(n))(f(a+\delta k_1) \leq f(a+\delta k) \leq f(a+\delta k_2))$ for all $n \in \mathbb{N}$. That is, $f(a+\delta k_1)$ and $f(a+\delta k_2)$ are respectively the minimum and the maximum of ${}^*P(n)$.

By Transfer, the following is also true in ${}^*\mathbb{R}$:

${}^*\sigma : (\exists k_1, k_2 \in {}^*J(n))(\forall k \in {}^*J(n))({}^*f(a+\delta k_1) \leq {}^*f(a+\delta k) \leq {}^*f(a+\delta k_2))$, where $\delta = \frac{b-a}{n}$ and $n \in {}^*\mathbb{N}$. That is, ${}^*f(a+\delta k_1)$ and ${}^*f(a+\delta k_2)$ are respectively the minimum and the maximum of ${}^*P$. In particular, when n is infinite, ${}^*\sigma$ still holds; and we have $\delta \approx 0$.

Given $u \in {}^*\mathbb{R}$ such that $a \leq u \leq b$, we have: $a+\delta k \leq u \leq a+\delta(k+1)$ for some $k \in {}^*J(n)$. But $(a+\delta(k+1)) - (a+\delta k) = \delta \approx 0$. So $(a+\delta(k+1)) \approx (a+\delta k)$. Then $(a+\delta(k+1)) \approx u \approx (a+\delta k)$. So ${}^*f\,(a+\delta k) \approx {}^*f(u)$ (since f is continuous on I and therefore uniformly continuous by Proposition 4.3.5). Thus ${}^*f(a+\delta k_1) \leq {}^*f(u) \leq {}^*f(a+\delta k_2)$ (from sentence ${}^*\sigma$) up to some infinitesimal error.

Let $x = st(u)$, $c = st(a+\delta k_1)$ and $d = st(a+\delta k_2)$. We have ${}^*f(c) \leq {}^*f(x) \leq {}^*f(d)$ for all $x \in [a, b]$. Therefore, there exist $c, d \in [a, b]$ such that $f(c) \leq f(x) \leq f(d)$ for all $x \in [a, b]$. □

4.4 Differentiability

4.4.1 Definition (Standard definition)**.** (Davis, 2009) A function $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $a \in \mathbb{R}$ if the limit $\lim\limits_{h \to 0} \dfrac{f(a+h) - f(a)}{h}$ exists and is finite. This limit is called the derivative of f at a and is denoted by $f'(a)$.

Equivalently, the definition can be formulated as follows: f is differentiable at a with derivative $f'(a)$ if $(\forall\epsilon \in \mathbb{R}^+)(\exists\delta \in \mathbb{R}^+)[(\forall h \in \mathbb{R})(0 < |h| < \delta \longrightarrow |\frac{f(a+h)-f(a)}{h} - f'(a)| < \epsilon)]$.

4.4.2 Proposition (Equivalent definition to 4.4.1)**.** A function $f : \mathbb{R} \to \mathbb{R}$ is differentiable at $a \in \mathbb{R}$ with derivative $f'(a)$ if and only if its $*-$transform $^*f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ satisfies $\frac{^*f(a+h)-^*f(a)}{h} \approx f'(a)$ whenever h is a non-zero infinitesimal.

Or equivalently, f is differentiable at $a \in \mathbb{R}$ with derivative $f'(a)$ if and only if $\frac{^*f(x)-^*f(a)}{x-a} \approx f'(a)$ whenever $x \approx a$ $(x \neq a)$.

Proof. Suppose f is differentiable at $a \in \mathbb{R}$. Let $f'(a)$ be the derivative of f at a. Then $(\forall\epsilon \in \mathbb{R}^+)(\exists\delta \in \mathbb{R}^+)[(\forall h \in \mathbb{R})(0 < |h| < \delta \longrightarrow |\frac{f(a+h)-f(a)}{h} - f'(a)| < \epsilon)]$.

Take $\epsilon \in \mathbb{R}^+$ and a corresponding $\delta \in \mathbb{R}^+$ such that the following sentence is true : $(\forall h \in \mathbb{R})(0 < |h| < \delta \longrightarrow |\frac{f(a+h)-f(a)}{h} - f'(a)| < \epsilon)$. By transfer, we have $(\forall h \in {}^*\mathbb{R})(0 < |h| < \delta \longrightarrow |\frac{^*f(a+h)-^*f(a)}{h} - f'(a)| < \epsilon)$. Now, let us choose h such that $h \approx 0$ (but $h \neq 0$). Then $|h| < \delta$ (since $\delta \in \mathbb{R}^+$). This implies that $|\frac{^*f(a+h)-^*f(a)}{h} - f'(a)| < \epsilon$. But ϵ was chosen arbitrarily. So $|\frac{^*f(a+h)-^*f(a)}{h} - f'(a)| \approx 0$. Therefore, $\frac{^*f(a+h)-^*f(a)}{h} \approx f'(a)$ whenever h is a non-zero infinitesimal.

Conversely, let us suppose that $\frac{^*f(a+h)-^*f(a)}{h} \approx f'(a)$ whenever h is a non-zero infinitesimal.

Let $\epsilon \in \mathbb{R}^+$. Let δ be a positive infinitesimal. Then $\forall h \in {}^*\mathbb{R}$, we have:

If $0 < |h| < \delta$, then $h \approx 0$. Whence $|\frac{^*f(a+h)-^*f(a)}{h} - f'(a)| \approx 0$. In particular, $|\frac{^*f(a+h)-^*f(a)}{h} - f'(a)| < \epsilon$. Then the following sentence is true in ${}^*\mathbb{R}$: $(\exists\delta \in {}^*\mathbb{R}^+)(\forall h \in {}^*\mathbb{R})(0 < |h| < \delta \longrightarrow |\frac{^*f(a+h)-^*f(a)}{h} - f'(a)| < \epsilon)$. It follows by transfer that $(\exists\delta \in \mathbb{R}^+)(\forall h \in \mathbb{R})(0 < |h| < \delta \longrightarrow |\frac{f(a+h)-f(a)}{h} - f'(a)| < \epsilon)$. But ϵ was chosen arbitrarily and δ was chosen accordingly. So $(\forall\epsilon \in \mathbb{R}^+)(\exists\delta \in \mathbb{R}^+)[(\forall h \in \mathbb{R})(0 < |h| < \delta \longrightarrow |\frac{f(a+h)-f(a)}{h} - f'(a)| < \epsilon)]$. Therefore f is differentiable at a with derivative $f'(a)$. □

4.4.3 Theorem (Continuity and differentiability)**.** *Consider a function $f : \mathbb{R} \to \mathbb{R}$. If f is differentiable at $a \in \mathbb{R}$, then f is continuous at a.*

Proof. Suppose f is differentiable at $a \in \mathbb{R}$. Let $f'(a)$ be the derivative of f at a. Then $\frac{^*f(a+h)-^*f(a)}{h} \approx f'(a)$ for any non-zero infinitesimal h. So $h\left(\frac{^*f(a+h)-^*f(a)}{h} - f'(a)\right) \approx 0$. Then $^*f(a+h) - {}^*f(a) \approx h.f'(a)$. Then $^*f(a+h) - {}^*f(a) \approx 0$. So $^*f(a+h) \approx {}^*f(a)$ whenever $h \approx 0$ $(h \neq 0)$. Hence f is continuous at a. □

4.4.4 Theorem (Chain rule)**.** *Consider $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$. If g is differentiable at a and f is differentiable at $g(a)$, then $f \circ g$ is differentiable at a with derivative $(f \circ g)'(a) = g'(a).f'(g(a))$.*

Proof. Let $x \approx a$ such that $x \neq a$. Suppose g is differentiable at a and f is differentiable at $g(a)$.

We need to prove that

$$\frac{^*f(^*g(x)) - {}^*f(^*g(a))}{x-a} \approx g'(a)f'(g(a)) \tag{4.4.1}$$

- If $^*g(x) = {}^*g(a)$, then $^*f(^*g(x)) - {}^*f(^*g(a)) = 0$ and $^*g'(a) = 0$. So Equation 4.4.1 trivially holds.
- Now assume $^*g(x) \neq {}^*g(a)$. Then $\frac{^*f(^*g(x))-^*f(^*g(a))}{x-a} = \frac{^*f(^*g(x))-^*f(^*g(a))}{^*g(x)-^*g(a)}.\frac{^*g(x)-^*g(a)}{x-a}$. we have $\frac{^*g(x)-^*g(a)}{x-a} \approx g'(a)$ since g is differentiable at a and $x \approx a$. But, since g is differentiable at a, it is also continuous at a. So $x \approx a$ implies $^*g(x) \approx {}^*g(a)$. Then $\frac{^*f(^*g(x))-^*f(^*g(a))}{^*g(x)-^*g(a)} \approx {}^*f'(^*g(a))$.

Therefore, $\frac{{}^*f({}^*g(x))-{}^*f({}^*g(a))}{x-a} \approx {}^*g'(a){}^*f'({}^*g(a))$. Hence $f \circ g$ is differentiable at a with derivative $(f \circ g)'(a) = g'(a).f'(g(a))$.

□

4.4.5 Theorem (Critical Point Theorem)**.** *Suppose $f : (a,b) \to \mathbb{R}$ is differentiable. If f achieves a maximum or a minimum at some $c \in (a,b)$ then $f'(c) = 0$.*

Proof. Suppose f is differentiable at c. Then $\forall h \approx 0$ ($h \neq 0$), we have $\frac{{}^*f(c+h)-{}^*f(c)}{h} \approx f'(c)$.

Suppose f achieves a minimum at c. Then ${}^*f(c+h) - {}^*f(c) \geq 0$ for all h. If h is negative infinitesimal, then $f'(c) \leq 0$. If h is positive infinitesimal, then $f'(c) \geq 0$. Therefore, $f'(c) = 0$.

Likewise, assuming that f achieves a maximum at c leads to $f'(c) = 0$ as well. □

4.5 Sequences

4.5.1 Definition (Standard definition of Convergence)**.** Let $(s_n)_{n\in\mathbb{N}}$ be a real-valued sequence. Then, $(s_n)_{n\in\mathbb{N}}$ converges to L if and only if $(\forall\epsilon \in \mathbb{R}^+)(\exists m \in \mathbb{N})(\forall n \in \mathbb{N})(n > m \longrightarrow |s_n - L| < \epsilon)$.

4.5.2 Definition (Non-standard extension of a sequence)**.** Consider the real-valued sequence $(s_n)_{n\in\mathbb{N}}$. $(s_n)_{n\in\mathbb{N}}$ can be represented as a function $f : \mathbb{N} \to \mathbb{R}$ defined by $f(n) = s_n$ for all $n \in \mathbb{N}$. This allows us to define a non-standard extension of $(s_n)_{n\in\mathbb{N}}$.

The non-standard extension of $(s_n)_{n\in\mathbb{N}}$ (denoted by $({}^*s_n)_{n\in\mathbb{N}}$) is represented by non-standard extension of f. So $({}^*s_n)_{n\in\mathbb{N}}$ is defined by ${}^*f : {}^*\mathbb{N} \to {}^*\mathbb{R}$ such that ${}^*f(n) = {}^*s_n$ for all $n \in {}^*\mathbb{N}$.

4.5.3 Theorem (Equivalent definition of convergence)**.** *$(s_n)_{n\in\mathbb{N}}$ converges to L if and only if $|{}^*s_N - L| \approx 0$ for all infinite hypernatural numbers N ($N \in {}^*\mathbb{N} \setminus \mathbb{N}$).*

Proof. Suppose $(s_n)_{n\in\mathbb{N}}$ converges to L. Let $\epsilon \in \mathbb{R}^+$. Take $m \in \mathbb{N}$ such that $(\forall n \in \mathbb{N})(n > m \longrightarrow |s_n - L| < \epsilon)$. By transfer, we have $(\forall n \in {}^*\mathbb{N})(n > m \longrightarrow |{}^*s_n - L| < \epsilon)$. Take $n = N$ where $N \in {}^*\mathbb{N} \setminus \mathbb{N}$. Then we have $(N > m \longrightarrow |{}^*s_N - L| < \epsilon)$. But, clearly, $N > m$ since N is infinite. So $|{}^*s_N - L| < \epsilon$. Since ϵ was taken arbitrarily and N was independent of ϵ, then $|{}^*s_N - L| \approx 0$.

Conversely, suppose $|{}^*s_N - L| \approx 0$ for any $N \in {}^*\mathbb{N} \setminus \mathbb{N}$. Let $\epsilon \in \mathbb{R}^+$. Take $m \in {}^*\mathbb{N}$ such that m is infinite.

For all $n \in {}^*\mathbb{N}$, if $n \geq m$ then n is infinite. So $|{}^*s_n - L| \approx 0$. In particular, $|{}^*s_n - L| < \epsilon$ for some $\epsilon \in \mathbb{R}^+$. There exists $m \in {}^*\mathbb{N}$ such that $(\forall n \in {}^*\mathbb{N})(n > m \longrightarrow |{}^*s_n - L| < \epsilon)$. By transfer, there exists $m \in \mathbb{N}$ such that $(\forall n \in \mathbb{N})(n > m \longrightarrow |s_n - L| < \epsilon)$. Hence $(s_n)_{n\in\mathbb{N}}$ converges to L. □

4.6 Integration

4.6.1 Definition (Riemann integration)**.** Let $a, b \in \mathbb{R}$, $n \in \mathbb{N}$ and let $f : [a,b] \to \mathbb{R}$ be a continuous function. The n-th Riemann sum of f is defined by

$$R_n = \sum_{i=0}^{n-1} f(x_i)\left(\frac{b-a}{n}\right) \tag{4.6.1}$$

where $x_i = a + i\left(\frac{b-a}{n}\right)$. The Riemann integral of f is defined by $\int_a^b f(x)dx = \lim_{n\to\infty} R_n$ (Ponstein, 2001).

Using the non-standard definition of convergence, we have: $\int_a^b f(x)dx \approx {}^*R_N$ for all $N \in {}^*\mathbb{N} \setminus \mathbb{N}$. Therefore

$$\int_a^b f(x)dx = st(R_N) \quad \text{for all} \quad N \in {}^*\mathbb{N} \setminus \mathbb{N}. \tag{4.6.2}$$

5. Further applications of non-standard analysis

In this chapter, we will explore more advanced theorems proved using non-standard analysis.

5.1 Peano Existence Theorem

5.1.1 Theorem. *(O'Neill, 2014) Let $f:\mathbb{R}\times[0,1]\to\mathbb{R}$ be a bounded continuous function. Then, for any $y_0\in\mathbb{R}$, there is a solution to the initial value problem $y'(t)=f(y(t),t),\ y(0)=y_0$.*

Proof. Let N be an infinite hypernatural number. Let ${}^*T=\{\frac{k}{N}:k\leq N\}$ where k is an hypernatural. We can inductively define a function ${}^*Y:{}^*T\to{}^*\mathbb{R}$ by:

$$ {}^*Y(\frac{k}{N})=y_0+\sum_{i=0}^{k-1}{}^*f\left(Y(\frac{i}{N}),\frac{i}{N}\right)\frac{1}{N}, \tag{5.1.1}$$

for some $k\in\{0,1,\ldots,N\}$, with ${}^*Y(0)=y_0$ (Cutland, 1988).

Claim: Y is continuous on T.

Proof of Claim: Let $\frac{k_1}{N},\frac{k_2}{N}\in{}^*T$ such that $\frac{k_1}{N}\approx\frac{k_2}{N}$ where $k_1,k_2\in\{0,1,\ldots,N\}$. We have ${}^*Y(\frac{k_1}{N})=y_0+\sum_{i=0}^{k_1-1}{}^*f\left({}^*Y(\frac{i}{N}),\frac{i}{N}\right)\frac{1}{N}$ and ${}^*Y(\frac{k_2}{N})=y_0+\sum_{i=0}^{k_2-1}{}^*f\left({}^*Y(\frac{i}{N}),\frac{i}{N}\right)\frac{1}{N}$.

Without loss of generality, we may assume $k_1<k_2$. Then we have ${}^*Y(\frac{k_2}{N})-{}^*Y(\frac{k_1}{N})=\sum_{i=k_1}^{k_2-1}{}^*f\left({}^*Y(\frac{i}{N}),\frac{i}{N}\right)\frac{1}{N}$. Since f is bounded (say below by some $m\in\mathbb{R}$ and above by some $M\in\mathbb{R}$), then ${}^*f$ is bounded (by m and M) as well. That is: $m\leq{}^*f\left({}^*Y(\frac{i}{N}),\frac{i}{N}\right)\leq M$.

$$\begin{aligned} m\leq{}^*f\left({}^*Y(\frac{i}{N}),\frac{i}{N}\right)\leq M &\implies \frac{m}{N}(k_2-k_1)\leq{}^*Y(\frac{k_2}{N})-{}^*Y(\frac{k_1}{N})\leq\frac{M}{N}(k_2-k_1)\\ &\implies 0\approx m\left(\frac{k_2}{N}-\frac{k_1}{N}\right)\leq{}^*Y(\frac{k_2}{N})-{}^*Y(\frac{k_1}{N})\leq M\left(\frac{k_2}{N}-\frac{k_1}{N}\right)\approx 0\\ &\implies {}^*Y(\frac{k_2}{N})-{}^*Y(\frac{k_1}{N})\approx 0\end{aligned}$$

Thus Y is continuous on T.

Now we can define a function $y:[0,1]\to\mathbb{R}$ by $y(t)=st({}^*Y(\frac{k}{N}))$ where $\frac{k}{N}$ is the largest element of ${}^*T$ such that $\frac{k}{N}\leq t$ (Cutland, 1988). The function y is continuous on $[0,1]$ since Y is continuous on T.

Also, since the function $s\mapsto y(s)$ and $s\mapsto s$ are continuous on $[0,1]$, so is $s\mapsto f(y(s),s)$ because f is continuous on $\mathbb{R}\times[0,1]$.

Then, $\int_0^t f(y(s),s)ds\approx\sum_{i=0}^{k-1}{}^*f\left({}^*y(\frac{i}{N}),\frac{i}{N}\right)\frac{1}{N}$. Since $y(t)=st({}^*Y(\frac{k}{N}))$ and y is continuous, we have ${}^*y(\frac{i}{N})\approx{}^*Y(\frac{i}{N})$. So ${}^*f\left({}^*y(\frac{i}{N}),\frac{i}{N}\right)\approx{}^*f\left(Y(\frac{i}{N}),\frac{i}{N}\right)$ (since f is continuous). Thus $\int_0^t f(y(s),s)ds\approx\sum_{i=0}^{k-1}{}^*f\left({}^*Y(\frac{i}{N}),\frac{i}{N}\right)\frac{1}{N}$.

So Equation 5.1.1 becomes: ${}^*Y(\frac{k}{N}) \approx y_0 + \int_0^t f(y(s), s)ds$. But ${}^*Y(\frac{k}{N}) \approx y(t)$ (since $y(t) = st({}^*Y(\frac{k}{N})))$. So $y(t) \approx y_0 + \int_0^t f(y(s), s)ds$. So $y(t) = y_0 + \int_0^t f(y(s), s)ds$. Hence $y'(t) = f(y(t), t)$ with $y(0) = y_0$. This completes the proof. □

5.2 Baker's theorem

5.2.1 Lemma. $\frac{\log 2}{\log 3}$ is irrational.

Proof. Suppose $\frac{\log 2}{\log 3}$ is rational. Then there exist $m, n \in \mathbb{N}$ such that $\frac{\log 2}{\log 3} = \frac{m}{n}$. This implies $2^n = 3^m$ which is absurd from the Fundamental Theorem of Arithmetic. Therefore $\frac{\log 2}{\log 3}$ is irrational. □

5.2.2 Proposition (Special case of Baker's theorem)**.** (Terence Tao)

For any integers p, q with $q > 1$, one has $|\frac{\log 2}{\log 3} - \frac{p}{q}| \geq \exp(-c \log^{c'} q)$ for some absolute constants $c, c' > 0$.

The standard proof of this proposition requires a lot of ϵ-δ management. However, the non-standard formulation of the proposition is simpler to prove. In order to make this formulation clear, we will first define the related terminology.

5.2.3 Definition. (Terence Tao) Let H be an infinite positive real number. Relative to this H, we can define various notions of size:

- An hyperreal z is polynomial size if one has $|z| \leq CH^C$ for some real $C > 0$.
- A hyperreal z is said to be polylogarithmic size if one has $|z| \leq C \log^C H$ for some real $C > 0$.
- A hyperreal z is said to be quasipolynomial size if one has $|z| \leq \exp(C \log^C H)$ for some real $C > 0$.
- A hyperreal z is said to be quasiexponentially small if one has $|z| \leq \exp(-C \log^C H)$ for every real $C > 0$.

Given two hyperreals X, Y with Y non-negative, we write $X \ll Y$ or $X = O(Y)$ if $|X| \leq CY$ for some real $C > 0$. We write $X \lll Y$ or $X = o(Y)$ if $|X| \leq cY$ for all real $c > 0$.

Now, we are ready to formulate the non-standard version of Proposition 5.2.2.

5.2.4 Proposition (Non-standard formulation of Proposition 5.2.2)**.** (Terence Tao) Let H be an infinite hypernatural number, and let $\frac{p}{q}$ be a rational of height at most H (i.e. non-standard $|p|, |q| \leq H$). Then $\frac{\log 2}{\log 3} - \frac{p}{q}$ is not quasiexponential small (relative to H).

Proof. A proof of this proposition is given in Terence Tao. □

We need to prove that Proposition 5.2.4 is equivalent to Proposition 5.2.2.

Proof. We first prove that Proposition 5.2.4 implies Proposition 5.2.2.

Suppose Proposition 5.2.2 fails. So, for any positive constants c, c', there exist integers p, q with $q > 1$ such that $|\frac{\log 2}{\log 3} - \frac{p}{q}| < \exp(-c \log^{c'} q)$. Take $c = c' = k > 0$, where k is an integer. There exist some corresponding p_k and q_k such that $|\frac{\log 2}{\log 3} - \frac{p_k}{q_k}| < \exp(-k \log^k q_k)$. But k was chosen arbitrarily. So we have $|\frac{\log 2}{\log 3} - \frac{p_k}{q_k}| < \exp(-k\log^k q_k)$ for all $k \in \mathbb{N}$. Let $p = [(p_k)_{k\in\mathbb{N}}]$ and $q = [(q_k)_{k\in\mathbb{N}}]$. So we have $|\frac{\log 2}{\log 3} - \frac{p}{q}| < \exp(-K\log^K q)$ for all $K \in \mathbb{R}^+$. Thus q must be infinite. Otherwise, $\exp(-K\log^K q) \approx 0$ when K is infinite which would imply $\frac{\log 2}{\log 3} \approx \frac{p}{q}$ and contradict the fact that $\frac{\log 2}{\log 3}$ is not rational (Lemma 5.2.1). Since q is infinite, let $q = H$. So we have $|\frac{\log 2}{\log 3} - \frac{p}{q}| < \exp(-K\log^K H)$ for all $K \in \mathbb{R}^+$. Thus $|\frac{\log 2}{\log 3} - \frac{p}{q}|$ is quasiexponentially small (contradiction with Proposition 5.2.4). Hence Proposition 5.2.4 implies Proposition 5.2.2.

Now let us prove that Proposition 5.2.2 implies Proposition 5.2.4.

Suppose Proposition 5.2.4 fails. Then there exist some $p, q \in {}^*\mathbb{N}$ such that $|\frac{\log 2}{\log 3} - \frac{p}{q}| \leq \exp(-d\log^d H)$ (where H is a infinite hyperreal) for all $d \in \mathbb{R}^+$.

In particular, we have $|\frac{\log 2}{\log 3} - \frac{p}{q}| \leq \exp(-d\log^d H)$, where $d = \max(c, c') + 1$ for some constants $c, c' > 0$.

Moreover, consider a function f from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$ defined by $f(c, c') = \exp(-ct^{c'})$, for some $t > 1$.

Claim: f is decreasing in c and c'.

Proof of Claim: We have $\frac{\partial f}{\partial c}(c, c') = -\exp(c'\log t)\exp(-ct^{c'}) < 0$. So f is decreasing in c. Also, $\frac{\partial f}{\partial c'}(c, c') = -ct^{c'}\log t\exp(-ct^{c'}) < 0$. So f is decreasing in c'. This completes the proof of Claim.

It follows that $\exp(-d\log^d H) < \exp(-c\log^{c'} H)$ since $d > c$ and $d > c'$. Thus $|\frac{\log 2}{\log 3} - \frac{p}{q}| < \exp(-c\log^{c'} H)$ (contradiction with Proposition 5.2.2). Therefore, Proposition 5.2.2 implies Proposition 5.2.4.

Hence Proposition 5.2.4 is equivalent to Proposition 5.2.2. □

6. Conclusion

In this essay, we have shown how non-standard analysis works. We first extended the set of usual real numbers to include infinitesimal and infinite numbers using an ultrapower construction. We then introduced the Transfer Principle which is the main theorem of non-standard analysis. This theorem allows us to convert a standard proof to a non-standard one and vice versa. In order to see how helpful is non-standard analysis, we reformulated the usual definitions of well-known mathematical concepts such as the notions of limits, continuity, differentiability. We also gave non-standard proofs of basics theorems in calculus such as the Intermediate Value Theorem, Extreme Value Theorem, and Chain Rule. We finally explored more advanced theorems of Calculus namely and Peano's Theorem. All these have helped to show that non-standard analysis is not only useful in making definitions more comprehensible but also giving nicer proofs of theorems.

Acknowledgements

I am very thankful towards my supervisor Dr Gareth Boxall for his permanent availability and guidance throughout this research project. I would like to thank my co-supervisor and tutor Georg Anegg who has inspired me during my studies at AIMS and particularly during this project.

Finally, I would like to express my deepest gratitude to AIMS staff, lecturers and tutors who have contributed to this amazing experience at AIMS.

References

A. 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.

G. Boxall. Model theory, 2016. Lecture notes.

N. Cutland. *Nonstandard Analysis and its Applications*, volume London Mathematical Society Student Texts 10. Cambridge University Press, 1988.

I. Davis. An Introduction to Nonstandard Analysis. *Internet source, publication date August*, 2009. URL http://www.math.uchicago.edu/~may/VIGRE/VIGRE2009/REUPapers/Davis.pdf.

R. Goldblatt. *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis*, volume Graduate Text in Mathematics 188. Springer Science & Business Media, 2012.

I. Goldbring. Lecture Notes on Nonstandard Analysis, UCLA Summer School in Logic. 2014. URL http://homepages.math.uic.edu/~isaac/NSA%20notes.pdf.

C. W. Henson. A Gentle Introduction to Nonstandard Extensions. In *Foundations of Nonstandard Analysis*, pages 1–49. University of Illinois at Urbana-Champaign, 1997. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.75.8855&rep=rep1&type=pdf.

H. J. Keisler. *Elementary Calculus: An Infinitesimal Approach*. 1986. URL https://www.math.wisc.edu/~keisler/calc.html.

H. J. Keisler. *Foundations of Infinitesimal Calculus*. 2009. URL http://www.enzoexposito.it/Analisi_Non_Standard/Keisler_Elementary_Calculus_Foundations.pdf.

K. O'Neill. An Introduction to Non-Standard Analysis and its Applications. 2014. URL https://math.berkeley.edu/~jhicks/links/SOTS/koneill030514.pdf.

J. Ponstein. Nonstandard Analysis: A naive Way to the Infinitesimals (an unorthodox treatment of Nonstandard Analysis). 2001. URL http://www.rug.nl/research/gradschool-economics-and-business/organization/publications/ponstein.pdf.

Terence Tao. Hilbert's seventh problem, and powers of 2 and 3. Terence , Tao, https://terrytao.wordpress.com/2011/08/21/hilberts-seventh-problem-and-powers-of-2-and-3/, Written 21 August 2011, Accessed April 2017.