Nonstandard Analysis via Alpha Theory

Jane Mabesa (janem@aims.ac.za)
African Institute for Mathematical Sciences (AIMS)

Supervised by: Dr Gareth Boxall
Co-supervised by: Dr Taboka Prince Chalebgwa
University of Stellenbosch, South Africa

23 May 2019

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

Abstract

Abraham Robinson developed nonstandard analysis in the 1960's. Since then it has been applied in areas such as measure and probability theory and mathematical economics to mention a few. This theory has been investigated for a considerable amount of time with the aim of simplifying the formalism Robinson used to present its methods. In this essay, we present the methods of nonstandard analysis by using Alpha theory, the elementary axiomatics for nonstandard analysis. This approach requires understanding postulated axioms for an infinite "ideal number" α and requires no formal first order logic background as in the case of most approaches such as the use of ultrapowers. Motivated by the results, we finally discuss the model theory approach of how we can formalize Alpha Theory, and give a few nonstandard formulations of results from classical calculus.

Keywords: nonstandard analysis, alpha theory, hyperreals.

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.

Jane Mabesa, 23 May 2019

Contents

1. INTRODUCTION

Classical calculus involves defining concepts using the notion of limit, which was formalised as the ϵ-δ approach by Weierstrass in the nineteenth century. The limit concept uses the notion of quantities being close to others, and it replaced the need for infinitely small non zero quantities called the infinitesimals which were used by Newton and Leibniz in the early development of calculus. The reason for this replacement was that they were not rigorously founded, and there was a challenge to Newton and Leibniz to develop a consistent theory of analysis using infinitesimals.

In the 1960's, Abraham Robinson introduced the nonstandard analysis. He developed the rigorous foundations for extending (in a consistent way) the set of real numbers $\mathbb{R}$ to include the infinite and infinitesimal elements by using the methods of model theory and mathematical logic (Robinson, 1974).

However, Robinson's work is not easily accessible without formal training in mathematical logic. This then led to many attempts by others such as Wilhelmus Luxemburg, who used ultrapowers an attempt to simplify matters whilst obtaining Robinson's new structure (Luxemburg, 1966). Later on, in the mid seventies Edward Nelson developed nonstandard analysis by introducing its axiomatic formulation referred to as the Internal Set Theory (IST). This is simply an extension of Zermelo-Fraenkel set theory (ZF) by additional axioms, being the transfer principle, the principle of idealization, and the principle of standardization (Nelson, 1977).

In this essay, we study another approach to nonstandard analysis called Alpha theory, based on the paper by (Benci and Di Nasso, 2003). Their approach is to reach nonstandard analysis whilst side-stepping technical notions such as formal logic and ultrapowers. Broadly speaking, there are two general approaches towards nonstandard analysis. These are the synctactic and the semantic approach. Alpha theory falls in the former and the model theoretic approach is an example of the latter.

Following Benci and Di Nasso (2003), we begin by classifying mathematical objects into atoms and nonatoms and generalise them as entities; we then study a few axioms postulated to govern the use of an "ideal number" denoted by α which can be thought of as an "infinitely large" natural number. We state and expand on the proofs of some propositions which immediately follow from the axioms. We then consider the star-operator which helps us in construction of hyper extensions, most importantly the hyperreal set $\mathbb{R}^*$. We then define the most important tool of nonstandard analysis, the transfer principle, which is useful in transferring first order properties from any entity to its hyper extension.

We then move on to nonstandard calculus wherein we begin by defining the infinitesimals and infinite numbers, and state the shadow theorem. Proceeding we also consider some relationships between the ideal values and the notion of limits, showing the relationship between nonstandard analysis and calculus; then consider internal sets which will be useful in defining saturation. We then introduce an alternative approach using model theory on how to formally generate the alpha theory.

Finally, by a way of example, we demonstrate nonstandard techniques in calculus by looking at the usual notions such as continuity and some of their well known consequences.

Formal definitions of all the concepts mentioned above are given in the following chapters, but generally, theorem statements and proofs have been customized to the context of this essay, hence may be stated differently from the original sources cited.

2. The Alpha Theory

Benci and Di Nasso (2003) begin by defining a new mathematical object which is central to their approach to non-standard analysis, namely α, considered to be an infinite ideal number, and give five axioms together with their explanations and some of their first consequences. In this Chapter, we elaborate on their work and then give proofs to a few of the consequences.

Let α be the mathematical object that captures the numerosity of $\mathbb{N}$, to be treated as an ideal natural number. The introduction of the number α to natural numbers can be related (in spirit) to the introduction of i to the real numbers to obtain the complex numbers.

2.1 The five axioms

The Zermelo-Fraenkel set theory ZFC approach to the foundations of mathematics implies that every entity in the mathematical universe can be viewed as a set. As an example, we will demonstrate on how we can encode an ordered triple as a set. We first give the following definitions.

2.1.1 Definition. (Holmes, 1998) The ordered pair (a, b) can be encoded as a set $\{\{a\}, \{a, b\}\}$ called the Kuratowski ordered pair of a and b.

2.1.2 Definition. (Holmes, 1998) An n-tuple $(a_1, a_2, \cdots, a_n)$ can be thought of as $(a_1, (a_2, \cdots, a_n))$; this is an inductive definition, based on the prior definition of the (n-1)-tuple. For example $(a_1, a_2, a_3, a_4) = (a_1, (a_2, a_3, a_4))$. Note that any n-tuple $(a_1, a_2, \cdots, a_{n-1}, a_n)$ is also an (n-1)-tuple, $(a_1 \cdots, (a_{n-1}, a_n))$.

Applying the above definitions to the ordered triple, we have

$$(x_1, x_2, x_3) = (x_1, (x_2, x_3)).$$

We observe that if we let $q = (x_2, x_3)$ we have an ordered pair (x_1, q) which, by the definition of a Kuratowski ordered pair becomes

$$(x_1, q) = \{\{x_1\}, \{x_1, q\}\}. \tag{2.1.1}$$

Applying the Kuratowski ordered pair definition again to (x_2, x_3) and substituting in equation (2.1.1) we have that

$$(x_1, x_2, x_3) = \{\{x_1\}, \{x_1, \{\{x_2\}, \{x_2, x_3\}\}\}\}.$$

As already evident, in practice, and for convenience, there is still a need to view certain objects not just as mere sets but as "atoms". For instance, numbers are taken to be atoms; even though real numbers can be defined as equivalence classes of Cauchy sequences, the complex numbers as the ordered pairs of the reals to mention a few. The set of atoms will be denoted by $\mathcal{A}$, and we have that $\mathbb{R} \subset \mathcal{A}$. We will also use the usual language of mathematics, such as referring to a sequence as a function whose domain is the set of natural numbers; and the empty set being the only set with no elements. We would like to make a distinction between sets and atoms whereby nonatoms are referred to as sets. We now give the following list of axioms from (Benci and Di Nasso, 2003) which will govern the use of α.

(1) Extension Axiom

Every sequence $\phi : \mathbb{N} \longrightarrow A$ for some set A has a unique ideal value (value at infinity), denoted by $\phi[\alpha]$. That is ϕ is "extendible" to $\mathbb{N} \cup \{\alpha\}$.

Furthermore, if two sequences ϕ, θ differ at all points, then $\phi[\alpha] \neq \theta[\alpha]$.

(2) Composition Axiom

If ϕ and θ are sequences and if f is any function such that compositions $f \circ \phi$ and $f \circ \theta$ are well defined, then

$$\phi[\alpha] = \theta[\alpha] \implies (f \circ \phi)[\alpha] = (f \circ \theta)[\alpha].$$

Composing a function f with a sequence produces a new sequence. So, if we compose f with two sequences with the same value at infinity then the output sequences have the same value at infinity also.

(3) Number axiom

Let the sequence $c_r : n \mapsto r$ be the constant sequence $\{r\}_n$ for all $n \in \mathbb{N}$ and $r \in \mathbb{R}$. Then $c_r[\alpha] = r$. If we let $r = n \in \mathbb{N}$, then the sequence $1_{\mathbb{N}} : n \mapsto n$ is defined on $\mathbb{N}$ and we have that $1_{\mathbb{N}}[\alpha] = \alpha \notin \mathbb{N}$.

The condition $c_r[\alpha] = r$ actually shows preservation of the notion of constant real sequences at infinity. The last part of the axiom says that the identity sequence has α as its value at infinity. That is, α is a new number. Thus, $1_{\mathbb{N}}$ is an example of a sequence $\varphi : \mathbb{N} \longrightarrow A$ such that $\alpha \notin A$.

(4) Pair Axiom

For all sequences ϕ, θ and ϑ :

$$\vartheta[\alpha] = \{\phi[\alpha], \theta[\alpha]\} \quad \text{whenever} \quad \vartheta(n) = \{\phi(n), \theta(n)\} \quad \text{for all n.}$$

This means if we present a sequence as a set of sequences that have ideal values, then the ideal value of this new sequence is the set of ideal values of those sequences. That is if for sequence $\xi(n) \in \vartheta(n)$, we either have for all n, $\xi(n) = \phi(n)$ or $\xi(n) = \theta(n)$ then, either $\xi[\alpha] = \phi[\alpha]$ or $\xi[\alpha] = \theta[\alpha]$ at infinity as well.

(5) Internal Set Axiom

If θ is a sequence of sets, then $\theta[\alpha]$ is also a set and,

$$\theta[\alpha] = \{\phi[\alpha] | \phi(n) \in \theta(n) \quad \text{for all n}\}.$$

All sequences that are pointwise members of θ have their values at infinity being elements of $\theta[\alpha]$, because they are obtained by preserving membership relation. That is, if $\phi(n) \in \theta(n)$ for all n, then $\phi[\alpha] \in \theta[\alpha]$.

2.2 First Consequences of the Axioms

The first proposition gives preservation of basic set operations at infinity, this of course excludes the power set, as operations on power sets belong to higher order logic, and nonstandard analysis is carried over first order logic. We will give proofs to selected parts of the propositions that will be considered as we move along and refer the reader to (Benci and Di Nasso, 2003) for the remaining proofs.

2.2.1 Proposition. (Benci and Di Nasso, 2003) For the sequences ϕ, θ and ϑ of nonempty sets, we have that for all $n \in \mathbb{N}$,

(1) Union: If $\vartheta(n) = \phi(n) \cup \theta(n)$ for all n, then $\vartheta[\alpha] = \phi[\alpha] \cup \theta[\alpha]$;

(2) Subset: If $\phi(n) \subseteq \theta(n)$ for all n, then $\phi[\alpha] \subseteq \theta[\alpha]$;

(3) Ordered pair: If $\vartheta(n) = \langle \phi(n), \theta(n) \rangle$ for all n, then $\vartheta[\alpha] = \langle \phi[\alpha], \theta[\alpha] \rangle$;

(4) Difference: If $\phi(n) \neq \theta(n)$ and $\phi(n) \notin \theta(n)$ for all n, then $\phi[\alpha] \neq \theta[\alpha]$ and $\phi[\alpha] \notin \theta[\alpha]$;

(5) Setminus: If $\vartheta(n) = \phi(n) \setminus \theta(n)$ for all n, then $\vartheta[\alpha] = \phi[\alpha] \setminus \theta[\alpha]$;

(6) Intersection: If $\vartheta(n) = \phi(n) \cap \theta(n)$ for all n, then $\vartheta[\alpha] = \phi[\alpha] \cap \theta[\alpha]$;

(7) Cartesian product: If $\vartheta(n) = \phi(n) \times \theta(n)$ for all n, then $\vartheta[\alpha] = \phi[\alpha] \times \theta[\alpha]$;

Proof. We give expanded versions of the proofs of (1), (2), and (3).

(1) Suppose that for all $n \in \mathbb{N}$, $\vartheta(n) = \phi(n) \cup \theta(n)$. Then, we have that $\phi(n) \cup \theta(n) \subseteq \vartheta(n)$. Let $\xi[\alpha] \in \phi[\alpha] \cup \theta[\alpha]$, then $\xi = \xi[\alpha]$ for some sequence $\xi(n) \in \phi(n) \cup \theta(n)$, which implies that $\xi(n) \in \vartheta(n)$ since $\phi(n) \cup \theta(n) \subseteq \vartheta(n)$. So, by internal set axiom, we have that $\xi[\alpha] \in \vartheta[\alpha]$. Thus, $\phi[\alpha] \cup \theta[\alpha] \subseteq \vartheta[\alpha]$.

Conversely, Suppose that for all $n \in \mathbb{N}$, $\vartheta(n) = \phi(n) \cup \theta(n)$. Let $\xi = \xi[\alpha]$ such that $\xi[\alpha] \in \vartheta[\alpha]$. Now, for all n, let the sequence $\zeta(n)$ be

$$\zeta(n) = \begin{cases} \phi(n), & \text{if } \xi(n) \in \phi(n); \\ \theta(n), & \text{if } \xi(n) \in \theta(n) \setminus \phi(n). \end{cases}$$

Applying the internal set axiom on this definition, it follows that if $\xi(n) \in \zeta(n)$, then $\xi[\alpha] \in \zeta[\alpha]$. But we can write $\zeta(n) \in \{\phi(n), \theta(n)\}$ which implies that $\zeta[\alpha] \in \{\phi[\alpha], \theta[\alpha]\}$ by pair axiom. In other words,

$$\zeta[\alpha] = \phi[\alpha] \quad \text{or} \quad \zeta[\alpha] = \theta[\alpha].$$

Thus, $\xi[\alpha] \in \phi[\alpha] \cup \theta[\alpha]$. Therefore, $\vartheta[\alpha] \subseteq \phi[\alpha] \cup \theta[\alpha]$. Hence, $\vartheta[\alpha] = \phi[\alpha] \cup \theta[\alpha]$. □

Proof. (2) Suppose that for all $n \in \mathbb{N}$, $\phi(n) \subseteq \theta(n)$. Then by Definition, $\theta(n) = \phi(n) \cup \theta(n)$. Now, let $\xi = \xi[\alpha]$ such that $\xi[\alpha] \in \phi[\alpha]$, then it follows that $\xi[\alpha] \in \phi[\alpha] \cup \theta[\alpha]$. But, we found from hypothesis that $\theta(n) = \phi(n) \cup \theta(n)$. So, by (1) it follows that $\xi[\alpha] \in \theta[\alpha]$. Therefore, $\phi[\alpha] \subseteq \theta[\alpha]$. □

Proof. (3) For all $n \in \mathbb{N}$, assume that $\vartheta(n) = \langle \phi(n), \theta(n) \rangle$. Recall that by Definition (2.1.1), the ordered pair $\langle a, b \rangle$ can be expressed by the Kuratowski pair $\{\{a\}, \{a, b\}\}$. So, we have that

$$\begin{aligned} \vartheta(n) &= \{\{\phi(n)\}, \{\phi(n), \theta(n)\}\} \\ &= \{\xi(n), \zeta(n)\}, \end{aligned}$$

where $\xi(n) = \{\phi(n)\}$ and $\zeta(n) = \{\phi(n), \theta(n)\}$. Then, it follows directly by pair axiom that $\vartheta[\alpha] = \{\xi[\alpha], \zeta[\alpha]\}$, where $\phi[\alpha] \in \xi[\alpha]$ and $\zeta[\alpha] = \{\phi[\alpha], \theta[\alpha]\}$ by pair axioms also. So, $\vartheta[\alpha] = \{\{\phi[\alpha]\}, \{\phi[\alpha], \theta[\alpha]\}\}$, and hence $\vartheta[\alpha] = \langle \phi[\alpha], \theta[\alpha] \rangle$ by rewriting the Kuratowski pair as an ordered pair. □

The next proposition tells us that if we have a sequence of a particular entity, then value at infinity of the sequence belongs to the clan of that entity, and a converse of this property.

2.2.2 Proposition. (Benci and Di Nasso, 2003)

(1) The ideal value $\phi[\alpha]$ is an atom (or a set, a nonempty set) whenever the sequence ϕ is of atoms (or sets, or nonempty sets, respectively);

(2) Let the ideal value $\phi[\alpha]$ be an atom (or a set, or a nonempty set). Then there exists a sequence $\theta(n)$ of atoms (or sets, nonempty sets, respectively) such that $\phi[\alpha] = \theta[\alpha]$;

(3) If the ideal value $\phi[\alpha]$ is a set, then its elements are values at infinity.

Proof. We give as an example and explain in detail the proof for (3):

(3) Suppose that the value at infinity $\phi[\alpha]$ is a nonempty set. Then, by (2) there exists a sequence $\{\theta(n)\}_{n\in\mathbb{N}}$ of nonempty sets such that $\phi[\alpha] = \theta[\alpha]$. Now, let $\xi = \xi[\alpha]$ such that $\xi[\alpha] \in \phi[\alpha]$, then $\xi[\alpha] \in \theta[\alpha]$ since $\phi[\alpha] = \theta[\alpha]$. But, by the internal set axiom, $\xi[\alpha] \in \theta[\alpha]$ for some sequence $\xi(n) \in \theta(n)$. And thus, $\xi[\alpha]$ is the value at infinity for the sequence $\{\xi(n)\}_{n\in\mathbb{N}}$. □

□

The next proposition tells us that the value at infinity is not affected by a change in finitely many values for a given sequence.

2.2.3 Proposition.

(1) If $\varphi(n) = \psi(n)$ eventually (i.e for all but finitely many n), then $\phi[\alpha] = \theta[\alpha]$;

(2) If $\phi(n) \neq \theta(n)$ eventually, then $\phi[\alpha] \neq \theta[\alpha]$.

The next proposition says that the set of indices at which two sequences with same value at infinity are identical is large enough to identify whether two other given sequences have the same value at infinity. Similarly even when they have different values at infinity.

2.2.4 Proposition. Assume that $\phi[\alpha] = \theta[\alpha]$ (or $\phi[\alpha] \neq \theta[\alpha]$) and let $\Lambda = \{n|\phi(n) = \theta(n)\}$ $(\Lambda = \{n|\phi(n) \neq \theta(n)\})$. Then, for all sequences ξ, ζ :

(1) If $\xi(n) = \zeta(n)$ (or $\xi(n) \neq \zeta(n)$, respectively) for all $n \in \Lambda$, then $\xi[\alpha] = \zeta[\alpha]$ (or $\xi[\alpha] \neq \zeta[\alpha]$, respectively);

(2) If $\xi(n) \in \zeta(n)$ or $\xi(n) \notin \zeta(n)$ for all $n \in \Lambda$, then $\xi[\alpha] \in \zeta[\alpha]$ or $\xi[\alpha] \notin \zeta[\alpha]$.

The following remark justifies the simplifying assumption that if $\phi[\alpha] \neq \theta[\alpha]$, then we can always assume that $\phi(n) \neq \theta(n)$ for all $n \in \mathbb{N}$. The same hold, if $\phi[\alpha] = \theta[\alpha]$.

2.2.5 Remark. Assuming $\phi[\alpha] \neq \theta[\alpha]$, let the set $\Gamma = \{n|\phi(n) \neq \theta(n)\}$. Define the new sequence ϕ' as $\phi'(n) = \phi(n)$ if $n \in \Gamma$ and $\phi'(n) \neq \theta(n)$ if $n \notin \Gamma$. Then, $\phi'(n) \neq \theta(n)$ for all $n \in \mathbb{N}$. By (1) of Proposition (2.2.4), we have that $\phi'[\alpha] = \phi[\alpha]$. By our assumption it follows that $\theta[\alpha] \neq \phi[\alpha] = \phi'[\alpha]$, and thus $\phi'[\alpha] \neq \theta[\alpha]$. Therefore, we assume that for any two sequences such that $\phi[\alpha] \neq \theta[\alpha]$, then $\phi(n) \neq \theta(n)$ for all n.

3. The Star-Operator

The star-operator ($*$) is the operation through which we can obtain a hyper extension of any entity, not necessarily just for pre-defined sequences as in the previous chapter. So, from this we will define the hyperreal numbers $\mathbb{R}^*$, which is the extension of the ordered field $\mathbb{R}$; then give the related extensions of its subsets and conclude the Chapter.

In constructing $\mathbb{R}^*$ or in general A^* by extending A, the first order relations for entities in A are preserved. For example, if the first order statement

$$\forall a, b \in A, \quad a \cdot b = 0 \quad \text{if and only if} \quad a = 0 \quad \text{or} \quad b = 0$$

holds for a and b in A, then the same is true for a^* and b^* in A^*. As a side note, this expresses being an integral domain for rings.

So, the process of carrying first order statements over is called the *transfer principle*, which actually states that a first order formula holds in A if and only if the corresponding formula holds in A^*.

Before specializing to hyperreal numbers $\mathbb{R}^*$, we consider the star-transforms of sets, functions, relations in general and use the *transfer principle* to prove some important statements that will be needed in the future.

3.1 Star-transforms of sets, functions, relations

3.1.1 Definition. (Benci and Di Nasso, 2003) The nonstandard extension of any entity A is defined as $A^* = c_A[\alpha]$, this is often referred to as the star transform of A. This is the value taken by the constant sequence $c_A : n \longmapsto A$.

3.1.2 Remark. We have seen that by the number axiom, $x = x^*$ for all $x \in \mathbb{R}$. This is not always the case because for the natural numbers, $\alpha^* \neq \alpha$ since for all n, $1_{\mathbb{N}}(n) = n \neq \alpha$.

However, when $A = \emptyset$ then by the internal set axiom $A^* = \emptyset$ whereas for $A \neq \emptyset$ we have that

$$A^* = \{\phi[\alpha] | \phi : \mathbb{N} \longrightarrow A\}.$$

Using the internal set axiom, if we let $A = \theta(n)$, then it follows that $\theta[\alpha] = A^*$, which is a way of using alpha theory to form nonstandard extensions for the sets.

One might ask whether all the set relations that hold for any two arbitrary sets A and B still hold for their star transforms A^* and B^* respectively. It turns out that this is the case. In the next proposition, we illustrate this for some set operations which we will need later for the rest we refer the reader to (Benci and Di Nasso, 2003); and all the proofs follow directly from the definition of the star transform of nonempty sets.

3.1.3 Proposition. (Benci and Di Nasso, 2003) Suppose that A, B, are non-empty sets. Then it follows that

(1) $A = B$ if and only if $A^* = B^*$;

(2) $(A \cap B)^* = A^* \cap B^*$;

(3) $(A \times B)^* = A^* \times B^*$;

(4) $A \subseteq B$ if and only if $A^* \subseteq B^*$.

We give as examples the proofs for (2) and (4), which we begin as follows:

Proof. (2) Assume that $\phi[\alpha]$ is a value at infinity for the sequence $\{\phi(n)\}_{n\in\mathbb{N}}$. Let $\phi[\alpha] \in (A \cap B)^*$. Then, by definition

$$(A \cap B)^* = \{\phi[\alpha] | \phi : \mathbb{N} \longrightarrow A \cap B\}.$$

But since $A \cap B \subseteq A$, it follows that

$$\{\phi[\alpha] | \phi : \mathbb{N} \longrightarrow A \cap B\} \subseteq \{\phi[\alpha] | \phi : \mathbb{N} \longrightarrow A\}.$$

Therefore, by definition, we have that $\phi[\alpha] \in A^*$. Similarly, it can be shown that $\phi[\alpha] \in B^*$. Thus, $\phi[\alpha] \in A^* \cap B^*$. Hence, $(A \cap B)^* \subseteq A^* \cap B^*$.

Conversely, assume that $\phi[\alpha] \in A^* \cap B^*$. Then, by definition, $\phi[\alpha] \in A^*$ and $\phi[\alpha] \in B^*$ which implies that

$$\phi[\alpha] \in \{\psi[\alpha] | \psi : \mathbb{N} \longrightarrow A\} \quad \text{and} \quad \phi[\alpha] \in \{\psi'[\alpha] | \psi' : \mathbb{N} \longrightarrow B\}.$$

Then, it turns out that $\phi[\alpha] \in \{\psi[\alpha] | \psi : \mathbb{N} \longrightarrow A\} \cap \{\psi'[\alpha] | \psi' : \mathbb{N} \longrightarrow B\}$. So, $\phi[\alpha] = \psi[\alpha] = \psi'[\alpha]$ for some $\psi : \mathbb{N} \longrightarrow A$ and $\psi' : \mathbb{N} \longrightarrow B$. By Remark (2.2.5), $\phi = \psi : \mathbb{N} \longrightarrow A$ and $\phi = \psi' : \mathbb{N} \longrightarrow B$. But in other words this means $\phi(n) \in A$ and $\phi(n) \in B$ for all $n \in \mathbb{N}$ which implies that $\phi(n) \in A \cap B$ for all $n \in \mathbb{N}$. So, $\phi[\alpha] \in \{\phi : \mathbb{N} \longrightarrow A \cap B\}$. Thus, $\phi[\alpha] \in (A \cap B)^*$. Hence, $A^* \cap B^* \subseteq (A \cap B)^*$. Therefore, $(A \cap B)^* = A^* \cap B^*$. □

Proof. (4) Assume that $A \subseteq B$. By definition, we have that $A = A \cap B$. Using (2) we have that $A^* = (A \cap B)^* = A^* \cap B^*$ and thus $A^* \subseteq B^*$.

Conversely, suppose that $A^* \subseteq B^*$. By definition, it follows that $A^* = A^* \cap B^*$. Using (1) and (2) we have that $A = A \cap B$. Then, it follows that $A \subseteq B$. Hence, $A \subseteq B \Leftrightarrow A^* \subseteq B^*$. □

Recall the Cartesian product of two sets A, B is given by

$$A \times B = \{\langle a, b\rangle : a \in A \quad \text{and} \quad b \in B\}.$$

A binary relation $\mathcal{R}$ over two sets A and B is a subset of $A \times B$. A binary relation on a set A is a set $\mathcal{R} \subseteq A \times A$. Using (3) of Proposition (3.1.3), we deduce that the star-transform $\mathcal{R}^*$ is also a binary relation on A^*, hence $\mathcal{R}^* \subseteq A^* \times A^*$. Consequently, it follows that the nonstandard extension of a function $f : A \longrightarrow B$ is a set $f^* \subseteq A^* \times B^*$ since a function is a relation. Thus, $f^* : A^* \longrightarrow B^*$.

The next proposition gives a characterisation of the star-transform of the function and we refer the reader to (Benci and Di Nasso, 2003) for its proof.

3.1.4 Proposition. (Benci and Di Nasso, 2003) Let A and B be sets. Then the star transform of a function f from A to B is defined to be a function f^* from A^* to B^* such that for every sequence $\phi : \mathbb{N} \longrightarrow A$,

$$f^*(\phi[\alpha]) = (f \circ \phi)[\alpha].$$

Moreover, f is injective (or surjective,respectively) if and only if f^* is injective (or surjective, respectively).

3.1.5 Remark. Since $\phi^*(\alpha) = \phi^*(1_{\mathbb{N}}[\alpha]) = (\phi \circ 1_{\mathbb{N}})[\alpha] = \phi[\alpha]$, we will write "$\phi^*(\alpha)$" in place of $\phi[\alpha]$. But we will drop $*$ to avoid confusion, and $\phi^*(\alpha)$ becomes $\phi(\alpha)$.

3.2 The hyperreal line

3.2.1 Definition. (Benci et al., 2010) Let $\mathbb{R}$ be the set of real numbers, then its star transform $\mathbb{R}^*$ is the set

$$\mathbb{R}^* = \{\phi(\alpha) | \phi : \mathbb{N} \longrightarrow \mathbb{R}\},$$

the set of ideal values of all real valued sequences. This is called the set of hyperreals.

3.2.2 Remark. To show that $\mathbb{R}$ and $\mathbb{R}^*$ are "genuinely" different, note that there exist $\alpha \in \mathbb{R}^*$ such that,

$$2 < \alpha, 2^2 < \alpha, 2^3 < \alpha, \cdots .$$

However, by Remark (3.2.2) and the number axiom, we have that $\mathbb{R} \subset \mathbb{R}^*$. The examples of hyperreal numbers are:

$$cos\frac{8}{4\alpha+3}, \quad (\alpha - \alpha^2)!$$

given as values at infinity for the real sequences

$$\left\{cos\frac{8}{4n+3}\right\}_n, \quad \{(n-n^2)!\}_n$$

respectively.

3.2.3 Remark. We have that $\mathbb{R}^*$ has the same first order properties as the field of real numbers and so it is a real-closed field.

We will denote all the binary function symbols in the language of $\mathbb{R}^*$ in a similar way with those of $\mathbb{R}$. That is, for example instead of using the star transform $+^*$ for addition in $\mathbb{R}^*$, we will use $+$ for sum function. So for all sequences φ, ψ we have the following as true

$$\xi(\alpha) = \phi(\alpha) + \theta(\alpha) \quad \text{with } \xi \text{ being sequence} \quad \xi(n) = \phi(n) + \theta(n).$$

Similarly, it will hold for all other binary functions even though for the product inverse function one has to be careful because it is not trivial that when $\varphi(\alpha) \neq 0$ it is always the case with $\varphi(n)$. But we address this by using Proposition (2.2.4) whereby if $\phi(\alpha) \neq 0$ then $\xi(\alpha) = \frac{1}{\phi(\alpha)}$ for some sequence ξ such that for all n with $\varphi(n) \neq 0$, $\xi(n) = \frac{1}{\phi(n)}$.

3.2.4 Remark. Benci and Di Nasso (2003) define an ordering on $\mathbb{R}^*$ to be given by the nonstandard extension $<^*$ of the ordering on $\mathbb{R}$. That is for any $\xi, \zeta \in \mathbb{R}^*$

$$\xi <^* \zeta \Leftrightarrow \zeta - \xi \in (\mathbb{R}_+)^*, \quad \text{where } \mathbb{R}_+ \text{ is the set of positive reals.}$$

The next proposition shows how values at infinity can be ordered, and we refer the reader for the proof to (Benci and Di Nasso, 2003) and we will drop $*$ so that $<^*$ becomes $<$ because the inclusion $\mathbb{R} \subset \mathbb{R}^*$ implies that both $<$ and $<^*$ are ordering relation symbols in $\mathbb{R}^*$.

3.2.5 Proposition. (Benci and Di Nasso, 2003) Let ϕ, θ be real sequences. Then $\phi(\alpha) < \theta(\alpha)$ whenever $\phi(n) < \theta(n)$ for all n.

3.2.6 Theorem. (Benci et al., 2010) The set of hyperreal numbers $\mathbb{R}^*$ is an ordered field.

Proof. Since $\mathbb{R}$ is an ordered field and ordering is a first order property, then as a consequence of Remark (3.2.4), we can deduce that $\mathbb{R}^*$ is also an ordered field. □

3.3 The hypernatural numbers

Definition (Benci and Di Nasso, 2003) The set of hypernatural numbers

$$\mathbb{N}^* = \{\phi(\alpha)|\phi : \mathbb{N} \longrightarrow \mathbb{N}\}$$

is the star-transform of the set of natural numbers.

The next proposition describes some few natural properties of the set of hypernatural numbers.

3.3.1 Proposition. (Benci and Di Nasso, 2003)

(1) The set of natural numbers is a proper subset of the set of hypernatural numbers;

(2) The set of hypernatural numbers is not bounded above in the hyperreal line.

(3) There are no hypernatural numbers strictly between two consecutive hypernatural numbers. That is, for every $\xi \in \mathbb{N}^*$ there is no $\zeta \in \mathbb{N}^*$ such that $\xi < \zeta < \xi + 1$.

We refer the reader to (Benci and Di Nasso, 2003) for the proofs of (1) and (2), and prove as an example (3). Before we give the proof we consider the following lemma.

3.3.2 Lemma. If $n \in \mathbb{N}$ and $n < m \in \mathbb{N}$, then $n + 1 \leq m$.

Proof. We proceed by induction on n as follows:

Let $m \in \mathbb{N}$ be arbitrarily chosen and $n = 1$. Then, if $1 < m$, we have that $m - 1 \in \mathbb{N}$. So, $1 \leq m - 1$ and thus, $1 + 1 \leq m$.

Assume that for arbitrarily chosen $m \in \mathbb{N}$, if we choose $n = k \in \mathbb{N}$ such that $k < m$, then $k + 1 \leq m$. We claim that if $(k + 1) < m$, then $(k + 1) + 1 \leq m$. To prove our claim, we assume that $k + 1 < m$ which implies that $1 < m$, and thus $m - 1 \in \mathbb{N}$. Also $k + 1 < m$ implies that $k < m - 1$. So since $m - 1 \in \mathbb{N}$ is arbitrary, then by inductive hypothesis, $k + 1 \leq m - 1$. Hence, $(k + 1) + 1 \leq m$. □

Proof. To prove (3) of Proposition (3.3.1), let $n \in \mathbb{N}$ and assume there exists a number m such that $n < m < n + 1$. If $m \in \mathbb{N}$, then by Lemma (3.3.2) applied on $n < m$, it follows that $n + 1 \leq m$ which contradicts the fact that $m < n + 1$. Thus, $m \notin \mathbb{N}$. So, by the transfer principle on ordering we have that there is no $\zeta = \zeta(\alpha) \in \mathbb{N}^*$ such that for any $\xi(\alpha) \in \mathbb{N}^*$, $\xi(\alpha) < \zeta(\alpha) < \xi(\alpha) + 1$. □

We know that $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$. We can introduce also hyper extensions for the set of integers and rational numbers to have $\mathbb{Z}^*$ and $\mathbb{Q}^*$ in the similar way as we did with the set of natural numbers and real numbers. So, we have that

$$\mathbb{N}^* \subset \mathbb{Z}^* \subset \mathbb{Q}^* \subset \mathbb{R}^*,$$

which follows by one of the consequences of the star transform on set operations (4) of Proposition (3.1.3) that says $A \subseteq B \Leftrightarrow A^* \subseteq B^*$.

3.3.3 Proposition. (Benci and Di Nasso, 2003)

(1) Given any hyperreal ζ, there exists a hyperinteger ξ such that $\xi \le \zeta < \xi + 1$;

(2) (Density of hyperrationals in hyperreals). If ζ and η are hyperreal numbers such that $\zeta < \eta$, then there exists a hyperrational number ξ such that $\zeta < \xi < \eta$.

Proof. (1) Let $\zeta = \zeta(\alpha) \in \mathbb{R}^*$ and $\xi(n)$ be the largest integer such that $\xi(n) \le \zeta(n)$ for all $n \in \mathbb{N}$. Then $\xi(n)$ is said to be the floor of $\zeta(n)$ and thus by properties of floor function, we must have

$$\zeta(n) < \xi(n) + 1$$

Thus, $\xi(\alpha) \le \zeta(\alpha) < \xi(\alpha) + 1$ by the transfer principle on ordering. □

Before we can prove (2) we give the following result as a corollary to (1) and a lemma that will be useful in proving this corollary.

3.3.4 Lemma. For all $x \in \mathbb{R}$ there exists $k \in \mathbb{Z}$ such that $k < x$.

Proof. Let $x \in \mathbb{R}$. Then, there exists $n \in \mathbb{N}$ such that $n > -x$. Hence, $k = -n < x$. □

3.3.5 Corollary. For every hyperreal $\zeta = \zeta(\alpha)$ there exists a unique hyperinteger $\xi = \xi(\alpha)$ with $\zeta - 1 \le \xi < \zeta$.

Proof. Let $x \in \mathbb{R}$. By Lemma (3.3.4), there exists $n \in \mathbb{Z}$ such that $n < x$. Since, a set of natural numbers $\mathbb{N}$ is unbounded above there exists $m \in \mathbb{N}$ such that $x < m$. Hence

$$n < x < m.$$

Now, choose the largest $k \in \mathbb{Z}$ from the finite collection $n, n+1, \cdots, m$ such that

$$k < x.$$

Then, $k + 1 \ge x$, and consequently, $x - 1 \le k < x$. Therefore, by the transfer principle on ordering $\zeta(\alpha) - 1 \le \xi(\alpha) < \zeta(\alpha)$. □

We now give the proof for (2) of Proposition (3.3.3),

Proof. (2) Suppose $\zeta = \zeta(\alpha) > 0$ and let $\eta = \eta(\alpha) > \zeta(\alpha)$. Then there exists $\varphi(\alpha) \in \mathbb{N}^*$ such that

$$\varphi(\alpha) > \frac{1}{\eta(\alpha) - \zeta(\alpha)}.$$

It follows that $\varphi(\alpha) \cdot \eta(\alpha) > 1 + \varphi(\alpha) \cdot \zeta(\alpha)$ which can be written as

$$\zeta(\alpha) < \eta(\alpha) - \frac{1}{\varphi(\alpha)}. \tag{3.3.1}$$

By Corollary (3.3.5) there exists $\xi(\alpha) \in \mathbb{Z}^*$ such that $\varphi(\alpha) \cdot \eta(\alpha) - 1 \le \xi(\alpha) < \varphi(\alpha) \cdot \eta(\alpha)$ which implies that

$$\eta(\alpha) - \frac{1}{\varphi(\alpha)} \le \frac{\xi(\alpha)}{\varphi(\alpha)} < \eta(\alpha). \tag{3.3.2}$$

Combining equations (3.3.1) and (3.3.2) yields

$$\zeta(\alpha) < \eta(\alpha) - \frac{1}{\varphi(\alpha)} \le \frac{\xi(\alpha)}{\varphi(\alpha)} < \eta(\alpha),$$

hence $\zeta(\alpha) < \frac{\xi(\alpha)}{\varphi(\alpha)} < \eta(\alpha)$.

□

3.4 Hyperfinite sets

Besides the set of hyperreal numbers another structure that one can consider is the so called hyperfinite set.

3.4.1 Definition. (Benci et al., 2010) Let A^* be the star transform of a nonempty set A. We define the set $\Lambda \subset A^*$ to be a hyperfinite set if

$$\Lambda = \{\phi(\alpha) : \phi(n) \in A_n \quad \text{for all } n\},$$

where $A_n \subset A$ is sequence of finite sets.

By internal set axiom, elements of Λ are the values at infinity for the elements of A_n. The cardinality of Λ is defined as follows:

$$|\Lambda| = \theta(\alpha) \in \mathbb{N}^*$$

with $\theta(n) = |A_n|$ being the cardinality of the finite set A_n.

Hyperfinite sets preserve some of the properties of finite sets. We know that any ordered finite set has a largest and smallest element. The next proposition shows that ordered hyperfinite sets retain the property of ordered finite sets having maximum and minimum elements.

3.4.2 Proposition. (Benci et al., 2010) Let $\Lambda \subset \mathbb{R}^*$ be a nonempty hyperfinite set. Then Λ has greatest and smallest element.

Proof. Let $\theta \subset \mathbb{R}^*$ be a hyperfinite set. Then by definition

$$\theta(\alpha) = \{\phi(\alpha) : \phi(n) \in \theta(n)\},$$

where $\{\theta(n)\}_n$ is a sequence of non-empty finite subsets of $\mathbb{R}$. For all $n \in \mathbb{N}$, we let $\mu(n) \in \theta(n)$ to be the smallest element. That is, for all $\phi(n) \in \theta(n)$, we have that $\mu(n) \leq \phi(n)$, which means

$$\mu(n) = min(\theta(n)).$$

So, since the statement 'for all $\phi(n) \in \theta(n)$' and ordering are first order statements, then by the transfer principle, it follows that, for all $\phi(\alpha) \in \theta(\alpha)$, we have that $\mu(\alpha) \leq \phi(\alpha)$. Thus, it follows that

$$\mu(\alpha) = min(\theta(\alpha))$$

since $\mathbb{R}^*$ is an ordered field. □

The proof for the greatest element proceeds analogously.

An example of hyperfinite set is the hyperfinite set grid $\mathbb{H}$ (Benci et al., 2010). The hyperfinite grid $\mathbb{H}_\alpha$ is defined as the α-value of the set

$$\mathbb{H}_n = \{\frac{k}{2} : k \in \mathbb{Z}, \frac{-n^2}{2} < k < \frac{n^2}{2}\};$$

namely,

$$\mathbb{H}_\alpha = \{\frac{k}{2} : k \in \mathbb{Z}, \frac{-\alpha^2}{2} < k < \frac{\alpha^2}{2}\}.$$

As in the similar manner with finite sets in the classical mathematics, we can find the norm of our hyperfinite set $\mathbb{H}_\alpha$ given by

$$\begin{aligned}|\mathbb{H}_\alpha| &= \left|\frac{\alpha^2}{2} - \frac{-\alpha^2}{2}\right| \\ &= \alpha^2.\end{aligned}$$

4. Introducing Nonstandard Calculus

In this section, we introduce the notion of infinitely small and infinitely large numbers and take a closer look at many of their properties and give a few examples. We then briefly look at the connection between ideal values and the notion of limits. The concept of countable saturation is usually used to prove the existence of infinitesimals, however, Benci and Di Nasso (2003) obtain it as a consequence of α-theory. We briefly discuss this and outline their proof of this result. We define an Archimedean field, then continue by stating and giving the proof of the shadow theorem and conclude the Chapter.

4.1 Infinitely small and Infinitely large numbers

We adopt the usual absolute value notation:

$$|\xi| = \begin{cases} \xi & \text{if} \quad \xi \geq 0; \\ -\xi & \text{if} \quad \xi < 0 \end{cases}$$

4.1.1 Definition. (Benci and Di Nasso, 2019) For $\xi \in \mathbb{F}$, where $\mathbb{F}$ is an ordered field, we make the following definition:

(1) If $|\xi| < \frac{1}{n}$ for all $n \in \mathbb{N}$ then ξ is called an infinitesimal;

(2) If there exists $k \in \mathbb{N}$ such that $|\xi| < k$ then ξ is said to be finite (or bounded)

(3) If $|\xi| > k$ for all $k \in \mathbb{N}$ then ξ is infinite (or unbounded) or (equivalently, if ξ is not finite).

Consequently, we have that all infinitesimal numbers are finite. Also infinitesimals and infinite numbers are multiplicative inverses of each other, apart from zero.

4.1.2 Remark. We note that in $\mathbb{R}, \epsilon = 0$ is the only number that satisfies $-\frac{1}{n} < \varepsilon < \frac{1}{n}$ for all $n \in \mathbb{N}$, hence, the only infinitesimal.

We refer the reader to (Benci and Di Nasso, 2019) for the proofs of the following propositions.

4.1.3 Proposition. (Benci and Di Nasso, 2019) An ordered field is Archimedean if and only if its only infinitesimal number is 0.

4.1.4 Proposition. (Benci and Di Nasso, 2019) Any ordered field $\mathbb{F}$ that properly extends the real field $\mathbb{R}$ is non-Archimedean.

It follows from these Propositions (4.1.3) and (4.1.4) that $\mathbb{R}$ and $\mathbb{R}^*$ are archimedean and nonarchimedean fields respectively. In light of the transfer principle, we can immediately deduce that being archimedean is not a first order property.

We have that $\mathbb{R}^*$ is an ordered field with infinitesimal and infinite elements. Since by Remark (3.2.3) and Theorem (3.2.6), $\mathbb{R}^*$ was classified as real-closed and ordered field respectively, then we can say it is a real-closed ordered field.

4.1.5 Proposition. (Benci and Di Nasso, 2003)

(1) The product of any infinitesimal number ε and any finite number ξ is infinitesimal;

(2) For any infinitesimal number $\varepsilon \neq 0$ and any number ξ which is not infinitesimal, then $\frac{\xi}{\varepsilon}$ is infinite.

Proof. (1) Suppose ξ is finite, then there exists $k \in \mathbb{N}$ such that $|\xi| < k$. Let $n \in \mathbb{N}$. Then, it follows that $nk \in \mathbb{N}$. Since, ε is infinitesimal, then $|\varepsilon| < \frac{1}{nk}$ and so $|\xi \cdot \varepsilon| = |\xi| \cdot |\varepsilon| < k \cdot \frac{1}{nk} = \frac{1}{n}$. Since this is true for all $n \in \mathbb{N}$, then it follows that $\xi \cdot \varepsilon$ is infinitesimal. □

Proof. (2) Suppose ξ is not an infinitesimal, then there exists $n \in \mathbb{N}$ such that $|\xi| > \frac{1}{n}$. Let $k \in \mathbb{N}$. Then, it follows that $nk \in \mathbb{N}$. Since $\varepsilon \neq 0$ is infinitesimal, certainly $|\varepsilon| < \frac{1}{nk}$ and so $|\varepsilon| < \frac{1}{nk} < \frac{1}{n} \cdot \frac{1}{k}$. So, $|\varepsilon| < |\xi| \cdot \frac{1}{k}$. Therefore, $k < \frac{|\xi|}{|\varepsilon|} = \left|\frac{\xi}{\varepsilon}\right|$. Since this is true for all $k \in \mathbb{N}$, it follows that $\frac{\xi}{\varepsilon}$ is infinite. □

We refer the reader to (Benci and Di Nasso, 2003) for other properties that have not been considered here such as infinitesimals being closed under addition. In fact, algebraically, the set of infinitesimals is a maximal ideal of $\mathbb{R}^*$. And in fact, $\mathbb{R}^*/I \cong \mathbb{R}$.

4.1.6 Definition. (Benci and Di Nasso, 2003) We say that two hyperreal numbers ξ and η are infinitely close if $\xi - \eta$ is infinitesimal. In this case we write $\xi \sim \eta$.

We have seen in the previous chapter that the hyperfinite sets have greatest and smallest element. We relate this to the subsets of $\mathbb{R}$ in the next proposition.

4.1.7 Proposition. (Benci and Di Nasso, 2003) Let A be a nonempty set such that $A \subseteq \mathbb{R}$, and let $l \in \mathbb{R}$. Then:

(1) Sup $A = l \Leftrightarrow l \geq A$, that is, $l \geq a$ for all $a \in A$ and $l \sim \xi$ for some $\xi \in A^*$;

(2) inf $A = l \Leftrightarrow l \leq A$ and $l \sim \xi$ for some $\xi \in A^*$;

(3) sup $A = +\infty$ [or inf $A = -\infty$] if and only if there exists an infinite $\xi \in A^*$ which is positive [negative, respectively].

Proof. We give the proof for (2) as follows:

(2) Assume that inf $A = l$. For every $n \in \mathbb{N}$ there exists $\xi(n) \in A$ with

$$l \leq \xi(n) < l + \frac{1}{n}.$$

Then, $\xi = \xi(\alpha) \in A^*$ is such that $l \sim \xi$.

Conversely, let $l \leq A$ and $l \sim \xi$ for some $\xi = \xi(\alpha) \in A^*$. If $l \neq$ inf A, then there exists $n \in \mathbb{N}$, with $a > l + \frac{1}{n}$ for all $a \in A$. It follows that $\xi > l + \frac{1}{n}$ which implies that $\xi - l > \frac{1}{n}$. Hence, $\xi \nsim l$ for all $\xi \in A^*$, which contradicts the hypothesis. □

The next theorem is a more general shadow theorem to the one found in (Benci and Di Nasso, 2003) about the hyperreal field, as it is stated based on any field that properly extends the real field $\mathbb{R}$. Thus, the shadow theorem specifically for the field of hyperreals can be taken as a corollary to this theorem.

4.1.8 Theorem. (Shadow Theorem) (Benci and Di Nasso, 2003) Let $\mathbb{F}$ be an ordered field that properly extends $\mathbb{R}$. Then, every finite number $\xi \in \mathbb{F}$ is infinitely close to a unique real number $r \sim \xi$ called the shadow of ξ. Symbolically $r = sh(\xi)$.

Proof. Since ξ is finite, the set of its real lower bounds

$$A = \{a \in \mathbb{R} | \xi > a\}$$

is non-empty and bounded above. Let $r =$ sup $A \in \mathbb{R}$ be its greatest lower bound. We claim that $r \sim \xi$.

We proceed to prove the claim. Let $n \in \mathbb{N}$ be fixed. Then, by properties of the greatest lower bounds, we have that

(1) $\xi \leq r + \frac{1}{n}$ which implies $\xi - r \leq \frac{1}{n}$

(2) $\xi > r - \frac{1}{n}$ which implies $\xi - r > -\frac{1}{n}$

Thus, we conclude that $\xi - r$ is infinitesimal. Hence, $\xi \sim r$.

Finally, as for the uniqueness, two real numbers are necessarily equal if they are infinitely close. Indeed, since the only infinitesimal real number is 0, if $r \sim \xi \sim x$, then $r - x = 0$. Hence, $r = x$. □

4.1.9 Corollary. Every finite hyperreal number ξ is infinitely close to a unique real number r, called the shadow of ξ. symbolically $r = sh(\xi)$.

4.1.10 Remark. The shadow of the hyperreal number ξ is commonly referred to in non-standard analysis literature as the "standard part" of ξ.

4.1.11 Remark. (Benci and Di Nasso, 2019) Let $\mathbb{F}$ be an ordered field that properly extends $\mathbb{R}$. Then every finite number $\xi \in \mathbb{F}$ has a unique representation in the canonical form:

$$\xi = sh(\xi) + \varepsilon$$

where $sh(\xi) \in \mathbb{R}$ is a real number, and $\varepsilon = \xi - sh(\xi)$ is infinitesimal.

Example:

$$\begin{aligned} \frac{\alpha + 8}{2\alpha + 2} &= \frac{(\alpha + 1) + 7}{2(\alpha + 1)} \\ &= \frac{1}{2} + \frac{7}{\alpha + 1}. \end{aligned}$$

This implies that $sh(\xi) = \frac{1}{2}$ and $\epsilon = \frac{7}{\alpha+1}$

We observe that

$$\xi = \frac{\alpha + 8}{2\alpha + 2}$$

is an example of a bounded hyperreal number that is neither real nor infinitesimal.

For unbounded hyperreal numbers ξ, we set $sh(\xi) = +\infty$ and $sh(\xi) = -\infty$ for unbounded positive and negative ξ respectively. So, Proposition (4.1.7) can now be restated as

1. sup $A = l \Leftrightarrow l \geq A$ and $sh(\xi) = l$ for some $\xi \in A^*$;
2. inf $A = l \Leftrightarrow l \leq A$ and $sh(\xi) = l$ for some $\xi \in A^*$.

4.2 Ideal values at α and the notion of limit.

In this subsection, we relate the notion of limits of sequences to the concept of ideal values. The limit of a real sequence $\{\varphi(n)\}_n$ is related to its ideal value $\varphi(\alpha)$ (Benci and Di Nasso, 2003). The limit of a sequence $\{\varphi(n)\}_n$ can be expressed as the shadow of the value $\varphi(\alpha)$ of the hyperreal sequence at the ideal value α. Thus,

$$\lim_{n \to \infty} \varphi(n) = sh(\varphi(\alpha)).$$

This works only when $\varphi(n)$ has a limit. Since the shadow function "sh" "rounds" each finite hyperreal number ξ to the nearest real number, that is, since $\xi - sh(\xi)$ is infinitesimal.

This formalises the natural intuition that for the ideal value α, the terms in the sequence are "very close" to the limit value of the sequence. In other words, this is to formally say, regardless which $\epsilon > 0$ we have, for a convergent sequence a_n there exists $N \in \mathbb{N}$ large enough, so that the sequence lies afterwards completely in the ϵ-neighbourhood of a_n $(a_n - \epsilon, a_n + \epsilon)$.

Before stating and proving the saturation principle, we introduce one last notion, that of internal sets.

4.3 Internal Sets

Definition (Benci and Di Nasso, 2003) An entity is *internal* if it is the ideal value of some sequence. An entity is *external* if it is not *internal*.

It follows that all the hyperreal numbers and the ideal number α are internal by the number axiom. In general, by definition the nonstandard extension A^* is internal. Consequently, its elements are also internal which means *transitivity* holds in the collection of internal sets.

4.3.1 Remark. It can be shown that the collection of internal sets is closed under sets operations. For example, if A and B are internal sets then $A \cap B$ is also an internal set.

Proof. Indeed, if we let $A = \{\phi(\alpha) | \phi : \mathbb{N} \longrightarrow C\}$ and $B = \{\theta(\alpha) | \theta : \mathbb{N} \longrightarrow D\}$. Then, by definition A and B are internal sets, and by (2) of Proposition (3.1.3) in Chapter 3, we have that $A \cap B$ is also internal. □

We refer the reader to (Benci and Di Nasso, 2003) for further reading on internal sets. We finally state the *saturation principle*.

4.3.2 Theorem(Countable Saturation Principle). (Benci and Di Nasso, 2003)

Let $\{A_k | k \in \mathbb{N}\}$ be a countable family of internal sets such that for each finite $n \in \mathbb{N}$, we have that $A_1 \cap \cdots \cap A_n \neq \emptyset$. Then, $\cap_k A_k \neq \emptyset$.

Proof. Let $\varphi_k(\alpha) = A_k$ for every k. For fixed n, we construct the following intersections

$$\varphi_1(n), \quad \varphi_1(n) \cap \varphi_2(n), \quad \varphi_1(n) \cap \varphi_2(n) \cap \varphi_3(n), \cdots, \varphi_1(n) \cap \varphi_2(n) \cap \cdots \cap \varphi_{n-1}(n) \cap \varphi_n(n).$$

Let $\psi(n)$ be such that

$$\psi(n) \in \varphi_1(n) \cap \varphi_2(n) \cap \cdots \cap \varphi_{n-1}(n) \cap \varphi_n(n)$$

if this intersection is nonempty. Otherwise, let $\psi(n) \in \varphi_1(n) \cap \varphi_2(n) \cdots \cap \varphi_{n-1}(n)$ if this intersection is nonempty. Otherwise, repeat iteratively until we get an element.

If $\varphi_1(n) = \emptyset$, let $\psi(n)$ be an arbitrary element. It follows by relabelling our sets that if $n \geq k$, then

$$\varphi_1(n) \cap \varphi_2(n) \cap \cdots \cap \varphi_k(n) \neq \emptyset.$$

So, $\psi(n) \in \varphi_1(n) \cap \varphi_2(n) \cap \cdots \cap \varphi_k(n)$. Recall that Proposition (2.2.1) the Difference (4) states that if $\phi(n) \neq \varphi(n)$ for all $n \in \mathbb{N}$, then $\phi[\alpha] \neq \varphi[\alpha]$, and the Intersection (6) states that if $\vartheta(n) = \phi(n) \cap \varphi(n)$ for all n, then $\vartheta[\alpha] = \phi[\alpha] \cap \varphi[\alpha]$. So, if we let $\vartheta(n) = \emptyset$ then we have that $\varphi_1(\alpha) \cap \varphi_2(\alpha) \cap \cdots \cap \varphi_k(\alpha) \neq \emptyset$.

So, if

$$n \in \{m \geq k | \varphi_1(m) \cap \varphi_2(m) \cap \cdots \cap \varphi_k(m) \neq \emptyset\}$$

then it follows that $\psi(n) \in \varphi_1(n) \cap \varphi_2(n) \cap \cdots \cap \varphi_k(n)$. Hence, by (2) of Proposition (2.2.4) which says if $\phi(n) \in \theta(n)$ for all n, then $\phi(\alpha) \in \theta(\alpha)$, it turns out that

$$\psi(\alpha) \in \varphi_1(\alpha) \cap \varphi_2(\alpha) \cap \cdots \cap \varphi_k(\alpha)$$

for all k. □

In an effort to introduce nonstandard analysis whilst avoiding the machinery of model theory which requires a formal training in first order logic, Benci and Di Nasso (2003) gave a system of axioms governing the use of an "infinitely large" natural number α. At first these axioms were stated informally. They then proceed to build nonstandard analysis, and obtain the *transfer principle* as a consequence of their α-theory. Then, they were able to define concepts of nonstandard analysis and compare results of their α-theory to that of others such as comparing Strong Cauchy's Infinitesimals Principles to some of their statements.

In the last Chapter of their paper, they expressed their axioms in first order language, demonstrating that α-theory can in principle be approached in purely formal terms. For example, the composition axiom reads as

$J2$ *Composition Axiom.* If ϕ and θ are sequences and if f is any function such that compositions $f \circ \phi$ and $f \circ \theta$ are well defined, then

$$\forall x[(J(\phi, x) \wedge J(\theta, x)) \longrightarrow \exists y(J(f \circ \phi, y) \wedge J(f \circ \theta, y)].$$

where J is a function defined on the class of all sequences.

As all axioms were formalised similarly, they were able to give a formal definition of α-theory, hence proving the transfer principle as a consequence of Łos theorem of ultrapowers given as a remark below.

4.3.3 Remark. (Benci and Di Nasso, 2003) The Alpha-Theory proves the transfer principle. That is, for every bounded formula $\sigma(x_1, \cdots, x_k)$ in the first order language $\mathcal{L} = \{\in\}$ of set theory, and for every $a_1, \cdots, a_k$, $\sigma(a_1, \cdots, a_k) \Leftrightarrow \sigma(a_1^*, \cdots, a_k^*)$.

5. Foundations and the Construction of Alpha

In this Chapter we show how to construct a model of non-standard reals via model theory approach and show how it relates with the α-theory approach. Our treatment closely follow the set of lecture notes on a course of model theory by Boxall (2018) which is our main reference.

5.1 Ultrafilters

5.1.1 Definition. (Boxall, 2018) Let X be a set. Then the powerset of X is the set $\mathbb{P}(X) = \{A : A \subseteq X\}$. Let J be a set. A **filter** on J is a subset $\mathcal{F} \subseteq \mathbb{P}(J)$ such that

(1) $\emptyset \notin \mathcal{F}$;

(2) For all $A, B \in \mathcal{F}, \quad A \cap B \in \mathcal{F}$;

(3) For all $A \in \mathcal{F}$ and $B \subseteq J$, if $A \subseteq B$ then $B \in \mathcal{F}$.

5.1.2 Definition. (CODY, 2015) We say that $\mathcal{F}$ is an **ultrafilter** over a set J if $\mathcal{F}$ is a filter on J and the following property also holds: For every $B \subseteq J$, either $B \in \mathcal{F}$ or $J \setminus B \in \mathcal{F}$.

The next example is a solution to an exercise from (Boxall, 2018) and it is picked as it will be needed later.

Example: Suppose that $J = \mathbb{N}$ and let $\mathcal{F} = \{A \subseteq \mathbb{N} : \mathbb{N} \setminus A \quad \text{is finite}\}$. Then,

1) $\mathbb{N} \setminus \emptyset = \mathbb{N}$ is infinite. Thus, $\emptyset \notin \mathcal{F}$.

2) Let $A, B \in \mathcal{F}$. Then, $\mathbb{N} \setminus (A \cap B) = (\mathbb{N} \setminus A) \cup (\mathbb{N} \setminus B)$. Thus the union of finite sets. Therefore $A \cap B \in \mathcal{F}$.

3) Let $A \in \mathcal{F}$ and $B \subseteq J$. Suppose $A \subseteq B$. Then, $\mathbb{N} \setminus B \subseteq \mathbb{N} \setminus A$ which is finite. So, $\mathbb{N} \setminus B$ is finite. So, $B \in \mathcal{F}$.

   Therefore, $\mathcal{F}$ is a filter on J.

The next proposition tells us that any given filter $\mathcal{F}$ can be "extended" to an ultrafilter. We state it without proof, for a discussion one may consult (Boxall, 2018).

5.1.3 Proposition. (Boxall, 2018) Let J be a set and let $\mathcal{F}$ be a filter on J. Then there is an ultrafilter U on J such that $\mathcal{F} \subseteq U$.

5.2 Ultraproducts

Let $J \neq \emptyset$ be an indexing set, and $(A_j)_{j \in J} \neq \emptyset$ be a family of non-empty sets. Let $\mathcal{U}$ be an ultrafilter on J.

5.2.1 Definition. (Boxall, 2018) For all $j \in J$, we define the Cartesian product of the sets A_j with ultrafilter $\mathcal{U}$ as follows, $T = \prod_{j \in J} A_j = \{(a_j)_{j \in J} : a_j \in A_j, \forall j \in J\}$. We define a relation $\sim$ on T as follows: For all $(a_j)_{j \in J}, (b_j)_{j \in J} \in T$, then $\quad (a_j)_{j \in J} \sim (b_j)_{j \in J}$ if $\{j \in J : a_j = b_j\} \in \mathcal{U}$.

The relation $\sim$ is an equivalence relation on T. As an example, we demonstrate reflexivity of $\sim$. Let $(a_j)_{j\in J} \in T$. Then

$$\{j \in J : a_j = a_j\} = J.$$

Now, $J \in \mathcal{U}$ or $(J \setminus J) \in \mathcal{U}$. But, since $(J \setminus J) = \emptyset \notin \mathcal{U}$. It follows that $J \in \mathcal{U}$. Therefore, $(a_j)_{j\in J} \sim (a_j)_{j\in J}$. □

We can therefore look at equivalence classes $[(a_j)_{j\in J}] = \{(b_j)_{j\in J} : (b_j)_{j\in J} \sim (a_j)_{j\in J}\}$. The **ultraproduct** $(\prod_{j\in J} A_j)/\sim = (\prod_{j\in J} A_j)/\mathcal{U}$ is the set of $\sim$ equivalence classes

$$(\prod_{j\in J} A_j)/\mathcal{U} = \{[(a_j)_{j\in J}] : (a_j)_{j\in J} \in \prod_{j\in J} A_j\} = T/\mathcal{U}.$$

For the rest of our discussion on this section unless otherwise specified we will always take $T = \prod_{j\in J} A_j$.

5.2.2 Definition. (Boxall, 2018) For each $j \in J$, let $B_j \subseteq A_j$. We define

$$[(B_j)_{j\in J}] = \{[(a_j)_{j\in J}] \in T/\mathcal{U} : \{j \in J : a_j \in B_j\} \in \mathcal{U}\}.$$

We would like to verify that this definition is well defined. To that end, let $(a_j)_{j\in J}, (b_j)_{j\in J} \in T$ such that $[(a_j)_{j\in J}] = [(b_j)_{j\in J}]$. We claim that $\{j \in J : a_j \in B_j\} \in \mathcal{U}$ if and only if $\{j \in J : b_j \in B_j\} \in \mathcal{U}$.

Proof. To prove our claim, suppose $\{j \in J : a_j \in B_j\} \in \mathcal{U}$. Then, we have that $\{j \in J : a_j = b_j\} \in \mathcal{U}$. So, $\{j \in J : a_j = b_j \quad \text{and} \quad a_j \in B_j\} \in \mathcal{U}$ since $\{j \in J : a_j = b_j \quad \text{and} \quad a_j \in B_j\} \subseteq \{j \in J : b_j \in B_j\}$. Thus, $\{j \in J : b_j \in B_j\} \in \mathcal{U}$.

For the converse, it can be shown similarly that if $\{j \in J : b_j \in B_j\} \in \mathcal{U}$, then $\{j \in J : a_j \in B_j\} \in \mathcal{U}$. Therefore, $[(B_j)_{j\in J}]$ is a well defined subset of $T/\mathcal{U}$. □

5.2.3 Definition. (Boxall, 2018) Let $(A_j)_{j\in J}$ be a non-empty family of non-empty sets. Let $\mathcal{U}$ be an ultrafilter on J. Let $n \in \mathbb{N}$. For each $j \in J$, let $B_j \subseteq A_j^n$, we define

$$[(B_j)_{j\in J}] = \{([(a_j^1)_{j\in J}], \cdots, [(a_j^n)_{j\in J}]) \in (A/\mathcal{U})^n : \{j \in J : (a_j^1, \cdots, a_j^n) \in B_j\} \in \mathcal{U}\}.$$

It can be shown that this is well defined by the same reasoning as in proof given under Definition (5.2.2). We would like to check whether square bracket-operator preserves set operations. We demonstrate this for intersections.

5.2.4 Proposition. (Boxall, 2018) For all $j \in J$, let $B_j, C_j \subseteq A_j$ that is $n = 1$. Then, $[(B_j)_{j\in J}] \cap [(C_j)_{j\in J}] = [(B_j \cap C_j)_{j\in J}]$.

Proof. Let $[(a_j)_{j\in J}] \in [(B_j)_{j\in J}] \cap [(C_j)_{j\in J}]$. Then, $\{j \in J : a_j \in B_j\} \in \mathcal{U}$ and $\{j \in J : a_j \in C_j\} \in \mathcal{U}$. Let $Y_1 = \{j \in J : a_j \in B_j\} \in \mathcal{U}$ and $Y_2 = \{j \in J : a_j \in C_j\} \in \mathcal{U}$. Suppose $(Y_1 \cap Y_2) \notin \mathcal{U}$. Since $\mathcal{U}$ is an ultrafilter on J, it follows that $J \setminus (Y_1 \cap Y_2) \in \mathcal{U}$. We then have that $J \setminus (Y_1 \cap Y_2) \in \mathcal{U}$. But $J \setminus (Y_1 \cap Y_2) = (J \setminus Y_1) \cup (J \setminus Y_2)$. So, we then have both $Y_1 \in \mathcal{U}$ and $J \setminus Y_1 \in \mathcal{U}$. This implies that $\emptyset \in \mathcal{U}$, thus a contradiction. Therefore, $(Y_1 \cap Y_2) \in \mathcal{U}$. This implies that $\{j \in J : a_j \in B_j \quad \text{and} \quad a_j \in C_j\} \in \mathcal{U}$. So, $\{j \in J : a_j \in B_j \cap C_j\} \in \mathcal{U}$. Thus, $[(a_j)_{j\in J}] \in [(B_j \cap C_j)_{j\in J}]$. Hence, $[(B_j)_{j\in J}] \cap [(C_j)_{j\in J}] \subseteq [(A_j \cap B_j)_{j\in J}]$.

Conversely, Let $[(a_j)_{j\in J}] \in [(B_j \cap C_j)_{j\in J}]$. Then, $\{j \in J : a_j \in B_j \cap C_j\} \in \mathcal{U}$. By definition,$B_j \cap C_j \subseteq B_j$. So, we have that $\{j \in J : a_j \in B_j \cap C_j\} \subseteq \{j \in J : a_j \in B_j\}$. Thus, $\{j \in J : a_j \in B_j\} \in \mathcal{U}$, which implies that $[(a_j)_{j\in J}] \in [(B_j)_{j\in J}]$.

Similarly, it can be shown that $[(a_j)_{j\in J}] \in [(C_j)_{j\in J}]$. Thus, $[(a_j)_{j\in J}] \in [(B_j)_{j\in J}] \cap [(C_j)_{j\in J}]$. Hence, $[(B_j \cap C_j)_{j\in J}] \subseteq [(B_j)_{j\in J}] \cap [(C_j)_{j\in J}]$. Therefore, $[(B_j \cap C_j)_{j\in J}] = [(B_j)_{j\in J}] \cap [(C_j)_{j\in J}]$ as required. □

5.2.5 Definition. (Boxall, 2018) If $A_j = A_i$ for all $i, j \in J$, then the ultraproduct $T/\mathcal{U}$ is called an **ultrapower.** In this case, we write $A_j = A$ for all $j \in J$, and let $T/\mathcal{U}$ be $A^{\mathcal{U}}$.

The []-operator specializes to ultrapowers in the following manner. If we let $(A_j)_{j\in J}$ be a non-empty family of non-empty sets such that $A_j = A$ for all $j \in J$. Then for the ultrafilter $\mathcal{U}$ on J, and $n \in \mathbb{N}$, let $B \subseteq A^n$, such that we have a family $(B_j)_{j\in J}$, where $B_j = B$ for all $j \in J$. Then we have $[(B_j)_{j\in J}] \subseteq (A^{\mathcal{U}})^n$. So, we now have

$$[B] = \{([(a_j^1)_{j\in J}], \cdots, [(a_j^n)_{j\in J}]) \in (A^{\mathcal{U}})^n : \{j \in J : (a_j^1, \cdots, a_j^n) \in B\} \in \mathcal{U}\}.$$

We give the following Proposition and leave the proof for the reader as it follows directly from the definition of ultrapower and observe its consequences that follows at the end of this Chapter.

5.2.6 Proposition. (Boxall, 2018) Let $B_j \subseteq A_j^n$, $\mathcal{U}$ be an ultrafilter on J, and $n = 0$. Then,

(1) $[(B_j)_{j\in J}] = \emptyset \Leftrightarrow \{j \in J : B_j = \emptyset\} \in \mathcal{U}$;

(2) $[(B_j)_{j\in J}] = ((\prod_{j\in J} A_j)/\mathcal{U})^0 \Leftrightarrow \{j \in J : B_j = A_j^0\} \in \mathcal{U}$.

5.3 Structures

We think of $\mathbb{R}$ as an ordered field $\langle \mathbb{R}, +, \cdot, 1, 0\rangle$. Defining the operations such as $+$ on $\mathbb{R}$ now has added structure rather than being just a set. In this section we study the notions of first order languages and structures.

5.3.1 Definition. (Halper, 2010) A **first order language** $\mathcal{L}$ is a set that contains the sets $R_{\mathcal{L}}$, $F_{\mathcal{L}}$, and $C_{\mathcal{L}}$ of relation symbols, function symbols, and constant symbols respectively, together with the variables $x_1, x_2, \cdots$, logical symbols such as $\wedge, \longrightarrow$, and binary equality relation such as $=, \equiv$. As our study is based on first order logic, we will consider the languages that are first order.

5.3.2 Definiton. (Halper, 2010) We define an **interpretation function** I of $\mathcal{L}$ to be a correspondence between $\mathcal{L}$ and a universe set M so that I maps n-ary relations $R_{\mathcal{L}}$ to n-ary relations $R \subseteq M^n$ on M, m-ary functions $F_{\mathcal{L}}$ to m-ary functions $G : M^m \longrightarrow M$ on M, and each constant $c \in R_{\mathcal{L}}$ to a constant $c \in M$. This is to give our language the meaning so that we can interpret what the elements of our language mean.

5.3.3 Definition. (Halper, 2010) An $\mathcal{L}$-structure is the pair $\langle M, I\rangle$ of a universe set M and an interpretation function I. We usually write $\mathcal{M} = \langle M, I\rangle$ to say that M is the underlying set of the $\mathcal{L}$-structure $\mathcal{M}$ and we call $\mathcal{M}$ a model.

We now connect the notions of ultraproducts and that of $\mathcal{L}$-structures.

5.3.4 Ultraproduct $\mathcal{L}$-structures. (Boxall, 2018) For each $j \in J$, let $\mathcal{M}_j = \langle M_j, I_j\rangle$ be a model. Let $\mathcal{U}$ be an ultrafilter on non-empty set J. If we leave out the interpretations it follows that $(M_j)_{j\in J}$ is a family of non empty sets, so it is possible to form the ultraproduct $(\prod_{j\in J} M_j)/\mathcal{U} = T/\mathcal{U}$, where $T = (\prod_{j\in J} M_j)$. So, it follows that we can define an interpretation I of $\mathcal{L}$ in $T/\mathcal{U}$. To that end, let

$R_{\mathcal{L}}, F_{\mathcal{L}}$ and $C_{\mathcal{L}}$ be as defined in Definition (5.3.1) and define a function $P_{\mathcal{L}} : R_{\mathcal{L}} \cup F_{\mathcal{L}} \longrightarrow \mathbb{Z}_+$ of the language $\mathcal{L}$ where $\mathbb{Z}_+ = \mathbb{N} \setminus \{0\}$. Then, the following definition holds for the interpretation I of $\mathcal{L}$ in $A/\mathcal{U}$:

5.3.5 Definition. (Boxall, 2018)

(1) Let $s \in R_{\mathcal{L}}$ with $P_{\mathcal{L}}(s) = n$. Then $I(s) = [(I_j(s))_{j\in J}] \subseteq (T/\mathcal{U})^n$;

(2) Let $s \in C_{\mathcal{L}}$. Then $I(s) = [(I_j(s))_{j\in J}] \in T/\mathcal{U}$.

For function symbols, suppose $s \in F_{\mathcal{L}}$ with $P_{\mathcal{L}}(s) = n$. Then, for each $j \in J$, define $G_j(s) = \{(a_1, \cdots, a_n, b) \in M_j^{n+1} : I_j(s)(a_1, \cdots, a_n) = b\}$. So, it follows that $G_j(s)$ is the graph of the function $I_j(s)$.

It can be proved that $[(G_j(s))_{j\in J}]$ is the graph of a function from $(T/\mathcal{U})^n$ to $T/\mathcal{U}$.

(3) Define $I(s) : (T/\mathcal{U})^n \longrightarrow T/\mathcal{U}$ to be a function whose graph is $[(G_j(s))_{j\in J}]$.

We shall always refer to this interpretation I of $\mathcal{L}$ in $T/\mathcal{U}$ as $I_{\mathcal{U}}$ and call it the ultraproduct interpretation with respect to $(M_j)_{j\in J}$ and $\mathcal{U}$. The $\mathcal{L}$-structure $(T/\mathcal{U}, I_{\mathcal{U}})$ is called the ultraproduct of $(M)_{j\in J}$ with respect to $\mathcal{U}$.

5.3.6 Definition. (Boxall, 2018) The terms of language $\mathcal{L}$, which are referred to as $\mathcal{L}$-terms are defined as follows:

(1) A variable is an $\mathcal{L}$-term. That is every element in $\{x_0, x_1, \cdots\}$;

(2) A constant $c \in C_{\mathcal{L}}$ is an $\mathcal{L}$-term;

(3) Given $f \in F_{\mathcal{L}}$ with $P_{\mathcal{L}} = n$ and if $x_1, x_2, \cdots, x_n$ are $\mathcal{L}$-terms, then $f(x_1, x_2, \cdots, x_n)$ is an $\mathcal{L}$-term.

We now move on to defining the idea of a first order formula. To define the concept of an $\mathcal{L}$-formula of $\mathcal{L}$, we begin by using the relation symbols, $\mathcal{L}$-terms, and equality to define atomic $\mathcal{L}$-formula.

5.3.7 Definition. (Boxall, 2018) We define atomic $\mathcal{L}$-formulas as follows

(1) Let t_1 and t_2 be $\mathcal{L}$-terms, then $t_1 = t_2$ is an atomic $\mathcal{L}$-formula;

(2) Let $t_1, \cdots, t_k$ be $\mathcal{L}$-terms and $s \in R_{\mathcal{L}}$ with $P_{\mathcal{L}}(s) = k$, then $(s(t_1, \cdots, t_k))$ is an atomic $\mathcal{L}$-formula.

Finally, we can say $\mathcal{L}$-formulas are defined as follows

(1) An atomic $\mathcal{L}$-formula is a formula.

(2) If ϕ and ψ are $\mathcal{L}$-formulas and x is a variable, then $(\neg\phi)$, $\phi \wedge \psi$ and $(\forall x)\phi$ are $\mathcal{L}$-formulas.

5.3.8 Remark. We can define the other $\mathcal{L}$-formulas by making some relations with the given $\mathcal{L}$-formulas in Definition (5.3.7). For instance, if ϕ and ψ are formulas then so is $\phi \vee \psi = \neg((\neg\phi) \wedge (\neg\psi))$, and thus also $\phi \longrightarrow \psi = \neg(\phi) \vee \psi$.

5.3.9 Definition. (Halper, 2010) We define free variables of an $\mathcal{L}$-formula to be variables that have no quantifier. An $\mathcal{L}$-formula that contains no free variables is called an $\mathcal{L}$-sentence.

Let k be an $\mathcal{L}$-term, φ be an $\mathcal{L}$-formula and $\mathcal{M} = \langle M, I\rangle$ be a model. Then we can define a function $k^{\mathcal{M}} : M^n \longrightarrow M$ where n is the number of variables in k. It should be noted that k is defined by first considering its subterms see (Boxall, 2018) for further reading on definable sets.

The next definition shows how an atomic $\mathcal{L}$-formula is used to define subsets of M^n using functions defined by $\mathcal{L}$-terms and we show only one case as we have already referred the reader on definable sets.

5.3.10 Definition. (Boxall, 2018) Let $\mathcal{M} = \langle M, I\rangle$ be a model, k_1, k_2 be some $\mathcal{L}$-terms and $j_1 < j_2 < \cdots, j_n$ for all $j_1, j_2, \cdots, j_n \in \mathbb{N}$. Let φ to be an $\mathcal{L}$-formula that takes the form $(k_1 = k_2)$ and $x_{j_1}, x_{j_2}, \cdots, x_{j_n}$ to be a complete list consisting of all variables that appear in φ. Then, $\varphi(\mathcal{M})$ is defined to be the set of all $\bar{x} = (x_{j_1}, x_{j_2}, \cdots, x_{j_n}) \in M^n$ such that $k_1^M(\bar{x}_{k_1}) = k_2^M(\bar{x}_{k_2})$. Here, $\bar{x}_{k_i} = (x_{j_1}, x_{j_2}, \cdots, x_{j_n})_{k_i}$ for $i = 1, 2$ refers to a complete list corresponding to an $\mathcal{L}$-term k_i. So, we say φ defines the set $\varphi(\mathcal{M})$ in M and so it is a definable set.

5.3.11 Remark. If we let σ to be an $\mathcal{L}$-sentence and $\mathcal{M} = \langle M, I\rangle$ to be a model. Then, it follows that

$$\sigma(\mathcal{M}) \subseteq M^0.$$

Therefore, $\sigma(\mathcal{M}) = \emptyset$ or $\sigma(\mathcal{M}) = M^0 = \{()\}$. We say σ is true in M if and only if $\sigma(\mathcal{M}) \neq \emptyset$. We write $\mathcal{M} \vDash \sigma$ and say $\mathcal{M}$ is a model of σ. That is σ is true in $\mathcal{M}$.

5.3.12 Definition. (Halper, 2010) Let $\mathcal{L}$ be a language and $\mathcal{M}$ be a model. Then, a set of $\mathcal{L}$-sentences of $\mathcal{L}$ is called a *theory* T of $\mathcal{L}$. Furthermore, the set of all $\mathcal{L}$-sentences that are true in $\mathcal{M}$ is called *theory* of $\mathcal{M}$, denoted $Th(\mathcal{M})$.

5.3.13 Definition. (Halper, 2010) Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. Then, $\mathcal{M} \equiv \mathcal{N}$ if and only if $Th(\mathcal{M}) = Th(\mathcal{N})$ That is, if a sentence is true in $\mathcal{M}$, then the same is true in $\mathcal{N}$, and from that we conclude that $\mathcal{M}$ and $\mathcal{N}$ are *elementarily equivalent*.

We state without proof $Łos$ lemma and focus on its consequence which shall be important later.

5.3.14 $Łos$ **lemma.** (Boxall, 2018) Let $\mathcal{L}$ be a language, φ be an $\mathcal{L}$-formula and $(\mathcal{M}_j)_{j\in J} = (M_j, I_j)$ be non-empty family of $\mathcal{L}$-structures so that we can obtain the family $(\varphi(\mathcal{M}_j))_{j\in J}$. If we further let

$$\mathcal{M} = (\prod_{j\in J} M_j/\mathcal{U}, I_{\mathcal{U}})$$

with $\mathcal{U}$ being an ultrafilter on J, we obtain $\varphi(\mathcal{M}) \subseteq (\prod_{j\in J} M_j/\mathcal{U})^n$. Then

$$\varphi(\mathcal{M}) = [(\varphi(\mathcal{M}_j))_{j\in J}]$$

5.3.15 Corollary. (Boxall, 2018) Let $\mathcal{L}, (\mathcal{M}_j)_{j\in J}, \mathcal{U}$ and $\mathcal{M}$ be defined as in Lemma (5.3.14). Let σ be an $\mathcal{L}$-sentence. Let $\mathcal{M}$ be the ultraproduct of $(M_j)_{j\in J}$ with respect to $\mathcal{U}$. Then, $\mathcal{M} \vDash \sigma$ if and only if $\{j \in J : \mathcal{M}_j \vDash \sigma\} \in \mathcal{U}$.

Proof. To prove this lemma we give expanded version of the proof found in (Boxall, 2018).

Let $\mathcal{M} = (\prod_{j\in J} M_j/\mathcal{U}, I_{\mathcal{U}})$ be a model and σ to be an $\mathcal{L}$-sentence. Suppose $\mathcal{M} \vDash \sigma$. That is σ is true in $\mathcal{M}$. Then, by Remark (5.3.11), it follows that $\sigma(\mathcal{M}) \neq \emptyset$ which implies that $\sigma(\mathcal{M}) = \{()\}$. Using $Łos's$ lemma, it follows that $\sigma(\mathcal{M}) = [(\sigma(\mathcal{M}_j))_{j\in J}]$. So, this means $[(\sigma(\mathcal{M}_j))_{j\in J}] = \{j \in J : () \in \sigma(\mathcal{M}_j)\} \in \mathcal{U}$. It turns out that $[(\sigma(\mathcal{M}_j))_{j\in J}] = \{j \in J : \sigma(\mathcal{M}_j) \neq \emptyset\} \in \mathcal{U}$. Since $\sigma(\mathcal{M}_j)$ is the set defined by $\mathcal{L}$-sentence σ in M_j for $j \in J$, we have that by Remark (5.3.11) $\{j \in J : \mathcal{M}_j \vDash \sigma\} \in \mathcal{U}$.

Conversely, suppose $\{j \in J : \mathcal{M}_j \vDash \sigma\} \in \mathcal{U}$. Using Remark (5.3.11) it follows that $\{j \in J : \sigma(\mathcal{M}_j) \neq \emptyset\} \in \mathcal{U}$. But, since σ is an $\mathcal{L}$-sentence the subset $\sigma(\mathcal{M}_j)$ defined in M_j by σ contains empty tuple $()$. So, it follows that $\{j \in J : () \in \sigma(\mathcal{M}_j)\} \in \mathcal{U}$. Hence, $() \in [(\sigma(\mathcal{M}_j))_{j\in J}]$. By $Łos's$ lemma we have that $() \in \sigma(\mathcal{M})$. So, $\sigma(\mathcal{M}) \neq \emptyset$. Therefore, $\mathcal{M} \vDash \sigma$. □

Let $\mathcal{M} = (\mathcal{M}_j, I_j)$, where $\mathcal{M}_j = \mathbb{R}$ and I_j to be the interpretation of $\mathcal{L}_{\mathbb{R}} = \{+, \times, -, <, 0, 1\}$ in $\mathbb{R}$, for all $j \in \mathbb{N}$. So, if $\mathcal{U}$ is an ultrafilter on $\mathbb{N}$ such that $\{C \subseteq \mathbb{N} : \mathbb{N} \setminus C \quad \text{is finite}\} \subseteq \mathcal{U}$ and $\mathbb{R}^U = (\prod_{j\in\mathbb{N}} \mathcal{M}_j)/\mathcal{U}$, then we can study the model $(\mathbb{R}^{\mathcal{U}}, I_{\mathcal{U}})$. So, using Corollary (5.3.15) to compare $(\mathbb{R}^U, I_U)$ and $(\mathbb{R}, I_{\mathbb{R}})$, we note the following:

Let σ to be an $\mathcal{L}_{\mathbb{R}}$-sentence. Then,

$$\begin{aligned}(\mathbb{R}^{\mathcal{U}}, I_{\mathcal{U}}) \vDash \sigma &\Leftrightarrow \{j \in \mathbb{N} : (\mathcal{M}_j, I_j) \vDash \sigma\} \in \mathcal{U}\\ &\Leftrightarrow (\mathcal{M}_j, I_j) \vDash \sigma \quad \forall j \in \mathbb{N}\\ &\Leftrightarrow (\mathbb{R}, I_{\mathbb{R}}) \vDash \sigma\end{aligned}$$

Hence, $\mathbb{R}$ and $\mathbb{R}^{\mathcal{U}}$ are elementarily equivalent by Definition (5.3.13).

5.3.16 Definition. (Davis, 2009) For any given elements $[(a_j)_{j\in\mathbb{N}}], [(b_j)_{j\in\mathbb{N}}] \in \mathbb{R}^{\mathcal{U}}$, the relation $\geq_{\mathcal{U}}$ is defined by

$$[(a_j)_{j\in\mathbb{N}}] \geq_{\mathcal{U}} [(b_j)_{j\in\mathbb{N}}] \Leftrightarrow \{j \in \mathbb{N} : a_j \geq_{\mathcal{U}} b_j\} \in \mathcal{U}.$$

5.3.17 Definition. (Davis, 2009) Let $x \in \mathbb{R}^{\mathcal{U}}$, we say x is infinite if $x \geq_{\mathcal{U}} n$ for all $n \in \mathbb{N}^{\sigma}$, where $\mathbb{N}^{\sigma}$ is a set of constant sequences with natural number values.

It follows that we can be able to generate infinite numbers in the set $\mathbb{R}^{\mathcal{U}}$, and thus we claim that $[(j)_{j\in\mathbb{N}}]$ is infinite.

Proof. Since for any fixed $n \in \mathbb{N}$, the set $A = \{j \in \mathbb{N} : j > n\} \in \mathcal{U}$, this implies that $[(j)_{j\in J}] \geq_{\mathcal{U}} [(n)_{j\in\mathcal{N}}]$. Since $n \in \mathbb{N}$ was chosen arbitrarily, this holds for all $n \in \mathbb{N}$. Hence, making it infinite element of $\mathbb{R}^{\mathcal{U}}$. □

We have that the α in the Alpha theory is precisely the number $[(j)_{j\in\mathbb{N}}]$ as it was the aim of this Chapter to construct such a number.

5.3.18 Remark. It follows that in principle, the axioms governing α can be now proven as consequence (or theorems) in the ultraproduct and we refer that as the future work. However, as to give an example with one of the propositions being (2) of Proposition (2.2.3), it can be stated as follows:

Let $\varphi(n)$ and $\psi(n)$ for all $n \in \mathbb{N}$ be sequences such that $\varphi(n) \not\approx \psi(n)$. Then,

$$\varphi[\alpha] = [(\varphi(n))_{n\in\mathbb{N}}] \neq [(\psi(n))_{n\in\mathbb{N}}] = \psi[\alpha].$$

The next theorem is the model theoretic analogue of the saturation principle $\mathbb{R}^{\mathcal{U}}$. We state it without proof and focus on its consequence.

5.3.19 Theorem. (Boxall, 2018) Let $A^{\mathcal{U}}$ be as above. Assume $J = \mathbb{N}$ and $\{X \subset \mathbb{N} : \mathbb{N}\setminus X \quad \text{is finite}\} \subseteq \mathcal{U}$. Let $n \in \mathbb{N}$. For each $m \in \mathbb{N}$, let $B_m \subseteq A^n$. Suppose

(1) $[B_m] \neq \emptyset$ for all $m \in \mathbb{N}$;

(2) $[B_k] \subseteq [B_m]$ for all $m, k \in \mathbb{N}$ such that $m \leq k$.

Then

$$\bigcap_{m\in\mathbb{N}} [B_m] \neq \emptyset.$$

This theorem can be used to formally generate the infinitesimals numbers. For example:

Let $n \in \mathbb{N}$, $B_n = (0, \frac{1}{n})$ and $A = \mathbb{R}$. Then $A^{\mathcal{U}} = \mathbb{R}^{\mathcal{U}}$, and we claim that $[B_n] \neq \emptyset$ and $[B_m] \subseteq [B_n]$ for all $m, n \in \mathbb{N}$ such that $n \leq m$.

Proof. To that end, suppose for the sake of contradiction that $[B_n] = \emptyset$. By Proposition (5.2.6), $\{j \in \mathbb{N} : B_n = \emptyset\} \in \mathcal{U}$. But for all $n \in \mathbb{N}$, $B_n \neq \emptyset$. Hence, a contradiction. Therefore, $[B_n] \neq \emptyset$.

For the case of subset, it is a well known fact that for all n, m with $n \leq m$, then $0 < \frac{1}{m} \leq \frac{1}{n}$. It turns out that $(0, \frac{1}{m}) \subseteq (0, \frac{1}{n})$. Therefore, we conclude that $[B_m] \subseteq [B_n]$, and refer the reader to check the converse part of the proof to Proposition (5.2.4) as our conclusion is based on that similar reasoning. □

We have shown that for $B_n = (0, \frac{1}{n})$ with $n \in \mathbb{N}$, then $[B_n] \neq \emptyset$ and $[B_m] \subseteq [B_n]$ for all $m, n \in \mathbb{N}$ such that $n \leq m$. So, by Theorem (5.3.19) it follows that

$$\bigcap_{m \in \mathbb{N}} [B_m] \neq \emptyset.$$

Hence, by definition of infinitesimals it follows that $\mathbb{R}^{\mathcal{U}}$ contains infinitesimal elements. So $\mathbb{R}^{\mathcal{U}}$ is a model of the hyperreals as it contains the infinitesimals and infinite numbers.

6. Applications

In this Chapter we consider some applications of nonstanrdard analysis to calculus.

6.1 Continuity

Classical calculus is based on the concept of a limit, which is formalized by ϵ-δ formalism. This was developed by Weierstrass, before then the use of infinitesimals to define concepts in calculus was met with scepticism, since there were no rigorous justifications. It was Robinson who first developed Nonstandard analysis rigorously. Many proponents of nonstandard analysis claim that it is pedagogically better than the classical approach to teaching calculus. In this Chapter we look at a few nonstandard formulations of classical calculus.

6.1.1 Definition. (Jerome, 2000) L is the **limit** of $f(x)$ as x approaches x_0 if whenever x is infinitely close but not equal to x_0, $f(x)$ is infinitely close to L. That is, if whenever $x \sim x_0$ but $x \neq x_0$, then $f(x) \sim L$, where $x_0, L \in \mathbb{R}$.

6.1.2 Remark. We say the limit of $f(x)$ does not exist as x approaches x_0 if there is no L satisfying the above criterion. Since $x \neq x_0$, $f(x_0)$ can be undefined whilst the limit exists.

6.1.3 Definition. (Benci and Di Nasso, 2003) Suppose that A contains a neighborhood of x_0, and $f : A \longrightarrow \mathbb{R}$. Then f is **continuous** at x_0 if for every $\xi \in A^*, \xi \sim x_0 \implies f^*(\xi) \sim f(x_0)$.

The next proposition relates the formal definition of a limit in the classical calculus to that of the nonstandard approach. We refer the reader to (Benci and Di Nasso, 2003) for the proof.

6.1.4 Proposition. (Benci and Di Nasso, 2003) Let $f : A \longrightarrow \mathbb{R}$ be a function, $x_0, L \in \mathbb{R}$, and suppose that A is a neighborhood of x_0. Then the following are equivalent:

(1) For every real number $\epsilon > 0$, there exists a real $\delta > 0$, such that for all reals $x \neq x_0$, $x_0 - \delta < x < x_0 + \delta \implies L - \epsilon < f(x) < L + \epsilon$;

(2) If $\epsilon \sim 0$ with $\epsilon \neq 0$, then $f(x_0 + \epsilon) \sim L$.

We now give the nonstandard definition of derivative of a function at a point x_0.

6.1.5 Definition. (Benci and Di Nasso, 2003) We say that the function f defined on a neighborhood of x_0 has **derivative** at x_0 if there exists a finite real number $f'(x_0)$ such that for all non-zero infinitesimal ϵ,

$$\frac{f^*(x_0 + \epsilon) - f(x_0)}{\epsilon} \sim f'(x_0)$$

6.1.6 Theorem. (Jerome, 2000) If f is differentiable at x_0 then f is continuous at x_0.

Proof. Suppose that the function $f(x)$ is differentiable at the point x_0. By definition, there exists an infinitesimal ϵ and a real number $f'(x_0)$ such that

$$\frac{f^*(x_0 + \epsilon) - f(x_0)}{\epsilon} \sim f'(x_0)$$

So, it follows that $\frac{f^*(x_0+\epsilon)-f(x_0)}{\epsilon} - f'(x_0)$ is infinitesimal, and thus using definition of infinitesimal we have that for all $n \in \mathbb{N}$,

$$\left|\frac{f^*(x_0+\epsilon)-f(x_0)}{\epsilon} - f'(x_0)\right| < \frac{1}{n}$$

which implies that $|(f^*(x_0+\epsilon)-f(x_0)) - \epsilon \cdot f'(x_0)| < \frac{|\epsilon|}{n}$. Thus, $(f^*(x_0+\epsilon)-f(x_0)) - \epsilon \cdot f'(x_0)$ is an infinitesimal.

But, since ϵ is infinitesimal and $f'(x_0)$ is finite real number, then $\epsilon \cdot f'(x_0)$ is infinitesimal. This follows from (1) of Proposition (4.1.5) that says if ξ is an infinitesimal number and ζ is a finite number, then $\xi \cdot \zeta$ is infinitesimal. So, it follows that $f^*(x_0+\epsilon) - f(x_0)$ is an infinitesimal also and we write $f^*(x_0+\epsilon) \sim f(x_0)$, which by Definition (6.1.3) gives us the conclusion. □

6.1.7 Definition. (Jerome, 2000) We say that f is continuous on an open interval **I** if f is continuous at every point x_0 in **I**. If in addition f has a derivative at every point of **I**, we say that f is differentiable on **I**.

6.1.8 Definition. (Jerome, 2000) f is said to be continuous on the closed interval [a, b] if f is continuous at each point x_0 where $a < x_0 < b$, upper continuous at a and lower continuous at b.

6.1.9 Rolle's Theorem. Suppose f is a continuous function on a closed interval $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists a number $x_0 \in (a, b)$ such that $f'(x_0) = 0$.

Proof. If f is constant on $[a, b]$, then $f'(x_0) = 0$, for all $x_0 \in [a, b]$ since for all $x_0 + \epsilon \in (a, b)^*$ $f^*(x_0+\epsilon) = f(x_0)$ and thus

$$f'(x_0) \sim \frac{f^*(x_0+\epsilon)-f(x_0)}{\epsilon} = 0.$$

Suppose there exists $x \in (a, b)$ such that $f(x) > f(a)$. Then, there exists $x_0 \in (a, b)$ with $x_0 \neq a$ and $x_0 \neq b$ such that $f(x)$ attains maximum value at x_0 since $f(a) = f(b)$. Since $(a, x_0) \subset (a, b)$ it follows that f is differentiable on (a, x_0). Let $\epsilon \sim 0, \epsilon \neq 0$ and $x_0 - \epsilon \in (a, x_0)^*$. Then,

$$f'(x_0) \sim \frac{f^*(x_0-\epsilon)-f(x_0)}{-\epsilon} \geq 0.$$

We have that f is differentiable on $(x_0, b) \subset (a, b)$ also, and so for fixed nonzero infinitesimal ϵ such that $x_0 + \epsilon \in (x_0, b)^*$ it follows that

$$f'(x_0) \sim \frac{f^*(x_0+\epsilon)-f(x_0)}{\epsilon} \leq 0.$$

Since $f'(x_0)$ is a finite real number, and we find that it is infinitely close to both negative and positive numbers it must be zero. Hence, we have that $f'(x_0) = 0$. Similarly, the same conclusion can be made with $f(x) < f(a)$. □

We would like to remind the reader of **Cauchy's mean value theorem** which is a generalization of the **mean value theorem**. It states that, if f and g are continuous functions on $[a, b]$, differentiable on (a, b) then, there exists some $x_0 \in (a, b)$ such that

$$(f(b) - f(a))g'(x_0) = (g(b) - g(a))f'(x_0).$$

In the case where $g'(x_0) \neq 0$ and $g(b) \neq g(a)$, we have

$$\frac{f'(x_0)}{g'(x_0)} = \frac{f(b) - f(a)}{g(b) - g(a)}$$

In particular, if $g(x) = x$, then $g'(x) = 1$ for all $x \in (a, b)$ and we get the classical mean value theorem.

In the light of Cauchy's mean value theorem, we provide a nonstandard proof of the following "additive" variant.

6.1.10 Theorem. Let f, g be continuous functions on $[a, b]$, differentiable on (a, b) such that $g(a)f(b) = g(b)f(a)$ and $g(a) + g(b) \neq 0$. Then, there exists $x_0 \in (a, b)$ such that

$$(f(b) + f(a))g'(x_0) = (g(b) + g(a))f'(x_0).$$

Proof. Let

$$G(x) = f(x) - \frac{f(b) + f(a)}{g(b) + g(a)} \cdot g(x).$$

We note that

$$\begin{aligned} G(a) &= f(a) - \frac{f(b) + f(a)}{g(b) + g(a)} g(a) \\ &= \frac{f(a) \cdot g(b) + f(a) \cdot g(a) - f(b) \cdot g(a) - f(a) \cdot g(a)}{g(b) + g(a)} \\ &= 0, \quad \text{since} \quad g(a)f(b) = g(b)f(a). \end{aligned}$$

Similarly, $G(b) = 0$ and thus $G(a) = G(b)$.

Since $G(x)$ is linear in $f(x)$ and $g(x)$, which are continuous functions on $[a, b]$ and differentiable on (a, b), $G(x)$ is continuous on $[a, b]$ and differentiable on (a, b). By the definition of differentiability on (a, b), for every $\mu \in [a, b]$ and $\mu + \epsilon \in [a, b]^*$, there exists an infinitesimal $\epsilon \neq 0$ such that

$$\begin{aligned} G'(\mu) &\sim \frac{G^*(\mu + \epsilon) - G(\mu)}{\epsilon} \\ &= \frac{f^*(\mu + \epsilon) - \frac{f(b)+f(a)}{g(b)+g(a)} g^*(\mu + \epsilon) - \left[f(\mu) - \frac{f(b)+f(a)}{g(b)+g(a)} \cdot g(\mu) \right]}{\epsilon} \\ &= \frac{f^*(\mu + \epsilon) - f(\mu) - \frac{f(b)+f(a)}{g(b)+g(a)} \cdot (g^*(\mu + \epsilon) - g(\mu))}{\epsilon} \end{aligned}$$

Thus,

$$\begin{aligned} G'(\mu) &\sim \frac{f^*(\mu + \epsilon) - f(\mu)}{\epsilon} - \frac{f(b) + f(a)}{g(b) + g(a)} \cdot \frac{g^*(\mu + \epsilon) - g(\mu)}{\epsilon} \\ &= f'(\mu) - \frac{f(b) + f(a)}{g(b) + g(a)} \cdot g'(\mu). \end{aligned}$$

Hence,

$$G'(\mu) \sim f'(\mu) - \frac{f(b) + f(a)}{g(b) + g(a)} \cdot g'(\mu) \tag{6.1.1}$$

Since $G(a) = G(b)$ and G is continuous on $[a, b]$ and differentiable on (a, b), by **Rolle's theorem** there exists $x_0 \in [a, b]$ such that $G'(x_0) = 0$. For such x_0 by equation (6.1.1) above, we have that

$$f'(x_0) - \frac{f(b) + f(a)}{g(b) + g(a)} \cdot g'(x_0) = 0.$$

Therefore, $(g(b) + g(a))f'(x_0) = (f(b) + f(a))g'(x_0)$ as required. □

Out of curiosity, given the geometrical interpretation of the usual mean value theorem, we would like to see what geometrical property this variant implies.

6.1.11 Remark. In the case where $g'(x_0) \neq 0$ and $g(b) + g(a) \neq 0$, we get $\frac{f'(x_0)}{g'(x_0)} = \frac{f(b)+f(a)}{g(b)+g(a)} = \frac{f(b)-(-f(a))}{g(b)-(-g(a))}$. Then, geometrically this means there is some tangent to the graph of the curve $\begin{cases} [a, b] \longrightarrow \mathbb{R}^2 \\ x \longmapsto (f(x), g(x)) \end{cases}$ which is parallel to the line defined by the points (-g(a),-f(a)) and (g(b),f(b)).

6.1.12 Example. Let $f(\theta) = \theta sin(2\theta)$ and $g(\theta) = \theta cos(2\theta)$ and $[a, b] = [\frac{\pi}{6}, \frac{7\pi}{6}]$. Then, it follows that $f(\theta)$ and $g(\theta)$ are continuous on $[a, b]$ and differentiable on (a, b) since they are products of polynomial and trigonometric functions which are continuous and differentiable everywhere.

We note that the end point b can be rewritten as $b = \frac{7\pi}{6} = \frac{\pi}{6} + \pi$. Using trigonometric identities of sine and cosine functions on any point $\theta + \pi$, it follows that

$$cos(2(\theta + \pi)) = cos(2\theta + 2\pi) = cos(2\theta)cos(2\pi) - sin(2\theta)sin(2\pi) = cos(2\theta)$$

and

$$sin(2(\theta + \pi)) = sin(2\theta + 2\pi) = cos(2\theta)sin(2\pi) + sin(2\theta)cos(2\pi) = sin(2\theta)$$

Taking $\theta = \frac{\pi}{6}$, we claim that $f(\theta)g(\theta + \pi) = f(\theta + \pi)g(\theta)$. To prove our claim, we proceed as follows

$$f(\theta)g(\theta + \pi) = \frac{\pi}{6} sin(\frac{2\pi}{6})) \cdot \frac{7\pi}{6} cos(\frac{2\pi}{6}) = \frac{7\pi^2}{36} sin(\frac{\pi}{3}))cos(\frac{\pi}{3}) \qquad (6.1.2)$$

and

$$f(\theta + \pi)g(\theta) = \frac{7\pi}{6} sin(\frac{2\pi}{6}) \cdot \frac{\pi}{6} cos(\frac{2\pi}{6})) = \frac{7\pi^2}{36} sin(\frac{\pi}{3}))cos(\frac{\pi}{3}) \qquad (6.1.3)$$

By comparing equations (6.1.2) and (6.1.3) we find that they are equal and hence $f(\theta)g(\theta + \pi) = f(\theta + \pi)g(\theta)$. □

With $\theta = \frac{\pi}{6}$ it follows that

$$g(\frac{\pi}{6}) + g(\frac{7\pi}{6}) = \frac{\pi}{6} cos(\frac{2\pi}{6})) + \frac{7\pi}{6} cos(\frac{2\pi}{6}) = (\frac{\pi}{6} + \frac{7\pi}{6})cos(\frac{\pi}{3}) = \frac{2\pi}{3} \neq 0 \qquad (6.1.4)$$

Therefore, by Theorem (6.1.10) there exists $x_0 \in [\frac{\pi}{6}, \frac{7\pi}{6}]$ such that $(f(\frac{7\pi}{6}) + f(\frac{\pi}{6}))g'(x_0) = (g(\frac{7\pi}{6}) + g(\frac{\pi}{6}))f'(x_0)$.

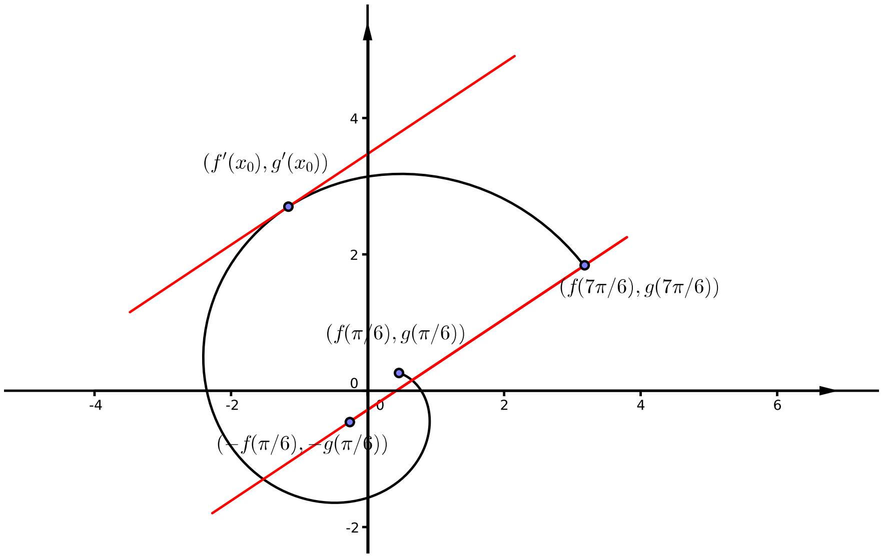

Figure 6.1: The graph of the given curve in example showing that tangent line to the curve is parallel to the line connecting the endpoints of the interval.

7. Conclusion

We began this essay by defining a new mathematical object namely α. We then postulated five axioms which governed the use of α, and stated in a form of propositions their consequences which acted as the building blocks for the alpha theory. We proceeded by adopting the notion of star operator to construct hyper extensions; stated the transfer principle which helped us in transferring the first order logic statements between entities and their nonstandard extensions.

We then constructed the set of hyperreal numbers $\mathbb{R}^*$ which extends the set of reals. We studied a few of its properties together with its subsets such as the set of hypernatural numbers. We went on to introduce nonstandard calculus where we defined notion of infinitesimals and infinite numbers which act as the basic tools for redefining concepts in calculus, and we gave a short comparison between notions of ideal values and that of limit concept.

We proceeded by defining one of the topics of nonstandard analysis, the internal set to obtain the saturation principle which can be used in proving the existence of the infinitesimal numbers. We then gave the formal foundation of the ideal value α using the model theory approach so as to formalise the alpha theory. Finally, we gave an application of nonstandard analysis by selecting and redefining some concepts from calculus.

Acknowledgements

I wish to express my gratitude to Gareth Boxall and Taboka Prince Chalebgwa who gave form to this work with all their love and dedication.

I would also like to thank Jeff Sanders, Kenneth Dadedzi and Lebeko Poulo who gave their time, and their support on all my challenges, and helped me to gain strength and focus on my studies until the realization of this essay. To Barry, Jan, Igsaan, and the entire AIMS 2018-19 family, thank you!

Finally, I would also like to express my gratitude to our Creator for the inspiration and beauty that gave this essay life. To my parents, my siblings, the rest of my family and my friends thank you very much for your support.

References

Benci, V. and Di Nasso, M. Alpha-theory: An elementary axiomatics for nonstandard analysis. *Expositiones Mathematicae*, 21:355–386, 12 2003. doi: 10.1016/S0723-0869(03)80038-5.

Benci, V. and Di Nasso, M. *How to Measure the Infinite: Mathematics with Infinite and Infinitesimal Numbers*. World Scientific Publishing Company Pte. Limited, 2019.

Benci, V., Galatolo, S., and Ghimenti, M. G. An elementary approach to stochastic differential equations using the infinitesimals. American Mathematical Society (AMS), 2010.

Boxall, G. Model theory. Available from https://sites.google.com/a/aims.ac.za/model-theory-2018/, 2018.

CODY, B. *ULTRAPRODUCTS, THE COMPACTNESS THEOREM AND APPLICATIONS*. Virginia Commonwealth University, 2015.

Davis, I. An introduction to nonstandard analysis. *Internet source, publication date August*, 14:25, 2009.

Halper, A. Ultraproducts and model theory. Available from http://www.math.uchicago.edu/~may/VIGRE/VIGRE2010/REUPapers/Halper.pdf, 2010.

Henson, C. W. Foundations of nonstandard analysis. In *Nonstandard analysis*, pages 1–49. Springer, 1997.

Holmes, R. Elementary set theory with a universal set. *Elementary Set Theory with a Universal Set*, 1998.

Jerome, K. H. Elementary calculus an infinitesimal approach second edition, 2000.

Keisler, H. J. *Elementary calculus: An infinitesimal approach*. Courier Corporation, 2013.

Luxemburg, W. A. J. Non-standard analysis: lectures on a. robinson's theory of infinitesimals and infinitely large numbers. 1966.

Nelson, E. Internal set theory: a new approach to nonstandard analysis. *Bulletin of the American Mathematical Society*, 83(6):1165–1198, 1977.

Robinson, A. *Non-standard analysis*. Princeton University Press, 1974.