On the Foundations of Nonstandard Analysis and Applications

Julio Pérez García

9th of June 2020

Abstract

This paper is a self-contained introduction to nonstandard analysis from set-theoretical foundations. The aim is to build from basic mathematical knowledge the machinery required to understand the powerful tool nonstandard analysis is and to make use of it in examples of modern research.

Contents

1 Introduction

It could be argued that the origins of nonstandard analysis date back to Leibniz and his conception of calculus through infinitesimals. Contrary to the usual epsilon-delta approach, Leibniz introduced infinitely small and infinitely large quantities he called infinitesimals; similar to how in the natural sciences and engineering $\frac{d}{dx}$ is often intuitively seen as a small change in x. This approach, forgotten for centuries in the development of analysis, was reclaimed by Abraham Robinson in the 1960's. Robinson developed a rigorous nonstandard analysis in a model-theoretic formulation, studying saturated extensions of models. Since then, nonstandard analysis has proven to be a versatile tool that has been used to tackle questions in a variety of areas: stochastic processes, calculus, Ramsey theory...

There are many ways to first approach nonstandard analysis, many of them requiring sound knowledge of model theory: nonstandard extensions are presented as ultrapowers of given models, which satisfy Łoś's theorem. In this paper I have decided to follow a different route similar to Henson's (Henson C.W. (1997) Foundations of Nonstandard Analysis), presenting a general set-theoretic definition for nonstandard extensions. This formulation might seem slightly arbitrary at first, as at its core is a enumeration of the sufficient conditions for the extension to satisfy the transfer theorem, but it presents a number of advantages. First, it gives a more flexible definition of nonstandard extension, of which ultrapowers will be a concrete example. Secondly, this definition gives a clearer sense of the motivation of nonstandard extensions themselves and it will useful in giving a solid foundation for nonstandard extensions of structures more complicated than sets, such as collections of sets.

The first sections of the paper will consist of an overview of the elemental knowledge required to employ nonstandard methods. I will begin by presenting the aforementioned definition of nonstandard extensions, the immediate goal after this will be to explicitly construct one. In order to achieve this, basic properties of filters and ultrafilters will be studied and then used to build the ultrapower nonstandard extension. After this, the transfer theorem will be introduced, again, following Henson's approach. To end this first part, the nonstandard real numbers ${}^*\mathbb{R}$ will be constructed as an example of the concepts developed so far. These would constitute the basic toolkit for the rest of the paper.

The following sections will mostly mimic the structure of the first part by first revisiting ultrafilters and then, nonstandard extensions on a deeper level with certain applications in mind. Most of the work done in this part is a based on the research done by Di Nasso et. al (Di Nasso et al. (2017) Nonstandard Methods in Ramsey Theory and Combinatorial Number Theory). The main contributions of this paper is to reformulate some of the less clear concepts in a more foundational approach, filling the gaps and correcting typos and errors in the aforementioned piece of research. I will first tackle König's lemma, a result in graph theory, by employing results about ${}^*\mathbb{R}$ from the first part. In the ultrafilters sections, my main goal will be to prove Hindman's theorem via the properties of the semitopological semigroup of ultrafilters on a set. In contrast to Di Nasso et. al, I will avoid u-idempotency and follow Hindman's own proof. I will briefly discuss nonstandard extensions of I-sets in a general context as done by Henson and then turn to the $(X, \mathcal{P}(X))$ setting, that is, nonstandard extensions of families of sets. The notions of internal and external objects will be introduced, as well as the internal definition principle, that will allow us to translate properties of subsets of X to subsets of ${}^*X$. Finally I will exploit this principle with hyperfinite sets, which will preserve in ${}^*X$ predicates about finite subsets of X, to prove some results about orderings of sets and graph-coloring.

2 Nonstandard Extensions

Nonstandard extensions are the key concept of nonstandard analysis. They are a expanded version of a given universe and in the right conditions, they will contain special elements that our original set did not have. Think of the real numbers $\mathbb{R}$, when asked to give a real number greater than zero but smaller than any other positive real number we are faced with an impossible task. Now, in its nonstandard extension we will be able to find such infinitely small element. These elements will allow us to make use of "infinite" quantities in our proofs.

Definition 1. (Nonstandard Extension of a Set) Let X be a non-empty set, a **nonstandard extension of** X is a tuple $({}^*X, {}^*)$ such that * is a collection of maps $\mathcal{P}(X^m) \to \mathcal{P}({}^*X^m)$ for all $m \in \mathbb{N}$ that sends every subset A of X^m to a non-empty subset ${}^*A \subseteq {}^*X^m$ and satisfies the following conditions $\forall m, n \in \mathbb{N}$:

- (E1) * preserves Boolean operations on subsets of X^m: if $A, B \subseteq X^m$ then

$$ {}^*A \subseteq {}^*X^m, {}^*(A \cup B) = {}^*A \cup {}^*B, {}^*(A \cap B) = {}^*A \cap {}^*B, {}^*(A \backslash B) = {}^*A \backslash {}^*B $$

- (E2) * preserves basic diagonals: let $m \geq 2$

$$ \Delta := \{(x_1, x_2, ..., x_m) : \forall i \in \{1, ..., m\} x_i \in X \text{ and } x_j = x_k \text{ for some } j \neq k \in \{1, ..., m\} $$

then

$$ {}^*\Delta := \{(x_1, x_2, ..., x_m) : \forall i \in \{1, ..., m\} x_i \in Y x_j = x_k \text{ for the same } j \neq k \in \{1, ..., m\}\}. $$

- (E3) * preserves Cartesian products: if $A \subseteq X^m$ and $B \subseteq X^n$, then

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

- (E4) * preserves projections onto the first m coordinates: for the projection on the m first coordinates

$$ \pi_m : X^{m+1} \to X^m $$

$$ (x_1, ..., x_m, x_{m+1}) \mapsto (x_1, ..., x_m). $$

If $A \subseteq X^{m+1}$ with $\pi(A) := \{\pi_m(x) : x \in A\}$, then

$$ {}^*(\pi(A)) = \pi({}^*A) = \{\pi_m(x) : x \in {}^*A\}. $$

[1]

Example 1. *It is easy to see that for any non-empty set X, X itself will be a non-standard extension, the trivial one. This is not a very interesting example, but it does suffice to show that nonstandard extension, do, indeed, exist. We will come back to this at the end of the section.*

Proposition 1. ${}^*\emptyset = \emptyset$ *and for all* $m \in \mathbb{N}$ *we have* ${}^*(X^m) = ({}^*X)^m$.

Proof. Let X be a non-empty set ${}^*\emptyset = {}^*(X \backslash X)$, by (E1) ${}^*(X \backslash X) = {}^*X \backslash {}^*X = \emptyset$, so we get the result. We will prove the second part of the proposition by induction. Let $\varphi(m)$ be the formula ${}^*(X^m) = ({}^*X)^m$.

Basic case: ${}^*(X) = {}^*(X \backslash \emptyset) = {}^*X \backslash {}^*\emptyset = {}^*X \backslash \emptyset = {}^*X$ by (E1) and the first part of the proposition.

Now let $d > 1$ and, we assume $\varphi(d)$ is true. $\varphi(d+1)$: ${}^*(X^{d+1}) = ({}^*X)^{d+1}$, ${}^*(X^{d+1}) = {}^*(X^d) \times {}^*(X)$, $\varphi(d)$ and the basic case are true, so this equals $({}^*X)^d \times {}^*\mathbb{X}$ and applying (E3) we get the result. Therefore $\varphi(d+1)$ is true and so $\varphi(m)$ is true for all $m \in \mathbb{N}$ by induction. $\square$

From now on we will be able to refer to Cartesian products of * images and * images of Cartesian products interchangeably as ${}^*X^m$.

Proposition 2. *If $A \subseteq X^m$ is non-empty, then ${}^*A \subseteq {}^*(X^m)$ is also non-empty.*

Proof. We will prove this by induction, wlog. $X \neq \emptyset$. Let $\varphi(m)$: if $A \subseteq X^m$ is non-empty, then ${}^*A \subseteq {}^*(X^m)$ is also non-empty. We will use the projection map as defined in **definition 1** in the following steps.

Basic case: let $A \subseteq X$ non-empty, let us consider the product $X \times A$, ${}^*(\pi(X \times A) = {}^*X$, which is non-empty. On the other hand, by applying *(E4)* and (E3) we find ${}^*(\pi(X \times A) = \pi({}^*(X \times A)) = \pi({}^*X \times {}^*A)$, now if ${}^*A$ is empty then $\pi({}^*X \times {}^*A) = {}^*X$ is also empty, which is a contradiction, therefore, ${}^*A$ is non-empty.

Let $d > 1$, assume as our induction hypothesis that $\varphi(d)$ is true. Let $A \subseteq X^{d+1}$ be non-empty, then $\pi_d({}^*A) = {}^*\pi_d(A)$. We know that $\pi_d(A) \subseteq X^d$ and it is non-empty, therefore by our induction hypothesis ${}^*\pi_d(A)$ is non-empty, hence ${}^*A$ is non-empty. We conclude by induction that $\varphi(m)$ is true for all $m \in \mathbb{N}$. $\square$

Proposition 3. *For all $A, B \in X^m$, $A \subseteq B \Leftrightarrow {}^*A \subseteq {}^*B$.*

Proof. Suppose $A \subseteq B$, then $A = A \cap B$ so by (E1) ${}^*A = {}^*(A \cap B) = {}^*A \cap {}^*B$, therefore ${}^*A \subseteq {}^*B$.

Now let ${}^*A \subseteq {}^*B$, suppose that $A \cap B = C \neq B$, then ${}^*A \cap {}^*B = {}^*(A \cap B) = {}^*C \neq {}^*B$, a contradiction.

$\square$

Corollary 1. ${}^*A = {}^*B \Leftrightarrow A = B$.

Definition 2. Let $f : X^m \to X^n$ be a map, the graph of f is

$$\Gamma = \{\bar{x} \in X^{m+n} : x_1, ..., x_m \in X^m, x_{m+1}, ..., x_n = f(x_1, ..., x_m)\}.$$

Proposition 4. *Let $A \subseteq X^m$, B^n and $f : A \to B$ a map with graph Γ, then ${}^*\Gamma$ is the graph of a function from ${}^*A$ to ${}^*B$.*

Proof. We can characterise Γ by $\Gamma \subseteq A \times B$, $\pi_m(\Gamma) = A$ and $\Gamma^2 \cap \{(x, y, u, v) \in X^{2(m+n)} : x = u\} \subseteq \{(x, y, u, v) \in X^{2(m+n)} : y = v\}$.

It is easy to see that ${}^*\Gamma \subseteq {}^*(A \times B) = {}^*A \times {}^*B$ by (E1) and (E3).

$\pi_m({}^*\Gamma) \subseteq {}^*A$ by iteration of the projection map and the final condition is again satisfied by **proposition 3**, (E1) and (E3).

$\square$

Note 1. Let $f : X^m \to X^n$ with graph Γ, we denote the function whose graph is ${}^*\Gamma$, ${}^*f$.

Remark 1. *It can be shown that a function f is injective/surjective/bijective if and only if $^*f$ is injective/surjective/bijective.*

Remark 2. *There is an embedding of X in $^*X$ by the map*

$$\Psi : X \to {}^*X$$

$$x \mapsto {}^*\{x\}.$$

*This map is injective due to **corollary 1** and $^*\{x\}$ will be non-empty by **proposition 2**, so we can identify its unique element with $x \in X$, ie. denote $y \in {}^*\{x\}$ by x. Therefore we find an embedding of X into $^*X$ of the form $\{y \in {}^*X : y \in {}^*\{x\}$ for some $x \in X\}$ and abusing notation we will simply call it $X \subseteq {}^*X$. When notation might suppose an issue it will be stated whether it is the original set X or its embedding that is being talked about, although it will not generally suppose an issue.*

Definition 3. (Standard and nonstandard elements) Let $y \in {}^*X^m$, we say it is **standard** if there exists $x \in X^m$ such that $^*x = y$. Otherwise, we say it is **nonstandard**.

Going back to our first example, it is natural to ask oneself whether any non-trivial nonstandard extensions exist. If a nonstandard extension of a set X is non-trivial, as we know X is embedded in $^*X$, the only option is that it contains some extra elements that are not originally in X. This leads to the following definition.

Definition 4. (Proper nonstandard extension) Let X be a non-empty set and let $^*X$ be a nonstandard extension, we say $^*X$ is **proper** if for every infinite subset A of X, $^*A$ contains a nonstandard element.

Our goal now will be to find a way to construct proper nonstandard extensions for a set X, as they are the ones that have special characteristic our original set does not have.

3 Filters and Ultrafilters

In order to construct a proper nonstandard extension it is vital that we first study filters. We will be interested in a particular type of filter, ultrafilters. Heuristically, ultrafilters can be thought of as gates that only contain "big subsets". Filters and ultrafilters are also of interest for their own sake, and later on in the paper we will develop the theory of filters further and prove some interesting results.

Definition 5. (Filter) let X be a non-empty set, a **filter on** X is a map $\mu : \mathcal{P}(X) \to \{0, 1\}$ that satisfies the following conditions: Let $A, B \in \mathcal{P}(X)$

- (F1) If $\mu(A) = 1$ and $A \subseteq B$, then $\mu(B) = 1$.
- (F1) If $\mu(A) = \mu(B) = 1$, then $\mu(A \cap B) = 1$.
- (F2) $\mu(\emptyset) = 0$.

Note 2. In literature and throughout this paper we will use interchangeably the concept of a filter over a set X to be defined as above or the set $\mathcal{U} = \{A \subseteq X : \mu(A) = 1\}$. Either version will be used depending on the context.

Definition 6. (Principal filter) let $A \in \mathcal{P}(X)$ be a non-empty set and let $\mu_A : \mathcal{P}(X) \to \{0, 1\}$ be a map such that for any $B \in \mathcal{P}(X)$, $\mu_A(B) = 1 \Leftrightarrow A \subseteq B$, we call this function the **principal filter** generated by A.

Remark 3. *It can be easily shown that the function in* ***definition 6*** *is, indeed, a filter.*

So now we can find a filter for any non-empty set X, particularly we can see that filters do exist.

Example 2. *Let* $X = \mathbb{R}$*, consider the closed interval* $[0, 1]$ *and let* $\mu : \mathcal{P}(\mathbb{R}) \to \{0, 1\}$ *be such that* $\mu(A) = 1 \Leftrightarrow [0, 1] \subseteq A$*. This is the principal filter generated by* $[0, 1]$*, as a set it is defined by* $\mathcal{U} = \{A \subseteq \mathbb{R} : \mu(A) = 1\}$*.*

Proposition 5. *If* X *is a non-empty set containing a finite subset* $A \subseteq X$ *and* μ *is a filter on* X *such that* $\mu(A) = 1$*, then* μ *is principal.*

Proof. Let μ be a filter and A a finite subset with $\mu(A) = 1$, we will find a finite recursive algorithm that will either show the result in one of the intermediate steps or will result in a contradiction in the last step:

Evidently either *i)* $\mu = \mu_A$ or *ii)* $\mu \neq \mu_A$. In case *i)* we are done, as μ is the principal ideal of A. In case *ii)* this means there exists $Y \subseteq X$ such that *a)* $\mu(Y) = 1$ and *b)* $A \not\subseteq Y$, call $A \cap Y = A_1$ and by *a)* and (F2), $\mu(A_1) = 1$. Now we can consider cases *i)* and *ii)* for μ_{A_2} and repeat the process. We iterate finitely many times until μ is the principal filter for some A_i or $A_i = \emptyset$, a contradiction as $\mu(\emptyset) = 0$.

$\square$

Corollary 2. *If* X *is a non-empty finite set, every filter is principal.*

Proof. Let X be a non-empty finite set and let μ be a filter, we can apply **definition 2** setting $A = X$ and the result follows. $\square$

Definition 7. (Fréchet filter) Let X be an infinite set, we define $\mu_F : \mathcal{P}(X) \to \{0, 1\}$ such that for all $A \in \mathcal{P}(X)$ we have $\mu_F(A) = 1 \Leftrightarrow X \backslash A$ is finite. We call this function the **Fréchet filter**.

Remark 4. *It can be easily shown that the function in* ***definition 7*** *is, indeed, a filter.*

Proposition 6. *Let X be an infinite set and let μ_F be a Fréchet filter on X, then μ_F is not principal.*

Proof. Let X be an infinite set and let μ_F be a Fréchet filter. Suppose for a contradiction that $\mu_F = \mu_A$ for some $A \in \mathcal{P}(X)$. If A is finite, $X \backslash A$ is infinite, a contradiction, so A has to be infinite. We know that $\mu_A(A) = 1 = \mu_F(A) \Rightarrow X \backslash A$ is finite, as A is infinite $A \backslash \{a\}$ for some $a \in A$ is well defined and $X \backslash (A \backslash \{a\})$ is also finite. This is a contradiction, as $\mu_F(A \backslash \{a\}) = 1$ but $A \backslash \{a\} \not\supseteq A \Rightarrow \mu_A(A \backslash \{a\}) = 0$. □

So we have finally found a non-principal filter, namely, the Fréchet filter.

Let us go back now to the picture of filters in general. It is our goal to somehow use filters to construct an equivalence relation on sets so that then we can quotient our set by such relation. In order to do this we would require that for any subset, it is either contained in the filter or its complement is. It is known that for any filter by (F2) that if $\mu(A) = 1 \Rightarrow \mu(X \backslash A) = 0$. In other words, once we know that a set in in a filter we can deduce that its compliment is not in the filter. The converse is not always true, so knowing that a set is not in a filter is not enough to deduce that its complement is.

Example 3. *We give an example of such case. Let $\mu_{\{1,2\}}$ be the filter generated by $\{1, 2\}$ on $\mathbb{N}$. Consider the sets $\{1\}$ and $\mathbb{N} \backslash \{1\}$, it is clear that neither of them is in the filter, as they do not contain the set $\{1, 2\}$.*

The fact that not every filter satisfies this property leads to the following definition.

Definition 8. (Ultrafilter) Let X be a non-empty set and let μ be a filter, if $\mu(A) + \mu(X \backslash A) = 1$ for all $A \in \mathcal{P}(X)$ we call it an **ultrafilter**.

We know by (F1) that for any filter, if a set in an element of the filter so is its union with any other set. Another particularly important property of ultrafilters is that the converse of this property is true. So if the union of two sets is in the ultrafilters, at least one of those sets will be in the ultrafilter too.

Proposition 7. *Let X be a set, let μ be an ultrafilter on X and let $X = A \cup B$ for some $A, B \subseteq X$, then $\mu(A) = 1$ or $\mu(B) = 1$ (not mutually exclusive).*

Proof. Suppose that $\mu(A) = 0$, then $\mu(X \backslash A) = 1$, $X \backslash A \subseteq B$, therefore $\mu(B) = 1$. □

Proposition 8. *Let X be a non-empty set and let μ be a principal filter on X, then μ is an ultrafilter if and only if it is generated by a singleton.*

Proof. Let X be a non-empty set, A be a non-empty, non-singleton subset of X and μ_A the principal filter generated by A. We can find two distinct $a_1, a_2 \in A$, then $\mu_A(\{a_1\}) = \mu_A(\{a_2\}) = 0$, so it is not an ultrafilter.

On the other hand, if we consider $\mu_{\{x\}}$ for x, for all $A \subseteq X$ we have that if $\mu_{\{x\}}(X \backslash A) = 0 \Rightarrow \{a\} \not\subseteq X \backslash A \Rightarrow a \not\in X \backslash A \Rightarrow a \in A \Rightarrow \{a\} \subseteq A \Rightarrow \mu_{\{x\}}(A) = 1$. The other direction comes from the definition of filter. □

We now know how to find principal ultrafilters, but the question of whether there are any non-principals ultrafilter arises again. Ultrafilters seem to be "bigger" than general filters, as they always contain a set or its complement, therefore it is promising to examine whether given a filter we can extend it to an ultrafilter by adding more elements to it.

Definition 9. (Extension of a filter) Let μ_1, μ_2 be filters on X, we say that μ_2 is an extension of μ_1 if for all $A \subseteq X$ we have $\mu_1(A) = 1 \Rightarrow \mu_2(A) = 1$.We write $\mu_1 \leq \mu_2$.

In order to prove a major result about filter extensions, we will make use of a well-known result in set theory.

Lemma 1. *(Zorn's Lemma) Let $(P, \leq)$ be a partially ordered set where every chain has a least upper bound, then there exists at least one maximal element in P.*

Theorem 1. *Let X be a non-empty set and let μ be a filter, then there exists an ultrafilter μ' that extends μ: $\mu \leq \mu'$.*

Proof. Let X be nonempty and let μ be a filter on X. Consider the set $P = \{\text{extensions of } \mu\}$, it is easy to show that the extension relation $\leq$ is a partial order on this set. Consider a chain C in P, then the map μ_C defined by: for all $A \subseteq X$, if $(A) = 1$ for some $\lambda \in C$ then $\mu_C(A) = 1$ is its upper bound. We can now apply Zorn's lemma and find a maximal extension μ' in P, which can be shown to be an ultrafilter. □

Remark 5. *We do not need the use of Zorn's lemma in order to extend principal filters to ultrafilters. Let μ_A be the principal filter generated by $A \subseteq X$, A is non-empty, so by **proposition 8** there exists $a \in A$ such that $\mu_{\{a\}}$ is an ultrafilter and it is easy to see that $\mu_A \leq \mu_{\{a\}}$.*

Thus, we can extend any filter to an ultrafilter; concretely we can extend a non-principal filter to a non-principal ultrafilter.

4 Construction of Nonstandard Extensions via Ultrapowers

In the last section, we explored some basic properties of filters and we showed that for any non-empty set we can always find an ultrafilter via extension. So now we are ready to develop a method to construct nonstandard extensions. Once we know how to do this, we will be in a position to identify the criteria to find a proper nonstandard extension, which was our goal.

Definition 10. Let X, Y be a non-empty sets, the we define X^Y to be the set of functions from Y to X.

Example 4. *Let X be a non-empty set and let $Y = \mathbb{N}$, then $X^{\mathbb{N}}$ will be the set of functions from $\mathbb{N}$ to X. That is, the set of sequences in X indexed by $\mathbb{N}$, we will denote these sequences by $\bar{x} \in X^{\mathbb{N}}$.*

Definition 11. Let X be a non-empty set, let $X^{\mathbb{N}}$ be the set of all sequences in X indexed by $\mathbb{N}$ and let μ be an ultrafilter on $\mathbb{N}$. Let $\bar{a} = (a_i)_{i\in\mathbb{N}}, \bar{b} = (b_i)_{i\in\mathbb{N}} \in X^{\mathbb{N}}$, we define $\sim_\mu$ by

$$\bar{a} \sim_\mu \bar{b} \Leftrightarrow \mu(\{i \in \mathbb{N} : a_i = b_i\}) = 1.$$

Proposition 9. *$\sim_\mu$ is an equivalence relation.*

Proof. Let X be a non-empty set, let μ be an ultrafilter on $\mathbb{N}$, let $\sim_\mu$ be defined as above for μ and let $\bar{a}, \bar{b}, \bar{c} \in X^{\mathbb{N}}$. Reflexivity and symmetry are immediate, we show transitivity:

Let $\bar{a} \sim_\mu \bar{b}, \bar{b} \sim_\mu \bar{c}$, then

$$\mu(\{i \in \mathbb{N} : a_i = b_i\}) = 1 \wedge \mu(\{i \in \mathbb{N} : b_i = c_i\}) = 1 \Rightarrow \mu(\{i \in \mathbb{N} : a_i = b_i\} \cap \{i \in \mathbb{N} : b_i = c_i\}) = 1$$

by (F2). Therefore $\mu(\{i \in \mathbb{N} : a_i = b_i = c_i\}) = 1$, now $\{i \in \mathbb{N} : a_i = b_i = c_i\} \subseteq \{i \in \mathbb{N} : a_i = c_i\}$ so by (F1) $\mu(\{i \in \mathbb{N} : a_i = c_i\}) = 1 \Rightarrow \bar{a} \sim_\mu \bar{c}$ □

Now that we have shown $\sim_\mu$ is an equivalence relation we can construct well-defined quotient spaces with it.

Definition 12. (Ultrapower) Let X be a non-empty set, let $X^{\mathbb{N}}$ be the set of all sequences in X indexed by $\mathbb{N}$ and let μ be an ultrafilter on $\mathbb{N}$. We define the **ultrapower of** X to be $X^{\mathbb{N}}/\sim_\mu$, where $\sim_\mu$ is as in **definition 11**.

Definition 13. (Ultrapower nonstandard extension) Let X be a non-empty set and μ an ultrafilter on $\mathbb{N}$, we define the collection of functions

$$^* : \mathcal{P}(X^m) \to \mathcal{P}(X^m)^{\mathbb{N}}/\sim \mu$$

$$^*A = \{\bar{a} \in (X^m)^{\mathbb{N}} : \mu(\{i \in \mathbb{N} : a_i \in A\}) = 1\}/\sim_\mu .$$

Remark 6. *With the above definition$^*(X^m) = (X^m)^{\mathbb{N}}/\sim_\mu$, the ultrapower of X.*

Note 3. As $^*\mathbb{X}$ is a quotient map, we will denote its elements by $[\bar{x}]$, where [ ] refers to the equivalence class of $\bar{x}$.

It still rests to show that this structure does, indeed, satisfy the definition of a nonstandard extension.

Theorem 2. *$(X^{\mathbb{N}}/\sim_\mu, ^*)$ as defined above is a nonstandard extension of X.*

Proof. Let $^*X$ be defined as above, we will check (E1)-*(E4)*:

- (E1) Let $B \subseteq A \subseteq X^m$

$$^*B = \{\bar{b} : \mu(\{i \in \mathbb{N} : b_i \in B \subseteq A\}) = 1\}/\sim_\mu$$

so by (F1) $\mu(\{i : b_i \in A\}) = 1$. Therefore

$$\{\bar{b} : \mu(\{i \in \mathbb{N} : b_i \in B \subseteq A\}) = 1\}/\sim_\mu \subseteq \{\bar{b} : \mu(\{i \in \mathbb{N} : b_i \in A\}) = 1\}/\sim_\mu .$$

Let $A, B \subseteq X^m$

$$^*(A \cup B) = \{\bar{a} : \mu(\{i \in \mathbb{N} : a_i \in (A \cup B)) = 1\}/\sim_\mu .$$

By **proposition 7** we get that

$$\mu(\{i : a_i \in (A \cup B)\}) = 1 \Leftrightarrow \mu(\{i : a_i \in A\})\}) = 1 \text{ or } \mu(\{i : a_i \in B\}) = 1.$$

Hence

$$^*(A \cup B) = \{\bar{a} : \mu(\{i : a_i \in A\}) = 1 \text{ or } \mu(\{i : a_i \in B\}) = 1\}/\sim_\mu =$$
$$= \{\bar{a} : \mu(\{i : a_i \in A\}) = 1\}/\sim_\mu \cup \{\bar{a} : \mu(\{i : a_i \in B\}) = 1\}/\sim_\mu = {}^*A \cup {}^*B.$$

Let $A, B \subseteq X^m$

$^*(A \cap B) \subseteq {}^*A \cap {}^*B$: let $[\bar{a}] \in {}^*(A \cap B)$, then $\mu(\{i : a_i \in A \cap B\}) = 1 \Rightarrow \mu(\{i : a_i \in A\}) = 1$ and $\mu(\{i : a_i \in B\}) = 1$ by (F1), so $[\bar{a}] \in {}^*A \cap {}^*B$.

${}^*A \cap {}^*B \subseteq {}^*(A \cap B)$: let $[\bar{a}] \in {}^*A \cap {}^*B$, then $\mu(\{i : a_i \in A\}) = \mu(\{i : a_i \in B\}) = 1 \Rightarrow \mu(\{i : a_i \in A \cap B\} = 1$ by (F2), therefore $[\bar{a}] \in {}^*(A \cap B)$.

Let $A, B \subseteq X^m$

$^*(A \backslash B) \subseteq {}^*A \backslash {}^*B$: let $[\bar{a}] \in {}^*(A \backslash B)$, then

$$\mu(\{i : a_i \in A \backslash B = 1\} \Rightarrow \mu(i : a_i \in A) = 1 \text{ and } \mu(\{i : a_i \in X^m \backslash B\}) = 1$$

both by (F1), therefore $[\bar{a}] \in {}^*A \backslash {}^*B$ and so ${}^*(A \backslash B) \subseteq {}^*A \backslash {}^*B$.

${}^*A \backslash {}^*B \subseteq {}^*(A \backslash B)$: let $[\bar{a}] \in {}^*A \backslash {}^*B$, then

$$\mu(\{i : a_i \in A\}) = 1 \text{ and } \mu(\{i : a_i \in X^m \backslash B\} = 1$$

this implies that $\mu(\{i : a_i \in A \backslash B) = 1$ by (F2), therefore $[\bar{a}] \in {}^*(A \backslash B)$ and ${}^*A \backslash {}^*B \subseteq {}^*(A \backslash B)$

For practical purposes we will check (E3) before (E2).

- (E3) Let $A \subseteq X^m, B \subseteq X^n$, then

$$^*(A \times B) = \{[\bar{x}] \in {}^*(X^{m+n}) : \mu(\{i : x_i = (x_i^1, x_i^2) \in A \times B\}) = 1\} =$$
$$= \{[\bar{x}] \in {}^*(X^{m+n}) : \mu(\{i : x_i^1 \in A \text{ and } x_i^2 \in B\}) = 1\}$$

since

$$\{i : x_i^1 \in A \text{ and } x_i^2 \in B\} \subseteq \{i : x_i^1 \in A\} \text{ and } \{i : x_i^1 \in A \text{ and } x_i^2 \in B\} \subseteq \{i : x_i^2 \in B\}.$$

We conclude by (F1) and (F2) that $\mu(\{i : x_i^1 \in A \text{ and } x_i^2 \in B\}) = 1 \Leftrightarrow \mu(\{i : x_i^1 \in A\}) = 1$ and $\mu(\{x_i^2 \in B\}) = 1$. Now,

$$\bar{x} \sim_\mu \bar{y} \Leftrightarrow \mu(\{i : x_i = y_i\}) = 1 \Leftrightarrow \mu(\{i : x_i^1 = y_i^1 \text{ and } x_i^2 = y_i^2\}) = 1 \Leftrightarrow$$

$$\Leftrightarrow \mu(\{i : x_i^1 = y_i^1\}) = 1 \text{ and } \mu(\{i : x_i^2 = y_i^2\}) = 1 \Leftrightarrow \bar{x}^1 \sim_\mu \bar{y}^1 \text{ and } \bar{x}^2 \sim_\mu \bar{y}^2.$$

Therefore

$$^*(A \times B) = \{[\bar{x}] \in {}^*(X^{m+n})^{\mathbb{N}} : \mu(\{i : x_i^1 \in A \text{ and } x_i^2 \in B) = 1\} =$$

$$= \{([\bar{x}^1], [\bar{x}^2]) \in {}^*(X^{\mathbb{N}})^{m+n}) : \mu(\{i : x_i^1 \in A\}) = 1 \text{ and } \mu(\{i : x_i^2 \in B) = 1\} =$$

$$= \{[\bar{x}^1] \in {}^*(X^{\mathbb{N}})^m : \mu(\{i : x_i^1 \in A\}) = 1\} \times \{[\bar{x}^2] \in {}^*(X^{\mathbb{N}})^n) : \mu(\{i : x_i^2 \in B) = 1\} = {}^*A \times {}^*B.$$

Remark: in the spirit of **proposition 1** it can be shown that $(X^m)^{\mathbb{N}}/\sim_\mu = (X^{\mathbb{N}}/\sim_\mu)^m$.

- (E2) Let $m \geq 2$, $\Delta := \{x = (x^1, x^2, ..., x^m) : \forall i \in \{1, ..., m\} x^i \in X \text{ and } x^j = x^k \text{ for some } j \neq k \in \{1, ..., m\}\}$ then

$$^*\Delta = \{[\bar{x}] \in (X^m)^{\mathbb{N}}/\sim_\mu : \mu(\{i \in \mathbb{N} : x_i \in \Delta\}) = 1\} =$$

$$= \{[\bar{x}] \in (X^m)^{\mathbb{N}}/\sim_\mu : \mu(\{i \in \mathbb{N} : x_i^j = x_i^k) = 1\}.$$

Using the remark that $(X^m)^{\mathbb{N}}/\sim_\mu = (X^{\mathbb{N}}/\sim_\mu)^m$ the former equals to

$$= \{([x^1], .., [x^m]) \in (X^{\mathbb{N}}/\sim_\mu)^m : \mu(\{i \in \{1, .., m\} : x^i \sim_\mu x^k\} =$$

$$= \{([x^1], .., [x^m]) \in (X^{\mathbb{N}}/\sim_\mu)^m : [x^j] = [x^k]\}.$$

- (E4) Let $A \subseteq X^{m+1}$, then

$$^*\pi_m(A) = {}^*\{\pi_m(x) : x \in A\} = \{[\bar{x}] \in {}^*X^m : \mu(\{i : x_i \in \{\pi(x) : x \in A^{\mathbb{N}}\}) = 1\}$$

and by (E2) it equals $\{[\bar{x}] \in {}^*X^m : \mu(\{i : x_i \in A) = 1\} = \pi_m(\{\bar{x} \in {}^*X^{m+1} : x \in A^{\mathbb{N}}\}) = 1\} = \pi_m({}^*A)$.

□

Proposition 10. *If $Y \subseteq X$ is finite, then $^*Y$ is finite. Moreover, $Y = {}^*Y$, so it has no nonstandard elements.*

Proof. It suffices to show that for any $\bar{y} \in \{\bar{a} \in X^{\mathbb{N}} : \mu(\{i : a_i \in Y\}) = 1\}$ we have $\bar{y} \sim_\mu (x, ..., x, ...)$ for some $x \in Y$. Say $Y = \{x_1, ..., x_n\}$ for some $n \in \mathbb{N}$, then $\{i : y_i \in Y\} = \bigcup_{j=1}^n \{i : y_i = x_j\}$, therefore as μ is an ultrafilter by **proposition 7** we have that for some $j \in \{1, ..., n\}$ $y \sim_\mu (x_j, ..., x_j, ...)$.

Now, for $\bar{y}_1, \bar{y}_2$ in distinct equivalence classes, they will each be related to distinct x_i, x_j, hence $Y = {}^*Y$. □

We are finally ready to enunciate the criteria that will allow us to discern whether a nonstandard extension constructed through ultrapowers is proper or not.

Proposition 11. *If X is an infinite set and μ is a filter on $\mathbb{N}$, then:*

- *If μ is principal, ${}^*X = X^{\mathbb{N}}/\sim_\mu$ is **not** a proper nonstandard extension.*
- *If μ is non-principal, ${}^*X = X^{\mathbb{N}}/\sim_\mu$ is a proper nonstandard extension.*

Proof. Let X be an infinite set, let ${}^*A$ be an infinite subset and μ is an ultrafilter on $\mathbb{N}$.

We start with the first case, so let μ be the principal filter generated by $I \subseteq \mathbb{N}$, let ${}^*A$ be an infinite subset. Let $[\bar{a}] \in {}^*A$, then $\mu(\{i : x_i \in A\}) = 1$ therefore $I \subseteq \{i : x_i \in A\}$ as μ is principal. Now suppose for a contradiction that for all $a \in A$ we have that $I \not\subseteq \{i : x_i = a\}$, hence

$$I \subseteq \bigcup_{a \in A} \{i : x_i \neq a\} = \{i : x_i \not\in A\} \Rightarrow \mu(\{i : x_i \not\in A\}) = 1$$

which is a contradiction. Therefore, for all $[\bar{x}] \in {}^*A$ there exists $a \in A$ such that $\mu(\{i : x_i = a\}) = 1$, so it has no nonstandard element.

Now, let μ be a non-principal ultrafilter, then by **proposition 5**, if $A \subseteq \mathbb{N}$ is finite $\mu(A) = 0$, so only infinite subsets will have $\mu(A) = 1$. Let $[\bar{b}]$ be such that for all $i \in \mathbb{N}$ $b_i \in A$ and for all $i \neq j \in \mathbb{N}$ $b_i \neq b_j$. Clearly $\{i : a = b_i\}$ is finite for all $a \in A$ (at most contains one element), therefore it is nonstandard.

□

Proposition 12. *Let $A = \{a_1, ..., a_n\} \subset X$ be finite and I be an infinite index set, then $\Pi_{i \in I} A/\sim_\mu$ has finite size.*

Proof. It suffices to show that for any $\bar{x} = (x_1, ..., x_i, ...) \in \Pi_{i \in I} A$ we have that $(x_1, ..., x_i, ...) \sim (a)$ for some $(a) = (a, ..., a, ...) \in \Pi_{i \in I} A$, as there are finite of the latter.

$$\bar{x} \in \Pi_{i \in I} A \Leftrightarrow \mu(\{i \in I : x_i \in A\}) = 1 \Leftrightarrow$$

$$\Leftrightarrow \mu(\{i : x_i \in \{a_1, ..., a_n\}\}) = 1 \Leftrightarrow \mu(\{i : x_i = a_1\} \cup \{i : x_i \in \{a_2, ..., a_n\}\}) = 1.$$

As μ is an ultrafilter, either $\mu(\{i : x_i = a_1\}) = 1$ or $\mu(\{i : x_i \in \{a_2, ..., a_n\}\}) = 1$. In the first case we are done as $\bar{x} \sim (a_1)$. In the second case we iterate the process, as $\{a_2, ..., a_n\}$ is finite, the iteration will end at some point, so we can conclude that $\bar{x} \sim (a_i)$ for some $i \in \{2, ..., n\}$. □

Proposition 13. *Let $A = \{a_1, ..., a_n, ...\} \subset X$ be countably infinite, I be an infinite index set and μ be a non-principal ultrafilter, then the size of $\Pi_{i \in I} A/\sim_\mu$ is uncountable.*

Proof. We will use an argument similar to Cantor's, assume for a contradiction that there is a bijection $f : \mathbb{N} \to \Pi_{i \in I} A/\sim_\mu$ such that $f(i) = [\bar{s}_i] = [(s_i(1), ..., s_i(n), ...)]$ for a representative $(s_i(1), ..., s_i(n), ...) \in \Pi_{i \in I} A$, $i \in \mathbb{N}$. We construct an element $\bar{s} \in \Pi_{i \in I} A$ that is not in the equivalence class of any element of this list:

Let $s(1)$ be distinct from $s_1(1)$.

Let $s(2)$ be distinct from $s_2(2)$ and $s_1(2)$.

In general, for $i \in I$ let $s(i)$ be distinct from $s_k(i)$ for all $k \leq i$. We can do this as A is countably infinite.

Claim: $\bar{s} \not\sim \bar{s}_i$ for all $i \in I$.

Fix $i \in I$, we can observe that $\{j : s(j) = s_i(j)\} \subseteq \{j : j \leq i\}$, which is finite, therefore $\mu(\{j : s(j) = s_i(j)\}) = 0$, hence $\bar{s} \not\sim \bar{s}_i$. So $[\bar{s}]$ is not in the list, a contradiction as f is a bijection. $\square$

Proposition 14. *Let I be an infinite index set, let $\{A_i\}_I$ be a collection of sets and let μ be non-principal. Then, $\Pi_{i \in I} A_i / \sim_\mu$ is uncountable if $\mu(\{i : |A_i| > i\}) = 1$.*

Proof. Let $\mu(J = \{i : |A_i| > i\}) = 1$ and suppose for a contradiction that there exists a bijection $f : \mathbb{N} \to \Pi_{i \in I} A_i \backslash \sim_\mu$ such that $f(i) = [\bar{s}_i] = [(s_i(1), ..., s_i(n), ...)]$ for a representative $(s_i(1), ..., s_i(n), ...) \in \Pi_{i \in I} A$, $i \in \mathbb{N}$. We construct an element $\bar{s} \in \Pi_{i \in I} A$ that is not in the equivalence class of any element of this list:

For $i \in I$: if $i \in J$, then $|A_i| > i$ so we can let $s(i)$ be distinct from $s_j(i)$ for all $k < i$. If $i \not\in J$, let $s(i) = s_i(i)$.

Claim: $\bar{s} \not\sim \bar{s}_i$ for all $i \in I$.

Fix $i \in I$, we can observe that $\{j : s(j) = s_i(j)\}$ is at most $\{j : j < i\} \cup I \backslash J$, but both of these sets are finite, so their union will be finite and $\mu(\{j : s(j) = s_i(j)\}) = 0$, hence $\bar{s} \not\sim \bar{s}_i$. So $[\bar{s}]$ is not in the list, a contradiction as f is a bijection. $\square$

5 Transfer Theorem

Before we exploring concrete examples of proper nonstandard extensions, it is vital that we present the transfer theorem. This theorem will be the tool that will let us translate first order formulas from our standard model to the nonstandard one. Its relevance will be established when looking at applications.

One could approach the following is a first order logic flavour, considering a very rich language that includes a constant for every element of the domain, a predicate for every subset, for every function...This would not add too much to our discussion, so we will instead follow Henson's formulation. In the following we will be considering formulas not in a meta-mathematical fashion, but rather as realised formulas within our domain X. This is similar to how the formula for continuity of a function is regarded when doing mathematics in $\mathbb{R}$.

Definition 14. (Formulas over X) Let X be a non-empty set, the **set of logical formulas over** X, $\mathcal{F}_X$, is the smallest set of logical formulas which satisfies the following conditions:

- i) For every $A \subseteq X^m$ $(\bar{x} \in A) \in \mathcal{F}_X$.
- ii) For every function $f : A \to B$, where $A \subseteq X^m, B \subseteq \mathbb{X}^n$ and $\bar{x} \in A, \bar{y} \in B$

$$(f(\bar{x}) = \bar{y}) \in \mathcal{F}_X.$$

- iii) If $\varphi(x_1, ..., x_m, y_1, .., y_n) \in \mathcal{F}_X$ and $a_1, ..., a_m \in X^n$, then the formulas

$$\varphi(x_1, ..., x_m, a_1, .., a_n) \in \mathcal{F}_X.$$

- iv) If $\varphi, \psi \in \mathcal{F}_X$, then $\neg\varphi$,$\varphi \vee \psi$,$\varphi \wedge \psi$,$\varphi \Rightarrow \psi$,$\varphi \Leftrightarrow \psi$,$\exists x\varphi$,$\forall x\varphi$ are all in $\mathcal{F}_X$.

We call the elements of $\mathcal{F}_X$ **formulas over** X. [2]

Definition 15. (Truth of a sentence over X) Let $\varphi(x_1, ..., x_m)$ be a sentence over X, φ is true if the set defined by φ is non-empty, that is, $B = \{\bar{x} \in X^m : \varphi(\bar{x}) \text{ is true in } X\} \neq \emptyset$. Otherwise it is false.

Definition 16. (*-Transform of a Formula over X) Let X be a non-empty set and let $^*X$ be a fixed nonstandard extension. Let $\varphi(x_1, .., x_n) \in \mathcal{F}_X$, the *-transform is a function:

$$\mathcal{F}_X \to \mathcal{F}_{^*X}$$

$$\varphi(x_1, .., x_n) \mapsto {}^*\varphi(X_1, .., X_n)$$

that satisfies the following inductive conditions on the complexity of formulas:

- (T1) Atomic formulas:

  Let $\{\alpha(1), ..., \alpha(m)\} \subseteq \{1, ..., n\}$, $A \in \mathcal{P}(X^m)$ and f a function, then:

$${}^*((x_{\alpha(1)}, ..., x_{\alpha(m)}) \in A) = (X_{\alpha(1)}, ..., X_{\alpha(m)}) \in {}^*A)$$

$${}^*(f(x_{\alpha(1)}, ..., x_{\alpha(k)}) = (x_{\alpha(k+1)}, ..., x_{\alpha(m)})) = ({}^*f(X_{\alpha(1)}, ..., X_{\alpha(k)}) = (X_{\alpha(k+1)}, ..., X_{\alpha(m)})).$$

- (T2) If $\varphi(x_1, .., x_m, y_1, ..., y_n)$ is an atomic formula and $(a_1, ..., a_n) \in X$, then:

$${}^*(\varphi(x_1, .., x_m, a_1, ..., a_n)) = {}^*\varphi(X_1, ..., X_m, {}^*A_1, ..., {}^*A_n).$$

  For the inductive steps we will assume that of $^*\varphi, {}^*\psi$ of $\varphi, \psi \in \mathcal{F}_X$ are well defined, then:

- (T3) $^*(\neg\varphi) = \neg^*\varphi$
- (T4) $^*(\varphi \vee \psi) = {^*\varphi} \vee {^*\psi}$
- (T5) $^*(\varphi \wedge \psi) = {^*\varphi} \wedge {^*\psi}$
- (T6) $^*(\exists x\varphi) = \exists X^*\varphi$
- (T7) $^*(\forall x\varphi) = \forall X^*\varphi$ [3]

So the *-transform takes a formula over X and extends the objects involved to their nonstandard extension in $^*X$. Now we will prove the transfer theorem, which will give us the connection between nonstandard extensions and the *-transform. We will prove it expanding Henson's proof.[4]

Theorem 3. *(Transfer Theorem) Let X be a non-empty set and consider a fixed nonstandard extension of X, then:*

- *i) Let $\varphi(x_1, ..., x_m) \in \mathcal{F}_X$ and let $^*\varphi(x_1, ..., x_m) \in {^*\mathcal{F}_X}$ be its *-transform. Let $B = \{\bar{x} \in X^m : \varphi(\bar{x})$ is true in $X\}$ be the set defined by $\varphi(x_1, ..., x_m)$, then the set defined by $^*\varphi(X_1, ..., X_M)$ is $^*B = \{\bar{x} \in {^*X^m} : {^*\varphi(\bar{x})}$ is true in $^*X\}$.*
- *ii) Let $\varphi \in \mathcal{F}_X$ be a sentence and let $^*\varphi \in {^*\mathcal{F}_X}$ be its *-transform, then φ is true in X if and only if $^*\varphi$ is true in $^*X$. [4]*

Proof. We will first show that *i)* implies *ii)*:

Let *i)* be true and let $\varphi \in \mathcal{F}_X$ be a sentence with *-transform $^*\varphi$.

Let φ be true in X, as sentences have no free variables, the set B defined by φ is either $\emptyset$ if φ is false or X^0 (the whole set of 0-length variables) if it is true. By part *i)* $^*B$ is the set defined by $^*\varphi$ and it is either $\emptyset$ or $^*X^m$ when it is false and true respectively. By **corollary 1** $B = A \iff {^*B} = {^*A}$, so $B = \emptyset \iff {^*B} = \emptyset$ and $B = X^0 \iff {^*B} = {^*X^0}$.

Now we will prove *i)* by induction on the complexity of formulas in X.

The first induction step will be the atomic formulas:

$\varphi(x_1, ..., x_n)$: $(x_{\alpha(1)}, ..., x_{\alpha(m)}) \in A \subseteq X^m$, where α is a map from $\{1, ..., m\}$ into $\{1, ..., n\}$. This lets us cover any combination of any length of the possibly distinct n variables of $\varphi(x_1, ..., x_n)$. We will prove by induction on n that the *-transform of the solution set $B = \{(x_1, ..., x_n) \in X^n : (x_{\alpha(1)}, ..., x_{\alpha(m)})\} \in A \subseteq X^m\}$ is $^*B = \{(X_1, ..., X_n) \in {^*X^n} : (X_{\alpha(1)}, ..., X_{\alpha(m)})\} \in {^*A} \subseteq {^*X^m}\}$:

Consider $C = \{(x_1, ..., x_n, x_{\alpha(1)}, ..., x_{\alpha(m)}) \in X^{n+m} : (x_{\alpha(1)}, ..., x_{\alpha(m)}) \in A \subseteq X^m\}$. We can observe that $B = \pi_n(C)$, as α is a function, the $\alpha(i)$ will equal some $i \in \{1, ..., n\}$, so C is the intersection of m diagonal subsets of X^{n+m}: $\Delta_1, ..., \Delta_m$, and $X^n \times A$. Therefore if we apply * we get

$$^*B = {^*\pi_n}(C) = \pi_n({^*C}) = \pi_n({^*(\Delta_1 \cap \Delta_2 \cap, ..., \cap \Delta_m \cap (X^{n+m} \times A))}) =$$
$$= \pi_n({^*\Delta_1} \cap {^*\Delta_2} \cap, ..., \cap {^*\Delta_m} \cap {^*(X^{n+m} \times A)}) = \pi_n({^*\Delta_1} \cap {^*\Delta_2} \cap, ..., \cap {^*\Delta_m} \cap ({^*X^{n+m}} \times {^*A})) =$$
$$= \pi_n(\{(X_1, ..., X_n, X_{\alpha(1)}, ..., X_{\alpha(m)}) \in {^*X^{n+m}} : (X_{\alpha(1)}, ..., X_{\alpha(m)}) \in {^*A} \subseteq {^*X^m}\}) =$$
$$= \{(X_1, ..., X_n) \in {^*X^n} : (X_{\alpha(1)}, ..., X_{\alpha(m)})\} \in {^*A} \subseteq {^*X^m}\}$$

by *(E4)*, (E1), (E3), (E2) in that order.

We also have to consider the case where we substitute variables $x_{\alpha(i)}$ for constant elements a_i of X : $B(a) = \{(x_{\alpha(1)}, ..., x_{\alpha(m)}) \in B : x_{\alpha(i)} = a_j \in X$ for some $i \in \{1, ..., m\}$.

Suppose we substitute $k \leq m$ variables, then up to reordering $B(a) = B \cap (\{a_1\} \times .... \times \{a_k\} \times X^{m-k})$, so by (E1), (E3) and the previous result ${}^*B({}^*A) = \{(X_{\alpha(1)}, ..., X_{\alpha(m)}) \in B : X_{\alpha(i)} = {}^*A_j \in {}^*X$ for some $i \in \{1, ..., m\}\}$, the desired result.

The other atomic formula we have to consider is $f(x_{\alpha(1)}, ..., x_{\alpha(k)}) = (x_{\alpha(k+1)}, .., x_{\alpha(m)})$ where f is a function from $A \subseteq X^k$ to $B \subseteq X^{m-k}$. The set defined by this proposition is its graph Γ, we know from **proposition 4** that ${}^*\Gamma$ is well defined and by definition ${}^*f$ is the function whose graph is ${}^*\Gamma$. In other words, the set defined by ${}^*f(X_{\alpha(1)}, ..., X_{\alpha(k)}) = (X_{\alpha(k+1)}, .., X_{\alpha(m)})$ is ${}^*\Gamma$ which is what we wanted.

In this case, we also want to handle the substitution by constant elements of A. Like in the previous case, we define $\Gamma(a) = \{(x_{\alpha(1)}, ..., x_{\alpha(k)}, x_{\alpha(k+1)}, ..., x_{\alpha(m)}) \in \Gamma : x_{\alpha(i)} = a_j$ for some $i \in \{1, ..., k\}\}$. Suppose we substitute $h \leq k$ variables, then up to reordering $\Gamma(a) = \Gamma \cap (\{a_1\} \times ... \times \{a_j\} \times X^{m-h})$ and we apply the same method as the previous case.

Now we will take the inductive steps:

1. Let $\varphi \in \mathcal{F}_X$ of m variables that defines a set B. The set defined by $\neg\varphi$ is $C = \{x \in X : \neg\varphi \text{ is true in } X\} = X^m \backslash B$, by (E1) ${}^*C = {}^*(X^m \backslash B) = {}^*X^m \backslash {}^*B$. Now, we know that the set defined by ${}^*\varphi$ is ${}^*B$, therefore the set defined by $\neg{}^*\varphi$ is also ${}^*X^m \backslash {}^*B$.

2. Let $\varphi, \psi \in \mathcal{F}_X$ of m, n variables respectively, and let A, B be the sets defined by each of them respectively. The set defined by $\varphi \vee \psi$ is $C = \{x \in X^{m+n} : \varphi \vee \psi \text{ is true in } X\} = \{x \in X^m : \varphi \text{ is true in } X\} \cup \{x \in X^n : \psi \text{ is true in } X\} = A \cup B$, by (E1) ${}^*C = {}^*(A \cup B) = {}^*A \cup {}^*B$. We know that the set defined by ${}^*\varphi$ is ${}^*A$ and the one defined by ${}^*\psi$ is ${}^*B$, therefore the set defined by ${}^*\varphi \vee {}^*\psi$ is also ${}^*A \cup {}^*B$.

3. Let $\varphi, \psi \in \mathcal{F}_X$ of m, n variables respectively, and let A, B be the sets defined by each of them respectively. The set defined by $\varphi \wedge \psi$ is $C = \{x \in X^{m+n} : \varphi \wedge \psi \text{ is true in } X\} = \{x \in X^m : \varphi \text{ is true in } X\} \cap \{x \in X^n : \psi \text{ is true in } X\} = A \cap B$, by (E1) ${}^*C = {}^*(A \cap B) = {}^*A \cap {}^*B$. We know that the set defined by ${}^*\varphi$ is ${}^*A$ and the one defined by ${}^*\psi$ is ${}^*B$, therefore the set defined by ${}^*\varphi \wedge {}^*\psi$ is also ${}^*A \cap {}^*B$.

4. Let $\varphi \in \mathcal{F}_X$ with variables $(x_1, .., x_m, y_1, ..., y_n)$, defining a set B, then the solution set for $\exists x \varphi(x, y)$ is true in X is $C = \{y \in X : \varphi(x, y)\} = \pi(B)$, by (E1) ${}^*C = {}^*\pi(B) = \pi({}^*B) = \{y \in {}^*X^n : {}^*\varphi(x, y)$ is true in ${}^*X$. We know that the set defined by ${}^*\varphi(x, y)$ is ${}^*B$, so the set defined by $\exists x {}^*\varphi(x, y)$ is also $\pi({}^*B)$.

5. Let $\varphi \in \mathcal{F}_X$ with variables $(x_1, .., x_m, y_1, ..., y_n)$, defining a set B, then the solution set for $\forall x \varphi(x, y)$ is $C = \{x \in X^m : \{x \times \{y : y \in Y^n \subseteq B\}$, therefore ${}^*C = \{x \in {}^*X^m : \{x\} \times \{y : y \in {}^*Y^m\} \subseteq {}^*B\}$. The set defined by ${}^*\varphi$ is ${}^*B$, therefore the set defined by $\forall y \varphi(x, y)$ is also equal to $\{x \in {}^*X^m : \{x\} \times \{y : y \in {}^*Y\} \subseteq {}^*B\}$. $\square$

6 The Nonstandard Real Numbers

In this section we will construct a proper nonstandard extension of the real numbers via the ultrapower method developed in **section 4**. For this task we will assume the existence of the real numbers and all its properties (the fact that it is a totally ordered field etc.). At the end of the section we will see some simple examples of the transfer principle applied to $\mathbb{N}$.

Definition 17. (Nonstandard Real Numbers) Assume $\mathbb{R}$ and let μ be a non-principal ultrafilter, let $\sim_\mu$ as in **definition 11** , the **nonstandard real numbers** are the pair $(^*, {}^*\mathbb{R})$, where

$$^* : \mathcal{P}(\mathbb{R}^m) \to \mathcal{P}((\mathbb{R}^m)^{\mathbb{N}} / \sim \mu)$$

$$^*A = \{\bar{a} \in (\mathbb{R}^m)^{\mathbb{N}} : \mu(\{i \in \mathbb{N} : a_i \in A\}) = 1\} / \sim_\mu .$$

And so ${}^*\mathbb{R} = \mathbb{R}^{\mathbb{N}} / \sim_\mu$, the ultrapower of $\mathbb{R}$.

Remark 7. *We want this nonstandard extension to be proper (ie. it has nonstandard elements that $\mathbb{R}$ does not have), this is why the ultrafilter μ chosen in the previous construction is non-principal (see **proposition 11**).*

We can now proceed to define a nonstandard version of the usual operations and relations on $\mathbb{R}$.

Definition 18. (Addition) Let $[\bar{a}], [\bar{b}] \in {}^*\mathbb{R}$, we define

$$^*+ : {}^*\mathbb{R} \times {}^*\mathbb{R} \to {}^*\mathbb{R}$$

$$[\bar{a}]^*{+}[\bar{b}] = [(a_1 + b_1, a_2 + b_2, ..., a_i + b_i, ...)].$$

Where for all $i \in \mathbb{N}$ $a_i + b_i$ is the usual addition in the reals.

Definition 19. (Multiplication) Let $[\bar{a}], [\bar{b}] \in {}^*\mathbb{R}$, we define

$$^*. : {}^*\mathbb{R} \times {}^*\mathbb{R} \to {}^*\mathbb{R}$$

$$[\bar{a}]^*.[\bar{b}] = [(a_1.b_1, a_2.b_2, ..., a_i.b_i, ...)].$$

Where for all $i \in \mathbb{N}$ $a_i.b_i$ is the usual multiplication in the reals.

Proposition 15. *The operations $^*+$ and $^*.$ are well defined.*

Proof. Let $[\bar{a}], [\bar{b}], [\bar{c}] \in {}^*\mathbb{R}$ and $[\bar{a}] = [\bar{c}]$. For $^*+$ to be well defined $[(a_1 + b_1, ..., a_i + b_i, ...)] = [c_1 + b_1, ..., c_i + b_i, ...)]$ and this is if and only if

$$(a_1 + b_1, ..., a_i + b_i, ...) \sim_\mu (c_1 + b_1, ..., c_i + b_i, ...) \Leftrightarrow \mu(\{i : a_i + b_i = c_i + b_i\}) = 1$$

by the definition of $+$ in the reals $\{i : a_i + b_i = c_i + b_i\} = \{i : a_i = c_i\}$ and we know this set has measure one, so it is well defined.

Let $[\bar{a}], [\bar{b}], [\bar{c}] \in {}^*\mathbb{R}$ and $[\bar{a}] = [\bar{c}]$. for $+$ to be well defined $[(a_1.b_1, ..., a_i.b_i, ...)] = [c_1.b_1, ..., c_i.b_i, ...)]$ and this is if and only if

$$(a_1.b_1, ..., a_i.b_i, ...) \sim_\mu (c_1.b_1, ..., c_i.b_i, ...) \Leftrightarrow \mu(\{i : a_i.b_i = c_i.b_i\}) = 1$$

by the definition of . in the reals $\{i : a_i.b_i = c_i.b_i\} = \{i : a_i = c_i\}$ and we know this set has measure one, so it is well defined. $\square$

Definition 20. Let $[\bar{a}], [\bar{b}] \in {}^*\mathbb{R}$, we define the relation ${}^* \leq$ by

$$[\bar{a}]^* \leq [\bar{b}] \Leftrightarrow \mu(\{i : a_i \leq b_i\}) = 1.$$

Where $\leq$ has the usual definition in the reals.

Proposition 16. *$({}^*\mathbb{R}, {}^*+, {}^*., {}^* \leq)$ is a totally ordered field, and extends $\leq$ in $\mathbb{R}$, that is, if $a \leq b$ in $\mathbb{R}$, then ${}^*a^* \leq {}^*b$ in ${}^*\mathbb{R}$.*

Proof. Anti-symmetry and reflexivity are immediate.

Let $[\bar{a}]^* \leq [\bar{b}]$ and $[\bar{b}]^* \leq [\bar{c}]$

$$\{i : a_i \leq b_i\} \cap \{i : b_{ii}\} = \{i : a_i \leq b_i \leq c_i\} \subseteq \{i : a_i \leq c_i\}$$

so by (F2) and (F1) $\mu(\{i : a_i \leq c_i\} = 1 \Rightarrow [\bar{a}]^* \leq [\bar{c}]$. So it is transitive.

It follows from the ultrafilter property that either $\mu(\{i : a_i \leq b_i\}) = 1$ or $\mu(\{i : b_i < a_i\}) = 1 \Rightarrow \mu(\{i : b_i \leq a_i\}) = 1$ by (F1), so ${}^* \leq$ is a total ordering.

Let $a, b \in \mathbb{R}$ and $a \leq b$, we have that

$$\{i : a_i = a\} \cap \{i : b_i = b\}) \subseteq \{i : a_i \leq b_i\}$$

so by (F2) and (F1), $\mu(\{i : a_i \leq b_i\}) = 1 \Rightarrow [\bar{a}]^* \leq [\bar{b}]$.

${}^*+$ is a binary operation: let $[\bar{a}], [\bar{b}] \in {}^*\mathbb{R}$, as $+$ is a binary operation on $\mathbb{R}$, we have that for all $i \in \mathbb{N}$

$$a_i + b_i \in \mathbb{R} \Rightarrow (a_1 + b_1, .., a_i + b_i, ...) \in \mathbb{R}^{\mathbb{N}} \Rightarrow [(a_1 + b_1, ..., a_i + b_i, ...)] = [(\bar{a} + \bar{b})] \in {}^*\mathbb{R}$$

Associativity of ${}^*+$: let $[\bar{a}], [\bar{b}], [\bar{c}] \in {}^*\mathbb{R}$, by associativity of $+$ in $\mathbb{R}$

$$[((a_1 + b_1) + c_1, ..., (a_i + b_i) + c_i, ...)] = [\bar{a} + \bar{b}]^* + [\bar{c}] = ([\bar{a}]^* + [\bar{b}])^* + [\bar{c}]$$

Commutativity of ${}^*+$: let $[\bar{a}] \in {}^*\mathbb{R}$, then $[\bar{a}]^* + [\bar{b}] = [(a_1 + b_1, .., a_i + b_i, ...)]$ which by commutativity of $+$ in $\mathbb{R}$ equals $[(b_1 + a_1, ..., b_i + a_i, ...)] = [\bar{b}]^* + [\bar{a}]$.

Additive identity: let $[\bar{a}] \in {}^*\mathbb{R}$, then can easily be checked that $0 = [(0, ..., 0, ...)]$ is the additive identity.

Additive inverses: let $[\bar{a}], [\bar{b}] \in {}^*\mathbb{R}$ such that for all $i \in \mathbb{N}$ we have $b_i = -a_i$, then

$$[\bar{a}]^* + [\bar{b}] = [(a_1 + b_1, ..., a_i + b_i, ..)] = [(0, 0, ..., 0, ...)] = [0] = [\bar{b} + \bar{a}]$$

by commutativity. For all $a \in {}^*\mathbb{R}$ we will denote such $[\bar{b}]$ as $-[\bar{a}]$.

Similarly for $({}^*\mathbb{R}, {}^*.)$, with $[1]$ as the multiplicative identity and the inverse of $[(a_1, ...)]$ being $[(b_i)_{i \in \mathbb{N}}]$ where $b_i = 1/a_i$ if $a_i \neq 0$, $b_i = a_i$ otherwise.

Distributive property: Let $[\bar{a}], [\bar{b}], [\bar{c}] \in {}^*\mathbb{R}$, then

$$[\bar{a}]^*.([\bar{b}]^* + [\bar{c}]) = [\bar{a}]^*.([(b_1 + c_1, ..., b_i + c_i, ...)]) = [(a_1(b_1 + c_1), ..., a_i.(b_i + c_i), ...)]$$

and by the distribute property of . over +, that equals

$$[(a_1.b_1 + a_1.c_1, ..., a_i.b_i + a_i.c_i, ...)] = [\bar{a}]^*.[\bar{b}]^* + [\bar{a}]^*.[\bar{c}]$$

□

Proposition 17. *There exists a field homomorphism from* $(\mathbb{R}, +, .)$ *into* $({}^*\mathbb{R}, {}^*+, {}^*.)$.

Proof. We already know from **remark 2** that there exists a canonical embedding from a set into a nonstandard extension, namely $\Psi : \mathbb{R} \to {}^*\mathbb{R}$ sending $r \mapsto {}^*\{r\} = {}^*r$. We showed implicitly in the previous proposition that both operations are homomorphisms when proving they are binary operations. $\square$

In the following we will explore some straightforward applications of the transfer theorem, proving that many properties of the standard real numbers are preserved under their nonstandard extension.

Proposition 18. *Every element in* ${}^*\mathbb{N}$ *has an immediate successor and every element in* ${}^*\mathbb{N}\backslash\{[0]\}$ *has an immediate predecessor.*

Proof. We know the statement that every element in $\mathbb{N}$ has an immediate successor

$$\varphi : \forall x \in \mathbb{N} \exists y \in \mathbb{N}(x < y \wedge \forall z \in \mathbb{N} x < z \to y \leq z)$$

is true and a well formed formula over $\mathbb{N}$, therefore applying the transfer theorem:

$${}^*\varphi : \forall x^* \in {}^*\mathbb{N} \exists y^* \in {}^*\mathbb{N}(x^* < y \wedge \forall z^* \in {}^*\mathbb{N} x^* < z \to y^* \leq z).$$

Is a true well formed formula over ${}^*\mathbb{N}$, as we can handle relations like ${}^*\in, {}^*<$ by their graphs, which are subsets of ${}^*\mathbb{N}^m$. Therefore, every element in ${}^*\mathbb{N}$ has an immediate successor. The second statement can be proved almost identically. $\square$

Proposition 19. ${}^*\mathbb{Q}$ *is dense in* ${}^*\mathbb{R}$

Proof. We formulate the statement as a sentence over $\mathbb{R}$ and apply transfer

$$\varphi : \forall x, y \in \mathbb{R} x < y \exists z \in \mathbb{Q}(x < z < y).$$

$\square$

Proposition 20. ${}^*\mathbb{R}$ *is not complete (ie. every bounded set has a supremum).*

Proof. We know that $\mathbb{R}$ is bounded in ${}^*\mathbb{R}$ by $[(1, 2, 3, ..., m, m+1, ...)]$, but $\mathbb{R}$ does not have a supremum. $\square$

Remark 8. *One could have been tempted to apply transfer theorem to the previous proposition, but density is not a first order property.*

Proposition 21. *(Nonstandard Archimedean Principle)* $\forall r \in {}^*\mathbb{R}_{>[0]} \exists n \in {}^*\mathbb{N}([0] < [1/n] < r)$

Proof. We apply transfer theorem as usual to the Archimedean principle in the real numbers and get the desired result.

$\square$

7 Application to Graph Theory and König's Lemma

We will prove König's lemma employing machinery from the previous sections.

Definition 21. ([Simple] graph) A **graph** G is a pair (V, E) where V is a set of vertices and a subset $E \subseteq V^2$ whose elements are called edges.

We can build a first order language $\mathcal{L}_G = \{E\}$ for graphs, where E is a binary relation satisfying that: $\forall x \neg(Exx)$ and $\forall x, y(Exy \leftrightarrow Eyx)$. A model $\mathcal{G}$ for such language will be a graph.

Definition 22. (Connected graph) We say a graph is **connected** when V is non-empty and for every pair of distinct vertices v, w there exists some $n \in \mathbb{N}$ such that $E(v_1, v_2), E(v_2, v_3), ..., E(v_{n-1}, v_n)$ and $v_1 = v, v_n = w$. Formally:

$$\forall v, w \in V \exists n \in \mathbb{N} \forall i < n \exists v_i \in V(E(v_i, v_{i+1})) \wedge (v_1 = v \wedge v_n = w))$$

Definition 23. Let $v, w \in V$ be distinct, we define the distance function of a graph $d : V^2 \to \mathbb{N}$ to be the minimum number of distinct vertices $v_2, .., v_n = w$ that connect v and w.

It can be shown using compactness theorem that graph connectedness is not a first-order property. So by limiting ourselves to the language of graphs, with graphs as our $\mathcal{L}_G$-structures we cannot work nicely with connectedness. One could be tempted to enrich $\mathcal{L}_G$ in order to avoid this issue, but we have developed in the previous sections a very powerful method that might be helpful instead. We built the nonstandard reals ${}^*\mathbb{R}$ and showed that some of the usual operations and relations are preserved, so now we can use this nonstandard extension to prove results about graphs. In the following will see a countable graph G as a subset of $\mathbb{N}$, identifying $v_i \in V$ with $i \in \mathbb{N}$, and adapting the edge relation accordingly.

Definition 24. (Nonstandard graph) Let $G = (V, E)$ be a graph, where $V \subseteq \mathbb{N}, E \subseteq \mathbb{N}^2$, we define its **nonstandard graph** to be ${}^*G = ({}^*V, {}^*E)$ such that:

${}^*V$ is the nonstandard extension of V in ${}^*\mathbb{R}$.

${}^*E$ is the binary relation on ${}^*V$ defined by ${}^*E(v, w) \Leftrightarrow \mu(\{i : E(v_i, w_i)\}) = 1$, where μ is the non-principal ultrafilter used to build ${}^*\mathbb{R}$.

Definition 25. (Locally finite graphs) We say that a graph G is **locally finite** if every vertex $v \in V$ has a finite neighborhood $N_v(G) := \{w \in V : E(v, w)\}$ in G.

In order to work with nonstandard methods it is convenient to characterise local finiteness of graphs by whether they contain standard elements only.

Lemma 2. *G is locally finite if and only if ${}^*(N_v(G)) \subset V$ for every $v \in V$.*

Proof. $\Rightarrow$ Let G be locally finite, then for every $v \in V$, the set $N_v(G)$ is finite, therefore its nonstandard extension is finite by **proposition 10**.

$\Leftarrow$ Now let ${}^*(N_v(G)) \subset V$ for every $v \in V$, therefore ${}^*(N_v(G))$ has no nonstandard elements, so its preimage must be finite, as μ is non-principal. By the way V is embedded in ${}^*V$, the preimage of ${}^*N_v(G)$ is just $N_v(G)$, therefore G is locally finite.

□

Definition 26. (Path) Let G be a graph, a **path** in G is a set of distinct vertices $\{v_i : i \in I \subseteq \mathbb{N}\}$ such that $E(v_i, v_{i+1})$ for all $i \in I$.

Proposition 22. *Every infinite, connected, locally finite graph has an infinite path.*

Proof. We will consider the proposition in the nonstandard extension first. Let G be an infinite, connected, locally finite graph and let ${}^*G$ be its nonstandard model.

By connectedness of G

$$\forall v, w \in V \exists n \in \mathbb{N} \forall i < n \exists v_i \in V(E(v_i, v_{i+1})) \wedge (v_1 = v \wedge v_n = w)$$

holds and is a sentence over G, therefore by applying transfer theorem we get

$$\forall v, w \in {}^*V \exists n \in {}^*\mathbb{N} \forall i < n \exists v_i \in {}^*V({}^*E(v_i, v_{i+1})) \wedge (v_1 = v \wedge v_n = w)$$

is true. Take vertices $v \in V$ and $w = \sigma \in {}^*V \backslash V$ (such σ exists as V is infinite). There exists $\lambda \in {}^*\mathbb{N}$ such that $\forall i < \lambda \exists v_i \in {}^*V({}^*E(v_i, v_{i+1})) \wedge (v_1 = v \wedge v_n = \sigma)$.

Claim: $\lambda \in {}^*\mathbb{N} \backslash \mathbb{N}$.

Suppose not, then $\lambda \in \mathbb{N}$, so λ is finite. Now v_1 is standard, so its neighborhood is finite, therefore has no nonstandard elements and because ${}^*E(v_1, v_2)$, $v_2 \in V$ ie. it is standard. We repeat this process $\lambda - 2$ times and finally deduce that $\sigma \in V$, a contradiction, therefore $\lambda \in {}^*\mathbb{N} \backslash \mathbb{N}$.

We can now assert that $\mathbb{N} \in \lambda$ and so $(v_i)_{i \in \mathbb{N}} \subset (v_i)_{1 \leq i \leq \lambda}$. The * preimage of $(v_i)_{i \in \mathbb{N}}$ is well defined and forms an infinite path in G, as the relation E between standard vertices is preserved under * and its inverse. $\square$

Theorem 4. *(König's Lemma) An infinite finitely branched tree has an infinite path.*

Proof. An infinite finitely branching tree is an infinite, connected, locally finite graph such that any two vertices are uniquely connected, therefore we can apply **proposition 22** and get the desired result. $\square$

8 The Space of Ultrafilters βS and Stone-Čech Compactification

We now come back to ultrafilters. We will delve further into their properties and prove some interesting results. We start by discussing the topological properties of the set of ultrafilters, which will conveniently be well behaved.

Note 4. For the rest of the paper the notation of filters as sets will be mostly used, going back to our functional notation when dealing with ultrapowers.

Definition 27. Let S be an infinite set, we denote the **set of all ultrafilters on** S by βS.

In the following we will consider S as an infinite topological space with the discrete topology (ie. every subset of S is open).

Definition 28. We define the topology τ on βS to be the collection of sets generated by the basic open sets $U_A := \{\mathcal{U} \in \beta S : A \in \mathcal{U}\}$ for each $A \subseteq S$.

Proposition 23. *$U_A \cap U_B = U_{A\cap B}$, $U_A \cup U_B = U_{A\cup B}$ for any $A, B \subseteq S$ and the basic open sets of the topology τ are closed.*

Proof. For $A, B \subseteq S$, we have $U_A \cap U_B = \{\mathcal{U} \in \beta S : A \in \mathcal{U}\} \cap \{\mathcal{U} \in \beta S : B \in \mathcal{U}\} = \{\mathcal{U} \in \beta S : A \in \mathcal{U}$ and $B \in \mathcal{U}\} = \{\mathcal{U} \in \beta S : A \cap B \in \mathcal{U}\}$ by the property (F1) of filters, so $\{U_A : A \in S\}$ is basis for the topology.

Let $A, B \subseteq S$, then $U_A \cup U_B = \{\mathcal{U} \in \beta S : A \in \mathcal{U}\} \cup \{\mathcal{U} \in \beta S : B \in \mathcal{U}\} = \{\mathcal{U} \in \beta S : A \in \mathcal{U}$ or $B \in U\} = \{\mathcal{U} \in \beta S : A \cup B \in \mathcal{U}\}$ as $A, B \subseteq A \cup B$.

Let $A \in S$, $\beta S \backslash U_A = \beta S \backslash \{\mathcal{U} \in \beta S : A \in \mathcal{U}\} = \{\mathcal{U} \in \beta S : S\backslash A \in \mathcal{U}\} = U_{S\backslash A}$, which is also a basic open set. □

Lemma 3. *(Finite intersection property) Let S be a set and $\mathcal{A}$ a collection of subsets of S such that every finite intersection of elements in $\mathcal{A}$ is non-empty, then there exists an ultrafilter $\mathcal{U}$ such that $\mathcal{A} \subseteq \mathcal{U}$.*

Proof. Let S be a set and $\mathcal{A}$ a collection of subsets of S such that every finite intersection of elements of $\mathcal{A}$ is non-empty . Let $\mathcal{A}_1$ be the set consisting of and all the finite intersections of elements of $\mathcal{A}$ (note that this set trivially contains $\mathcal{A}$), we can see that $\mathcal{A}_1$ satisfies *(F2),(F3)*, but (F1) is not clear. Construct $\mathcal{A}_2$ to be the set of supersets of elements in $\mathcal{A}_1$. It can be easily shown that $\mathcal{A}_2$ is a filter on S and by **theorem 1** we can extend it to an ultrafilter $\mathcal{U}$. □

Proposition 24. *The topological space βS is compact Hausdorff.*

Proof. It is sufficient to show that every cover by the basic open sets U_A has a finite subcover. Let $I \subseteq \mathbb{N}$ be an infinite index set and let $\bigcup_{i\in I} U_{A_i}$ be a cover of βS. Suppose for a contradiction that it has no finite subcover, then for any finite $J \subset I$ we have that there exists an ultrafilter $\mathcal{U}$ such that $\{s : s \in A_i\} \not\in \mathcal{U}$ for all $i \in J$, so by the ultrafilter property $\{s : s \in S\backslash A_i\} \in \mathcal{U}$ for all $i \in J$, hence $S\backslash A_i \in \mathcal{U}$ for all $i \in J$ and $\bigcap_{i\in J} S\backslash A_i \in \mathcal{U}$. We conclude that any finite intersection is non-empty and by **lemma 3** there exists an ultrafilter $\mathcal{V} \in \beta S$ such that $\{S\backslash A_i : i \in I\} \subseteq \mathcal{V}$, so $\mathcal{V} \in U_{S\backslash A_i}$ for all $i \in I$, a contradiction as U_{A_i} cover βS.

Let $\mathcal{U}, \mathcal{V}$ be distinct ultrafilters on S, then there exists at least one $A \subseteq S$ such that $A \in \mathcal{U}$ and $S \backslash A \in \mathcal{V}$. The open sets U_A and $U_{S \backslash A}$ contain $\mathcal{U}$ and $\mathcal{V}$ respectively and are disjoint, therefore βS is Hausdorff. $\square$

Lemma 4. *Let Y be a compact Hausdorff topological space, S be an infinite set and $(y_s)_S$ a sequence in Y indexed by S. For any $\mathcal{U} \in \beta S$ there exists a unique $y \in Y$ such that for any open neighborhood U of y we have $\{s : y_s \in U\} \in \mathcal{U}$.*

Proof. Let Y be a compact Hausdorff topological space, S be an infinite set and $(y_s)_S$ a sequence in Y indexed by S and $U \in \beta S$.

Uniqueness: suppose for a contradiction it is not unique. Then there exists $y, y' \in Y$ distinct, such that for any open neighborhood U of y we have $\{s : y_s \in U\} \in \mathcal{U}$ and for any open neighborhood U' of y' we have $\{s : y_s \in U'\} \in \mathcal{U}$. As $\mathcal{U}$ is a filter the intersection $\{s : y_s \in U \cap U'\} \in \mathcal{U}$, therefore, it is non-empty. But Y is Hausdorff and y, y' distinct, therefore at least a pair U, U' must have an empty intersection, a contradiction.

Existence: suppose again, for a contradiction that it does not exist. Then for any $y \in Y$ there exists an open neighborhood U_y such that $\{s : y_s \in U_y\} \not\in \mathcal{U}$. $\bigcup_{y \in Y} U_y$ is a cover of Y, as Y is compact there is a finite subcover $Y = \bigcup_{i \in J} U_{y_i}$ for a finite index set J, then by **proposition 7** we know that at least one of the $U_{y_i} \in \mathcal{U}$, a contradiction. $\square$

Definition 29. (Ultralimit) We call such unique y the **ultralimit of** $(y_s)_S$ with respect to $\mathcal{U}$: $\lim_{s,\mathcal{U}} y_s$ or just $\lim_{\mathcal{U}} y_s$ if s is the only index.

Definition 30. (Stone-Čech Compactification) Let X be a topological space, its **Stone–Čech compactification** is a compact Hausdorff space βX together with a continuous map $i : X \to \beta X$ that has the following universal property: any continuous map $f : X \to Y$, where Y is a compact Hausdorff space, extends uniquely to a continuous map $\beta f : \beta X \to Y$.

Proposition 25. *βS is the Stone-Čech compactification of S*

Proof. We already showed that βS is a compact Hausdorff topological space, so it only remains to show the existence of the inclusion i and βf.

By **remark 8** we know that the principal filter generated by $\{s\}, s \in S$, is an ultrafilter $\mathcal{U}_s$, and distinct elements will be sent to distinct ultrafilters. Hence, we can embed S in βS by $i(s) = \mathcal{U}_s$.

Let Y be a compact Hausdorff topological space and $f : X \to Y$ continuous. Our claim is that $\beta f(\mathcal{U}) = \lim_{\mathcal{U}} f(s)$ is our desired extension.

By the definition of the ultralimit we know it is unique.

For $s \in S$, $\beta f(\mathcal{U}_s) = \lim_{\mathcal{U}_f} f(s) = f(s)$ as any open set around $f(s)$ is in the principal ultrafilter, therefore βf extends f.

It rests to show that βf is continuous. Let $U \subseteq Y$ be open and let $\mathcal{U} \in \beta f^{-1}(U)$ with $y \in U$ as its ultralimit. Consider the set $A = \{s \in S : f(s) \in W$ for every open $y \in W \subseteq U\}$, $A \in \mathcal{U}$, so $\mathcal{U} \in U_A$. For any $\mathcal{A} \in U_A$ we have that $\beta f(\mathcal{V}) = y \in U$, therefore $\beta f(U_A) \subseteq U$, hence βf is continuous. $\square$

9 Application to Semigroups

In this section we will see an application of the theory of ultralimits and the space of ultrafilters to semigroups.

Definition 31. (Semigroup) A set S together with an associative binary operation "." is called a **semigroup**.

Definition 32. (Right topological semigroup) let S be a semigroup with a topology τ, we say S is a **right topological semigroup** if the function $\rho : S \to S$, $t \mapsto ts$ is continuous for every $s \in S$.

We will now extend the semigroup structure to the set βS of ultrafilters on the semigroup S.

Definition 33. Let $(S, .)$ be a semigroup, for $\mathcal{U}, \mathcal{V} \in \beta S$, for $A \subseteq S$ define $\odot : \beta S^2 \to \beta S$ by

$$A \in \mathcal{U} \odot \mathcal{V} \Leftrightarrow \{s \in S : s^{-1}A \in \mathcal{V}\} \in \mathcal{U}.$$

Where $s^{-1}A = \{t \in S : st \in A\}$ [6]

Proposition 26. *$(\beta S, \odot)$ is a compact, right topological semigroup.*

Proof. We first show that $\odot$ is a binary operation, that is, $\mathcal{U} \odot \mathcal{V}$ is an ultrafilter. If $A \in \mathcal{U} \odot \mathcal{V}$ and $A \subseteq B$ we have that $s^{-1}A \subseteq s^{-1}B$, therefore $\{s \in S : s^{-1}A \in \mathcal{V}\} \subseteq \{s \in S : s^{-1}B \in \mathcal{V}\}$ and so $\{s \in S : s^{-1}B \in \mathcal{V}\} \in \mathcal{U}$ as $\{s \in S : s^{-1}A \in \mathcal{V}\} \in \mathcal{U}$.

If $A, B \in \mathcal{U} \odot \mathcal{V}$, then $\{s \in S : s^{-1}A \in \mathcal{V}\} \in \mathcal{U}$ and $\{s \in S : s^{-1}B \in \mathcal{V}\} \in \mathcal{U}$, therefore

$$\{s \in S : s^{-1}A \in \mathcal{V}\} \cap \{s \in S : s^{-1}B \in \mathcal{V}\} \in \mathcal{U} \Rightarrow$$

$$\Rightarrow \{s \in S : s^{-1}A \in \mathcal{V} \text{ and } s^{-1}B \in \mathcal{V}\} \in \mathcal{U} \Rightarrow \{s \in S : s^{-1}(A \cap B) \in \mathcal{V}\} \in \mathcal{U} \Rightarrow$$

$$\Rightarrow A \cap B \in \mathcal{U} \odot \mathcal{V}.$$

It is easy to see that the empty set is not an element, otherwise the empty set would be in the ultrafilter $\mathcal{U}$, which cannot be.

To check associativity it suffices to show that

$$A \in (\mathcal{U} \odot \mathcal{V}) \odot \mathcal{W} \Leftrightarrow$$

$$\Leftrightarrow \{s \in S : s^{-1}A \in \mathcal{U} \odot \mathcal{V}\} \in \mathcal{W} \Leftrightarrow \{s \in S : s^{-1}A \in \mathcal{U}\} \in (\mathcal{V} \odot \mathcal{W})$$

$$\{s \in S : s^{-1}A \in \mathcal{U}\} \in (\mathcal{V} \odot \mathcal{W}) \Leftrightarrow \{t \in S : t^{-1}\{s \in S : s^{-1}A \in \mathcal{U}\} \in \mathcal{V}\} \in \mathcal{W} \Leftrightarrow$$

$$\Leftrightarrow \{t \in S : t^{-1}s^{-1}A(\mathcal{U} \odot \mathcal{V}\} \in \mathcal{W} \Leftrightarrow \{s \in S : s^{-1}A \in (\mathcal{U} \odot \mathcal{V})\}\} \in \mathcal{W} \Leftrightarrow$$

$$\Leftrightarrow A \in \mathcal{U} \odot (\mathcal{V} \odot \mathcal{W}).$$

We have proved that βS is compact and that $\odot$ is binary associative operation, therefore, it is a compact semigroup. It rests to show that he map $\rho : \beta S \to \beta S$ such that $\mathcal{U} \mapsto \mathcal{U} \odot \mathcal{V}$ for a fixed $\mathcal{V} \in \beta S$, is continuous.

Proving continuity on basic open sets suffices, so fix $A \in S$, we have to show that $\rho^{-1}(U_A)$ is open. Let $B = \{x \in S : x^{-1}A \in \mathcal{V}\}$, claim that $\rho^{-1}(U_A) = U_B$. We can show that $\rho(U_B) = \rho(\{\mathcal{U} : \{x \in S : x^{-1}A \in \mathcal{V}\} \in \mathcal{U}\})$ and this is clearly U_A. $\square$

10 Idempotent Elements and Ellis-Nakamura Lemma

Definition 34. (Idempotent element) Let $(S, .)$ be a semigroup, we say that $s \in S$ is **idempotent** if $s.s = s$.

Lemma 5. *Let X be a compact topological space and let $\mathcal{A}$ be a collection of subsets of X with the finite intersection property, then $\bigcap_{A \in \mathcal{A}} A \neq \emptyset$.*

Proof. Suppose for a contradiction that $\bigcap_{A\in\mathcal{A}} A = \emptyset$, then $X \backslash (\bigcap_{A\in\mathcal{A}} A) = X \Rightarrow \bigcup_{A\in\mathcal{A}} X\backslash A = X$. Therefore $\bigcup_{A\in\mathcal{A}} X\backslash A$ is a cover, so by compactness it has a finite subcover $\bigcup_{A\in\mathcal{A}_0} X\backslash A = X \Rightarrow X \backslash \bigcap_{A\in\mathcal{A}_0} A = X$, a contradiction since that would mean that $\bigcap_{A\in\mathcal{A}_0} A = \emptyset$, contradicting the finite intersection property. □

Theorem 5. *(Ellis-Nakamura lemma) Let $(S, .)$ be a compact right semitopological semigroup, then S has an idempotent element. (We assume it as a fact).[5]*

Corollary 3. *Let S be a semigroup, βS be the semigroup of ultrafilters on S and $T \subseteq \beta S$ be a non-empty closed subsemigroup of βS, then T has an idempotent element.*

Proof. By **proposition 26** βS is a right topological semigroup, hence T will be a compact right topological semigroup. By **theorem 5** we know that there is an idempotent element in T. □

Definition 35. (Idempotent ultrafilters) We call idempotent elements in βS **idempotent ultrafilters**, in other words, $\alpha \in \beta S$ is idempotent if $\alpha \odot \alpha = \alpha$.

Proposition 27. *If S has an idempotent element $\alpha \in S$, then it generates a principal idempotent ultrafilter U_α.*

Proof. Let $\alpha \in S$ be idempotent and let $\mathcal{U}_\alpha$ be the principal ultrafilter generated by it. We want to show that $\mathcal{U}_\alpha \odot \mathcal{U}_\alpha = \mathcal{U}_\alpha$: let $A \in \mathcal{U}_\alpha \odot \mathcal{U}_\alpha$, this is if and only if $\{s : \{s \in S : s^{-1}A \in \mathcal{U}_\alpha\}\} \in \mathcal{U}_\alpha$

$$\Leftrightarrow \alpha \in \{s : s^{-1}A \in \mathcal{U}_\alpha\} \Leftrightarrow \alpha \in \{t : t\alpha \in A\} \Leftrightarrow \alpha\alpha = \alpha \in A \Leftrightarrow A \in \mathcal{U}_\alpha.$$

□

Lemma 6. *If S has no idempotent elements, an idempotent ultrafilter on S cannot be principal.*

Proof. Suppose for a contradiction that $\alpha \in \beta S$ is an idempotent principal ultrafilter, then there must exist $a \in S$ such that $A \in \alpha \Leftrightarrow \{a\} \subseteq A$ by **proposition 8**. Clearly $\{a\} \in \alpha$, hence $\{a\} \in \alpha \odot \alpha$ as α is idempotent. Therefore, $\{s \in S : s^{-1}\{a\} \in \alpha\} \in \alpha$, this means that $\{a\} \subseteq \{s \in S : \{a\}^{-1}\{a\}\}$, which implies that

$$\{a\}a^{-1} \in \alpha \Leftrightarrow \{a\} \subseteq \{t : ta \in \{a\}\} \Leftrightarrow aa = a.$$

But this would mean there is an idempotent element $a \in S$, so α cannot be idempotent. □

Corollary 4. *Idempotent ultrafilters on sets with no idempotent elements do not have finite elements.*

Proof. By **lemma 6** we have that that the idempotent ultrafilter cannot be principal and by **corollary 2** there cannot be a finite set in the ultrafilter. □

11 Hindman's Theorem

In this section we use the theory of ultrafilters and idempotent elements to prove Hindman's theorem.

Definition 36. Let $(c_n : n \in \mathbb{N})$ be a sequence of distinct elements in $A \subseteq \mathbb{N}$ and let $F \subset \mathbb{N}$ be finite. We define $c_F := \sum_{n \in F} c_n$.

Given (c_n) we define $FS((c_n)) := \{c_F : F \subset \mathbb{N} \text{ finite non-empty }\}$.

We say that $A \subseteq \mathbb{N}$ is an FS**-set** if there exists a sequence (c_n) of distinct elements in A such that $FS((c_n)) \subseteq A$. [7]

For $\theta, \gamma \in \beta S$ and $A \subseteq \mathbb{N}$ we will denote $(A - n) = \{m \in \mathbb{N} : m + n \in A\}$ and $A_\gamma = \{n \in \mathbb{N} : (A - n) \in \gamma\}$. With this notation we get

$$\theta \odot \gamma = \{A \subseteq \mathbb{N} : \{n : (A - n) \in \gamma\} \in \theta\} = \{A : A_\gamma \in \theta\}.$$

Proposition 28. *For $\theta, \gamma \in \beta S$ and $A, B \subseteq \mathbb{N}$:*

1. $A_\gamma \cap B_\gamma = (A \cap B)_\gamma$.
2. $A_\gamma \cup B_\gamma = (A \cup B)_\gamma$.
3. $\mathbb{N} \backslash A_\gamma = (\mathbb{N} \backslash A)_\gamma$.

Proof. 1.

$$A_\gamma \cap B_\gamma = \{n \in \mathbb{N} : (A - n) \in \gamma\} \cap \{n \in \mathbb{N} : (B - n) \in \gamma\}$$

$$= \{n \in \mathbb{N} : (A-n) \in \gamma \text{ and } (B-n) \in \gamma\} = \{n \in \mathbb{N} : \{m \in \mathbb{N} : m+n \in A\} \in \gamma\} \text{ and } \{m \in \mathbb{N} : m+n \in B\} \in \gamma\}$$

$$= \{n \in \mathbb{N} : \{m \in \mathbb{N} : m + n \in A \cap B\} \in \gamma\} = \{n \in \mathbb{N} : (A \cap B) - n \in \gamma\} = (A \cap B)_\gamma.$$

2.

$$A_\gamma \cup B_\gamma = \{n \in \mathbb{N} : (A - n) \in \gamma\} \cup \{n \in \mathbb{N} : (B - n) \in \gamma\}$$

$$= \{n \in \mathbb{N} : (A-n) \in \gamma \text{ or } (B-n) \in \gamma\} = \{n \in \mathbb{N} : \{m \in \mathbb{N} : m+n \in A\} \in \gamma\} \text{ or } \{m \in \mathbb{N} : m+n \in B\} \in \gamma\}$$

$$= \{n \in \mathbb{N} : \{m \in \mathbb{N} : m + n \in A \cup B\} \in \gamma\} = \{n \in \mathbb{N} : (A \cup B) - n \in \gamma\} = (A \cup B)_\gamma.$$

3.

$$\mathbb{N} \backslash A_\gamma = \mathbb{N} \backslash \{n \in \mathbb{N} : (A - n) \in \gamma\}$$

$$= \{n \in \mathbb{N} : (A - n) \not\in \gamma\} = \{n \in \mathbb{N} : \mathbb{N} \backslash (A - n) \in \gamma\}$$

$$= \{n \in \mathbb{N} : (\mathbb{N} \backslash A) - n \in \gamma\} = (\mathbb{N} \backslash A)_\gamma.$$

□

Proposition 29. *Let $(\beta(\mathbb{N} \backslash \{0\}), \odot)$ be the space of ultrafilters on $\mathbb{N} \backslash \{0\}$ and let $\alpha \in \beta(\mathbb{N} \backslash \{0\})$ be an idempotent ultrafilter, then every element of α is an FS-set.*

Note 5. In this proposition we do not include 0 in $\mathbb{N}$, as we will use the fact that idempotent ultrafilters over S do not have finite elements if S has no idempotent elements. See **lemma 6**.

Proof. Let $A \in \alpha$, as $\alpha = \alpha \odot \alpha$ we have that $\{s \in \mathbb{N}\backslash\{0\} : A_\alpha \in \alpha\} \in \alpha$, hence $A_\alpha \in \alpha$ and so $A \cap A_\alpha \in \alpha$ (hence non-empty), let $A_\alpha = A_1$.

Claim: $A_1 - n \in \alpha$ for any $n \in A_1$.

We know that $(A-n) = \{m \in \mathbb{N}\backslash\{0\} : m+n \in A\} \in \alpha$ and $A_\alpha - n = \{m \in \mathbb{N}\backslash\{0\} : m+n \in A_\alpha\} \in \alpha$ as $n \in A_1$. From these considerations we conclude that $(A-n) \cap A_\alpha - n = \{m \in \mathbb{N}\backslash\{0\} : m+n \in A\} \cap \{m \in \mathbb{N}\backslash\{0\} : m+n \in A_\alpha\} = \{m \in \mathbb{N}\backslash\{0\} : m+n \in A \cap A_\alpha\} = A_1 - n$, which is in α and $(A-n) \cap A_\alpha - n \in \alpha$, hence $A_1 - n \in \alpha$.

We choose $x_1 \in A_1$ and let $A_2 = A_1 \cap (A_1 - x_1)$, which by our previous claim is in α, so A_2 is not finite by **corollary 4**. In particular, there are infinitely many elements greater than x_1, pick $x_2 \in A_2$ such that $x_2 > x_1$.

Claim: $x_1 + x_2 \in A_1$.

As $x_2 \in A_1 - x_1$, we have that $x_2 + x_1 + n \in A \Rightarrow (x_1 + x_2) + n \in A$, therefore $x_1 + x_2 \in A_1$.

So again by our first claim we conclude that $A_1 - (x_1 + x_2) \in \alpha$. Let $A_3 = A_1 \cap (A_1 - x_1) \cap (A_1 - x_2) \cap (A_1 - (x_1 + x_2))$, clearly $A_3 \in \alpha$, so it has infinite elements and we can pick $x_3 \in A_3$ such that $x_3 > x_2$.

We iterate the above process and construct a sequence $(x_i : i \in \mathbb{N}\backslash\{0\})$ such that $x_1 < x_2 < x_3 < ....$ Note that every $x_i \in A_1 \subseteq A$ and finally that $FS((x_i)) = \{\sum_{I \subset \mathbb{N}} x_i : I$ finite subset of $\mathbb{N}\backslash\{0\}\}$ is a subset of A by construction of the x_i.Therefore A is an FS set. [13]

Theorem 6. *(Hindman's theorem) For any finite partition of* $\mathbb{N}\backslash\{0\}$*, there is at least one partition that is an* FS *set.*

Proof. $(\beta(\mathbb{N}\backslash\{0\}), \odot)$ is a compact Hausdorff semitopological semigroup, so by Ellis theorem it has an idempotent ultrafilter α on $\mathbb{N}\backslash\{0\}$. Let $\mathbb{N}\backslash\{0\}$ be partitioned by $C_1 \sqcup C_2 \sqcup ...C_n$, then $C_1 \sqcup C_2 \sqcup ...C_n \in \alpha$, so either $C_1 \in \mathcal{U}$ or $C_2 \sqcup C_3 \sqcup ...C_n \in \alpha$. In the first case we are done, as C_1 is an FS-set by **proposition 29**. In the second case, we iterate the process and as there are finitely many C_i it will end at some point, concluding that $C_i \in \alpha$ for some $i \in \{2, ..., n\}$. [13] ☐

12 Internal and External Objects

We will examine how nonstandard extensions work when applied to higher order objects such as power sets. We begin by giving some general definitions and results that are the higher order equivalent of **sections 1** and **5**.

Definition 37. (I-set) An ***I*-set** X_I is an indexed family $(X_i : i \in I)$ of sets. A **sort** of the I-set X_I is one of the sets X_i, where $i \in I$. We say that X_I is non-empty if X_i is non-empty for every $i \in I$. [8]

Example 5. *Let $I = \{1, 2\}$ and X be non-empty then $(X, \mathcal{P}(X))$ is a non-empty 2-set and X is a sort.*

Definition 38. (Nonstandard extension of an I-set) It is the same as **definition 1**, but replacing X by X_I and X^m by X^α, a finite sequence in X_I indexed by α. [8]

All of the propositions proved in section 1 and the definitions of nonstandard and standard elements also hold for nonstandard extensions of I-sets with the appropriate replacements mentioned above.

Definition 39. (Set of formulas over X_I) It is the same as **definition 14** , with the difference that variables can range over any element of X_I, in other words they can range over any sort X_i. [9]

Definition 40. (*-transform on I-sets) It is the same as **definition 16** having into account that the variables range over the sorts and the objects involved can be any subset or Cartesian product of subsets of the sorts. We basically substitute each object for its nonstandard extension in atomic formulas and build up inductively. [9]

Theorem 7. *(Transfer theorem for I-sets) It is the same as* ***theorem 3****, but variables ranging over sorts and objects being subsets or Cartesian products of subsets of sorts. [9]*

Proof. The proof is almost identical to the original proof of transfer theorem, so it is omitted. $\square$

Example 6. *Let $I = \{1, 2\}$ consider the 2-set $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$. The formula $\forall x \in [1, 2)\exists y \in \{\{4, 2\}, (3, 4)\}$ is a formula over $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ with *-transform $\forall x \in {}^*[1, 2)\exists y \in {}^*\{\{4, 2\}, (3, 4)\}$.*

We will turn our attention to the particular case we are interested in, unsurprisingly, $(X, \mathcal{P}(X))$.

Definition 41. Let X be a non-empty set, let μ be an ultrafilter on an infinite index set I, ${}^*X = X^I/\sim_\mu$ its nonstandard extension. Then for any sequence $T = (T_i, i \in I)$ where $T_i \in \mathcal{A} \subseteq \mathcal{P}(X)$ there is a well defined set

$$\hat{T} := (\Pi_{i\in I} T_i)/\sim_\mu = \{[x] \in {}^*X : \mu(\{i : x_i \in T_i\}) = 1\}.$$

Definition 42. (Nonstandard Extension of a collection of sets) Let X be a non-empty set, ${}^*X = X^I/\sim_\mu$ its nonstandard extension and $\mathcal{A} \subseteq \mathcal{P}(X)$ a collection of subsets of X, we define

$${}^*\mathcal{A} := \{\hat{T} : T = (T_i, i \in I) \text{ where } T_i \in \mathcal{A}\}.$$

[10]

Note 6. It is worth noticing that in our 2-set setting, we would require of a restriction of membership $E = \{(x, Y) \in X \times \mathcal{P}(X) : x \in A\}$ whose *-transform would have to be preserved in order to satisfy the transfer theorem and avoid fuzzy membership situations [11]. This restriction is encoded in our definition by the condition $\mu(\{i : x_i \in T_i\}) = 1$, which is precisely the membership relation in nonstandard extensions of subsets of X.

Proposition 30. *Let X be a non-empty set, let $\mathcal{A} \subseteq \mathcal{P}(X)$ let $T = (T_i, i \in I)$ where $T_i \in \mathcal{A}$ and let $T' = (T'_i : i \in I)$ where $\mu(\{i : T'_i = T_i\}) = 1$, then $\hat{T} = \hat{T}'$.*

Proof.

$$[\bar{x}] \in \hat{T}' \Leftrightarrow \mu(\{j : x_j \in T'_j\}) = 1 \text{ and } \mu(\{i : T'_i = T_i\}) = 1 \Leftrightarrow$$

$$\Leftrightarrow \mu(\{i : x_i \in T_i \in \mathcal{A}\} = 1 \Leftrightarrow [\bar{x}] \in \hat{T}.$$

□

The previous proposition means that even if two sequences of sets are distinct, if they have a number of elements in common in the ultrafilter, their ultraproduct will be the same.

Proposition 31. *The $*$-map in **definition 42** is a nonstandard extension.*

Proof. Having the previous proposition into account it is a matter of going through the definition of a nonstandard extension. □

Proposition 32. *If $\mathcal{A} = \{Y\}$, where $Y \subseteq X$, then $^*\mathcal{A} = \{^*Y\}$.*

Proof. As there is only of element Y, the only possible sequence is the constant sequence $T : i \mapsto Y$ and $\hat{T} = \{[x] \in {}^*X : \mu(\{x_i \in Y\}) = 1\}$, and this is $^*Y$. □

Like we did with nonstandard extensions of elements of sets where we identified $^*x = {}^*\{x\}$ for $x \in X$, we can identify nonstandard extensions of elements of $\mathcal{P}(X)$ as $^*Y = {}^*\{Y\}$. This is an embedding of $\mathcal{P}(X)$ into $^*\mathcal{P}(X)$.

Corollary 5. *If $\mathcal{A} = \{Y\}$, where $Y \subseteq X$ is finite, then $^*\mathcal{A} = Y$.*

Proof. By **proposition 32** we know that $^*\mathcal{A} = \{^*Y\}$ and by **proposition 10** we know that $^*Y = Y$, therefore $^*Y = \mathcal{A}$. □

Hence, we can identify $^*Y = Y$ for $Y \subseteq$ finite, similar with how $^*x = x$ for $x \in X$.

Definition 43. (Internal Object) We call A an **internal object** if $A \in {}^*\mathcal{P}(X)$. In other words, A is internal if $A = \hat{T}$ for some $T = (T_i : i \in I)$ where $T_i \in \mathcal{P}(X)$. We say that A is **external** if this is not the case.

Internal objects will be the sets involved in the formulas over $(X, \mathcal{P}(X))$, as they are the nonstandard extensions of subsets in X. This characteristic will make them very useful as we can translate properties between these objects and objects in X. External objects are not as well behaved, as they they no preimage in X, so we cannot say too much about them using transfer principle. The convenient nature of internal sets is reflected in the following proposition.

Theorem 8. *Let $\varphi(x, y_1, ..., y_n)$ be a formula over the 2-set $(X, \mathcal{P}(X))$. Let x range over X and the variables y_k range over $\mathcal{P}(X)$. Let $A, A_1, ..., A_n$ be internal subsets of ${}^*X$. Let B be the subset of ${}^*X$ defined by ${}^*\varphi(x, a_1, ..., a_m, A_1, ..., A_n)$:*

$$B = \{x \in A : {}^*\varphi(x, A_1, ..., A_n)\}.$$

Then B is internal.

Proof. We know that A and A_i are internal, so there exists $\mathcal{A}, \mathcal{A}_i \subseteq \mathcal{P}(X)$ such that ${}^*\mathcal{A} = A$ and ${}^*\mathcal{A}_i = A_i$. Let $\mathcal{C} = \bigcup_{C \in \mathcal{A}} \mathcal{P}(C)$, the the following sentence over $\mathcal{P}(X)$ is true:

$$\psi(\mathcal{C}, \mathcal{A}_1, ..., \mathcal{A}_n) = \forall x \in \mathcal{C} \forall y_1 \in \mathcal{A}_1 ... \forall y_n \in \mathcal{A}_n \exists z \in \mathcal{C}(z = \{t \in x : j(t, y_1, ..., y_n\})$$

by transfer theorem $\psi({}^*\mathcal{C}, {}^*\mathcal{A}_1, ..., {}^*\mathcal{A}_n)$ is also true, as $A \in {}^*\mathcal{C}$ and ${}^*\mathrm{A}_i \in \mathcal{A}_i$, we get explicitly get that there exists $B = \{x \in A : {}^*\varphi(x, A_1, ..., A_n)\} \in {}^*\mathcal{C}$

$\square$

13 Saturation and Hyperfinite Sets

In this final section, we will explore saturation and Hyperfinite sets in sets that satisfy convenient properties.

Definition 44. (Countable saturation principle) Let $\mathcal{A} \subseteq \mathcal{P}(X)$ and $\mathcal{B} = \{B_n\}_{n\in\mathbb{N}} \subseteq {}^*\mathcal{A}$ be a collection of internal sets satisfying the finite intersection property, we say that ${}^*X$ satisfies the **countable saturation property** if $\bigcap_{n\in\mathbb{N}} B_n \neq \emptyset$ for any such $\mathcal{B}$. [12]

Countable-saturation is an important property, as it roughly states that ${}^*X$ will realize formulas where objects of size at most countable are involved. This is consistent with the way formulas over sets and the transfer principle were defined and how they worked. We saw ${}^*R$ and how we could prove, using transfer theorem, properties in $\mathbb{N}$, this is because ${}^*R$ as we built it is countably-saturated. Furthermore, every proper nonstandard extension built via ultrapowers will be countably-saturated!

Proposition 33. *Let X be a non-empty set and let ${}^*X = X^{\mathbb{N}}/\sim_\mu$ be a proper nonstandard extension, then ${}^*X$ has the countable saturation property.*

Proof. Let $\mathbb{N} = \{B_n\}_{n\in\mathbb{N}}$ be a collection of internal sets of ${}^*X$ with the finite intersection property. As each B_n is internal, $B_n = \hat{T}_n = \{[x] \in \Pi_{i\in\mathbb{N}} T_n(i) : \mu(\{i : x(i) \in T_n(i)\}) = 1\}/\sim_\mu$ for some sequence $T_n = \{T_n(i)\}_{i\in\mathbb{N}}$. We will construct and element τ that is in $\bigcap_{n\in\mathbb{N}} B_n$:

For each $i \in \mathbb{N}$ find $max\{j \in \mathbb{N} : j \leq n$ and $T_1(i) \cap ... \cap T_j(i) \neq \emptyset\}$, we then choose $\tau(i) \in T_1(i) \cap ... \cap T_j(i)$.

Note that if $i \geq k$ and $T_1(i) \cap ... \cap T_k(i) \neq \emptyset$, then $\tau(i) \in T_1(i) \cap ... \cap T_k(i)$.

Fix $k \in \mathbb{N}$, by the finite intersection property $\hat{T}_1 \cap ... \cap \hat{T}_k \neq \emptyset$, so there exists $[\alpha] \in {}^*X$ such that $\mu(\Lambda_j = \{i : \alpha(i) \in T_j(i)\}) = 1$ for all $j \leq k$, in fact, $\mu(\{i : T_1(i) \cap ... \cap T_k(i) \neq \emptyset\}) = 1$, as $\Lambda_1 \cap ... \cap \Lambda_k \subseteq \{i : T_1(i) \cap ... \cap_k (i) \neq \emptyset\} = \Gamma$. Now $\Gamma = \{i \in \Gamma : i < k\} \cup \{i \in \Gamma : i \geq k\}$ and as $\{i \in \Gamma : i < k\}$ is finite we have that $\mu(\{i \in \Gamma : i \geq k\}) = 1$. Finally, $\{i \in \Gamma : i \geq k\} \subseteq \{i : \tau(i) \in T_1(i) \cap ... \cap T_k(i)\}$, therefore $\mu(\{i : \tau(i) \in T_1(i) \cap ... \cap T_k(i)\}) = 1$ and so $[\tau] \in \hat{T}_1 \cap ... \cap \hat{T}_k$ for any k, so we are done. □

Corollary 6. ${}^*\mathbb{R}$ *has the countable saturation property.*

Definition 45. (Hyperfinite set) We say a set is **hyperfinite** if it is an element of ${}^*\mathcal{A}$ for a collection of finite sets $\mathcal{A}$.

Proposition 34. *Every hyperfinite set H is internal.*

Proof. Let X be a non-empty set, let $\mathcal{A} \subseteq \mathcal{P}(X)$ be a collection of finite sets and let $H \in {}^*\mathcal{A}$. It is clear that H is of the form $\hat{T}$ for some $(T = (T_i, i \in I)$ where $T_i \in \mathcal{A} \subseteq \mathcal{P}(X)$, so it is internal. □

Proposition 35. *Let X be a non-empty set and let $H \subseteq {}^*X$ be hyperfinite, then $H \in {}^*Fin(X)$, where $Fin(X) = \{$ finite subsets of $X\}$.*

Proof. If H is hyperfinite, then $H \in {}^*\mathcal{A}$ for a collection of finite subsets of X. So we have that $H \in {}^*\mathcal{A} \cap {}^*\mathcal{P}(X) = {}^*(\mathcal{A} \cap \mathcal{P}(X)) \subseteq {}^*Fin(X)$, as $\mathcal{A}$ is a collections of finite sets. □

So hyperfinite subsets of ${}^*X$ are internal sets which come from the nonstandard extension of the family of finite sets of X. It is nice that by being internal we will be able to apply transfer theorem to sentences where they are involved. What makes them even more interesting is that by being an element of $Fin(A)$, we will be able to pass properties of finite sets to hyperfinite ones and vice versa. We will see this in action in the following results.

Definition 46. (κ-enlarging property) Let $\mathcal{F} \subseteq \mathcal{P}(X)$ satisfying the finite intersection property and $|\mathcal{F}| < \kappa$, we say that ${}^*X$ satisfies the **κ-enlarging property** if $\bigcap_{F \in \mathcal{F}} {}^*F \neq \emptyset$ for any such $\mathcal{F}$. [12]

Proposition 36. *Let ${}^*X$ have the κ-enlarging property, then for each $A \subseteq X$ with $|A| < \kappa$ there exists a hyperfinite set such that $H \subseteq {}^*A$ and $A \subseteq H$. [12]*

Proof. For each $a \in A$ define the set $F_a = \{B \subseteq A : a \in B \text{ finite }\}$ and let $\mathcal{F} = \{F_a : a \in A\}$. Since the $|Fin(X)| = |X|$, there exists a bijection $\gamma : Fin(X) \to X$, we define

$$\gamma' : \mathcal{F} \to \mathcal{P}(X)$$

$$\gamma'(F_a) = \{\gamma(B) : B \in F_a\} := \Gamma_{F_a}.$$

Consider the set $\Gamma = \{\Gamma_{F_a} : a \in A\}$, this set has the same size as A, hence less than κ and has the finite intersection property, so we can apply the κ-enlarging property and conclude that $\bigcap_{a \in A} {}^*\Gamma_{F_a} \neq \emptyset$, call this set α. We can take the ${}^*\gamma$ inverse of the elements of α and get the set $\{B \subseteq {}^*A : {}^*\gamma(B) \in \alpha \text{ and } B \in {}^*F_a \text{ for all } a \in A\}$ pick an element of this set and call it H.

Claim: H is hypefinite with $H \subseteq {}^*A$ and $A \subseteq H$.

H is hyperfinite as $H \in {}^*F_a$ for all $a \in A$ with F_a being a collection of finite sets. It is clear that $H \subseteq {}^*A$ and because for each $a \in A$ we have that $H \in F_a$, we can conclude that $A \subseteq H$. $\square$

In the following propositions we will assume the κ-enlarging property always holds, where κ is greater than the cardinality of the sets involved.

Corollary 7. *Every infinite set can be linearly ordered.*

Proof. Let X be an infinite set, by **proposition 36** we find a hyperfinite set $H \subseteq {}^*X$ and $X = \{{}^*x : x \in X\} \subseteq H$. It is known that any finite set can be linearly ordered and we can express this sentence as a formula over $(X, \mathcal{P}(X))$

$$\forall x \in Fin(X) \exists y \in x^2 (\text{ "x is linearly ordered" }).$$

By transfer, any set in ${}^*Fin(X)$ can be linearly ordered, therefore by **proposition 35** H can be linearly ordered, and so X inherits this ordering. $\square$

Proposition 37. *Every partial order on a set can be extended to a linear order.*

Proof. We prove the finite case by induction:

$n = 2$: Let $(X = \{x_1\}, \leq)$ be a partially ordered set, we define $\leq'$ by:

If $x_1 \leq x_2$ or x_{21} then either $x_1 \leq' x_2$ or $x_2 \leq' x_1$ respectively. Otherwise let $x_1 \leq' x_2$ and $x_1 \leq' x_1$, $x_2 \leq' x_2$. This is clearly a linear order.

Assume every partially ordered set of size $d-1$, for some $d \in \mathbb{N}$, can be extended to a linear order.Let $(X=\{x_1,...,x_d\},\leq)$ be a partially ordered set. As X is finite we can find its minimal element, say x. Once we have identified this minimal element we can define $\Sigma=\{x_i \in X : x \leq x_i\}$ and. We know by our induction hypothesis that $(X\backslash\{x\},\leq)$ can be extended to a linearly ordered set $(X\backslash\{x\}f \leq')$. We define the relation $\leq''$ on the set X by:

If $x_i \leq' x_j$ then $x_i \leq'' x_j$ for all $x_i, x_j \in X\{x\}$.

If $x_i \in \Sigma$ then $x \leq'' x_i$.

If $x_i \not\in \Sigma$ then $x \leq'' x_i$.

It is easy to check that this is a linear order.

Let $(X,\leq)$ be a partially ordered infinite set, by **proposition 36** we can find a hyperfinite $H \subseteq {}^*X$ and $X=\{{}^*x : x \in X\} \subseteq H$. Similarly to the previous proposition, we can express the sentence "every partial order on a finite set can be extended to a linear order" as a formula over $(X,\mathcal{P}(X))$, apply transfer and conclude that every partial ordering in H can be extended to a linear one that X will inherit. □

Theorem 9. *Let G be a graph, then G is k-colorable if and only if every subset of G is k-colorable.*

Proof. Being k-colorable means that we can assign a colour to each vertex such that vertices joint by an edge have different colours using at most k colours. As usual, after identifying the vertices of G with natural numbers, we can express k-colorability as a formula

$$\bigwedge_{i\in G}((\bigvee_{1\leq j\leq k} C_i^j)\bigwedge_{1\leq j\neq l\leq k}\neg(C_i^j \wedge C_i^l)))\wedge(\bigwedge_{(i,j)\in E}(\bigwedge_{1\leq l\leq k}\neg(C_i^l \wedge C_j^l))).$$

Where C_i^j are the constants that represent the coloring of the ith vertex by the jth colour. So we can also write the sentence "every finite subgraph of G is a k-colorable" as a $(G,\mathcal{P}(G))$ sentence. We find a hyperfinite $G \subseteq H \subseteq {}^*G$ by **proposition 36**, we apply transfer to "every finite subgraph of G is a k-colorable" and conclude that H is k-colorable, with G inheriting this coloring. The converse is clear. □

References

[1] C. Ward Henson *Foundations of Nonstandard Analysis*. pp.4,5. Available at https://pdfs.semanticscholar.org/86fe/d3b5ebc6ad2b593af24a6a196355d3b19be6.pdf

[2] C. Ward Henson *Foundations of Nonstandard Analysis*. pp.17,18. Available at https://pdfs.semanticscholar.org/86fe/d3b5ebc6ad2b593af24a6a196355d3b19be6.pdf

[3] C. Ward Henson *Foundations of Nonstandard Analysis*. pp.18-19. Available at https://pdfs.semanticscholar.org/86fe/d3b5ebc6ad2b593af24a6a196355d3b19be6.pdf

[4] C. Ward Henson *Foundations of Nonstandard Analysis*. pp.20-21. Available at https://pdfs.semanticscholar.org/86fe/d3b5ebc6ad2b593af24a6a196355d3b19be6.pdf

[5] Robert Ellis *Distal transformation groups*.pp. 401–405. Pacific J. Math. 8 (1958) Mauro Di Nasso and Isaac Goldbring and Martino Lupini

[6] Mauro Di Nasso, Isaac Goldbring, Martino Lupini *Nonstandard Methods in Ramsey Theory and Combinatorial Number Theory* (2017). p.3. Available at https://arxiv.org/abs/1709.04076

[7] Mauro Di Nasso, Isaac Goldbring, Martino Lupini *Nonstandard Methods in Ramsey Theory and Combinatorial Number Theory* (2017).p.50. Available at https://arxiv.org/abs/1709.04076

[8] C. Ward Henson *Foundations of Nonstandard Analysis*. p.23. Available at https://pdfs.semanticscholar.org/86fe/d3b5ebc6ad2b593af24a6a196355d3b19be6.pdf

[9] C. Ward Henson *Foundations of Nonstandard Analysis*. p.27,28. Available at https://pdfs.semanticscholar.org/86fe/d3b5ebc6ad2b593af24a6a196355d3b19be6.pdf

[10] Mauro Di Nasso, Isaac Goldbring, Martino Lupini *Nonstandard Methods in Ramsey Theory and Combinatorial Number Theory* (2017). pp.16,17. Available at https://arxiv.org/abs/1709.04076

[11] C. Ward Henson *Foundations of Nonstandard Analysis*. p.29,30. Available at https://pdfs.semanticscholar.org/86fe/d3b5ebc6ad2b593af24a6a196355d3b19be6.pdf

[12] Mauro Di Nasso, Isaac Goldbring, Martino Lupini *Nonstandard Methods in Ramsey Theory and Combinatorial Number Theory* (2017). pp.20,21. Available at https://arxiv.org/abs/1709.04076

[13] Neil Hindman *Algebra in the Stone-Čech Compactification and its Applications to Ramsey Theory*, Scientiae Mathematicae Japonicae 62 (2005). pp.321-329.

[14] Chen Chung Chang, H. Jerome Keisler *Model Theory*, Courier Corporation (2012).