Foundations of the Transfer Principle in Nonstandard Analysis

Jason Zhang

July 2026

Abstract

The purpose of this paper is to introduce the transfer principle and its model theoretic consequences in nonstandard analysis. Starting from fundamental ideas about nonstandard analysis, we analyze the properties behind the theory of the hyperreals in order to build justification for the transfer principle. Eventually, we develop several perspectives on the transfer principle and provide explicit constructions that satisfy the transfer principle. Notably, we develop the theory of ultrapowers and show how it serves as a model for the hyperreals. We also prove the famous Łoś Theorem, which is the mathematical certification that ultrapower models satisfy the transfer principle.

1 Introduction

Nonstandard Analysis is fundamentally about a reformulation of mathematics itself to accommodate infinitesimal elements in a way that does not violate any natural intuition concerning how numbers work. Founded largely by Abraham Robinson [Rob96] in the early 1960s, nonstandard analysis allows for a calculus with infinitesimal elements in a logically rigorous way. These infinitesimals (and also infinite elements) are collected in a set known as the "hyperreals", denoted $^*\mathbb{R}$. Partially inspired by Leibniz's "law of continuity", which states

"whatever succeeds for the finite, succeeds for the infinite,"

the cornerstone of nonstandard analysis is the *transfer principle*, which allows for statements true in the hyperreals to also be true for the reals (see [KK11] for more historical details). Thus, nonstandard analysis finds itself applications in unexpected places, such as Ramsey theory (as seen in [NGL18]), where mathematicians work with infinitesimals and use the transfer principle to justify that whatever work they have done is also true in the regular reals.

2 Theory of Hyperreals

To begin, we must understand what the hyperreals are from an intuitive perspective. As described previously informally, the hyperreals $^*\mathbb{R}$ are the reals with infinitesimal and infinite elements. We make these notions exact with the following definition.

Definition 2.1 (Nasso et al., [NGL18])**.** We call ϵ *infinitesimal* if $|\epsilon| < \frac{1}{n}$ for all $n \in \mathbb{N}$.

Definition 2.2 (Nasso et al., [NGL18]). We call δ an *infinite* element if $|\delta| > n$ for all $n \in \mathbb{N}$.

It should be noted that $\mathbb{R}$ actually admits an infinitesimal element, that being 0. This is true because $\frac{1}{n} > 0$ for any $n > 0$. However, we will be studying a structure with nontrivial infinitesimals and infinite elements, which $\mathbb{R}$ certainly does not have.

Proposition 2.2.1. *The set of infinitesimals is closed under addition and multiplication.*

Proof. Let a, b be infinitesimals.

(i) We know that $|a| < \frac{1}{n}$ and $|b| < \frac{1}{n}$ for any $n \in \mathbb{N}$. We can further push this to $|a| < \frac{1}{2n}$ and $|b| < \frac{1}{2n}$ since $2n \in \mathbb{N}$. By the triangle inequality, $|a + b| \le |a| + |b| < \frac{1}{2n} + \frac{1}{2n} = \frac{1}{n}$ as desired.

(ii) Again, we know that $|a| < \frac{1}{n}$ and $|b| < \frac{1}{n}$ for any $n \in \mathbb{N}$. Because $\sqrt{n} \le n$, we can say that $\frac{1}{\sqrt{n}} \ge \frac{1}{n}$. Therefore, $|a| < \frac{1}{\sqrt{n}}$ and $|b| < \frac{1}{\sqrt{n}}$ for any $n \in \mathbb{N}$. It follows that $|ab| = |a| \cdot |b| < \frac{1}{\sqrt{n}} \cdot \frac{1}{\sqrt{n}} = \frac{1}{n}$ as desired. $\square$

Now, we say that $\mathbb{R}$ is *Archimedean* because for every positive number $x \in \mathbb{R}^+$, there exists a natural number $n \in \mathbb{N}$ such that $nx > 1$.

Proposition 2.2.2. ${}^*\mathbb{R}$ *is not Archimedean.*

Proof. Assume for the sake of contradiction that ${}^*\mathbb{R}$ is Archimedean, and so for every positive $x \in {}^*\mathbb{R}^+$, there exists a natural number $n \in \mathbb{N}$ such that $nx > 1$. We know that there exists some infinitesimal positive $\epsilon \in {}^*\mathbb{R}^+$. Because we assumed ${}^*\mathbb{R}$ to be Archimedean, this means there exists some $n \in \mathbb{N}$ where $n \cdot \epsilon > 1$. However, this implies that $\epsilon = |\epsilon| > \frac{1}{n}$ for some $n \in \mathbb{N}$, contradicting the fact that $|\epsilon| < \frac{1}{n}$ for all $n \in \mathbb{N}$. $\square$

Every finite hyperreal ξ can be rewritten as $\xi = r + \epsilon$ where $r = \operatorname{st} \xi \in \mathbb{R}$ and $\xi - r$ is an infinitesimal. This is called the "standard part" of a hyperreal number, and it is similar to how $\operatorname{Re}(a + bi) = a$.

Proposition 2.2.3. *The function* $\operatorname{st} : {}^*\mathbb{R} \to \mathbb{R}$ *is a field homomorphism.*

Proof. As a reminder, this means that $\operatorname{st}(a + b) = \operatorname{st} a + \operatorname{st} b$, $\operatorname{st}(ab) = (\operatorname{st} a)(\operatorname{st} b)$, and $\operatorname{st} 1_{{}^*\mathbb{R}} = 1_{\mathbb{R}}$. Let $a = r_1 + \epsilon_1$ and $b = r_2 + \epsilon_2$. It should be noted that $\operatorname{st} a = r_1$ and $\operatorname{st} b = r_2$. We will go through every case:

(i) We can compute $a + b = (r_1 + r_2) + (\epsilon_1 + \epsilon_2)$. By Proposition 2.2.1, we know that the infinitesimals are closed under addition. Thus, $\epsilon = \epsilon_1 + \epsilon_2$ is an infinitesimal. It is then true that $\operatorname{st}(a + b) = r_1 + r_2$ as desired.

(ii) We can compute $ab = r_1 r_2 + r_1 \epsilon_2 + r_2 \epsilon_1 + \epsilon_1 \epsilon_2$. Note that for $r \in \mathbb{R}$ and infinitesimal ϵ, it follows that $|r||\epsilon| \le \lceil |r| \rceil \cdot |\epsilon|$. The right hand side can be interpreted as a sum of infinitesimals, so Proposition 2.2.1 applies again. Furthermore, we know that $\epsilon_1 \epsilon_2$ should also be infinitesimals because of Proposition 2.2.1. Therefore, $\operatorname{st}(ab) = r_1 r_2$.

(iii) Since $1 \in \mathbb{R}$ already, it is obvious that $\operatorname{st} 1 = 1$.

Having considered every case, we are done. $\square$

We can extract the hypernaturals, hyperintegers, and hyperrationals out of the hyperreals as one would extract the naturals, integers, and rationals out of the reals. For example, the hypernaturals ${}^*\mathbb{N}$ consists of the nonnegative integers in ${}^*\mathbb{R}$. The definition of "integers" (that is, an ordered subring satisfying certain properties) is not particularly useful to us here, but it is completely possible to define formally. The structure of ${}^*\mathbb{N}$ appears as follows:

$$ {}^*\mathbb{N} = \{ \underbrace{0, 1, 2, \dots}_{\text{finite elements } (\mathbb{N})}, \underbrace{\dots, N-1, N, N+1, \dots}_{\text{infinite elements}} \}. $$

While it may appear that ${}^*\mathbb{R}$ and $\mathbb{R}$ are vastly different structures, it turns out that they share many similarities. For example, while ${}^*\mathbb{R}$ has infinite elements, it is still true that for all $x \in {}^*\mathbb{R}$, $x < x + 1$. We do not typically expect of this from infinity; we would usually assume $\infty = \infty + 1$. However, there is no such $x \in {}^*\mathbb{R}$ such that $x = x + 1$. This particular property, known as the *transfer principle* is immensely useful in developing the hyperreals: both in defining further tools such as integrals or measures, and in applications to fields such as combinatorics. We will explore the transfer principle more in depth later.

3 Model Theory

We need to explore briefly concepts from model theory as nonstandard analysis is founded on model theoretic principles. In fact, the "nonstandard" comes from the idea of "nonstandard **models**".

3.1 Mathematical Logic Review

Mathematical logic fundamentally grounds model theory and set theory. These fields all share an intimate connection so it is important to establish mathematical logical principles.

Definition 3.1 (Jech, [Jec06]). An elementary formula or statement $\phi(A_1, \dots, A_n)$ is comprised of atomic formulas:

$$ x = y \quad \text{or} \quad x \in y, $$

and built up by connectives

$$ \varphi \wedge \psi,\ \varphi \vee \psi,\ \neg\varphi,\ \varphi \to \psi,\ \varphi \leftrightarrow \psi, $$

(logical and, logical or, negation, implication, bidirectional implication) as well as quantifiers

$$ \forall x\ \varphi, \text{and } \exists x\ \varphi $$

(where φ is some other formula).

This makes our notion of formula precise, which washes away any possibility of making an inherently contradictory formula (like "This sentence is false."). In practice, we will use a lot more symbols, constants, etc. but all of those can be built and re-expressed into an elementary formula.

Definition 3.2 (Kleene, [Kle02]). A logical statement is roughly called *first order* if and only if it quantifies over elements of sets. For example: $\forall x \in X\ (\dots)$

A non-example would be $\forall x \subseteq X \ (\dots)$. One workaround we will use is with that of the powerset of X, notated $\mathcal{P}(X)$, which is the set of all subsets of X. In which case we can quantify $\forall x \in \mathcal{P}(X) \ (\dots)$. This may seem ridiculous as it seems we are not doing anything different, but there is a subtle difference that we will actually explore later on.

We are inclined to explore specific theories T which are composed of axioms. An example of which is ZFC, which is the go-to theory for set theory.

Definition 3.3. A structure M is a model of a theory T if for every true statement ϕ in T, the corresponding statement ϕ^M is true in M (notated $M \vDash \phi^M$).

We will expand on this definition later.

3.2 Filters and Ultrafilters

Definition 3.4. A *filter* F on S is a subset of the powerset $\mathcal{P}(S)$ that satisfies:

(i) $\varnothing \notin F$ and $S \in F$;

(ii) if A, B in F then $A \cap B \in F$;

(iii) if $A \in F$ and $A \subseteq B \subseteq S$ then $B \in F$.

Intuitively, F is meant to capture the "large" subsets of S.

Definition 3.5. An *ultrafilter* U on S is a filter on S such that for all subsets $A \subseteq S$, either $A \in U$ or $S \setminus A \in U$.

An ultrafilter is called *principal* if it is generated by a single element. Specifically, for an ultrafilter U on S, if $U = \{A \subseteq S : a \in A\}$ for some fixed element a, then U is a principal ultrafilter. These will not be useful in our case, we desire nonprincipal ultrafilters. The existence of a nonprinicpal ultrafilter requires the axiom of choice. To clarify, the axiom of choice informs us that every family of nonempty sets has a choice function where one can select an element out of each element of the family.

Proposition 3.5.1. *Let S be an infinite set. The set $F = \{X \subseteq S : S \setminus X$ is finite$\}$ is a filter, known as the Fréchet filter.*

Proof. We will verify all the conditions of being a filter. Note that the Fréchet filter contains all of the *cofinite* sets by definition.

1. Because S is infinite, $S \setminus \varnothing = S$ is infinite so it is not in the filter, implying $\varnothing \notin F$ as desired. Furthermore, $S \setminus S = \varnothing$ is finite, so $S \in F$.

2. Let A and B are in F. Because we are operating on the powerset algebra of S (not entirely important to this discussion, but this behaves exactly like a boolean algebra), De' Morgan's laws hold. Let A^c be the complement of a subset A in S. Therefore, $(A \cap B)^c = A^c \cup B^c$. Both A^c and B^c are finite, so their union is and thus $A \cap B$ is cofinite and in F.

3. Let $A \in F$ and $A \subseteq B \subseteq S$. Because $A \subseteq B$, it follows that $S \setminus B \subseteq S \setminus A$ (this is true by the fact $S \setminus B$ is the same as S removing A, and then removing the leftover elements $B \setminus A$). Therefore, $S \setminus B$ is also finite as desired.

Having considered all conditions, we are done. $\square$

Proposition 3.5.2. *For every filter F on S, every finite subset $H = \{X_1, \ldots, X_n\} \subseteq F$ has a nonempty intersection $X_1 \cap \cdots \cap X_n \neq \varnothing$. This is known as the finite intersection property.*

Proof. Assume for the sake of contradiction that such an $H = \{X_1, \ldots, X_n\} \subseteq F$ exists where $X_1 \cap \cdots \cap X_n = \varnothing$. Recall that F is closed under intersections, so $\bigcap H = \varnothing \in F$, contradiction. $\square$

Lemma 3.6. *Every nonprincipal ultrafilter contains the corresponding Frèchet filter.*

Proof. The following proof is adapted from [Sco]. Let our ultrafilter be U operating on an infinite set S. For any $x \in S$, it cannot be the case that $\{x\} \in U$ or else our ultrafilter would be principal (since a principal ultrafilter is precisely the filter generated by singleton sets). Thus, $S \setminus \{x\} \in U$ for all $x \in S$. Now, notice that for any finite subset $H \subsetneq S$,

$$S \setminus H = \bigcap_{x \in H} (S \setminus \{x\}).$$

Every filter is closed under finite intersections, so $S \setminus H \in U$. Notice that $S \setminus H$ is cofinite for any finite H, and thus U has all the cofinite subsets of S, meaning it is a Frèchet filter. $\square$

3.3 The Incompleteness Phenomena

It turns out, if a theory is consistent then it has a model. However, most theories (like ZFC) have multiple non-isomorphic models. This is a result known as the *incompleteness phenomena*, and is the subject of the Incompleteness Theorems due to Gödel [Göd31].

One analogy to understand this is with vector spaces. Consider the vector spaces $\mathbb{R}^2$ and $\mathbb{R}^3$ (over scalar field $\mathbb{R}$). These both certainly satisfy the vector space axioms and common sentences like "$\vec{v} + \vec{u} = \vec{v} \implies \vec{u} = \vec{0}$". However, these two vector spaces are certainly not isomorphic. The vector space $\mathbb{R}^2$ can have a basis with two vectors, but this is not possible in $\mathbb{R}^3$.

Remark. Another analogy is with that of Euclidean Geometry as defined in Euclid's Elements [Euc56]. Let $\mathbf{EG}^-$ be Euclid's Geometry without the parallel postulate. Both spherical and hyperbolic geometries would be models of $\mathbf{EG}^-$, as would the classical flat Euclidean Geometry. If we were to consider regular Euclidean Geometry as "the standard correct model", then hyperbolic and spherical geometries could be seen as nonstandard models and statements like "every triangle has angles summing up to 180°" would be "true but unprovable".

We claim that ${}^*\mathbb{R}$ is a nonstandard model of $\mathbb{R}$. It satisfies a specific set of axioms that define $\mathbb{R}$. We are interested in a specific type of nonstandard model, that which obeys the *transfer principle*.

4 Star Maps and Transfer Principle

As hinted earlier in the paper, we would like to study the hyperreals because it obeys the *transfer principle*. This principle informally says for the hyperreals that "statements are true in the hyperreals if and only if they are true in the reals". However, we aim to find a more general version.

Definition 4.1. The *star map* $*$ is a function that sends "mathematical objects" (usually sets) to their nonstandard extension in a way that satisfies the *transfer principle*.

In other words, the star maps $X \mapsto {}^*X$. Notice that the star map's definition has the *transfer principle* built into it. In this way, the transfer principle can be seen as a property of the star map itself. The transfer principle shows up as a consequence or a theorem when we try to prove a specific mapping $*'$ actually works as a star map.

Definition 4.2 (Transfer Principle)**.** Let $\phi(A_1, \ldots, A_n)$ be a *first order bounded* elementary statement about the mathematical objects $A_1, \ldots, A_n$. Then $\phi(A_1, \ldots, A_n)$ is true if and only if $\phi({}^*A_1, \ldots, {}^*A_n)$ is true.

This intuitively means that statements true in the extension given by the star map are true in the original structure (and vice-versa). The value in this would be a mathematical proof strategy as follows: Suppose we are asked to prove (or disprove) conjecture P which is about set X. It might be difficult to do so. However, we can employ the following proof strategy:

(i) Formulate P as a statement-type that'll work by the transfer principle (first order elementary).

(ii) Analyze and prove P in the nonstandard extension ${}^*X$. We are free to take advantage of the nonstandard elements and behavior that ${}^*X$ might have (for example, infinitesimal elements in ${}^*\mathbb{R}$).

(iii) Because P is true in ${}^*X$, it must also be true in X which completes the proof.

Remark. It is important to note the "first order bounded" condition on the transfer principle. Because we demand ϕ be first order, a formula with $\forall x \subseteq X$ is not allowed. Although, $\forall x \in \mathcal{P}(X)$ is fine, this occurs because ${}^*\mathcal{P}(X) \neq \mathcal{P}({}^*X)$ in general (this is similarly true for sets of functions $A \to B$).

Note that the star map definition actually leaves a lot of room for interpretation: we will primarily consider star maps $* : V \to {}^*V$ where ${}^*V$ is a nonstandard model of V that satisfies transfer principle. But do note that there is no specific requirement that forces us to think of star maps in this way, and we will later revisit this to develop more interesting or general star maps.

5 Ultrapower Model of the Hyperreals

We are now equipped to formally define a star map $*$ that will give us a concrete definition of the hyperreals.

Definition 5.1 (Nasso et al. [NGL18])**.** Fix some infinite index set I and an ultrafilter $\mathcal{U}$ on I. The *ultrapower* of $\mathbb{R}$ modulo the ultrafilter $\mathcal{U}$ is the quotient of real I-sequences $\mathbb{R}^I$ modulo the equivalence relation $\equiv_{\mathcal{U}}$ which is defined by

$$\sigma \equiv_{\mathcal{U}} \tau \iff \{i \in I : \sigma(i) = \tau(i)\} \in \mathcal{U}.$$

We can succinctly summarize this definition with the equation

$${}^*\mathbb{R} = \mathbb{R}^I/\mathcal{U}.$$

More generally, fixing an ultrafilter $\mathcal{U}$ on some index set I, the star map that arises will be defined as ${}^*X = X^I/\mathcal{U}$. We will prove that this mapping $*$ is actually a star map (i.e., satisfies the transfer principle) in a later section.

It should be noted that for $\sigma, \tau \in {}^*\mathbb{R}$,

$$\sigma < \tau \iff \{i \in I : \sigma(i) < \tau(i)\} \in \mathcal{U}.$$

Analogous definitions follow for $>, \leq, \geq$. In measure theoretic terms, this can be thought of as "$\sigma < \tau$ if and only if $\sigma(i) < \tau(i)$ *almost everywhere.*" This is informal, but develops the correct intuition.

To motivate the fact that our construction $\mathbb{R}^I/\mathcal{U}$ really does behave like the intuitive definition of ${}^*\mathbb{R}$ defined in section 2, we will consider an example. Let $I = \mathbb{N}$ and $\mathcal{U}$ be a *nonprincipal* ultrafilter over I. Now, because $\mathcal{U}$ is nonprincipal, we cannot get an exact description of how it works (it arises nonconstructively by the axiom of choice). We note that if $\sigma \equiv_{\mathcal{U}} \tau$, then they belong to an equivalence class $[\sigma] = \{\tau \in {}^*\mathbb{R} : \sigma \equiv_{\mathcal{U}} \tau\}$.

The intuition is that each element of ${}^*\mathbb{R}$ is a sequence of reals. For example, $\sigma = \langle 1, 1, 1, \dots \rangle$ would be a sequence that would describe 1 in ${}^*\mathbb{R}$. This is described in function notation with each sequence $\sigma : I \to \mathbb{R}$. The equivalence relation $\sigma \equiv_{\mathcal{U}} \tau$ occurs if "the sequences σ and τ are very similar". The purpose of the ultrafilter is to capture *large* sets, and thus roughly two sequences are equal by $\mathcal{U}$ if they are equal in a large number (more precisely, infinite) of positions. Similarly, $\sigma < \tau$ if σ is less than τ in a large number of positions. This informal notion is made formal by the ultrafilter, which captures this notion nicely. We can define addition and multiplication with $[\sigma] + [\tau] = [\sigma + \tau]$ and $[\sigma] \cdot [\tau] = [\sigma \cdot \tau]$ where $\sigma + \tau$ and $\sigma \cdot \tau$ is the result of component-wise addition and multiplication of the sequences. We can now justify that infinitesimal and infinite elements do indeed exist.

Proposition 5.1.1. *Let $N = \langle 1, 2, 3, \dots \rangle \in {}^*\mathbb{R}$. Then N is an infinite element.*

Proof. Regardless of the ultrafilter $\mathcal{U}$ selected (as long as it is nonprincipal), for each $r \in \mathbb{R}$ we have $\langle r, r, r, \dots \rangle \in [r]$. In other words, the sequence $\langle r, r, r, \dots \rangle$ will always belong to the equivalence class $[r]$. Now, recall that a number is infinite if and only if it is greater than every natural $n \in \mathbb{N}$. Notice that for any $n \in \mathbb{N}$, there are an infinite number of positions $i \in I$ in N such that $N(i) > n$. More specifically, the positions where $N(i) \leq n$ forms a finite set, so the positions where $N(i) > n$ forms a cofinite set. By Lemma 3.6, we know that all cofinite sets are in our ultrafilter $\mathcal{U}$. It follows that $N > n$ for all $n \in \mathbb{N}$, concluding that N is indeed an infinite element. $\square$

A similar proof can follow for an infinitesimal element which would approach 0 (for example, $\langle 1, \frac{1}{2}, \frac{1}{4}, \dots \rangle$). Despite being infinite, N would still behave like a finite element. Certainly, $N < N + 1 = \langle 2, 3, 4, \dots \rangle$ because the positions $i \in I$ where $N(i) < (N+1)(i)$ is the same as I, which must be in our ultrafilter.

Lemma 5.2. *The structure $({}^*\mathbb{R}, +, \cdot, <, \mathbf{0}, \mathbf{1})$ forms an ordered field.*

Proof. We will aim to show just one property as the other properties all follow similar proofs. Let us attempt to prove that every $[\sigma] \neq \mathbf{0}$ has a multiplicative inverse. First, the sequence $\mathbf{0}$ will be defined as a sequence of all 0's (analogous for $\mathbf{1}$). It is obvious that it satisfies the additive identity by the definitions of $+$ and $\cdot$. Now, we know that

$$A = \{i \in I : \sigma(i) = 0\} \notin \mathcal{U}.$$

by the fact that $[\sigma] \neq \mathbf{0}$. Using the fact that $\mathcal{U}$ is an ultrafilter, we can conclude that the complement

$$A^c = \{i \in I : \sigma(i) \neq 0\}$$

is in $\mathcal{U}$. Now we choose an I-sequence τ such that $\tau(i) = \frac{1}{\sigma(i)}$ whenever $i \in A^c$ (that is, the index hits a nonzero element in the σ-sequence). Notice that

$$A^c \subseteq \{i \in I : \sigma(i) \cdot \tau(i) = 1\} \in \mathcal{U}$$

by the properties of a filter. This implies that $[\sigma] \cdot [\tau] = \mathbf{1}$ as desired. $\square$

Remark. One may wonder if different ultrafilters $\mathcal{U}$ and $\mathcal{V}$ would give rise to different ${}^*\mathbb{R}$. Under the continuum hypothesis, a statement that $|\mathcal{P}(\mathbb{N})| = |\mathbb{R}|$, then as ordered fields all ${}^*\mathbb{R}$ formed by ultrafilters on $I = \mathbb{N}$ are isomorphic. However, there would still be differences when more general functions and relations are considered. To see a more in depth treatment, please refer to [NGL18, Page 13].

6 Internal vs. External Sets

Previously, it had been said that quantifying $\forall x \subseteq X$ was not the same as $\forall x \in \mathcal{P}(X)$ because ${}^*\mathcal{P}(X) \neq \mathcal{P}({}^*X)$ in general. We aim to further develop this idea:

Lemma 6.1. *For sets X, ${}^*\mathcal{P}(X) \subseteq \mathcal{P}({}^*X)$.*

Proof. We need to show that $\forall x \in {}^*\mathcal{P}(X) \to x \in \mathcal{P}({}^*X)$. It is true that $\forall x \in \mathcal{P}(X)\ \forall y \in x\ (y \in X)$ by definition. Applying the transfer principle, we receive that $\forall x \in {}^*\mathcal{P}(X)\ \forall y \in x\ (y \in {}^*X)$. The $\forall y \in x\ (y \in {}^*X)$ part implies that x is a subset of ${}^*X$, which implies it is a member of $\mathcal{P}({}^*X)$. $\square$

Of course, this does not demonstrate that there is a case where ${}^*\mathcal{P}(X) \subsetneq \mathcal{P}({}^*X)$. For that, let us first consider the structure ${}^*\mathbb{N}$ (the set of nonnegative integers according to ${}^*\mathbb{R}$).

$${}^*\mathbb{N} = \{ \underbrace{0, 1, 2, \dots}_{\text{finite elements } (\mathbb{N})}, \underbrace{\dots, N-1, N, N+1, \dots}_{\text{infinite elements}} \}.$$

Now, it is true that "every nonempty subset of $\mathbb{N}$ has a least element." However, it is clear that the set ${}^*\mathbb{N} \setminus \mathbb{N} = \{\dots, N-1, N, N+1, \dots\}$ does not have this property. So, is the transfer principle false? Well, remember the previous remark and the discussion concerning subsets. The statement $\forall x \in \mathcal{P}({}^*\mathbb{N})\ (x \neq \varnothing \to x$ has least element) is false. But, by transfer principle, $\forall x \in {}^*\mathcal{P}(\mathbb{N})\ (x \neq \varnothing \to x$ has least element) works! This is because the set ${}^*\mathcal{P}(\mathbb{N})$ collects what is known to be the "internal" subsets of ${}^*\mathbb{N}$. That is, the subsets our nonstandard model determined to be "wholesome"-enough to exist (in other words, the subsets which actually still abide in accordance with the transfer principle). The subset ${}^*\mathbb{N} \setminus \mathbb{N}$ is "external", which means our nonstandard model either refuses its existence or declares that it is the empty set. This is also because the nonstandard model doesn't believe $\mathbb{N}$ is a set either (therefore, $\mathbb{N}$ is also external). Imagine if someone asked you "Which natural numbers in $\mathbb{N}$ are infinite?" You'd respond either "There are no such numbers," or "Absurd question!"

Generally, properties about the subsets of A tranfer to the *internal* subsets of ${}^*A$ and likewise the properties of the set of functions $A \to B$ transfer to the internal functions ${}^*A \to {}^*B$.

Lemma 6.2. *Let $\phi(x, y_1, \dots, y_n)$ be an elementary statement. If A is an internal set and $y_1, \dots, y_n$ were internal objects, then the set $\{x \in A : \phi(x, y_1, \dots, y_n)\}$ is internal.*

Proof. Our assumption grants us that there is a family $\mathcal{F}$ of sets such that $A \in {}^*\mathcal{F}$ and sets Y_i such that $y_i \in Y_i$ for each $i \in \{1, \dots, n\}$. Let $\mathcal{C} = \bigcup\{\mathcal{P}(X) : X \in \mathcal{F}\}$. First, notice that $\mathcal{F} \subseteq \mathcal{C}$ since for all $X \in \mathcal{F}$, $X \in \mathcal{P}(X)$ so $X \in \mathcal{C}$. The following property $\varphi(\mathcal{C}, Y_1, \dots, Y_n)$ is true about objects $\mathcal{C}, Y_1, \dots, Y_n$:

$$\forall x \in \mathcal{C}, x_1 \in Y_1, \dots, x_n \in Y_n \; \exists z \in \mathcal{C} \text{ so that “} z = \{t \in x : \phi(t, x_1, \dots, x_n)\}\text{” holds.}$$

The statement $z = \{t \in x : \phi(t, x_1, \dots, x_n)\}$ can indeed be formulated as an elementary statement as ϕ can be and $z \subseteq x$ can also be. By the transfer principle, $\varphi({}^*\mathcal{C}, {}^*Y_1, \dots, {}^*Y_n)$ is also true. Because $\mathcal{F} \subseteq \mathcal{C} \subseteq {}^*\mathcal{C}$, we can conclude that $A \in {}^*\mathcal{C}$. A similar argument concludes that $y_i \in {}^*Y_i$ for all $i \in \{1, \dots, n\}$. Now, φ tells us that for any x in $\mathcal{C}$ and $x_i \in Y_i$, there exists $z = \{t \in x : \phi(t, x_1, \dots, x_n)\}$. The transfer principle changes this to ${}^*\mathcal{C}$ and ${}^*Y_i$. We just proved that the elements we want are in those sets. This implies that there is an internal set in ${}^*\mathcal{C}$ where $C = \{t \in A : \phi(t, y_1, \dots, y_n)\}$. $\square$

6.1 Hyperfinite Sets

One useful thing about internal sets is that we can generalize our notion of finiteness to the hyperreals.

Definition 6.3. A set A is said to be *hyperfinite* if it is a member of the nonstandard extension ${}^*\mathcal{F}$ of some family of finite sets $\mathcal{F}$.

Hyperfinite sets are all internal objects, which makes them useful with respect to transfer. In our ultrapower model, hyperfinite sets would be defined by sequences of finite sets.

Proposition 6.3.1. *A set $A \subseteq {}^*X$ is hyperfinite if and only if $A \in {}^*\mathcal{P}_{\text{Fin}}(X)$ where $\mathcal{P}_{\text{Fin}}(X)$ is the finite powerset of X (the set of all finite subsets of X).*

Proof. If A is a hyperfinite subset, then it is internal, so $A \in {}^*\mathcal{P}(X)$. Let $\mathcal{F}$ be a family of finite sets for which $A \in {}^*\mathcal{F}$. Then, $A \in {}^*\mathcal{P}(X) \cap {}^*\mathcal{F} = {}^*(\mathcal{P}(X) \cap \mathcal{F}) \subseteq {}^*\mathcal{P}_{\text{Fin}}(X)$ as desired.

On the other hand, if $A \in {}^*\mathcal{P}_{\text{Fin}}(X)$, then A is hyperfinite because $\mathcal{P}_{\text{Fin}}(X)$ is a family of finite sets. $\square$

Now, just because we call a set “hyperfinite” does not mean that set is actually finite. In fact, it is sometimes far from the truth.

Proposition 6.3.2. *Every hyperfinite interval $[1, N] \subsetneq {}^*\mathbb{N}$ for $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ has cardinality greater than or equal to the continuum (the cardinality of $\mathbb{R}$).*

Proof. Note that the cardinality of the interval $(0, 1) \subsetneq \mathbb{R}$ is equal to the cardinality of $\mathbb{R}$ itself. Fix $N \in {}^*\mathbb{N} \setminus \mathbb{N}$. We wish to show that $|\mathbb{R}| \leq |[1, N]|$. In other words, there is an injective function $f : \mathbb{R} \to [1, N]$. But since $|\mathbb{R}| = |(0, 1)|$, we can instead find $f : (0, 1) \to [1, N]$. Let

$$f(x) = \min\left\{a \in [1, N] : x < \frac{a}{N}\right\}.$$

The minimum is well defined by the transfer principle and by Lemma 6.2. Now, we need to show that f is injective. Take $x, y \in \mathbb{R}$, and let $f(x) = f(y)$. We aim to show that $x = y$. For a moment, consider if we had taken $x', y' \in {}^*\mathbb{R}$ where $x' = x + \varepsilon_1$ and $y' = y + \varepsilon_2$ for infinitesimals $\varepsilon_1, \varepsilon_2$. It is no longer guaranteed that $f(x') = f(y') \implies x' = y'$. Indeed, the only thing we do get is that $\epsilon = |x' - y'| < \frac{1}{N}$. It is notable that ϵ is an infinitesimal because N is an infinite element. This can be a variety of infinitesimals in ${}^*\mathbb{R}$, but in $\mathbb{R}$, we care about $x = \operatorname{st} x'$ and $y = \operatorname{st} y'$. In which case, the ϵ collapses to the only infinitesimal available in $\mathbb{R}$. That is $0 = |x - y|$ which implies $x = y$. $\square$

We established that a lot of hyperfinite intervals have cardinality greater or equal to the continuum. One might wonder if we can have a subset with cardinality equal to the cardinality of $\mathbb{N}$.

Proposition 6.3.3. *Every countably infinite (with the size of $|\mathbb{N}|$) subset of ${}^*\mathbb{N}$ is external.*

Proof. Let $A \subseteq {}^*\mathbb{N}$. By the contrapositive, it is enough to show that if A is internal, then A must be either finite or greater than or equal to the cardinality of $\mathbb{R}$. By transfer, A is either hyperfinite and in bijection with some interval $[1, a] \subsetneq {}^*\mathbb{N}$, or A is not in which case there should be a function $f : {}^*\mathbb{N} \to A$ which is injective.

(i) For the first case, if $a \in \mathbb{N}$ then A is finite. If $a \in {}^*\mathbb{N} \setminus \mathbb{N}$, then by Proposition 6.3.2 A has cardinality greater than or equal to $\mathbb{R}$.

(ii) Since for all $N \in {}^*\mathbb{N} \setminus \mathbb{N}$, we have $|[1, N]| \geq |\mathbb{R}|$, it follows that $|{}^*\mathbb{N}| \geq |[1, N]| \geq |\mathbb{R}|$. So, for $f : {}^*\mathbb{N} \to A$ to be injective, we need $|A| \geq |{}^*\mathbb{N}| \geq |\mathbb{R}|$. ☐

6.2 Brief Note on Applications

The most obvious appeal of nonstandard analysis in its greater applications to mathematics is that of calculus.

Definition 6.4. The Hyperfinite Grid for some fixed infinite $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ is defined as $\mathbb{H}_N = \{\frac{a}{N} : a \in \{0, 1, \ldots, N\}\}$ which partitions the interval $[0, 1]$ into subintervals of length $\frac{1}{N}$.

We can now define an integral for functions $f : [0, 1] \to \mathbb{R}$ with this type of set. For some hyperfinite grid $\mathbb{H}$ and a integration domain $A \subseteq [0, 1]$, we define

$$\int_A f(x)\ \mathrm{d}_{\mathbb{H}}(x) = \mathrm{st}\left(\sum_{i \in {}^*A \cap \mathbb{H}} {}^*f(i) \cdot \frac{1}{N}\right).$$

This is known as the *grid integral*. It turns out, nonstandard analysis also has massive applications to fields like Ramsey Theory and combinatorics. All of these applications are explored in Nonstandard Methods in Ramsey Theory and Combinatorial Number Theory by Nasso et al. in [NGL18].

7 Saturation Principles

Another crucial part to nonstandard analysis is the *saturation principle*. It is important for applications to various in nonstandard analysis. For example, it is needed for defining certain measures which are crucial for analysis. This explored more in [NGL18, Page 31].

Definition 7.1 (Chang, [CK12])**.** A model is κ-saturated if and only if whenever a family of sets $\{A_i\}_{i \in I}$ satisfies the finite intersection principle and $|I| \leq \kappa$ then

$$\bigcap_{i \in I} A_i \neq \varnothing.$$

A special case is the $\aleph_0$-saturation principle. This is also known as the countable saturation principle, and occurs when $\kappa = |\mathbb{N}|$.

Lemma 7.2. *Every ultrapower model generated by nonprincipal ultrafilter $\mathcal{U}$ of $\mathbb{N}$ satisfies the countable saturation principle.*

Proof. We have fixed an index set $I = \mathbb{N}$. Now, let $A = \{a_i\}_{i\in\mathbb{N}}$ be a countable family of internal subsets of ${}^*\mathbb{R}$ with the finite intersection property. We will now create a family $B = \{b_i\}_{i\in\mathbb{N}}$ which is a family of sequences of subsets of $\mathbb{R}$. This is because the ultrapower works by sequences for any object. For $i \in \mathbb{N}$, set the sequence such that

$$a_i = \{[\sigma] \in {}^*\mathbb{R} : \{j \in \mathbb{N} : \sigma(j) \in b_i(j)\} \in \mathcal{U}\}.$$

The setup can be summarized as follows:

(i) We have the family of subsets $\{a_i\}_{i\in\mathbb{N}}$ which we want to prove some property about.

(ii) We construct a new family B of sequences of subsets. But, these are subsets of the original $\mathbb{R}$.

(iii) If some number (secretly a sequence) $\sigma \in {}^*\mathbb{R}$ is in the sequences of subsets b_i on the indices within the ultrafilter, then we say $[\sigma]$ is in the family a_i. So, in a way, each b_i is an approximation of a_i that occurs "before the ultrafilter". That is, b_i is a sequence that are subsets that is grounded in $\mathbb{R}$, but when interpreted by $\mathcal{U}$ it becomes a_i again.

Now, for any $n \in \mathbb{N}$, pick an element $\tau(n) \in b_0(n) \cap \cdots \cap b_n(n)$ if it is nonempty. If this is indeed empty, then we take $b_0(n) \cap \cdots \cap b_{n-1}(n)$ instead, and continue iterating down. If $b_0(n) = \varnothing$ (which would kill the intersection no matter what), then we just set $\tau(n) = 0$. Now, if $b_0(n) \cap \cdots \cap b_k(n) \neq \varnothing$, and $n \geq k$, then $\tau(n) \in b_0(n) \cap \cdots \cap b_k(n)$. This is true because if m is the stage where we picked $\tau(n)$ from (basically the first number that satisfies the nonempty condition by our descent from $n, n-1, \ldots$), then if $\tau(n) \in b_0(n) \cap \cdots \cap b_m(n)$ then $\tau(n) \in b_0(n) \cap \cdots \cap b_k(n)$ for when $k \leq m$, which are all the intersections that are nonempty.

Fix $k \in \mathbb{N}$. By the finite intersection property, $a_0 \cap \cdots \cap a_k \neq \varnothing$. Following this, there must be a sequence σ such that $\Lambda_j = \{i \in \mathbb{N} : \sigma(i) \in b_j(i)\} \in \mathcal{U}$ for every $j \in \{0, \ldots, k\}$ (the sequence σ is consistent throughout each j). This is due to the fact that each $a_0, \ldots, a_k$ are nonempty. Remembering that filters are closed over intersections, we can see that any $\Lambda_0 \cap \cdots \cap \Lambda_k \in \mathcal{U}$ as well. Furthermore, define Γ such that $\Gamma(k) = \{i \in \mathbb{N} : b_0(i) \cap \cdots \cap b_k(i) \neq \varnothing\}$. Consider $i \in \Lambda_0 \cap \cdots \cap \Lambda_k$. Remember that we have a sequence such that $\sigma(i) \in b_j(i)$ for all $j \in \{0, \ldots, k\}$, indicating that each b_j is nonempty at i specifically with element $\sigma(i)$. This means that $\sigma(i) \in b_0(i) \cap \cdots \cap b_k(i)$. It follows that $\Gamma(k)$ (the set of indices for which the intersection is nonempty) is a superset of $\Lambda_0 \cap \cdots \cap \Lambda_k$. Therefore, $\Lambda_0 \cap \cdots \cap \Lambda_k \subseteq \Gamma(k) \in \mathcal{U}$ because filters respect supersets.

Now, remember that $\{i \in \mathbb{N} : i \geq k\} \in \mathcal{U}$ as well because it is a cofinite set, which means it is in the nonprincipal ultrafilter $\mathcal{U}$ by Lemma 3.6. This also means that $\Gamma(k) \cap \{i \in \mathbb{N} : i \geq k\} = \{i \in \Gamma(k) : i \geq k\} \in \mathcal{U}$. Remember that we can construct a sequence τ such that if $b_0(n) \cap \cdots \cap b_k(n) \neq \varnothing$, and $n \geq k$, then $\tau(n) \in b_0(n) \cap \cdots \cap b_k(n)$. So, the set $\{i \in \mathbb{N} : \tau(i) \in b_0(i) \cap \cdots \cap b_k(i)\}$ is a superset to $\{i \in \Gamma(k) : i \geq k\}$ as we treat i as the n here, meaning the $n \geq k$ condition is satisfied. Since it is a superset, $\{i \in \Gamma(k) : i \geq k\} \in \mathcal{U}$. Therefore, $[\tau] \in a_0 \cap \cdots \cap a_k$. But we can push $k \to \infty$ since we never bounded $k \in \mathbb{N}$, which means it works for all countable intersections. □

Lemma 7.3. *For every infinite cardinal κ there is an ultrafilter $\mathcal{U}$ such that the corresponding ultrapower model satisfies κ-saturation.*

The above fact is much harder to prove. Its importance is that it demonstrates how ultrapower models actually do have some differences between themselves and that different ultrapower models are not uninteresting, despite the transfer principle. However, it is not entirely relevant to our discussion of how the transfer principle works itself and merely serves as an interesting example. A more complete proof and account can be found in [CK12, Section 6.1].

8 Ultrapower Models satisfy Transfer Principle

We now wish to prove that our star map defined by ${}^*X = X^I/\mathcal{U}$ does indeed satisfy the transfer principle (and thus is actually a star map). We will first prove a more general theorem about ultraproducts.

8.1 Łoś' Theorem on Ultraproducts

Fix some index set I and an ultrafilter $\mathcal{U}$ on I. Consider the set

$$X = \prod_{i \in I} X_i / \mathcal{U}$$

where

$$\sigma \equiv_{\mathcal{U}} \tau \iff \{i \in I : \sigma(i) = \tau(i)\} \in \mathcal{U}.$$

Note that the product $\prod$ here means cartesian product.

Definition 8.1 (Jech, [Jec06])**.** Let $\{\mathfrak{U}_i : i \in I\}$ be a system of models for some fixed language. We obtain a model $\mathfrak{U}$ (called the *ultraproduct* of $\{\mathfrak{U}_i : i \in I\}$) with universe X by interpreting the language as follows:

(i) If ϕ is some predicate or statement, then $\phi^{\mathfrak{U}}([A_1], \ldots, [A_n])$ if and only if $\{i \in I : \phi^{\mathfrak{U}_i}(A_1(i), \ldots, A_n(i))\} \in \mathcal{U}$.

(ii) If f is some function, then $f^{\mathfrak{U}}([x_1], \ldots, [x_n]) = [y]$ if and only if $y(i) = F^{\mathfrak{U}_i}(x_1(i), \ldots, x_n(i))$ for all $i \in I$.

(iii) If c is a constant, then $c^{\mathfrak{U}} = [a]$ where $a(i) = c^{\mathfrak{U}_i}$ for all $i \in I$.

Remark. If the ultrafilter $\mathcal{U}$ was just a regular filter F, then this would be called a reduced product instead. The point of the ultrafilter is to satisfy an analog of the law of the excluded middle (everything is either true or false). Furthermore, if all X_i were equal, then the universe X would be some X_0^I, which recovers the phrase "ultrapower" used in the discussion of the hyperreals.

Theorem 8.2 (Łoś, [Łoś55])**.** *Let $\mathcal{U}$ be an ultrafilter on I. Let $\mathfrak{U}$ be the ultraproduct $\{\mathfrak{U}_i : i \in I\}$ by $\mathcal{U}$. If ϕ is some formula, then for every $A_1, \ldots, A_n \in \prod_{i \in I} X_i$,*

$$\mathfrak{U} \vDash \phi([A_1], \ldots, [A_n]) \text{ if and only if } \{i \in I : \mathfrak{U}_i \vDash \phi(A_1(i), \ldots, A_n(i))\}.$$

Proof. We will induct on the complexity of the formulas ϕ. That is, we will start from atomic formulas and build our way up (tree induction style) by further proving the theorem for $\neg, \wedge, \exists$.

Note that similar proofs follow for $\vee, \rightarrow, \leftrightarrow, \forall$, but also they can be recovered from the three aforementioned connectives and quantifiers.

$$\begin{aligned}
\mathfrak{U} \models [\sigma] = [\tau] &\iff [\sigma] = [\tau] \\
&\iff \sigma \equiv_{\mathcal{U}} \tau \\
&\iff \{i \in I : \sigma(i) = \tau(i)\} \in \mathcal{U} \\
&\iff \{i \in I : \mathfrak{U}_i \models \sigma(i) = \tau(i)\} \in \mathcal{U}.
\end{aligned}$$

The last step is true because $\mathfrak{U}_i \models \sigma(i) = \tau(i) \iff \sigma(i) = \tau(i)$ (which is true because $\sigma(i)$ and $\tau(i)$ exist within $\mathfrak{U}_i$ and are dictated by $\mathfrak{U}_i$).

Let ϕ be some predicate or relation. Then we have

$$\begin{aligned}
\mathfrak{U} \models \phi([A_1], \dots, [A_n]) &\iff \phi^{\mathfrak{U}}([A_1], \dots, [A_n]) \\
&\iff \{i \in I : \phi^{\mathfrak{U}_i}(A_1(i), \dots, A_n(i))\} \in \mathcal{U} \\
&\iff \{i \in I : \mathfrak{U}_i \models \phi(A_1(i), \dots, A_n(i))\} \in \mathcal{U}.
\end{aligned}$$

This already encapsulates $\in$ or any other possible relation, and a trivial modification of this proof would allow for all atomic formulas.

Now, for our induction step. Assume that our theorem holds for some statement φ. Now, we do tree induction over the formulas.

(i) For $\neg\varphi$,

$$\begin{aligned}
\mathfrak{U} \models \neg\varphi &\iff \mathfrak{U} \not\models \varphi \\
&\iff \{i \in I : \mathfrak{U}_i \not\models \varphi\} \in \mathcal{U} \\
&\iff \{i \in I : \mathfrak{U}_i \models \varphi\} \in \mathcal{U}.
\end{aligned}$$

(ii) For φ and ψ,

$$\begin{aligned}
\mathfrak{U} \models \varphi \wedge \psi &\iff \{i \in I : \mathfrak{U}_i \models \varphi\} \in \mathcal{U} \text{ and } \{i \in I : \mathfrak{U}_i \models \psi\} \in \mathcal{U} \\
&\iff \{i \in I : \mathfrak{U}_i \models \varphi \wedge \psi\} \in \mathcal{U}.
\end{aligned}$$

The last fact is true by the fact for a filter F: $X \in F$ and $Y \in F$ implies $X \cap Y \in F$.

(iii) Assume we have proven the theorem for $\varphi(a, A_1, \dots, A_n)$ (abbreviated $\varphi(a, \dots)$) and we want to show it remains true for $\exists a\ \varphi(a, \dots)$. Assume $\mathfrak{U} \models \exists a\ \varphi(a, \dots)$. Then, there exists $u \in \prod_{i \in I} X_i = X$ such that $\mathfrak{U} \models \varphi(u, \dots)$ which implies

$$\{i \in I : \mathfrak{U}_i \models \exists a\ \varphi(a, \dots)\}.$$

Now, for the other part of the proof, assume the above statement is true. For each $i \in I$, let $a_i \in X_i$ be such that $\mathfrak{U}_i \models \varphi(a_i, \dots)$ if a proper a_i exists, and anything else otherwise. Define $u \in X$ as $u(i) = a_i$, which concludes $\mathfrak{U} \models \exists a\ \varphi(a, \dots)$ as desired. □

Thinking about this theorem on ultrapowers, we get the immediate corollary:

Corollary 8.2.1. *Consider the mapping d which sends $d(c) = [c]$ (where $[c]$ is the equivalence class of the constant function $\varsigma(i) = c$ for every $i \in I$). The mapping d satisfies an analog of the transfer principle.*

The proof of this corollary is an immediate application of Theorem 8.2. The fact that d is an analog is true in a model theoretic sense: we would usually not take d as the star map because we want additional properties to be satisfied that makes everything nicer. However, the "transfer principle" between structures N and M in the sense that

$$N \vDash \varphi \iff M \vDash \varphi$$

is indeed satisfied by d.

8.2 Łoś' Theorem on Superstructures

We now develop an analog of Łoś' Theorem on Superstructures, which are specific structures that are more accommodating to nonstandard analysis.

Let us define the cumulative hierarchy V in regular set theory. The intention of V is to be the collection (specifically, proper class) of ALL sets.

Definition 8.3. We define V in the following way:

(i) $V_0 = \varnothing$.

(ii) $V_{\alpha+1} = \mathcal{P}(V_\alpha)$.

(iii) $V_\alpha = \bigcup_{\beta<\alpha} V_\beta$ for limit ordinals α.

Then, $V = \bigcup_{i \in \mathrm{Ord}} V_i$.

Remark. What this definition means is that, starting with the empty set, we endlessly take subsets for all levels of infinity (recall that there are different sizes of infinity due to [Can91]). This is able to define any set imaginable, including $\mathbb{N}$ or $\mathbb{R}$. The reason why is because in set theory, everything is a set. That includes 4, π, and even sin.

Suppose that $*$ is some star map with nonstandard natural numbers ${}^*\mathbb{N}$. Sets can be nested up to any finite length, but ${}^*V$ believes that some infinite values are finite, which can create an infinite descent $\cdots \in \xi - 2 \in \xi - 1 \in \xi$ which is not permitted in usual set theories. There are several ways to account for this, but we will avoid this entirely by a different method.

Definition 8.4. Let X be some set of *atoms*. We define the structure $V(X)$ as:

(i) $V_0(X) = X$.

(ii) $V_{\alpha+1}(X) = V_\alpha(X) \cup \mathcal{P}(V_\alpha(X))$.

Then, $V(X) = \bigcup_{i \in \mathbb{N}} V_i(X)$.

We usually take $X = \mathbb{R}$. Now, we define the ultrapowers we care about in nonstandard analysis.

Definition 8.5. Let $\mathcal{U}$ be an ultrafilter on a set of indices I. The *bounded ultrapower* of the *superstructure* $V(X)$ is

$$V(X)^I_b/\mathcal{U} = \bigcup_{i \in \mathbb{N}} V_i(X)/\mathcal{U}.$$

Relations such as $\in_\mathcal{U}$ are defined via equivalence classes where we demand the indices that satisfy the relation are in the ultrafilter $\mathcal{U}$.

Lemma 8.6. *Fix an index set I and ultrafilter $\mathcal{U}$ on I. Let $V(X)^I_b/\mathcal{U}$ be the bounded ultrapower over the structure $V(X)$ operating over atomic set X. Let $\phi[(A_1], \ldots, [A_n])$ be some elementary statement/formula with $[A_1], \ldots, [A_n]$ from $V(X)^I_b/\mathcal{U}$. Then,*

$$V(X)^I_b/\mathcal{U} \vDash \phi([A_1], \ldots, [A_n]) \iff \{i \in I : \phi(A_1(i), \ldots, A_n(i)) \text{ holds}\} \in \mathcal{U}.$$

Proof. Follows immediately as an application of Theorem 8.2 with the universe being $V(X)$. □

Corollary 8.6.1 ("Almost" Transfer Principle)**.** *If $d : V(X) \to V(X)^I_b/\mathcal{U}$ is some embedding (called the diagonal embedding), then for every property $\phi(A_1, \ldots, A_n)$ where $A_1, \ldots, A_n \in V(X)$:*

$$\phi(A_1, \ldots, A_n) \iff V(X)^I_b/\mathcal{U} \vDash \phi(d(A_1), \ldots, d(A_n)).$$

This lemma and corollary are the analog of the theorem and corollary discussed in the previous section. In fact, the embedding d is the same one we were interested in last section. To complete into an actual star mapping, we need another mapping $\pi : V(X)^I_b/\mathcal{U} \to V({}^*X)$ where $V({}^*X)$ is the nonstandard universe (built by stage upon the nonstandard extension of X). We require π to satisfy the specific property of turning $\in_{\mathcal{U}}$ to an actual relationship (this is a property known as the *Mostowski Collapse* by Mostowski [Mos49]). That is, $[\sigma] \in_{\mathcal{U}} [\tau] \iff \pi([\sigma]) \in \pi([\tau])$. We define $* = \pi \circ d$, which gives rise to the following diagram:

$$\begin{tikzcd} & V(X)^I_b/\mathcal{U} \arrow[dr, "\pi"] & \\ V(X) \arrow[ur, "d"] \arrow[rr, "*"] & & V({}^*X) \end{tikzcd}$$

This completes the construction and justification that the star maps formed by ultrapowers are indeed star maps that satisfy the transfer principle.

9 Iterated Hyperextensions

Sometimes, we would like to consider an iterated star map application like ${}^{**}\mathbb{R}$. After all, if ${}^*\mathbb{R}$ believed itself to be the "true $\mathbb{R}$", then it would also believe that there are nonstandard $\mathbb{R}$ out there (from our perspective, that would be ${}^{**}\mathbb{R}$).

Proposition 9.0.1. *${}^*\mathbb{N}$ is an initial segment of ${}^{**}\mathbb{N}$.*

Proof. Since $\mathbb{N} \subsetneq {}^*\mathbb{N}$, transfer principle tells us that ${}^*\mathbb{N} \subsetneq {}^{**}\mathbb{N}$. Furthermore, apply transfer principle to the first order formula:

$$\forall x \in \mathbb{N}\ \forall y \in {}^*\mathbb{N} \setminus \mathbb{N}\ x < y$$

to obtain

$$\forall x \in {}^*\mathbb{N}\ \forall y \in {}^{**}\mathbb{N} \setminus {}^*\mathbb{N}\ x < y.$$

□

Remark. This might seem ill-formed because we had said that ${}^*\mathbb{N} \setminus \mathbb{N}$ was an external set, yet we are using transfer principle. However, by our explicit construction of ${}^*\mathbb{N}$, we can talk about it **from the perspective of** $\mathbb{N}$. Of course, ${}^*\mathbb{N}$ does not believe in ${}^*\mathbb{N} \setminus \mathbb{N}$, but it is perfectly fine with ${}^{**}\mathbb{N} \setminus {}^*\mathbb{N}$!

Proposition 9.0.2. *While it is true that $n = {}^*n$ for all $n \in \mathbb{N}$, this is not true for $n \in {}^*\mathbb{N}$ in general.*

Proof. The fact that $n = {}^*n$ for all $n \in \mathbb{N}$ follows from definition. We wanted ${}^*\mathbb{N}$ to contain a copy of $\mathbb{N}$ and by the ultrapower model we have $n \cong [n]$ anyways. Now, if $n \in {}^*\mathbb{N}\setminus\mathbb{N}$, then ${}^*n \in {}^{**}\mathbb{N}\setminus{}^*\mathbb{N}$. By Proposition 9.0.1, we conclude that $n < {}^*n$ which means $n \neq {}^*n$ for some $n \in {}^*\mathbb{N}$. $\square$

9.1 Iterated Ultrapower Constructions

Unfortunately, the ultrapower construction does not have an obvious way to give rise to iterated extensions. Let us fix an ultrafilter $\mathcal{U}$ on $I = \mathbb{N}$. Let ${}^*\mathbb{N} = \mathbb{N}^{\mathbb{N}}/\mathcal{U}$ as before. Let us take another ultrapower to obtain ${}^{**}\mathbb{N} = ({}^*\mathbb{N})^{\mathbb{N}}/\mathcal{U} = (\mathbb{N}^{\mathbb{N}}/\mathcal{U})/\mathcal{U}$. Consider the diagonal embedding $d_0 : \mathbb{N} \to \mathbb{N}^{\mathbb{N}}/\mathcal{U}$, specifically for this instance we want ${}^*a = d_0(a)$ (there is no concern for a π mapping back like before since we are not working with a superstructure). Now, consider the next diagonal embedding $d_1 : \mathbb{N}^{\mathbb{N}}/\mathcal{U} \to (\mathbb{N}^{\mathbb{N}}/\mathcal{U})/\mathcal{U}$. Recall that, by Proposition 9.0.2, $n < {}^*n$ for some $n \in {}^*\mathbb{N}$. How come $a \neq d_1(a)$ for some $a \in {}^*\mathbb{N}$ when it should behave like d_0 where $b = d_0(b)$ for $b \in \mathbb{N}$? This is because that mapping that does what we would want to see from d_0 is actually $d_0^{\mathcal{U}} : \mathbb{N}^{\mathbb{N}}/\mathcal{U} \to (\mathbb{N}^{\mathbb{N}}/\mathcal{U})/\mathcal{U}$. This mapping, being ${}^*\mathbb{N}$'s take on d_0, is a mapping that preserves $a = d_0^{\mathcal{U}}(a) = {}^*a$. The mapping d_1 is perceived from $\mathbb{N}$. With this difference aside, we have now recovered what we wanted from the ultrapower model.

9.2 Support for Iterated Superstructures

We will go over a proof sketch.

Theorem 9.1. *There can exist a star map for superstructures $* : V(X) \to V(X)$ in a nontrivial way.*

Proof. We need to verify transfer principle over this potential star map. Specifically, we are considering the structure $\langle V(X), V(X), *\rangle$ rather than $\langle V(X), V({}^*X), *\rangle$. Consider the following diagram:

$$\begin{tikzcd} & V(X)^I_b/\mathcal{U} \arrow[dr, "\Psi"] & \\ V(X) \arrow[ur, "d"] \arrow[rr, "*"] & & V(X) \end{tikzcd}$$

The first part in constructing d is the same. However, we now have this new map $\Psi : V(X)^I_b/\mathcal{U} \to V(X)$ that replaces the previous $\pi : V(X)^I_b/\mathcal{U} \to V({}^*X)$ mapping. We will then set $* = \Psi \circ d$. We want Ψ to map back in a nontrivial way.

First, we need to choose X and I carefully. In particular this doesn't work for general choices. We just need to show at least one exists. Fix two infinite cardinals λ and κ. For our atom set X, we will demand $|X| = \lambda^\kappa$. Note that $\lambda^\kappa \geq |\mathbb{R}|$, so let us also sneak in a copy of $\mathbb{R}$ into X because we can. We are just adding useful sets to our ground set for the sake of convenience. Next, we demand $I = \mathcal{P}_{\text{Fin}}(\kappa)$ (the finite subsets of κ). If you are uncomfortable with the fact that we are taking "subsets of a number", remember that everything in set theory is a set. Still, you can reinterpret this as $I = \mathcal{P}_{\text{Fin}}(Y)$ where Y is any set such that $|Y| = \kappa$. Let us now take a nonprincipal ultrafilter $\mathcal{U}$ over I. We aim to inductively define Ψ for each stage $n \in \mathbb{N}$.

(i) Consider $\Psi_0 : X^I/\mathcal{U} \to X$. Note that $\lambda^\kappa = |X| \leq |X^I/\mathcal{U}| \leq |X^I| = (\lambda^\kappa)^\kappa = \lambda^\kappa$, so $|X| = |X^I/\mathcal{U}|$. We can just pick any bijection such that $\Psi_0([c_x]) = x$ (a reversal of $d(x) = [c_x]$) where c_x is just a constant sequence with x in it for $x \in X \setminus \mathbb{R}$. As a side note, the reason we do $X \setminus \mathbb{R}$ is because this is a very special set of atoms X_0. In fact, our previous choice of $I = \mathcal{P}_{\text{Fin}}(\kappa)$ was actually motivated by this $X_0 = X \setminus \mathbb{R}$ set so we need to ensure it does not get messed up when passing through Ψ, which is what this definition does.

(ii) Let $\Psi_n : V_n(X)^I/\mathcal{U} \to V_n(X)$ be a function which is validly defined. If the set of indices i where $\sigma(i) \in V_n(X)$ is within the ultrafilter (or "almost-everywhere"), then simply $\Psi_{n+1}([\sigma]) = \Psi_n([\sigma])$. Otherwise, $\sigma(i) \in \mathcal{P}(V_n(X))$ (in the $V_{n+1}(X)$ stage), so we define it as:

$$\Psi_{n+1}([\sigma]) = \{\Psi_n([\tau]) : \{i \in I : \tau(i) \in \sigma(i)\} \in \mathcal{U}\}.$$

In other words, the places where $\tau \in \sigma$ happens almost everywhere by the ultrafilter. Since the codomain of Ψ_n is $V_n(X)$ by the induction hypothesis, it follows that this Ψ_{n+1} evaluation gives a subset of $V_n(X)$ which lives in $V_{n+1}(X)$.

From here, we can set $* = \Psi \circ d$ to recover a star map as desired. The proof that $\Psi \circ d$ satisfies the transfer principle is similar to one provided in [CK12, Theorem 4.4.5]. □

10 Nonstandard Set Theories

We develop an alternative approach to the star map $*$ and nonstandard analysis in general. Namely, from a philosophical standpoint, we are changing the fundamental ontology of our mathematical universe. Normally, set theories like ZFC operate over only sets and the relation $\in$ between them. In certain set theories designated for nonstandard approaches, we introduce additional predicates or symbols to quantify the "standard-ness" of the objects.

10.1 A Review of Traditional Set Theory

The most common set theory as discussed previously is of course ZFC. Below is a list of its axioms [Jec06]:

(i) Axiom of Extensionality: Two sets are equal if and only if they have the same elements.

(ii) Axiom of Regularity (or Foundation): Every non-empty set x contains a member y such that they are disjoint.

(iii) Axiom Schema[1] of Specification: Any subset made by set builder notation (i.e., $\{x \in z : \varphi(x)\}$ for some formula φ) exists as long as the set it originates from (i.e., z) does.

(iv) Axiom of Pairing: If a, b are sets, then the set $\{a, b\}$ exists.

(v) Axiom of Union: The union between two sets exists.

(vi) Axiom of Powerset: The powerset of any set exists.

[1] An axiom schema means that it is actually a group of infinite axioms that all share a similar form.

(vii) Axiom Schema of Replacement: The image of a set according to a definable function exists and is a set.

(viii) Axiom of Infinity: There exists an infinite set.

(ix) Axiom of Choice: Every family of functions has a choice function.

Definition 10.1 (Von Neumann, [Neu23])**.** The canonical way to define an *ordinal* is a set that is well ordered by $\in$. The more practical definition is:

(i) The smallest ordinal is $0 = \varnothing$.

(ii) $\alpha + 1 = \{\alpha\} \cup \alpha$.

(iii) If β is an ordinal that cannot be expressed as $\alpha + 1$ and is nonzero (called a limit ordinal), then $\beta = \bigcup_{\lambda<\beta} \lambda$.

In set theory, we notate $\omega = \mathbb{N}$. Previously, we discussed an issue with the nonstandard extension of V. That is, ${}^*\omega$ has infinite descent because of any nonstandard natural ξ has $\xi \ni \xi - 1 \ni \cdots$. This is why we opt to define the superstructure format, which bypasses this issue. However, there are alternative methods to rid of this issue.

Remark. It should be noted that model theory goes far beyond the quite tame models we have developed here for nonstandard analysis. Indeed, the fundamental philosophy that mathematics has multiple wild behaving models has spawned in new fields of research. For further reading, a good place to start would be referring to Thomas Jech's Set Theory (Third Millennium Edition) [Jec06].

10.2 Nelson's Internal Set Theory

Internal Set Theory (IST) is an extension of ZFC in which the standard and nonstandard elements live in the same universe together. There is an additional unary predicate standard(x) to mean "x is standard".

If ϕ is a formula where standard(x) does not appear, then it is also a formula of ZFC and is thus called an "internal formula". Otherwise ϕ is external. Note that we will introduce new pieces of notation: $\forall^{\text{st}} x\ \phi$ means $\forall x(\text{standard}(x) \to \phi)$, $\forall^{\text{fin}} x\ \phi$ means $\forall x(\text{finite}(x) \to \phi)$, and $\forall^{\text{st fin}} x\ \phi$ means $\forall x(\text{standard}(x) \wedge \text{finite}(x) \to \phi)$. Similar mirrors exist for $\exists^{\text{st}}$, $\exists^{\text{fin}}$, $\exists^{\text{st fin}}$.

Now, IST has three additional axiom schemas by Nelson [Nel77]:

(i) Transfer Principle: If $\phi(x, A_1, \ldots, A_n)$ is an internal formula, then:

$$\forall^{\text{st}} A_1, \ldots, A_n(\forall^{\text{st}} x\ \phi(x, A_1, \ldots, A_n) \to \forall x\ \phi(x, A_1, \ldots, A_n)).$$

(ii) Idealization Principle: Let $\phi(x, y)$ be an internal formula. Then,

$$\forall^{\text{st fin}} z \exists x \forall y \in z\ \phi(x, y) \leftrightarrow \exists x \forall^{\text{st}} y\ \phi(x, y).$$

(iii) Standardization Principle: Let $\phi(z)$ be some formula which may not be internal. Then,

$$\forall^{\text{st}} x \exists^{\text{st}} y \forall^{\text{st}} z(z \in y \leftrightarrow z \in x \wedge \phi(z)).$$

The specifics of these are not important; we will only develop a high level understanding of what IST is attempting to do. Indeed, instead of constructing a direct star map, it is already built into the the universe of sets. Specifically, the implicit star map implied by the transfer principle axiom schema is between the universe of standard objects and internal objects.

Working in IST is much like doing nonstandard analysis in ZFC. In fact, it is very much like ZFC! One of the cornerstones of IST is the *conservation theorem*.

Lemma 10.2 (Conservation Theorem)**.** *Let ϕ be an internal theorem. Then if ϕ is provable in* IST*, then it is provable in* ZFC*.*

Proof. A proof can be found in [Nel77]. □

Notate ϕ^{st} for a statement ϕ to be the statement identical to ϕ except all $\forall$ and $\exists$ to be relativized to $\forall^{\mathrm{st}}$ and $\exists^{\mathrm{st}}$.

Lemma 10.3 (Relativization Theorem)**.** *For an internal statement ϕ, ϕ is true if and only if ϕ^{st} is true.*

Proof. This proof idea is due to [Laz12]. First, we can establish that

$$\forall^{\mathrm{st}} A_1, \ldots, A_n (\exists x\ \phi(x, A_1, \ldots, A_n) \to \exists^{\mathrm{st}} x\ \phi(x, A_1, \ldots, A_n)).$$

is true by replacing ϕ with $\neg\phi$ in the Transfer Principle Axiom Schema and taking the contrapositive. Furthermore, we can also write every internal formula as $Q_1 x_1 \ldots Q_n x_n\ \psi(x_1, \ldots, x_n)$ where Q_i is either $\forall$ or $\exists$ and ψ is an internal formula without quantifiers. Notice that we have two nice facts, the Transfer Principle Axiom Schema and the contrapositive above. One can turn $\forall^{\mathrm{st}} \to \forall$ (and $\forall \to \forall^{\mathrm{st}}$, which is trivial because any standard object is also an object in general) and the other can turn $\exists \to \exists^{\mathrm{st}}$ (the other direction is once again trivial, any standard object is also a regular object). Therefore, we can apply the schema and its converse as so:

$$\begin{gathered} Q_1 x_1 \ldots Q_n x_n\ \psi(x_1, \ldots, x_n) \\ \iff \\ Q_1^{\mathrm{st}} x_1 \ldots Q_n x_n\ \psi(x_1, \ldots, x_n) \\ \iff \\ \cdots \\ \iff \\ Q_1^{\mathrm{st}} x_1 \ldots Q_{n-1}^{\mathrm{st}} x_{n-1}\ Q_n x_n\ \psi(x_1, \ldots, x_n) \\ \iff \\ Q_1^{\mathrm{st}} x_1 \ldots Q_n^{\mathrm{st}} x_n\ \psi(x_1, \ldots, x_n). \end{gathered}$$

It doesn't matter what quantifiers each Q_i is, we showed it was possible for $\forall$ and $\exists$. So, $\phi \iff \phi^{\mathrm{st}}$. □

The two above lemmas along with the transfer principle axioms can all be seen as perspectives of the transfer principle. In particular:

(i) The transfer principle axiom schema establishes the mapping between the class of standard sets and internal sets.

(ii) The Conservation Theorem can be seen as "transfer between ZFC and IST".

(iii) The Relativization Theorem can be seen as a strengthening of (i).

Remark. Internal Set Theory is not the only alternative set theory that tries to solve the issue of nonstandardness. For example, there are *anti-foundation* axioms which allow infinite descent chains like $x \in x$. To read more on these, please see [Vop89].

Acknowledgments

I would like to thank Soham Patil and Simon Rubinstein-Salzedo for their guidance and help with developing this paper.

References

[Can91] Georg Cantor. "Ueber eine elementare Frage der Mannigfaltigketislehre." In: *Jahresbericht der Deutschen Mathematiker-Vereinigung* 1 (1890/91), pp. 72–78. URL: `http://eudml.org/doc/144383`.

[CK12] C.C. Chang and H.J. Keisler. *Model Theory*. Dover Books on Mathematics. Dover Publications, 2012. ISBN: 9780486488219. URL: `https://books.google.com/books?id=EQZ_9z_FEQIC`.

[Euc56] Euclid. *The Elements of Euclid*. Ed. by Sir Thomas L. Heath. 2nd. New York: Dover Publications Inc., 1956. URL: `https://archive.org/details/euclid_heath_2nd_ed`.

[Göd31] Kurt Gödel. "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I". In: *Monatshefte für Mathematik und Physik* 38 (1931), pp. 173–198. URL: `https://api.semanticscholar.org/CorpusID:197663120`.

[Jec06] Thomas Jech. *Set Theory: The Third Millennium Edition, revised and expanded*. Springer Science and Business Media, 2006.

[KK11] Karin Usadi Katz and Mikhail G. Katz. "A Burgessian Critique of Nominalistic Tendencies in Contemporary Mathematics and its Historiography". In: *Foundations of Science* 17.1 (Mar. 2011), pp. 51–89. ISSN: 1572-8471. DOI: `10.1007/s10699-011-9223-1`. URL: `http://dx.doi.org/10.1007/s10699-011-9223-1`.

[Kle02] Stephen Cole Kleene. *Mathematical Logic*. en. Dover Books on Mathematics. Mineola, NY: Dover Publications, Dec. 2002.

[Laz12] Simon Lazarus. "INTERNAL SET THEORY AND AN INTUITIVE DEVELOPMENT OF THE CALCULUS". In: *UChicago REU* (2012).

[Loś55] Jerzy Loś. "Quelques Remarques, Théorèmes Et Problèmes Sur Les Classes Définissables D'algèbres". In: *Studies in logic and the foundations of mathematics* 16 (1955), pp. 98–113. URL: `https://api.semanticscholar.org/CorpusID:123785472`.

[Mos49] Andrzej Mostowski. "An undecidable arithmetical statement". eng. In: *Fundamenta Mathematicae* 36.1 (1949), pp. 143–164. URL: `http://eudml.org/doc/213187`.

[Nel77] Edward Nelson. "Internal set theory: a new approach to nonstandard analysis". In: *Bulletin of the American Mathematical society* 83.6 (1977), pp. 1165–1198.

[Neu23] John Von Neumann. "Zur Einführung der transfiniten Zahlen". In: *Sectio scientiarum mathematicarum* (1923).

[NGL18] Mauro Di Nasso, Isaac Goldbring, and Martino Lupini. *Nonstandard Methods in Ramsey Theory and Combinatorial Number Theory*. arXiv, 2018.

[Rob96] Abraham Robinson. *Non-standard Analysis*. Princeton University Press, 1996. URL: `http://www.jstor.org/stable/j.ctt1cx3vb6` (visited on 07/11/2026).

[Sco] Brian M. Scott. *Every non-principal ultrafilter contains a cofinite filter*. URL: `https://math.stackexchange.com/questions/476289/every-non-principal-ultrafilter-contains-a-cofinite-filter`.

[Vop89] Petr Vopěnka. *Mathematics in the Alternative Set Theory Teubner*. 1989.