SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Some Basic Results in Non-Standard Analysis

av

Milton Ernarp

2025 - No K10

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

Some Basic Results in Non-Standard Analysis

Milton Ernarp

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Gregory Arone

2025

Abstract

In this thesis, we lay the groundwork for non-standard analysis and derive some classic results from real analysis, through alternative methods. Non-standard analysis extends the real number line to the hyperreal number line, which we construct by introducing ultrafilters. The hyperreal number-line includes infinitesimal numbers, whose absolute value is less than any positive real number, and unlimited numbers, whose absolute value is greater than any real number. A fundamental result we prove is the transfer principle, which allows first-order statements about one structure, to be transferred to a corresponding statement about the other structure. Using transfer, we establish a basic theory of non-standard analysis, including proofs of classic theorems such as the Squeeze Theorem, the Intermediate Value Theorem, the Chain Rule, and the Fundamental Theorem of Calculus. We do this using infinitesimals — akin to the original methods of Leibniz — instead of the standard limit-based formulations. The thesis shows how non-standard methods can offer a more intuitive alternative for analysis.

Sammanfattning

I denna avhandling lägger vi grunden för icke-standardanalys och härleder några klassiska resultat från reell analys med hjälp av alternativa metoder. Icke-standardanalysen utvidgar den reella tallinjen till den hyperreella tallinjen, som vi konstruerar genom att introducera ultrafilter. Den hyperreella tallinjen innehåller infinitesimala tal, vars absolutbelopp är mindre än varje positivt reellt tal, samt obegränsade tal, vars absolutbelopp är större än varje reellt tal. Ett centralt resultat vi bevisar är överföringsprincipen, som möjliggör att första ordningens utsagor om den ena strukturen kan överföras till motsvarande utsagor om den annan struktur. Med hjälp av överföringsprincipen etablerar vi en grundläggande teori för icke-standardanalys, inklusive bevis för klassiska satser såsom instängningssatsen, satsen om mellanliggande värden, kedjeregeln och integralkalkylens fundamentalsats. Detta gör vi med hjälp av infinitesimaler – i Leibniz anda – snarare än genom standardmetoder baserade på gränsvärden. Avhandlingen visar hur icke-standardmetoder kan erbjuda ett mer intuitivt alternativ i analysens tjänst.

Acknowledgements

I would like to thank:

Gregory Arone, for generously letting me stop by his office so often, and for always asking insightful follow-up questions.

Sama Agahi, for lending me his copy of *Models and Ultraproducts*.

Annemarie Luger, for reminding me of the registration deadline, and letting me in despite a late registration.

Contents

1 Introduction

Our modern conceptions of the derivative and integral trace back to Newton and Leibniz, but if a modern math student went back to their seventeenth century writings, she would be very confused. Not only because of notational differences, but also because their methods differ fundamentally. In the seventeenth century, when Newton and Leibniz worked, our modern notion of limits - which we nowadays base analysis on - was still two hundred years away. Newton and Leibniz instead made use of non-zero *infinitesimals*, numbers that are infinitely small yet distinct from zero. Newton's view of infinitesimals was ambiguous because he was more interested in the physical applications of calculus; Leibniz by contrast tried to base calculus in a formal system which included infinitesimals [3]. For example, the derivative of a function $f(x)$, in modern notation, was defined as the ratio

$$\frac{f(x+\Delta x)-f(x)}{\Delta x}$$

where Δx is a non-zero infinitesimals. To demonstrate, the derivative of a function $f(x) = x^n$ would be computed as

$$\frac{(x+\Delta x)^n - x^n}{\Delta x} = \frac{x^n + nx^{n-1}\Delta x + \binom{n}{2}x^{n-2}\Delta x^2 + \ldots + \Delta x^n - x^n}{\Delta x} = nx^{n-1} + \binom{n}{2}x^{n-2}\Delta x + \ldots + \Delta x^{n-1} = nx^{n-1}$$

In the final step, since Δx is infinitesimal, any multiple of Δx would also be infinitesimal, hence all terms except nx^{n-1} are disregarded. This reasoning reaches the correct conclusion but the methods seem questionable. The first steps crucially uses that Δx is non-zero since it is in the denominator, yet in final step, all its multiples are disregarded, as if it were zero. Problems like this (and philosophical worries about the existence of infinitely small quantities) led to the eventual replacement of infinitesimals when the foundation of calculus was formalized in the early 19th century [3].

The infinitesimals were then largely abandoned, only to be used as pedagogical tools ... until Abraham Robinson introduced *non-standard analysis* in 1960 [1]. Robinson demonstrated that, with modern developments in mathematical logic, the real number line could be rigorously extended to the *hyperreal* number line, which includes infinitesimals (among other non-standard entities). Robinson also showed that within the hyperreals one can develop a theory of real-analysis, alternative to the standard way, hence the title.

In this paper, we present the construction of hyperreal numbers and show how the foundations of calculus can be developed with this framework. We begin by constructing the hyperreals via equivalence classes of infinite real-valued sequences, making crucial use of non-principal ultrafilters. We then develop the formal language and a natural translation form the reals to the hyperreals, in order to prove the Transfer Principle, a central result that ensures certain properties of the real numbers are preserved in the hyperreals and vice versa. Following this, we investigate the arithmetic and algebraic properties of the hyperreal number-line and its non-standard elements, to familiarize us with the structure we are working in. Finally, we apply these ideas to classical topics in analysis including convergence, continuity, differentiation, and integration, demonstrating how non-standard methods can be used to give alternative, and sometimes, more intuitive version of classic results in analysis.

2 Construction of The Hyperreals

In this section we provide a construction of the hyperreal numbers from the real numbers, based on the material in chapter 2 and 3 from [1] and chapter. We construct the hyperreals rather than simply postulating infinitesimals, because, as noted in the introduction, there is a risk that the concept of infinitesimals could be inconsistent. In that case, defining a structure containing infinitesimals might lead to contradictions, like defining N to be "the greatest integer" and deducing that $N \geq N+1$. If we however construct the hyperreals from structure we already "know" [1] are non-contradictory, we can be sure that it does not lead us to falsehood.

Robinson proved the existence of the hyperreals via the *compactness theorem*. A sketch of the construction goes as follows [1]. According to the compactness theorem, if for a set of (appropriately formalized) sentences Σ, for every finite subset $\Gamma \subseteq \Sigma$ there exists a structure which makes all the sentences in Γ true - a so called model of Γ - then the entirety of Σ has a model. Considered the set of true sentences about the real number $\Sigma_{\mathfrak{R}}$, and the infinite set of sentences $\{0 < \varepsilon\} \cup \{\varepsilon < \frac{1}{n} : n \in \mathbb{N}\}$. Every finite subset of $\Sigma_{\mathfrak{R}} \cup \{0 < \varepsilon\} \cup \{\varepsilon < \frac{1}{n} : n \in \mathbb{N}\}$ can be modeled by the real numbers, interpreting ε as $\frac{1}{k+1}$ where k is the largest n such that $\varepsilon < \frac{1}{n}$ is in the finite subset. Hence by the compactness theorem $\Sigma_{\mathfrak{R}} \cup \{0 < \varepsilon\} \cup \{\varepsilon < \frac{1}{n} : n \in \mathbb{N}\}$ has a model which satisfies the same sentences as the reals but also contains an infinitesimal ε.

If this is ones first interaction with the compactness theorem, this line of reasoning might seem like a trick; we are not shown this hyperreal structure, just ensured that it exists. We therefore opt for a more algebraic construction, based on equivalence classes of countably infinite sequences of real numbers, defined using a formalized notion of a property holding for 'most' indices in the sequence. Although this is more complicated than Robinson's' method, it provides an explicit construction of the hyperreals, giving us more insight into their structure. For the curious, this entire section is just a particular case of a broader topic called *Ultraproducts* in which one generalizes this to any structure over any index set. For more on the topic, see [2].

2.1 Filters

The goal of this and the next sub-section is to formalize an notion when a subset $X \subseteq I$ contains 'most' of I. The concept of a filters on sets is a beginning of formalizing this notion.

Definition 2.1 (Filter)**.** A filter $\mathcal{F}$ on a set I, is a subset of $\mathcal{P}(I)$ that is closed under finite intersections and supersets, i.e.

(i) If $X, Y \in \mathcal{F}$, then $X \cap Y \in \mathcal{F}$;

(ii) If $X \in \mathcal{F}$ and $X \subseteq Y \subseteq I$, then $Y \in \mathcal{F}$.

Additionally a filter $\mathcal{F}$ on I is called a *proper* filter if it's a proper subset of $\mathcal{P}(I)$

Closure under supersets is a desirable property: if a set contains most elements of a given set, then so should any of its supersets. The closure under finite intersection isn't quite as intuitive, but we see further on that it proves very useful.

Some trivial but useful properties of filters include

Proposition 2.2. For every filter $\mathcal{F}$, on a set I

(a) $\varnothing \in \mathcal{F}$ iff $\mathcal{F} = \mathcal{P}(I)$,

(b) $I \in \mathcal{F}$ iff $\mathcal{F} \neq \varnothing$,

(c) if $X \cap Y \in \mathcal{F}$, then $X, Y \in \mathcal{F}$.

Proof. The 'only if'-part of (a) follows from the fact that every set is a superset of the empty set, while the 'if'-part of (b) follows from the fact that every set in $\mathcal{P}(\mathcal{I})$ is a subset of I, by definition. For (c), it follows from $X \cap Y \subseteq X, Y$ for every X, Y. The converses of (a),(b) are trivial. □

Let's consider some examples of of filters on a set I to get a better grasp of filters

[1] We use quotation marks because, by Gödel's second incompleteness theorem, no system containing elementary arithmetic can prove its own consistency. Therefore, we can never be sure that the system in which we are working — ZFC, in this case — is consistent. However, by general consensus, it is assumed to be.

Example. 1. $\mathcal{P}(I)$ and $\varnothing$ are filters, all though degenerate cases.

2. Given an $i \in I$ let $\mathcal{F}^i = \{X \subseteq I : i \in X\}$. If $X, Y \in \mathcal{F}^i$, then $i \in X \cap Y$ so $X \cap Y \in \mathcal{F}^i$. Also if $X \subseteq Z$, then $i \in \mathbb{Z}$ so $Z \in \mathcal{F}^i$. Ergo $\mathcal{F}^i$ is a filter. Note that for any proper filter $\mathcal{H}$, if $i \in \mathcal{H}$, then $\mathcal{F}^i \subseteq \mathcal{H}$. Moreover if $\mathcal{H}$ contains a subset not contain i, then by being closed under intersection, it would contain $\varnothing$ hence not proper by **propostion 2.1**. Therefore $i \in \mathcal{H}$ iff $\mathcal{H} = \mathcal{F}^i$.

3. Let $\mathcal{F}^{co} = \{X \subseteq I : \overline{X} \text{ is finite}\}$, where $\overline{X}$ is the complement of X, i.e. $I - X$,. If $X, Y \in \mathcal{F}^{co}$, then $\overline{X \cap Y} = \overline{X} \cup \overline{Y}$ is finite so $X \cup Y \in \mathcal{F}^{co}$. In addition if $X \subseteq Z$, then $\overline{Z} \subseteq X$, hence $\overline{Z}$ is finite so $Z \in \mathcal{F}^{co}$. Ergo $\mathcal{F}^{co}$ is a filter. Note that $\mathcal{F}^{co}$ is only interesting on infinite sets; if I is finite, then $\mathcal{F}^{co} = \mathcal{P}(I)$.

4. Given $\mathcal{A} \subseteq \mathcal{P}(I)$ let $\mathcal{F}^{\mathcal{A}} = \{X \subseteq I : \bigcap_{A \in \mathcal{A}'} A \subseteq X \text{ for some finite } \mathcal{A}' \subseteq \mathcal{A}\}$ which is the closure of $\mathcal{A}$ under finite intersection and supersets, and is therefore a filter.

The idea behind $\mathcal{F}^{\mathcal{A}}$ is that for any $\mathcal{A} \subseteq I$, we can construct a minimal filter containing $\mathcal{A}$. To determine when $\mathcal{F}^{\mathcal{A}}$ is a proper filter we define the following

Definition 2.3 (Finite Intersection Property)**.** For $\mathcal{A} \subseteq \mathcal{P}(I)$ we say $\mathcal{A}$ has the *finite intersection property* (fip for short) if for every finite $\mathcal{A}' \subseteq \mathcal{A}$, the intersection $\bigcap_{A \in \mathcal{A}'} A$ is non-empty.

Proposition 2.4. $\mathcal{F}^{\mathcal{A}}$ is proper if and only if $\mathcal{A}$ has the fip

Proof. For the 'only if'-part, assume $\mathcal{F}^{\mathcal{A}}$ is proper. If $\mathcal{A}$ doesn't have the fip, then for some finite $\mathcal{A}' \subseteq \mathcal{A}$ we have $\bigcap_{A \in \mathcal{A}'} A = \varnothing \in \mathcal{F}^{\mathcal{A}}$. Therefore $\mathcal{F}^{\mathcal{A}}$ isn't proper by **proposition 2.1**, contradicting our origin assumption, so by reductio ad absurdum $\mathcal{F}^{\mathcal{A}}$ has the fip.

For the 'if'-part, assume $\mathcal{A}$ has the fip. Then $\bigcap_{A \in \mathcal{A}'} A \neq \varnothing$ for all finite $\mathcal{A}' \subseteq \mathcal{A}$, in particular every $A \in \mathcal{A}$ is thus non-empty. Therefore, since the empty set is the only superset of the empty set, $\mathcal{F}^{\mathcal{A}}$ never contains $\varnothing$. So, by **proposition 2.1**, $\mathcal{F}^{\mathcal{A}}$ must be proper. □

2.2 Ultrafilters

Further refining our notion of 'most' we consider a special kind of filters.

Definition 2.5 (Ultrafilter)**.** An *ultrafilter* $\mathcal{U}$ on I is an proper filter on I such that for every $X \subseteq I$, either $X \in \mathcal{U}$ or $\overline{X} \in \mathcal{U}$.

Remark. The 'or' is exclusive because if $X, \overline{X} \in \mathcal{U}$, then $X \cap \overline{X} = \varnothing \in \mathcal{U}$ which would contradict $\mathcal{U}$ being proper **by proportion 2.1**.

That ultrafilters are proper ensures that not every subset of I is considered to contain 'most' of the elements — that would render the notion meaningless. That either $X \in \mathcal{U}$ or $\overline{X} \in \mathcal{U}$ also aligns with our intuitions since one of them must contain most.

The following propositions we prove about ultrafilters correspond to logical inferences. We only note this here but there's a deep connection between ultrafilters and logic; for more see [2].

Proposition 2.6. Given an ultrafilter $\mathcal{U}$ the union $X \cup Y \in \mathcal{U}$ if and only if $X \in \mathcal{U}$ or $Y \in \mathcal{U}$.

Proof. For 'if'-part of assume $X \in \mathcal{U}$ or $Y \in \mathcal{U}$. Since both $X \subseteq X \cup Y$ and $Y \subseteq X \cup Y$, in both cases we have $X \cup Y \in \mathcal{U}$.

For the 'only if'-part assume $X \cup Y \in \mathcal{U}$ and assume, for the case of contradiction, that neither $X \in \mathcal{U}$ nor $Y \in \mathcal{U}$. Then $\overline{X}, \overline{Y} \in \mathcal{U}$ which implies $\overline{X} \cap \overline{Y} = \overline{X \cup Y} \in \mathcal{U}$. But then $(X \cup Y) \cap (\overline{X \cup Y}) = \varnothing \in \mathcal{U}$ which contradicts $\mathcal{U}$ being proper so, $X \in \mathcal{U}$ or $Y \in \mathcal{U}$. □

Proposition 2.7. Given an ultrafilter $\mathcal{U}$ on I and some pairwise disjoint sets $X_1, X_2, \ldots, X_n \subseteq I$ the following holds: $X_1 \cup X_2 \cup \cdots \cup X_n \in \mathcal{U}$ iff exactly one $X_i \in \mathcal{U}$.

Proof. For 'if'-part, assume exactly one $X_i \in \mathcal{U}$. Then since for each i, $X_i \subseteq X_1 \cup X_2 \cup \cdots \cup X_n$, we have $X_1 \cup X_2 \cup \cdots \cup X_n \in \mathcal{U}$ in every case.

For the 'only if'-part assume $X_1 \cup X_2 \cup \cdots \cup X_n \in \mathcal{U}$. By induction on the number of sets and **proposition 2.3** it follows that $X_1 \in \mathcal{U}$ or $X_2 \in \mathcal{U}$ or ... or $X_n \in \mathcal{U}$, so at least one $X_i \in \mathcal{U}$. If $X_i, X_j \in \mathcal{U}$ with $i \neq j$, then $X_i \cap X_j = \varnothing \in \mathcal{U}$ contradicting $\mathcal{U}$ being proper. Therefore at least 1 and at most 1 i.e. exactly one X_i is in $\mathcal{U}$. □

Let's return to our examples of filters on a set I to see if they are ultrafilters on I.

Example.

1. Both $\varnothing$ and $\mathcal{P}(I)$ are never ultrafilters.
2. $\mathcal{F}^i = \{X \subseteq I : i \in X\}$ is always an ultrafilter since for every $X \in I$ either $i \in X$ or $i \in \overline{X}$.
3. $\mathcal{F}^{co} = \{X \subseteq I : \overline{X} \text{ is finite}\}$ is not necessarily an ultrafilter since X and $\overline{X}$ may both be infinite e.g. the even and odd numbers in the integers.
4. $\mathcal{F}^{\mathcal{A}}$ is also not necessarily an ultrafilter. Take for example $\mathcal{A} = \{A\}$ it's not always the case that $A \subseteq X$ or $A \subseteq \overline{X}$ for every $X \subseteq I$.

An equivalent definition for ultrafilters is that they're maximal proper filters. This makes intuitive sense since the $X \in \mathcal{U}$ or $\overline{X} \in \mathcal{U}$ can be seen as 'stuffing a proper filter up to just bellow it's breaking point'. We formally prove this equivalence as follows.

Theorem 2.8. *A filter $\mathcal{U}$ is an ultrafilter if and only if $\mathcal{U}$ is a maximal proper filter with respect to $\subseteq$.*

Proof. For the 'only if'-part assume $\mathcal{U}$ is an ultrafilter, and for the sake of contradiction that $\mathcal{U} \subset \mathcal{F} \subset \mathcal{P}(\mathbb{N})$ such that $\mathcal{F}$ is a proper filter. Then for some $X \subseteq I$ we have $X \in \mathcal{F}$ and $X \notin \mathcal{U}$. It follows that $\overline{X} \in \mathcal{U}$ so by $\mathcal{U} \subset \mathcal{F}$ we get $\overline{X} \in \mathcal{F}$. But then $X \cap \overline{X} = \varnothing \in \mathcal{F}$, contradicting it being proper. Ergo $\mathcal{U}$ is a maximal proper filter with respect to $\subseteq$.

For the 'if'-part assume $\mathcal{U} \not\subset \mathcal{F}$ for every proper filter $\mathcal{F}$ and for the sake of contradiction, that $\mathcal{U}$ isn't an ultrafilter i.e. $X, \overline{X} \notin \mathcal{U}$ for some $X \subseteq I$. Consider $\mathcal{U} \cup \{X\}$, every finite intersection of $\mathcal{U} \cup \{X\}$ is either a finite intersection of $\mathcal{U}$ so non-empty by $\mathcal{U}$ being a proper filter, or $X \cap Y$ for some $Y \in \mathcal{U}$. If $X \cap Y = \varnothing$, then $Y \subseteq \overline{X}$ which implies $\overline{X} \in \mathcal{U}$ contradicting our assumption, so $\mathcal{U} \cup \{X\}$ must have the fip. Then $\mathcal{U} \subset \mathcal{F}^{\mathcal{U} \cup \{X\}}$ is a proper filter by **proposition 2** contradicting our assumption that $\mathcal{U}$ is maximal. Ergo $\mathcal{U}$ is an ultrafilter □

Although we have shown that every maximal proper filter is an ultrafilter, it is not obvious that every proper filter is contained in one. To establish this, we assume **Zorn's Lemma**.

Lemma 2.9 (*Zorn's Lemma*)**.** For any partially order set P, If every totally order subset (a so called chain) of P has an upper bound in P, then P has a maximal element.

With **Zorn's lemma** we can prove

Proposition 2.10. Every proper filter $\mathcal{F}$ over I is contained in some ultrafilter $\mathcal{U}$.

Proof. Let P be the set of proper filters on I and consider $P_{\mathcal{F}} = \{\mathcal{H} \in P : \mathcal{F} \subseteq \mathcal{H}\}$ which is a po-set under $\subseteq$. Let $\{\mathcal{H}_n\}_{n \in N} \subseteq P_{\mathcal{F}}$, where N is an index set, be a chain. Consider the union $\bigcup_{n \in N} \mathcal{H}_n$. It's clearly an upper bound since $\mathcal{H}_k \subseteq \bigcup_{n \in N} \mathcal{H}_n$ for every $k \in N$. If $X, Y \in \bigcup_{n \in N} \mathcal{H}_n$, then $X \in \mathcal{H}_k$, $Y \in \mathcal{H}_\ell$ for some $k, \ell \in N$. By $\{\mathcal{H}_n\}_{n \in N}$ being totally ordered we know $\mathcal{H}_k \subseteq \mathcal{H}_\ell$ in which case $X, Y \in \mathcal{H}_\ell$, or $\mathcal{H}_\ell \subseteq \mathcal{H}_k$ in which case $X, Y \in \mathcal{H}_k$. In either case $X \cap Y$ is in some $\mathcal{H}_i$ so $X \cap Y \in \bigcup_{n \in N} \mathcal{H}_n$. So $\bigcup_{n \in N} \mathcal{H}_n$ is closed under union. If $X \in \bigcup_{n \in N} \mathcal{H}_n$ and $X \subseteq Y \subseteq I$, then $X \in \mathcal{H}_k$ and consequently $Y \in \mathcal{H}_k$ for some $k \in N$. It then follows that $Y \in \bigcup_{n \in N} \mathcal{H}_n$ so $\bigcup_{n \in N} \mathcal{H}_n$ is closed under finite intersection and superset, hence a filter. It's also a proper filter since if $\varnothing \in \bigcup_{n \in N} \mathcal{H}_n$, then $\varnothing \in \mathcal{H}_k$ for some $k \in N$, contradicting $\mathcal{H}_k$ being proper. Because $\mathcal{F} \in \mathcal{H}_k \subseteq \bigcup_{n \in N} \mathcal{H}_n$ we then have $\bigcup_{n \in N} \mathcal{H}_n \in P_{\mathcal{F}}$. Applying *Zorn's lemma*, there is a maximal proper filter $\mathcal{U} \in P_{\mathcal{F}}$, which by the definition of $P_{\mathcal{F}}$ contains $\mathcal{F}$, and by **theorem 2.1** is an ultrafilter. □

For the final refinement of our 'most' formalization, we consider a special case of ultrafilters, namely non-principal ones. This is for their following properties

Theorem 2.11. *An ultrafilter $\mathcal{U}$ is principal if and only if it contains a finite set.*

Proof. For the 'if'-part, assume a finite set $\{a_1, \dots, a_n\}$ is in $\mathcal{U}$, then since $\{a_1, \dots, a_n\} = \{a_1\} \cup \dots \cup \{a_n\}$ is a finite disjoint union, by **proposition 4**, $\{a_i\} \in \mathcal{U}$, for exactly one a_i. Since $\mathcal{U}$ is proper, $\{a_i\} \in \mathcal{U}$ then implies $\mathcal{U} = \mathcal{F}^i$.

The 'only if'-part follows from $\{i\} \in \mathcal{F}^i$. □

Corollary 2.12. An ultrafilter $\mathcal{U}$ is non-principal if and only if $\mathcal{F}^{co} \subseteq \mathcal{U}$

Proof. For the 'if'-part, if $\mathcal{F}^{co} \subseteq \mathcal{U}$, then $\mathcal{U}$ contains all compliments of finite sets. Therefore $\mathcal{U}$ contains no finite sets, so by **theorem 2.2**, $\mathcal{U}$ is non principal.

For the 'only if'-part, assume an ultrafilter $\mathcal{U}$ is non-principal. If $X \in \mathcal{F}^{co}$, then $\overline{X}$ is finite so by **theorem 2** $\overline{X} \notin \mathcal{U}$ which implies $X \in \mathcal{U}$. Ergo $\mathcal{F}^{co} \subseteq \mathcal{U}$. □

That every co-finite set contains 'most' elements is very sensible property. We thus arrive at non-principal ultrafilter formalizing our notion of 'most'.

2.3 The Ring of Infinite Sequences Modulo $\mathcal{U}$

Let $\mathbb{R}^{\mathbb{N}}$ denote the set of functions $\mathbb{N} \to \mathbb{R}$, or as we view it, the set of infinite sequences of real numbers writing $r = (r_n) = (r_1, r_2, r_3, \ldots) \in \mathbb{R}^{\mathbb{N}}$. We also use bold font to denote constants sequence, so for every real r, $\mathbf{r}$ denotes the sequence $(r, r, r, \ldots) \in \mathbb{R}^{\mathbb{N}}$

Remark. In this paper we do not consider 0 to be a natural number so $\mathbb{N} = \{1, 2, 3, \ldots\}$.

Defining addition and multiplication component wise as

$$(r_1, r_2, r_3, \ldots) \oplus (s_1, s_2, s_3, \ldots) = (r_1 + s_1, r_2 + s_2, r_3 + s_3, \ldots) \text{ and}$$
$$(r_1, r_2, r_3, \ldots) \otimes (s_1, s_2, s_3, \ldots) = (r_1 \cdot s_1, r_2 \cdot s_2, r_3 \cdot s_3, \ldots),$$

makes $(\mathbb{R}^{\mathbb{N}}, \oplus, \otimes)$ a ring with zero $\mathbf{0}$, one $\mathbf{1}$ and $-(r_1, r_2, r_3, \ldots) = (-r_1, -r_2, -r_3, \ldots)$. Note that $(\mathbb{R}^{\mathbb{N}}, \oplus, \otimes)$ is not a field since there are zero-divisors e.g $(0, 1, 0, \ldots) \otimes (1, 0, 0, \ldots) = \mathbf{0}$ but neither $(0, 1, 0, \ldots)$ or $(1, 0, 0, \ldots)$ are $\mathbf{0}$.

By **theorem 2.5** $\mathcal{F}^{co}$ on $\mathbb{N}$ is contained in some ultrafilter $\mathcal{U}$ which, by **corollary 2.1** is non-principal. We define the relation $\equiv_{\mathcal{U}}$ on $\mathbb{R}^{\mathbb{N}}$ as

$$r \equiv_{\mathcal{U}} s \text{ if and only if } \{n \in \mathbb{N} : r_n = s_n\} \in \mathcal{U}.$$

Since the subsets of $\mathcal{U}$ contain most natural numbers, $r \equiv_{\mathcal{U}} s$ holds iff r and s are equal in most components. For ease of notation we introduce $[\![r = s]\!] = \{n \in \mathbb{N} : r_n = s_n\}$. One should think of equality between sequences as not binary, but rather measured on a spectrum from $[\![r = s]\!] = \varnothing$ (equal in no components) to $[\![r = s]\!] = \mathbb{N}$ (equal in all components).

Proposition 2.13. The relation $\equiv_{\mathcal{U}}$ is an equivalence relation.

Proof. The relation reflexive since $[\![r = r]\!] = \mathbb{N} \in \mathcal{U}$, so $r \equiv_{\mathcal{U}} r$ for all $r \in \mathbb{R}^{\mathbb{N}}$.

The relation is symmetric since $[\![s = r]\!] = [\![r = s]\!]$, so if $s \equiv_{\mathcal{U}} r$, then $r \equiv_{\mathcal{U}} s$ for any $r, s \in \mathbb{R}^{\mathbb{N}}$.

Finally for transitivity, assume $r \equiv_{\mathcal{U}} s$ and $s \equiv_{\mathcal{U}} t$, then $[\![r = s]\!], [\![s = t]\!] \in \mathcal{U}$. Because $[\![r = s]\!] \cap [\![s = t]\!] \subseteq [\![r = t]\!]$, it follows that $[\![r = t]\!] \in \mathcal{U}$, so $r \equiv_{\mathcal{U}} t$. □

Denoting the equivalence class of r as $[r] = \{s \in \mathbb{R}^{\mathbb{N}} : r \equiv_{\mathcal{U}} s\}$ we consider the quotient ring

$$\mathbb{R}^{\mathbb{N}} / \equiv_{\mathcal{U}} = \{[r] : r \in \mathbb{R}^{\mathbb{N}}\}$$

which we denote by ${}^*\mathbb{R}$. We define the operations $+$ and $\cdot$ on ${}^*\mathbb{R}$ as

$$[r] + [s] = [r \oplus s] \quad \text{and} \quad [r] \cdot [s] = [r \otimes s].$$

Remark. These operations are well-defined. Let $r \equiv_{\mathcal{U}} r', s \equiv_{\mathcal{U}} s'$, then $[\![r = r']\!], [\![s = s']\!] \in \mathcal{U}$ and consequently $[\![r = r']\!] \cap [\![s = s']\!] \in \mathcal{U}$. The intersection is both a subset of $[\![r \oplus s = r' \oplus s']\!]$ and $[\![r \otimes s = r' \otimes s']\!]$, it follows that $[\![r \oplus s = r' \oplus s']\!], [\![r \otimes s = r' \otimes s']\!] \in \mathcal{U}$. Thus $[r] + [s] = [r'] + [s']$ and $[r] \cdot [s] = [r'] \cdot [s']$, so addition and multiplication don't depend on equivalence class representative

We define an order on ${}^*\mathbb{R}$ as

$$[r] < [s] \text{ if and only if } [\![r < s]\!] := \{n \in \mathbb{N} : r_n < s_n\} \in \mathcal{U}.$$

Reusing the $[\![\cdot]\!]$ notation we again get intuition that $[r]$ is less than $[s]$ iff r is s for most indices. With this ordering we get

Proposition 2.14. The structure $({}^*\mathbb{R}, +, \cdot, <)$ is an order field

Proof. Since ${}^*\mathbb{R}$ is a quotient ring of $\mathbb{R}^{\mathbb{N}}$ by a congruent equivalence, it's also a ring with zero $[\mathbf{0}]$, one $[\mathbf{1}]$ and $-[r] = [-r]$. To show ${}^*\mathbb{R}$ is field suppose $[r] \neq [\mathbf{0}]$ and let

$$s_n = \begin{cases} \frac{1}{r_n} & \text{if } n \in [\![r \neq 0]\!] \\ 0 & \text{otherwise.} \end{cases}$$

Then $[\![r \otimes s = \mathbf{1}]\!] = [\![r \neq 0]\!] \in \mathcal{U}$ so $[r] \cdot [s] = [r \otimes s] = [\mathbf{1}]$, ergo $[s] = [r]^{-1}$.

The order $<$ induced on ${}^*\mathbb{R}$ is linear since for any $r, s \in \mathbb{R}^{\mathbb{N}}$, the union $[\![r < s]\!] \cup [\![r = s]\!] \cup [\![r > s]\!] = \mathbb{N}$ is finite and disjoint, so by **proposition 2.4** exactly one of them must be in $\mathcal{U}$. Thus $[r] < [s]$ or $[r] = [s]$ or $[r] > [s]$ for all $[r], [s] \in {}^*\mathbb{R}$ so the order is linear. Finally the positive Hyperreals are closed under addition and multiplication since $[\![0 < r]\!] \cap [\![0 < s]\!]$ is a subset of $[\![0 < r + s]\!]$ and $[\![0 < r \cdot s]\!]$ so if $[r]$ and $[s]$ are positive, then $[r] + [s]$ and $[r] \cdot [s]$ are positive. □

We call this ordered filed *the hyperreal numbers.*

2.4 Hyperreals Numbers

There's a natural inclusion of $\mathbb{R}$ in ${}^*\mathbb{R}$ by identifying x with ${}^*x = [\mathbf{x}]$. We call the map $x \mapsto {}^*x$ the *natural translation.* Note that

$$\begin{aligned}{}^*(x+y) &= [(x+y, x+y, x+y, \ldots)] \\ &= [(x,x,x,\ldots)] + [(y,y,y,\ldots)] \\ &= {}^*x + {}^*y \text{ and} \\ {}^*(x \cdot y) &= [(x\cdot y, x\cdot y, x\cdot y, \ldots)] \\ &= [(x,x,x,\ldots)] + [(y,y,y,\ldots)] \\ &= {}^*x \cdot {}^*y.\end{aligned}$$

Furthermore if $x < y$, then $[\![\mathbf{x} < \mathbf{y}]\!] = \mathbb{N} \in \mathcal{U}$ so ${}^*x < {}^*y$. Finally if ${}^*x = {}^*y$, then their equal for some component. Since both ${}^*x$ and ${}^*y$ are equivalence classes of constants sequences, it follows that they're equal for all components. Hence $x = y$ so the map is injective. We've thereby proven

Proposition 2.15. The map $x \mapsto {}^*x$ is a field embedding of $\mathbb{R}$ into ${}^*\mathbb{R}$. □

It follows that there is a isomorphic copy of $\mathbb{R}$ in ${}^*\mathbb{R}$. The question arises if ${}^*\mathbb{R}$ is proper extension of $\mathbb{R}$ or if just we've reinvented the reals. Here our assumption that $\mathcal{U}$ is non-principal becomes crucial.

Proposition 2.16. The quotient ring $\mathbb{R}^{\mathbb{N}}/\equiv_{\mathcal{U}}$ is isomorphic to $\mathbb{R}$ if and only if the ultrafilter $\mathcal{U}$ is principal.

Proof. For the 'if'-part, assume is $\mathcal{U}$ principal so $\mathcal{U} = \{X \subseteq \mathbb{N} : k \in X\}$ for some $k \in \mathbb{N}$. Then for any sequence $r = (r_1, r_2, r_3 \ldots) \in \mathbb{R}^{\mathbb{N}}$ we have $k \in [\![(r_1, \ldots, r_k, \ldots) = (r_k, r_k, r_k, \ldots)]\!] = [\![r = \mathbf{r}_k]\!]$, so $[\![r = {}^*r_k]\!] \in \mathcal{U}$ by $\mathcal{U}$ being principally generated by k. Consequently every $[r] \in {}^*\mathbb{R}$ equals ${}^*r_k$. Hence $x \mapsto {}^*x$ is surjective and thereby an isomorphism between $\mathbb{R}$ and $\mathbb{R}^{\mathbb{N}}/\equiv_{\mathcal{U}}$.

For the 'only if'-part, assume by contraposition that $\mathcal{U}$ is non principal. Consider the sequence $(\frac{1}{n}) = (1, \frac{1}{2}, \frac{1}{3}, \ldots) \in \mathbb{R}^{\mathbb{N}}$. The set $[\![(\frac{1}{n}) > \mathbf{0}]\!] = \mathbb{N} \in \mathcal{U}$ so $[0] < [(\frac{1}{n})]$, but for every $x \in \mathbb{R}^+$ there is a natural number m such that $x > \frac{1}{m}$ so $[\![\frac{1}{n} \geq \mathbf{x}]\!]$ is always finite. So $\overline{[\![(\frac{1}{n}) \geq \mathbf{x}]\!]} = [\![(\frac{1}{n}) < \mathbf{x}]\!] \in \mathcal{F}^{co}$ for every $x \in \mathbb{R}^+$. By **corollary 2.12**, $\mathcal{F}^{co} \subseteq \mathcal{U}$, we thus have an element $\varepsilon = [(\frac{1}{n})]$ such that $0 < \varepsilon < {}^*x$ for every $x \in \mathbb{R}^+$. This isn't the case any real number, so an isomorphism between $\mathbb{R}^{\mathbb{N}}/\equiv_{\mathcal{U}}$ and $\mathbb{R}$ cannot exist. □

By analogous argument, $\varepsilon^{-1} = [(n)]$ is greater than every $x \in \mathbb{R}$, so in addition to the *standard* real numbers, the hyperreals also contain numbers greater than zero but lesser than every positive, and numbers greater than every real number. We explore these *non-standard* numbers in **section 4**.

Remark. An attentive reader might have noticed that we've been sloppy by saying *the* hyperreal numbers when we haven't specified our non-principal ultrafilter $\mathcal{U}$. The quotient ring might be different depending on which $\mathcal{U}$ we chose. To guarantee that every hyperreal number-line is isomorphic we would need to assume *the Continuum Hypothesis* according to [1] section 3.16. However, we show in **section 3.4** that, for practical purposes, the choice of non-principal ultrafilter doesn't matter.

3 Transfer

Now that we've constructed the hyperreal number-line, one might ask: "Who cares about hyperreal numbers?! I only care about the *real* real numbers". At first glance, the hyperreals might appear to be just a neat demonstration that infinitesimals are not contradictory — but they are much more than that.

The goal of this section is to, based on material from chapter 3 of [1] and chapter 5 of [2], prove *the transfer principle.* The transfer principle allows one to transfer any (appropriately formalized) statement about the structure of the real numbers, to a corresponding statement about structure of the hyperreals, and vice versa; results in the hyperreals are therefore directly relevant to the real numbers. The transfer principle allows us to set aside the talk of ultrafilters, in place of a more natural way to talk about the hyperreals - analogous to how the supremum axiom let's one avoid talking about the real numbers in terms of Dedekind cuts. To get there we first need a formal notion of what a 'structure of the reals/hyperreals' and a 'sentence about the reals/hyperreals' means, which we get through a quick detour through first-order logic. We then extend the *natural translation* to apply to more than just numbers, before proving the transfer principle.

3.1 The Relational Structure and First-Order Language of the Reals

We begin by defining relation structures, within which sentences can be true or false.

Definition 3.1 (Relational Structure)**.** A relational structure $\mathfrak{A}$ is a pair $(A, Rel_{\mathfrak{A}})$ where A is a non-empty set, called the *domain* or *universe* of $\mathfrak{A}$, and $Rel_{\mathfrak{A}}$ is a set of k-ary relations on A. By a k-ary relation on A we mean a subset of A^k.

For our relational structure of the reals, denoted by $\mathfrak{R}$, we take $\mathbb{R}$ as the domain, and the crude approach of taking $Rel_{\mathfrak{R}} = \bigcup_{n\in\mathbb{N}} \mathcal{P}(\mathbb{R}^n)$, including all possible n-ary relations, for every $n \in \mathbb{N}$. This way we include all interesting relations, like the binary identity relation $\mathrm{Id}_{\mathbb{R}} = \{(x, x) : x \in \mathbb{R}\}$, all subset of $\mathbb{R}$ as unary relations, and for any real-valued k-ary function f, its graph $G_f = \{(x_1, \dots, x_k, f(x_1, \dots, x_k)) : x_1, \dots, x_k \in D_f\}$. We also have a lot 'junk' in $Rel_{\mathfrak{R}}$, for example the relation $\{(\pi, 42), (\frac{3}{5}, 100^{100})\}$ but too much is better then too little.

Based on $\mathfrak{R}$ we inductively define the set of $\mathcal{L}_{\mathfrak{R}}$-formulas in the following way

Definition 3.2 ($\mathcal{L}_{\mathfrak{R}}$-formulas)**.**

- For every k-ary $R \in Rel_{\mathfrak{R}}$, the expression $Rt_1 \dots t_k \in \mathcal{L}_{\mathfrak{R}}$-formulas where $t_1, \dots, t_k$ are either real numbers or variables $x, y, z, \dots$. We call formulas of this form *atomic.*
- If $\varphi \in \mathcal{L}_{\mathfrak{R}}$-formulas, then $\neg\varphi \in \mathcal{L}_{\mathfrak{R}}$-formulas
- If $\varphi, \psi \in \mathcal{L}_{\mathfrak{R}}$-formulas, then $(\varphi \wedge \psi) \in \mathcal{L}_{\mathfrak{R}}$-formulas.
- If $\varphi \in \mathcal{L}_{\mathfrak{R}}$-formulas and x is a variable, then $\exists x\varphi \in \mathcal{L}_{\mathfrak{R}}$-formulas.

Throughout this section we use $\varphi, \psi, \chi, \dots$ as variables for formulas.

Remark. Induction on $\mathcal{L}_{\mathfrak{R}}$-formulas is possible by proving that a property P holds for all atomic formulas and proving that, if P holds for φ and ψ, then P holds for $\neg\varphi, (\varphi \wedge \psi)$ and $\exists x\varphi$.

To give interpenetration rules we must first distinguish between *free* and *bound* variables. A variable x occurs bounded in φ if it's under the scope of a quantifier and free if it isn't bounded. If a formula contains free variables we cannot assign a truth-value to it since free variables do not refer to anything. Formally we can define the set of free variables occurring in a formula as follows

$$\begin{aligned}
\text{For any constant } c, &\quad Fv(c) = \varnothing,\\
\text{for any variable } x, &\quad Fv(x) = \{x\},\\
\text{for any atomic formula } Rt_1 \dots t_k &\quad Fv(Rt_1 \dots t_k) = Fv(t_1) \cup \dots \cup Fv(t_k),\\
\text{for any negation } \neg\varphi &\quad Fv(\neg\varphi) = Fv(\varphi),\\
\text{for any conjunction } \varphi \wedge \psi &\quad Fv(\varphi \wedge \psi) = Fv(\varphi) \cup Fv(\psi),\\
\text{for any existential formula } \exists x\varphi &\quad Fv(\exists x\varphi) = Fv(\varphi) - \{x\}.
\end{aligned}$$

We call a formula with no free variables a *sentence* and denote the set of sentences by $\mathcal{L}_{\mathfrak{R}}$-sentences. Given a formula φ, we write $\varphi(x_1, \dots, x_p)$ to indicate that *at most* the variables $x_1, \dots, x_p$ occur free in φ. We use $\varphi(t_1, \dots, t_p)$ to denote the resulting formula of *simultaneously* substituting each free occurrence of x_i in φ with the corresponding term t_i, for $i = 1, \dots, p$.
We define the truth of a sentence in a structure inductively in the following way.

Definition 3.3 (Truth of $\mathcal{L}_{\mathfrak{A}}$-sentences in $\mathfrak{A}$)**.**

- An atomic formula $Rt_1 \dots t_k$ is true in a structure $\mathfrak{A}$ iff $(t_1, \dots, t_k) \in R$.
- A negation $\neg\varphi$ is true in $\mathfrak{A}$ iff φ is not true, i.e. false, in $\mathfrak{A}$.
- A conjunction $(\varphi \wedge \psi)$ is true in $\mathfrak{A}$ iff φ is true in $\mathfrak{A}$ and ψ is true in $\mathfrak{A}$.
- An existential sentence $\exists x\varphi(x)$ is true in $\mathfrak{A}$ iff $\varphi(a)$ is true in $\mathfrak{A}$ for some $a \in A$.

By convention the other logical connectives are defined as

$$\begin{aligned}(\varphi \vee \psi) &:= \neg(\neg\varphi \wedge \neg\psi),\\ (\varphi \to \psi) &:= \neg\varphi \vee \psi,\\ (\varphi \leftrightarrow \psi) &:= (\varphi \to \psi) \wedge (\psi \to \varphi),\\ \forall x\varphi &:= \neg\exists x\neg\varphi.\end{aligned}$$

We also define expressions containing k-ary function-symbols $\varphi(f(t_1,\ldots,t_k))$, where $(t_1,\ldots,t_k) \in D_f$, as $\exists x(G_f t_1 \ldots t_k x \wedge \varphi(x))$. We also define bounded quantification as $(\forall\varphi(x_1,\ldots x_p))\psi := \forall x_1 \ldots \forall x_p(\varphi(x_1,\ldots x_p) \to \psi)$ and $\exists\varphi(x_1,\ldots x_p)\psi = \exists x_1 \ldots \exists x_p(\varphi(x_1,\ldots x_p) \wedge \psi)$.

We drop the parentheses when there's no risk of misunderstanding e.g. writing $\varphi \wedge \psi \wedge \chi$ instead of $((\varphi \wedge \psi) \wedge \chi)$. Since unary relations are just subsets of $\mathbb{R}$, we write $t \in R$ instead of Rt. When appropriate we use infix notation and familiar symbols for some relations e.g. writing $t = s$ instead of $\mathrm{Id}_{\mathbb{R}}st$.

3.2 Extending the Natural Translation

To get from $\mathfrak{R}$ to the relational structure of the hyperreals, denoted by ${}^*\mathfrak{R}$, we first extend our natural translation function from **proposition 2.8** to apply to real relations as-well.

Definition 3.4 (Translations of Relations)**.** Given a k-ary relation on $\mathbb{R}$ we define the its natural translation ${}^*R$, as the k-ary relation on ${}^*\mathbb{R}$ defined by

$$([r^1],\ldots,[r^k]) \in {}^*R \text{ if and only if } [\![(r^1,\ldots,r^k) \in R]\!] := \{n \in \mathbb{N} : (r^1_n,\ldots,r^k_n) \in R\} \in \mathcal{U}.$$

where $\mathcal{U}$ is the non-principal ultrafilter on $\mathbb{N}$ from **section 2.3**. Note that this definition is consistent with how we previously defined equality and less the ordering in the hyperreals and generalizes the 'true for most indecies'-intuition to any relation. It is well-defined since, if $([r^1],\ldots,[r^k]) \in {}^*R$ and $([r^1],\ldots,[r^k]) = ([s^1],\ldots,[s^k])$, then $[\![r^1 = s^1]\!],\ldots,[\![r^k = s^k]\!], [\![(r^1,\ldots,r^k) \in R]\!] \in \mathcal{U}$ and because $[\![(r^1,\ldots,r^k) \in R]\!] \cap [\![r^1 = s^1]\!] \cap \ldots \cap [\![r^k = s^k]\!] \subseteq [\![(s^1,\ldots,s^k) \in R]\!]$ it follows that $[\![(s^1,\ldots,s^k) \in R]\!] \in \mathcal{U}$ so $([s^1],\ldots,[s^k]) \in {}^*R$.

We shall now investigate the translation of unary relations, i.e real sets $X \subseteq \mathbb{R}$, but these results can be generalized to relation of an arbitrary arity. If $x \in X$, then $[\![\mathbf{x} \in X]\!] = \mathbb{N}$ so ${}^*x \in {}^*X$ while if $x \notin X$, then $[\![\mathbf{x} \in X]\!] = \varnothing$ so ${}^*x \notin {}^*X$. Identifying x with ${}^*x$, we get ${}^*X \cap \mathbb{R} = X$ and $X \subseteq {}^*X$. This does not rule out the possibility that ${}^*X$ is a strict superset of X - non-standard elements could be introduced during the translation. In fact, we know precisely for which subsets that is the case.

Proposition 3.5. For any $X \subseteq \mathbb{R}$, $X = {}^*X$ if and only if X is finite

Proof. For the 'if'-part assume $X = \{x_1,\ldots,x_n\}$ and assume $[r] \in {}^*X$. Then finite union $[\![r = \mathbf{x}_1]\!] \cup \cdots \cup [\![r = \mathbf{x}_n]\!] = [\![r \in X]\!] \in \mathcal{U}$ but $[\![r = \mathbf{x}_i]\!] \cap [\![r = \mathbf{x}_j]\!] = \varnothing$ for every $x_i \neq x_j$, so **proposition 2.7**, $[\![r = \mathbf{x}_i]\!] \in \mathcal{U}$ for exactly one $x_i \in X$. Thus ${}^*X \subseteq X$ and since the converse always holds, $X = {}^*X$.

For the 'only if'-part assume by contraposition that X is an infinite set $\{x_1, x_2, x_3, \ldots\}$. Since X is infinite we may construct a never repeating real sequence s, where each $s_n \in X$. It follows that $[\![s \in X]\!] = \mathbb{N} \in \mathcal{U}$ so $[s_n] \in X$ however, for any $x \in \mathbb{R}$, the set $[\![s = \mathbf{x}]\!]$ is either singleton, or empty, since s is non-repeating. In either case, $[\![\mathbf{x} = s]\!]$ is finite so $\overline{[\![\mathbf{x} = s]\!]} = [\![\mathbf{x} \neq s]\!] \in \mathcal{U}$, for every $x \in \mathbb{R}$. Therefore $[s_n] \in {}^*X$ is non-standard, so $X \subset {}^*X$ and $X \neq {}^*X$. □

The natural translation preserve most set theoretic operations.

Proposition 3.6. Let $X, Y \subseteq \mathbb{R}$, then

1. $X \subseteq Y$ iff ${}^*X \subseteq {}^*Y$
2. $X = Y$ iff ${}^*X = {}^*Y$
3. ${}^*(X \cup Y) = {}^*X \cup {}^*Y$
4. ${}^*(X \cap Y) = {}^*X \cap {}^*Y$
5. ${}^*(X - Y) = {}^*X - {}^*Y$, in particular ${}^*(\overline{X}) = {}^*(\mathbb{R} - X) = {}^*\mathbb{R} - {}^*X = \overline{({}^*X)}$, so we may write ${}^*\overline{X}$ without risk of confusion

6. ${}^*(X^+) = ({}^*X)^+$ where $X^+ = \{x \in X : x > 0\}$, so we may write ${}^*X^+$ without risk of confusion
7. ${}^*(X_1 \times \cdots \times X_k) = {}^*X_1 \times \cdots \times {}^*X_k$, in particular ${}^*(X^k) = ({}^*X)^k$, so we may write ${}^*X^k$ without risk of confusion.

Proof.

1. For the if-part assume ${}^*X \subseteq {}^*Y$. If $x \in X$, then $x \in {}^*X$ so by assumption $x \in {}^*Y$ which implies $x \in \mathbb{R} \cap {}^*Y = Y$. For only-if part, if $X \subseteq Y$, then $[\![r \in {}^*X]\!] \subseteq [\![r \in {}^*Y]\!]$ so if $x \in {}^*X$, then $x \in {}^*Y$.
2. Since $X = Y$ iff $X \subseteq Y$ and $Y \subseteq X$ this follows from 1.
3. Follows from the fact that $[\![r \in X \cup Y]\!] = [\![r \in X]\!] \cup [\![r \in Y]\!]$ and **proposition 2.7**.
4. Follows from the fact that $[\![r \in X \cap Y]\!] = [\![r \in X]\!] \cap [\![r \in Y]\!]$ and **proposition 2.2 (c)**.
5. Follows from the fact that $[\![r \in X - Y]\!] = [\![r \in X]\!] \cap [\![r \notin Y]\!] = [\![r \in X]\!] \cap \overline{[\![r \in Y]\!]}$ and **proposition 2.2 (c)**.
6. Follows from the fact that $[\![r \in X^+]\!] = [\![r \in X]\!] \cap [\![r > 0]\!]$ and **proposition 2.2 (c)**.
7. Follows from $[\![(r^1, \ldots, r^k) \in X_1 \times \cdots \times X_k]\!] = [\![r^1 \in X_1]\!] \cap \cdots \cap [\![r^k \in X^k]\!]$ and induction on **proposition 2.2 (c)**.

□

Remark. It follows from 3 and induction on the number of sets that ${}^*(\bigcup_{n=1}^N X_n) = \bigcup_{n=1}^N {}^*X_n$ for any $N \in \mathbb{N}$ but this doesn't hold in the infinite case. Take for example $\bigcup_{n=1}^\infty {}^*\{n\} = \bigcup_{n=1}^\infty \{n\} = \mathbb{N}$ but ${}^*(\bigcup_{n=1}^\infty \{n\}) = {}^*\mathbb{N}$ which is distinct by **proposition 3.1**.

Definition 3.7 (Translation of Total Functions)**.** Given a total k-ary function $f : \mathbb{R}^k \to \mathbb{R}$, we define its natural translation as the function

$$\begin{gathered} {}^*f : {}^*\mathbb{R}^k \to {}^*\mathbb{R}, \\ ([x^1], \ldots, [x^k]) \mapsto [((f(x_n^1, \ldots, x_n^k))] = [(f(x_1^1, \ldots, x_1^k), f(x_2^1, \ldots, x_2^k), \ldots)]. \end{gathered}$$

This definition aligns with our way of identifying functions with their graph.

Proposition 3.8. For any function $f : \mathbb{R}^k \to \mathbb{R}$ and any $[x^1], \ldots, [x^k], [y] \in {}^*\mathbb{R}$ we have

$$([x^1], \ldots, [x^k], [y]) \in_f \quad \text{if and only if } {}^*f([x^1], \ldots, [x^k]) = [y],$$

or put more succinctly, ${}^*G_f = G_{{}^*f}$.

Proof. It's follows form the definition of G_f that $[\![(x^1, \ldots, x^k, y) \in G_f]\!] = [\![f(x^1, \ldots x^k) = y]\!]$ so the left and right side coincide. □

Definition 3.7 is thus well-defined since it's equivalent with **definition 3.4**, which we've already proven well-defined.

Much like how the translation of a set preserves its standard elements, the translation of a function agrees with the original function on all real numbers.

Proposition 3.9. Given $f : \mathbb{R}^k \to \mathbb{R}$ and its natural translation ${}^*f : {}^*\mathbb{R}^k \to {}^*\mathbb{R}$, then for all $(x_1, \ldots, x_k) \in \mathbb{R}^k$ we have ${}^*f({}^*x_1, \ldots, {}^*x_k) = {}^*(f(x_1, \ldots, x_k))$.

Proof. By definition we have ${}^*f({}^*x_1, \ldots, {}^*x_k) = [(f(x_1, \ldots, x_k), f(x_1, \ldots, x_k), \ldots)]$ and since $f(x_1, \ldots, x_k) \in \mathbb{R}$ this equals ${}^*(f(x_1, \ldots, x_k)) = [(f(x_1, \ldots, x_k), f(x_1, \ldots, x_k), \ldots)]$. □

This proposition shows that ${}^*f$ is an extension of f to ${}^*\mathbb{R}$ so for convenience we denote both the real and hyperreal functions with the same symbol, unless we want to make the difference explicit.

The natural translation also preserve important function properties

Proposition 3.10. If $f : \mathbb{R}^k \to \mathbb{R}$ is injective, then ${}^*f$ is also injective. The same holds for surjectivity.

Proof. Assume $f : \mathbb{R}^k \to \mathbb{R}$ is injective and let ${}^*f([x^1], \ldots, [x^k]) = {}^*f([y^1], \ldots, [y^k])$ for $[x^1], \ldots [x^k], [y^1], \ldots, [y^k] \in {}^*\mathbb{R}$. Then $[\![f(x^1, \ldots x^k) = f(y^1, \ldots, y^k)]\!] \in \mathcal{U}$ and by f being injective $[\![f(x^1, \ldots x^k) = f(y^1, \ldots, y^k)]\!] \subseteq [\![(x^1, \ldots, x^k) = (y^1, \ldots, y^k)]\!]$ so $([x^1], \ldots, [x^k]) = ([y^1], \ldots, [y^k])$, ergo ${}^*f$ is injective.

Assume $f : \mathbb{R}^k \to \mathbb{R}$ is surjective and consider a general hyperreal $[y] \in {}^*\mathbb{R}$. By f being surjective we know $[y] = [(y_1, y_2, \ldots)] = [(f(x_1^1, \ldots, x_1^k), f(x_2^1, \ldots, x_2^k), \ldots]$ for some $(x_1^1, \ldots x_1^k), (x_2^1, \ldots x_2^k), \ldots \in \mathbb{R}^k$. Then $([(x_1^1, x_2^1, \ldots], \ldots, [(x_1^k, x_2^k, \ldots]) = ([x^1], \ldots, [x^k]) \in {}^*\mathbb{R}^k$ is such that ${}^*f([x^1], \ldots, [x^k]) = [y]$, ergo ${}^*f$ is surjective. □

Translating partial k-ary functions whose domain is $X = X_1 \times \cdots \times X_k \subseteq \mathbb{R}^k$ is possible but requires some finesse. Consider for example the function $f(x) = \sqrt{x}$ whose domain is $\mathbb{R}^{\geq 0}$. Since $[(-1, 0, 0, 0, \ldots)] = 0$ we should have $f([(-1, 0, 0, 0, \ldots)]) = f(0) = 0$ but $[(f(-1), f(1), f(1), f(1), \ldots)])$ is undefined in the first component. We circumvent this in the same way we defined the multiplicative inverse.

Definition 3.11 (Translation of a Partial Function)**.** Given a k-ary function $f : X \to \mathbb{R}$, with $X \subseteq \mathbb{R}^k$, we define its natural translation as the function

$$
\begin{aligned}
&{}^*f : {}^*X \to {}^*\mathbb{R}, \\
&([x^1], \ldots, [x^k]) \mapsto [(s_n)] \quad \text{where } s_n = \begin{cases} f(x_n^1, \ldots, x_n^k) & \text{if } n \in [\![(x^1, \ldots, x^k) \in X]\!], \\ 0 & \text{if } n \notin [\![(x^1, \ldots, x^k) \in X]\!]. \end{cases}
\end{aligned}
$$

3.3 The Relational Structure and First-Order Languages of the Hyperreals

Using our extended translation function we define the relational structure of the hyperreals, which we denote ${}^*\mathfrak{R}$, as $({}^*\mathbb{R}, \{{}^*R : R \in Rel_{\mathfrak{R}}\})$.

We extend natural translation map for the final time, to apply to any $\mathcal{L}_{\mathfrak{R}}$-formula by inductive definition.

Definition 3.12 (The Translation of a $\mathcal{L}_{\mathfrak{R}}$-formula)**.**

- For atomic formulas${}^*(Rt_1 \ldots t_k) = {}^*R^*t_1 \ldots {}^*t_k$ where variables are unaffected i.e. ${}^*x = x$.
- For negations negation ${}^*(\neg\varphi) = \neg{}^*\varphi$.
- The conjunctions ${}^*(\varphi \wedge \psi)$ is $({}^*\varphi \wedge {}^*\psi)$.
- For existential formulas ${}^*\big(\exists x\varphi\big) = \exists x^*\varphi$.

It follows that these definitions also define translations for the rest of the connectives since, given any $\varphi, \psi \in \mathcal{L}_{\mathfrak{R}}$-formulas

- ${}^*(\varphi \vee \psi) := {}^*\big(\neg(\neg\varphi \wedge \neg\psi)\big) = \neg(\neg{}^*\varphi \wedge \neg{}^*\psi) =: {}^*\varphi \vee {}^*\psi$
- ${}^*(\varphi \to \psi) := {}^*(\neg\varphi \vee \psi) = (\neg{}^*\varphi \vee {}^*\psi) =: {}^*\varphi \to {}^*\psi$
- ${}^*(\varphi \leftrightarrow \psi) := {}^*\big((\varphi \to \psi) \wedge (\psi \to \varphi)\big) = ({}^*\varphi \to {}^*\psi) \wedge ({}^*\psi \to {}^*\varphi) =: {}^*\varphi \leftrightarrow {}^*\psi$
- ${}^*(\forall x\varphi) := {}^*(\neg\exists x\neg\varphi) = \neg\exists x\neg{}^*\varphi =: \forall x^*\varphi$
- ${}^*\big(\varphi(f(t_1, \ldots, t_k))\big) := {}^*\big(\exists x(G_f t_1 \ldots t_k \wedge \varphi(x)\big) = \exists x\big({}^*G_f{}^*t_1 \ldots {}^*t_k \wedge {}^*\varphi(x)\big) = {}^*\varphi({}^*f(t_1, \ldots, t_k))$ by **propositions 3.8** and **3.9**.
- ${}^*\big((\forall\varphi(x_1, \ldots x_p))\,\psi\big) := {}^*\big(\forall x_1 \ldots \forall x_p(\varphi(x_1, \ldots x_p) \to \psi\big) = \forall x_1 \ldots \forall x_p({}^*\varphi(x_1, \ldots x_p) \to {}^*\psi) =: \big(\forall{}^*\varphi(x_1, \ldots x_p)\big)\,{}^*\psi$
- ${}^*\big((\exists\varphi(x_1, \ldots x_p))\,\psi\big) := {}^*\big(\exists x_1 \ldots \exists x_p(\varphi(x_1, \ldots x_p) \wedge \psi\big) = \exists x_1 \ldots \exists x_p({}^*\varphi(x_1, \ldots x_p) \wedge {}^*\psi) =: (\exists{}^*\varphi(x_1, \ldots x_p))\,{}^*\psi$

In effect, the translation of a formula changes all relations, functions and real numbers to their hyperreal counterpart but leaves all other logical symbols unaffected.

Example. The $\mathcal{L}_{\mathfrak{R}}$-sentence

$$(\forall\theta \in \mathbb{R})\big(\sin\theta = 0 \leftrightarrow (\exists k \in \mathbb{Z})(\theta = k \cdot \pi)\big)$$

express that the sinus functions is 0 iff the argument is a integer multiple of π. Its natural translation is then the $\mathcal{L}_{{}^*\mathfrak{R}}$-sentence

$$(\forall\theta \in {}^*\mathbb{R})\big({}^*\!\sin\theta = {}^*0 \leftrightarrow (\exists k \in {}^*\mathbb{Z})(\theta = k \,{}^*\!\cdot\, {}^*\pi)\big)$$

which says the hyperreal extension of sinus is 0 iff its argument is a hyper-integer multiple of π. We use the * for $\sin, \cdot, 0$ and π to show that they are affected by the translation but per our discussion of **proposition 3.9**, there is no risk of confusion, so we will omit the notation for functions and real numbers from now on.

We define the hyperreal languages as $\mathcal{L}_{^*\mathfrak{R}}$-formulas $= \{^*\varphi : \varphi \in \mathcal{L}_{^*\mathfrak{R}}\text{-formulas}\}$. Note that this language is weak, it only has terms for standard elements and, because of **proposition 3.5**, no infinite subsets of $\mathbb{R}$ like $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ can be referred to. In addition none of the interesting subsets of $^*\mathbb{R}$ we introduce in the next section section, can be referred to either. Despite theses weaknesses the language is interesting because of its natural correspondence to the language of the reals, which lays the groundwork for the transfer principle.

3.4 The Transfer Principle

The transfer principle actually follows from of a more powerful theorem, called **Łos Theorem**, which applies to the more general topic of ultraproducts we mentioned in the beginning of **section 2**. What we prove here is only a special case.

Theorem 3.13 (Łos Theorem)**.** *For every $\mathcal{L}_{\mathfrak{R}}$-formula $\varphi(x_1, \ldots, x_p)$ and every $[s^1], \ldots [s^p] \in {}^*\mathbb{R}$*

$$\varphi([s^1], \ldots [s^p]) \text{ is true in } {}^*\mathfrak{R} \text{ iff } [\![\varphi(s^1, \ldots, s^p)]\!] := \{n \in \mathbb{N} : \varphi(s^1_n, \ldots s^p_n) \text{ is true in } \mathfrak{R}\} \in \mathcal{U}.$$

This theorem states that the definition of truth for atomic sentences in **definition 3.4**, permeates throughout all sentences in $\mathcal{L}_{^*\mathfrak{R}}$-sentences. Although it may seem trivial, we actually need to make use of **Axiom of Choice** for the 'if'-part [2].

Proof. We prove this by induction on the length of φ.

Base case: If φ is atomic then it follows immediately from **definition 3.4**.

Induction step: Assume that the property holds for all formulas shorter than $\varphi(x_1, \ldots, x_p)$.
If $\varphi(x_1, \ldots, x_p)$ is a negation of the form $\neg\psi(x_1, \ldots, x_p)$, then

$$\begin{aligned}
\neg\psi([s^1], \ldots [s^p]) \text{ is true in } {}^*\mathfrak{R} &\text{ iff } \psi([s^1], \ldots [s^p]) \text{ is false in } {}^*\mathfrak{R} && [\neg \text{ definition}]\\
&\text{ iff } [\![\psi(s^1, \ldots s^p)]\!] \notin \mathcal{U} && [\text{induction hypothesis}]\\
&\text{ iff } \overline{[\![\psi(s^1, \ldots s^p)]\!]} \in \mathcal{U}. && [\mathcal{U} \text{ being an ultrafilter}]\\
&\text{ iff } [\![\neg\psi(s^1, \ldots s^p)]\!] \in \mathcal{U}. && [\overline{[\![\psi]\!]} = [\![\neg\psi]\!]]
\end{aligned}$$

If $\varphi(x_1, \ldots, x_p)$ is a conjunction of the form $(\psi \wedge \chi)(x_1, \ldots, x_p)$, then

$$\begin{aligned}
(\psi \wedge \chi)([s^1], \ldots, [s^p]) \text{ is true in } {}^*\mathfrak{R} &\text{ iff } \psi([s^1], \ldots, [s^p]) \text{ and } \chi([s^1], \ldots, [s^p]) \text{ are true in } {}^*\mathfrak{R} && [\wedge \text{ definition}]\\
&\text{ iff } [\![\psi(s^1, \ldots, s^p)]\!], [\![\chi(s^1, \ldots, s^p)]\!] \in \mathcal{U} && [\text{induction hypothesis}]\\
&\text{ iff } [\![\psi(s^1, \ldots, s^p)]\!] \cap [\![\chi(s^1, \ldots, s^p)]\!] \in \mathcal{U}. && [\textbf{proposition 2.2 (c)}]\\
&\text{ iff } [\![(\psi \wedge \chi)(s^1, \ldots, s^p)]\!] \in \mathcal{U}. && [[\![\psi]\!] \cap [\![\chi]\!] = [\![\psi \wedge \chi]\!]]
\end{aligned}$$

If $\varphi(x_1, \ldots, x_p)$ is a existential formula of the form $\exists x \psi(x_1, \ldots, x_p, x)$, then

$$\begin{aligned}
\exists x\psi([s^1], \ldots, [s^p], x) \text{ is true in } {}^*\mathfrak{R} &\text{ iff there exists a } [r] \in {}^*\mathbb{R} \text{ s.t } \psi([s^1], \ldots, [s^p], [r]) \text{ is true in } {}^*\mathfrak{R} && [\exists \text{ definition}]\\
&\text{ iff there exists a } r \in \mathbb{R}^{\mathbb{N}} \text{ s.t } [\![\psi(s^1, \ldots, s^p, r)]\!] \in \mathcal{U}. && [\text{induction hypothesis}]
\end{aligned}$$

If $[\![\psi(s^1, \ldots, s^p, r)]\!] \in \mathcal{U}$ for some $r \in \mathbb{R}^{\mathbb{N}}$, then since $[\![\psi(s^1, \ldots, s^p, r)]\!] \subseteq [\![\exists x\psi(s^1, \ldots, s^p, x)]\!]$, it follows that $[\![(\exists x\psi)(s^1, \ldots, s^p, x)]\!] \in \mathcal{U}$.

Conversely, suppose $[\![\exists x\psi(s^1, \ldots, s^p, x)]\!] \in \mathcal{U}$, then for each index $n \in [\![\exists x\psi(s^1, \ldots, s^p, x)]\!]$ there exists a $r_n \in \mathbb{R}$ s.t $\psi(s^1_n, \ldots, s^p_n, r_n)$ is true in $\mathfrak{R}$. With axiom of choice, which we assume implicitly with **Zorn's lemma**, we may construct the sequence q defined as $q_n = \begin{cases} r_n & \text{if } n \in [\![\exists x\psi(s^1, \ldots, s^p, x)]\!] \\ 0 & \text{if } n \notin [\![\exists x\psi(s^1, \ldots, s^p, x)]\!] \end{cases}$. By the construction of q, we have $[\![\exists x\psi(s^1, \ldots, s^p, x)]\!] \subseteq [\![\psi(s^1, \ldots, s^p, q)]\!]$, ergo there exists a $r \in \mathbb{R}^{\mathbb{N}}$ s.t $[\![\psi(s^1, \ldots, s^p, r)]\!] \in \mathcal{U}$. ☐

Before proving the transfer principle, we establish one final lemma.

Lemma 3.14. For all $\varphi \in \mathcal{L}_{\mathfrak{R}}$-sentence, if φ is true in $\mathfrak{R}$, then $[\![^*\varphi]\!] = \mathbb{N}$. If φ is false in $\mathfrak{R}$, then $[\![^*\varphi]\!] = \varnothing$.

Proof. This follows from the fact that for all $\varphi \in \mathcal{L}_{\mathfrak{R}}$-sentences, all the terms of $^*\varphi$ are equivalence classes of constant sequences. Hence it's the same in each index, so if it's true in one, it's true in all. ☐

Theorem 3.15 (The Transfer Principle)**.** *For any sentence $\varphi \in \mathcal{L}_{\mathfrak{R}}$-sentences*

$$\varphi \text{ is true in } \mathfrak{R} \text{ if and only if } {}^*\varphi \text{ is true in } {}^*\mathfrak{R}.$$

Proof. If φ is true in $\mathfrak{R}$, then by **lemma 3.14** $[\![\varphi]\!] = \mathbb{N} \in \mathcal{U}$ so by **Łos Theorem** ${}^*\varphi$ is true in ${}^*\mathfrak{R}$. If φ is false in $\mathfrak{R}$, then by **lemma 3.14** $[\![\varphi]\!] = \varnothing \notin \mathcal{U}$ so by **Łos Theorem** ${}^*\varphi$ is false in ${}^*\mathfrak{R}$. □

The transfer principle also solve our problem at the end of **section 2.4** where we worried that the choice of non-principal ultrafilter might effect ${}^*\mathbb{R}$. By transfer, if a ${}^*\varphi \in \mathcal{L}_{{}^*\mathfrak{R}}$-sentence was true in *a* hyperreal structure, then φ would be true in $\mathfrak{R}$, and then by transfer φ would be true in *any* hyperreal structure. The same holds if ${}^*\varphi$ would be false. Thus all hyperreal structure make the same set of $\mathcal{L}_{{}^*\mathfrak{R}}$-sentence true, which makes them equivalent for the purposes of real analysis.

Example. To demonstrate the usefulness of the **transfer principle** let's investigate the structure of the hypernatural numbers. With transfer we can avoid the tedious talk of infinite sequences and ultrafilter. Applying transfer on the following true $\mathcal{L}_{\mathfrak{R}}$-sentences

$$(\forall x \in \mathbb{N})(1 \leq x), \quad (\forall n \in \mathbb{N})\neg(\exists m \in \mathbb{N})(n < m < n+1) \text{ and } (\forall x \in \mathbb{R})(\exists n \in \mathbb{N})(x < n)$$

yields

$$(\forall x \in {}^*\mathbb{N})(1 \leq x), \quad (\forall n \in {}^*\mathbb{N})\neg(\exists m \in {}^*\mathbb{N})(n < m < n+1) \text{ and } (\forall x \in {}^*\mathbb{R})(\exists n \in {}^*\mathbb{N})(x < n)$$

which tells us that hypernatural are also all greater than 1, that they're still discretely distributed, each one 1 step apart, and that they occur infinitely often along the hyperreal number-line. Furthermore, the $\mathcal{L}_{\mathfrak{R}}$-sentence

$$(\forall m \in \mathbb{N})(m \leq n \to \big(m = 1 \vee m = 2 \vee \ldots \vee m = n)\big)$$

is true in $\mathfrak{R}$ for any fixed $n \in \mathbb{N}$, so by transfer

$$(\forall m \in {}^*\mathbb{N})(m \leq n \to \big(m = 1 \vee m = 2 \vee \ldots \vee m = n)\big)$$

is true in ${}^*\mathfrak{R}$, for every $n \in \mathbb{N}$. This means all non standard hypernatural are greater than every standard natural.

Remark. Transferring an infinite set sentences, one for each element in some real set, is a clever way to get around the fact that $\mathcal{L}_{{}^*\mathfrak{R}}$-sentences can't quantify over infinite real sets.

4 The Non-standard Elements of the Hyperreals

As we saw in **section 2.4**, the hyperreals contain non-standard numbers — some smaller than every standard real number, others larger than any standard real number. In this section, based on chapter 5 of [1], we use the **transfer principle** to develop a theory of arithmetic in the hyperreals, examining how addition, multiplication, and division behave when involving non-standard numbers. This naturally lead into the Algebraic structure of the hyperreals, which lays the ground works for non-standard analysis.

4.1 Hyperreal Arithmetic

To introduce some terminology, we call a $x \in {}^*\mathbb{R}$

- *infinitesimal* if $|x| < r$ for all $r \in \mathbb{R}^+$,
- *limited* if $r < x < s$ for some $r, s \in \mathbb{R}$,
- *unlimited* if $r < |x|$ for all $r \in \mathbb{R}$.

We shall denote the set of infinitesimals with $\mathbb{I}$, the set of limited numbers with $\mathbb{L}$ and, for any $X \subseteq {}^*\mathbb{R}$ we let $X_\infty = \{x \in X : x \text{ is unlimited}\} = X - \mathbb{L}$. Note that $\mathbb{I}$ and $\mathbb{R}$ are subsets of $\mathbb{L}$. By convention, we use Greek letters for infinitesimals, lowercase Latin letters for limited numbers and uppercase Latin letters for unlimited numbers - although ε will also be used for small positive real number when we reason in standard analysis.

One might already have some intuitions about the arithmetic properties of infinitesimals, limited and unlimited numbers. For example, the sum of an unlimited and limited number should still be unlimited. We prove that these intuitions hold in ${}^*\mathbb{R}$.

Proposition 4.1. Let α, β be infinitesimals, a, b be limited, but not infinitesimal and A, B be unlimited. Then

Sums:

$\alpha + \beta$ is infinitesimal,
$a + b$ is limited, possibly infinitesimal,
$\alpha + a$ is limited, never infinitesimal,
$\alpha + A$ and $a + A$ are unlimited,

Products:

$\alpha \cdot \beta$ and $\alpha \cdot a$ are infinitesimal,
$b \cdot c$ is limited, never infinitesimal,
$A \cdot a$ and $A \cdot B$ are unlimited,

Quotients:

$\dfrac{\alpha}{a}, \dfrac{\alpha}{A}$ and $\dfrac{a}{A}$ are infinitesimal,
$\dfrac{a}{b}$ is limited, never infinitesimal,
$\dfrac{a}{\alpha}, \dfrac{A}{\alpha}$ and $\dfrac{A}{a}$ are unlimited,

Undetermined:

$\dfrac{\alpha}{\beta}, \dfrac{A}{B}, \alpha \cdot \beta$ and $A + B$ can be infinitesimal, limited or unlimited.

By transfer,

$$|x + y| \leq |x| + |y|, \quad \big||x| - |y|\big| \leq |x + y| \quad \text{and} \quad |xy| = |x||y|$$

hold for all $x, y \in {}^*\mathbb{R}$. We do not prove all of the statements above but the proofs for the ones omitted use the same ideas.

Proof. Let α, β be infinitesimal, a, b be limited but not infinteismal and A, B be unlimited.

By the triangle inequality we have $|\alpha + \beta| \leq |\alpha| + |\beta|$. Because α and β are infinitesimals $|\alpha| + |\beta|$ must be less then every positive real, hence $\alpha + \beta$ is also infinitesimals.

By the reverse triangle inequality we have $\big||A| - |a|\big| \leq |A + a|$. Since a is limited we have $\big||A| - |a|\big| \geq \big||A| - r\big|$, and because A is unlimited $\big||A| - r\big| = |A| - r$ for some $r \in \mathbb{R}$. Because $\big||A| - |a|\big| = |A| - r$ is greater than every real, it follows that $|A + a|$ is as-well, which means $A + a$ is unlimited.

The product $|\alpha \cdot a| = |\alpha| \cdot |a| \leq r \cdot s$ for some positive real s and every positive real r, hence less then every positive real, so $\alpha \cdot a$ is infinitesimal.

By transfer we can rewrite the quotient $\frac{a}{A} = \frac{1}{A} \cdot a$. Since A is unlimited $\frac{1}{A} = A^{-1}$ must infinitesimal since inverting both sides of a difference by ${}^*\mathbb{R}$ being an ordered field. So $\frac{a}{A} = \frac{1}{A} \cdot a$ is an infinitesimal times a limited number so, by our previous proof, infinitesimal.

In the case $\frac{\alpha}{\beta}$ if $\alpha = \beta^2$, then $\frac{\alpha}{\beta} = \beta$ thus infinitesimal. If $\alpha = \beta$, then $\frac{\alpha}{\beta} = 1$ which is limited but not infinitesimal. If $\beta = \alpha^2$, then $\frac{\alpha}{\beta} = \frac{1}{\alpha}$ which is unlimited. So $\frac{\alpha}{\beta}$ is undetermined. $\square$

Given a hyperreal x and an a non-zero infinitesimal ε, the sum $x + \varepsilon$ is an infinitesimal distance from x on the number-line, yet distinct from x. From a real perspective they are indistinguishable since the difference between them is less than every real number. We therefore define the following relation on ${}^*\mathbb{R}$.

Definition 4.2. We say two hyperreals x and y, are *infinitely close*, denoted $x \simeq y$, if $x - y \in \mathbb{I}$.

This relation is reflexive, since $0 \in \mathbb{I}$, and symmetric because if $x - y \in \mathbb{I}$, then $y - x = (-1)(x - y) \in \mathbb{I}$ by **proposition 4.1**. If $x \simeq y$ and $y \simeq z$, then $x - z = x - (y - y) - z = (x - y) + (y - z)$ is a sum of infinitesimal so by **proposition 4.1** also an infinitesimal, so $x \simeq z$; being infinitely close is thus transitive, and thereby an equivalence. We denoted the equivalence class of x as $\mu(x) = \{y \in {}^*\mathbb{R} : y \simeq x\}$, calling it the *monad* of x. It follows that $\mu(x) = \{x + \varepsilon : \varepsilon \in \mathbb{I}\}$ and in particular $\mu(0) = \mathbb{I}$.

If $x \simeq y$, then $x + z \simeq y + z$ for any $z \in {}^*\mathbb{R}$ by $x + z - (y + z) = x - y$ so being infinitely close is a congruence relation with respect to addition. As for multiplication, $x \simeq y$ implies $a \cdot x \simeq a \cdot y$ if a is limited since $a \cdot x - a \cdot y = a(x - y)$ is then infinitesimal. This, however, fails if a is unlimited. Consider for example an infinitesimal ε. By **proposition 4.1** $\varepsilon \simeq \varepsilon^2$ but $\frac{1}{\varepsilon^2} \cdot \varepsilon^2 = 1 \not\simeq \frac{1}{\varepsilon} = \frac{1}{\varepsilon^2} \cdot \varepsilon$.

Any function defined on a subset of the real numbers can be extended to a function on the subset's natural translation. These real functions are also well-defined for non-standard elements (except in the case of finite domains, but those are uninteresting). To demonstrate, we consider the extensions of the trigonometric functions.

Example.

1. Since the sine function is defined for all real numbers, its natural extension is also defined for all hyperreal numbers. We investigate how it behaves for infinitesimal inputs. In introductory analysis courses, one proves that $\cos x \leq \frac{\sin x}{x} \leq 1$ for all $0 < x < \frac{\pi}{2}$ using a geometric argument involving the areas of a triangle and a circular sector. It follows that all terms are positive, so the second inequality implies $|\sin x| \leq |x|$ for all $x \in]0, \frac{\pi}{2}[$, so by transfer, it holds for all $x \in {}^*]0, \frac{\pi}{2}[$ which includes $\mathbb{I}^+$. Therefore, $\sin \varepsilon$ is infinitesimal for all positive infinitesimal ε. By transfer, sine is also odd for hyperreal numbers, consequentiality $\sin(-\varepsilon) = -\sin \varepsilon$ is

2. The cosine function is also defined for all hyperreal numbers. We also investigate its behavior for infinitesimal arguments. The trigonometric identity $\cos x = 1 - 2\sin^2 \frac{x}{2}$ holds for all real x, so by transfer it holds for all hyperreal x. Thus by our previous result, $\cos \varepsilon = 1 - 2\sin^2 \frac{\varepsilon}{2} \simeq 1$ for all infinitesimal ε.

3. The tangent function is defined for all real numbers except for odd integer multiples of $\frac{\pi}{2}$ so its hyperreal extension is defined for all hyperreals except for odd hyperinteger multiples of $\frac{\pi}{2}$. Consider $\tan(\frac{\pi}{2} + \varepsilon)$ where ε is infinitesimal. By transfer we know that $\tan = \frac{\sin}{\cos}$ and the addition formulas for sine and cosine, hold for all hyperreals. We then get

$$\tan\left(\frac{\pi}{2} + \varepsilon\right) = \frac{\sin\left(\frac{\pi}{2} + \varepsilon\right)}{\cos\left(\frac{\pi}{2} + \varepsilon\right)} = \frac{\sin\frac{\pi}{2}\cos\varepsilon + \cos\frac{\pi}{2}\sin\varepsilon}{\cos\frac{\pi}{2}\cos\varepsilon - \sin\frac{\pi}{2}\sin\varepsilon} = \frac{\cos\varepsilon}{-\sin\varepsilon} \simeq \frac{1}{-\sin\varepsilon}.$$

   We know the sine function is positive on $]0, \frac{\pi}{2}]$ and negative on $]-\frac{\pi}{2}, 0[$. Therefore, $\sin\varepsilon$ is a positive infinitesimal when ε is a positive infinitesimal and is a negative infinitesimal when ε is a negative infinitesimal. This means that $\tan\left(\frac{\pi}{2} + \varepsilon\right) \simeq \frac{1}{-\sin\varepsilon}$ is positive unlimited when ε is a negative infinitesimal, but negative unlimited when ε is a positive infinitesimal.

Shadows

We know all the standard real numbers are limited, we also know that they are closed under adding infinitesimals so $r + \varepsilon$ is limited for every real r and infinitesimal ε. The question arise if this exhaust all limited hyperreals or if there are limited hyperreal that are not the sum of a real number and an infinitesimal. The answer is no which we prove as follows.

Theorem 4.3. *For every $a \in \mathbb{L}$ there exists a unique $r \in \mathbb{R}$ such $a \simeq r$. Equivalently, for every $a \in \mathbb{L}$ there's one and only one real number $r \in \mu(a)$.*

Proof. Let $a \in \mathbb{L}$ and consider the set $A = \{x \in \mathbb{R} : x < a\}$ which is a real non-empty subset, embedded into the hyperreal number-line. Since a is limited there exists real numbers m, M such that $m < a < M$ and consequently A is non empty and bounded from above by M. By the supremum axiom, A must have a least upper-bound r which is a real number. Since r is an upper-bound $S + s \notin A$ for all positive real s which, by how A is defined, means $a \leq r + s$ or equivalently $a - r \leq s$ for all positive real s. Since r is the least upper-bound we also have $r - s \in A$ for all positive real s, otherwise $r - s$ would be the least real upper-bound. Again by A:s definition we get $r - s < a$ thus $-s < a - r$. And so $-s < a - r < s$ or equivalently $|a - r| < s$ for all positive real s i.e. $a \simeq r \in \mathbb{R}$. So every limited hyperreal is close to *at least one* real number.

For Uniqueness let r and s be real numbers such that $r \simeq s$. Since their difference is always a real numbers, it's only infinitesimal if $r - s = 0$ i.e $r = s$. Ergo hyperreal can be infinitely close to *at most one* real number. $\square$

One can thus think of as the real number-line were around every real number there's an 'infinitesimal cloud' that is its monad. In addition, by **theorem 4.3**, the map

$$\begin{aligned} \mathrm{sh}: \mathbb{L} &\to \mathbb{R}, \\ x &\mapsto y \text{ such that } x \simeq y \text{ and } y \in \mathbb{R}. \end{aligned}$$

is well defined. We call $\mathrm{sh}(a)$, the *the shadow of a*.

Remark. The shadow of a is sometimes called *standard part of a* since **theorem 4.3** implies that $a = r + \varepsilon$ for some real r and infinitesimal ε, so $\mathrm{sh}(a) = \mathrm{sh}(r + \varepsilon) = r$.

We prove some useful properties of the shadow map.

Proposition 4.4. Let $a, b \in \mathbb{L}$ then the following holds

(i) $\mathrm{sh}(a \pm b) = \mathrm{sh}(a) \pm \mathrm{sh}(b)$,

(ii) $\mathrm{sh}(a \cdot b) = \mathrm{sh}(a) \cdot \mathrm{sh}(b)$,

(iii) $\mathrm{sh}(a/b) = \mathrm{sh}(a)/\mathrm{sh}(b)$ assuming $\mathrm{sh}(b) \neq 0$,

(iv) if $\mathrm{sh}(a) < \mathrm{sh}(b)$, then $a < b$,

(v) if $a \leq b$, then $\mathrm{sh}(a) \leq \mathrm{sh}(b)$.

Proof. For (i) note that we can always write $\mathrm{sh}(a) = a + \alpha$ and $\mathrm{sh}(b) = b + \beta$ for some $\alpha, \beta \in \mathbb{I}$ the difference $\big(\mathrm{sh}(a) \pm \mathrm{sh}(b)\big) - (a \pm b) = (a+\alpha) \pm (b+\beta) - (a \pm b) = \alpha \pm \beta \in \mathbb{I}$ so $\mathrm{sh}(a) \pm \mathrm{sh}(b) \simeq a \pm b$. Because $\mathrm{sh}(a) \pm \mathrm{sh}(b)$ is a sum/difference of real numbers, it's also real so by **theorem 4.3** $\mathrm{sh}(a) \pm \mathrm{sh}(b) = \mathrm{sh}(a \pm b)$

For (ii) we again have $\mathrm{sh}(a) \cdot \mathrm{sh}(b) \in \mathbb{R}$ and $\mathrm{sh}(a) \cdot \mathrm{sh}(b) - a \cdot b = (a+\alpha)(b+\beta) = a \cdot b + a \cdot \beta + \alpha \cdot b + \alpha \cdot \beta - a \cdot b = a \cdot \beta + \alpha \cdot b + \alpha \cdot \beta \in \mathbb{I}$ so $\mathrm{sh}(a) \cdot \mathrm{sh}(b)$ is the unique real number infinitely close to $a \cdot b$ i.e. $\mathrm{sh}(a) \cdot \mathrm{sh}(b) = \mathrm{sh}(a \cdot b)$.

The proof of (iii) is analogous to the proof of (ii).

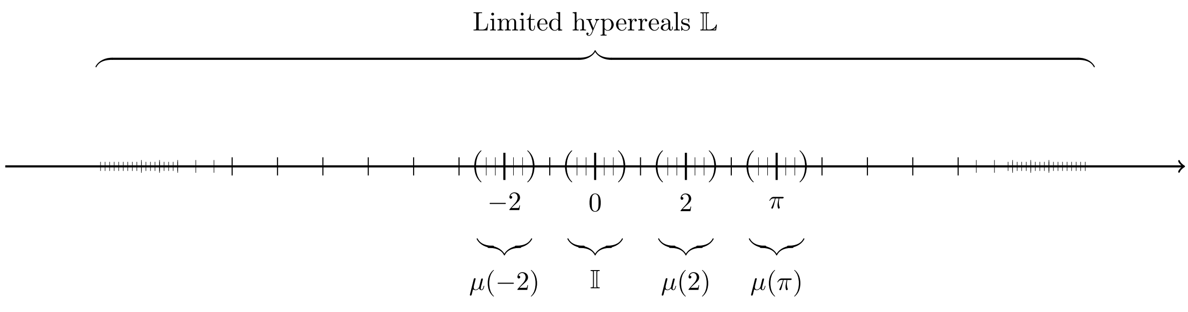
For (iv) if $\mathrm{sh}(a) < \mathrm{sh}(b)$, then as proven above $\mathrm{sh}(b - a) = \mathrm{sh}(b) - \mathrm{sh}(a) > 0$. The difference $b - a$ cannot then be non-positive since, if it were, $\mathrm{sh}(b - a) - (b - a) \geq \mathrm{sh}(b - a) > 0$ contradicting $\mathrm{sh}(b - a) \simeq b - a$. Therefore $b - a > 0$ i.e. $b > a$.

The proof of (v) is analogous to the proof of (iv). □

Remark. A strict inequality doesn't imply a strict inequality in (v). For example, if r is real and ε is a positive infinitesimal, then $r - \varepsilon < r + \varepsilon$ but $\mathrm{sh}(r - \varepsilon) = \mathrm{sh}(r + \varepsilon) = r$.

A curious reader might have noted that, by **proposition 4.1** $\mathbb{L}$ is a sub-rings of ${}^*\mathbb{R}$ and $\mathbb{I}$ is an ideal in $\mathbb{L}$, and wondered about the quotient ring $\mathbb{L}/\mathbb{I}$. In **proposition 4.2** we've proven that $\mathrm{sh}: \mathbb{L} \to \mathbb{R}$ is a ring homomorphism and it's easy to verify that sh is surjective with $\ker(\mathrm{sh}) = \mathbb{I}$. So by the first ring homomorphism theorem $\mathbb{L}/\mathbb{I} \simeq \mathbb{R}$ where the pre-image of a is $\mu(a)$.

Since 'infinitely close' is an equivalence we know by **proposition 4.3** that the monads of the real numbers partition the reals but, with **Proposition 4.4** (iv) and (v), we know that the monads never overlap. The fact that $\mathbb{L}/\mathbb{I} \simeq \mathbb{R}$ then shows that this partition has the ring structure of $\mathbb{R}$. With this information we sketch



Theorem 4.3 is actually equivalent to the notion that $\mathbb{R}$ has 'no holes' since (informally) if it had, a limited number in that hole wouldn't be infinitely close to any real. We prove this formally by proving

Theorem 4.5. *theorem 4.3 is equivalent with the completeness of* $\mathbb{R}$.

Proof. The the completeness of $\mathbb{R}$ implies **theorem 4.3** since we used the supremum axiom in its proof.

We prove the completeness of $\mathbb{R}$ by proving the supremum axiom. Let $X \subseteq \mathbb{R}$ be non-empty with a real upper-bound M. Consider the sequence $s : \mathbb{N} \to \mathbb{R}$ defined as $s_n = \min Z_n$ where $Z_n = \{k \in \mathbb{Z} : \frac{k}{n}$ is an upper-bound to $X\}$.. There's always $\lceil n \cdot M \rceil \in Z_n$ and $n \cdot x$ is always a lower bound of Z_n, where x is any element of X. Then because Z_n is a non-empty subset of $\mathbb{Z}$ bounded from below, it must have a minimum, we are therefore assured that s is well-defined.

There is then a hyperreal extension of the sequences from ${}^*\mathbb{N}$ to ${}^*\mathbb{R}$. Following sequence notation we write s_n instead of $s(n)$ for unlimited n as well. Consider $\frac{s_N}{N}$ for some $N \in {}^*\mathbb{N}_\infty$. By definition $\frac{s_N}{N}$ is an upper-bound of X, so it's greater then some real number $r \in X$. Also by definition we know $\frac{s_N - 1}{N}$ is not an upper-bound, since if it were, $\frac{s_N}{N}$ wouldn't be the least upper-bound. Hence $\frac{s_N-1}{N}$ is less then some real number $s \in X$. Then because $\frac{s_N-1}{N} = \frac{s_N}{N} - \frac{1}{N} < s$, we get $\frac{s_N}{N} < s + \frac{1}{N} < s + 1 \in \mathbb{R}$ since $\frac{1}{N}$ is infinitesimal by **proposition 4.1**. Consequently $r < \frac{s_N}{N} < s + 1$ which means $\frac{s_N}{N}$ is limited so by **theorem 4.3** we can find $\mathrm{sh}(\frac{s_N}{N}) = L \in \mathbb{R}$.

Because $\frac{s_N}{N}$ is greater then every element of X, it follows by **proportion 4.4** (iv) and that $X \subseteq \mathbb{R}$, that L is also a upper bound. For the sake of contradiction, let q be real upper-bound to X strictly less than L. Note that $\frac{s_N}{N} - \frac{s_N-1}{N} = \frac{1}{N} \in \mathbb{I}$ so $\mathrm{sh}(\frac{s_N-1}{N}) = L > q = \mathrm{sh}(q)$. Then by **proposition 4.4** (iv) we get $\frac{s_N-1}{N} > q$, but $\frac{s_N-1}{N}$ is not an upper-bound, resulting in contradiction. Ergo L is the least real upper-bound of X. □

Remark. Considering the restriction the shadow map to the limited hyperrationals, that is $\mathrm{sh} : {}^*\mathbb{Q} \cap \mathbb{L} \to \mathbb{R}$, we still get a surjective homomorphism with the kernel ${}^*\mathbb{Q} \cap \mathbb{I}$. This is because 'every real number is arbitrarily close to a rational number' is true in $\mathfrak{R}$ and can be express by the $\mathcal{L}_{\mathfrak{R}}$-sentence

$$(\forall r \in \mathbb{R})(\forall \varepsilon \in \mathbb{R}^+)(\exists q \in \mathbb{Q})(|r - q| < \varepsilon).$$

Applying transfer and then taking ε to be a positive infinitesimal means

$$(\forall r \in {}^*\mathbb{R})(\exists q \in {}^*\mathbb{Q})(|r - q| < \varepsilon)$$

is true in ${}^*\mathfrak{R}$. Ergo every hyperreal number and in particular every real number is infinitely close to a hyperrationals. Thus the restriction of the shadow map is still surjective so $({}^*\mathbb{Q} \cap \mathbb{L})/({}^*\mathbb{Q} \cap \mathbb{I}) \simeq \mathbb{R}$. This gives an alternative construction of $\mathbb{R}$ from $\mathbb{Q}$ by preforming the ultrafilter construction on $\mathbb{Q}^{\mathbb{N}}$. Then the limited numbers are ${}^*\mathbb{Q} \cap \mathbb{L}$ and the infinitesimal are ${}^*\mathbb{Q} \cap \mathbb{I}$), so taking the quotient group gives a structure isomorphic to $\mathbb{R}$.

5 Non-Standard Analysis

All the necessary building blocks to start doing analysis with the hyperreals are now in place. Throughout the rest of this paper we prove things about real sequences, series, functions *etc* by finding equivalent condition for their hyper-real counterparts, then carrying out the analysis within the hyperreal framework. These results are based on chapter 6 through 9 of [1].

5.1 Convergence of Sequences

In this subsection we show how the asymptotic behavior of a real sequence s, is determined by what values it takes for unlimited hypernatural. We sometimes call the image of ${}^*\mathbb{N}_\infty$ under s, the *extended tail* of s, painting a picture of the plot of s continuing in infinity.

To begin we remind ourselves that in $\mathbb{R}$, a real valued sequences s converges to a real number L if every open interval around L contains all s terms after some $\omega \in \mathbb{N}$. We then call L the *limit* of s. This can be express by the $\mathcal{L}_{\mathfrak{R}}$-sentence

$$(\forall \varepsilon \in \mathbb{R}^+)(\exists \omega \in \mathbb{N})(\forall n \in \mathbb{N})(\omega \leq n \to |s_n - L| < \varepsilon). \qquad (1)$$

Theorem 5.1. *A real sequence s converges to a real number L if and only if $s_N \simeq L$ for all $N \in {}^*\mathbb{N}_\infty$.*

Proof. For the 'only if'-part, assume s converges to a real number L so (1) is true in $\mathfrak{R}$. It then follows for every positive real ε, there exists a ω_ε such that

$$(\forall n \in \mathbb{N})(\,\omega_\varepsilon \leq n \rightarrow |s_n - L| < \varepsilon)$$

is true in $\mathfrak{R}$ and by transfer it holds for all hypernatural. Since ω_ε is real, it's limited so any $N \in {}^*\mathbb{N}_\infty$ is larger then it so $|s_N - L| < \varepsilon$. Since this can be done for all $\varepsilon \in \mathbb{R}^+$ it means $|s_N - L|$ is infinitesimal. Ergo $s_N \simeq L$ for all $N \in {}^*\mathbb{N}_\infty$.

For the 'if'-part, suppose $s_N \simeq L$ for all $N \in {}^*\mathbb{N}_\infty$. Fixing a $N \in {}^*\mathbb{N}_\infty$, every hyper natural n greater than N must also be unlimited. It follows that $(\forall n \in {}^*\mathbb{N})(N \leq n \rightarrow |s_n - L| < \varepsilon)$ is true in ${}^*\mathfrak{R}$ for any $\varepsilon \in \mathbb{R}^+$. Binding N with an existential quantification results in

$$(\exists \omega \in {}^*\mathbb{N})(\forall n \in {}^*\mathbb{N})(\omega \leq n \rightarrow |s_n - L| < \varepsilon).$$

which is a $\mathcal{L}_{{}^*\mathfrak{R}}$-sentence. Applying transfer means

$$(\exists \omega \in \mathbb{N})(\forall n \in \mathbb{N})(\omega \leq n \rightarrow |s_n - L| < \varepsilon).$$

is true in $\mathfrak{R}$ for each $\varepsilon \in \mathbb{R}^+$, which implies (1). Accordingly s converges to L. □

The pictures this theorem paints is if s converges to L if and only if its extended tails stays entirely in $\mu(L)$. If s is convergent, then $\text{sh}(s_N)$ for any unlimited N is its limit; together with **proposition 4.4** this means that the sum or product of convergent sequences converges to the sum/product of their limits.

Similarly, divergence to positive or negative infinity of a sequences s is also determined by the extended tail. As a reminder, s divergence to positive infinity if the $\mathcal{L}_{\mathfrak{R}}$-sentences

$$(\forall r \in \mathbb{R})(\exists \omega \in \mathbb{N})(\forall n \in \mathbb{N})(\omega \leq n \rightarrow r < s_n) \tag{2}$$

is true in $\mathfrak{R}$ and s divergence to negative infinity if the $\mathcal{L}_{\mathfrak{R}}$-sentences

$$(\forall r \in \mathbb{R})(\exists \omega \in \mathbb{N})(\forall n \in \mathbb{N})(\omega \leq n \rightarrow s_n < r) \tag{3}$$

is true in $\mathfrak{R}$.

Theorem 5.2. *A real sequence s*

(i) diverges to positive infinity if and only if s_N is positive unlimited for all $\mathbb{N} \in {}^*\mathbb{N}_\infty$,

(ii) diverges to negative infinity if and only if s_N is negative unlimited for all $N \in {}^*\mathbb{N}_\infty$.

Proof. For the 'only if'-part of (i), let s be real valued sequence which diverges to positive infinity which means (2) is true in $\mathfrak{R}$. Similarly to the previous proof we fixed an arbitrary r with its corresponding ω_r then apply to transfer to get

$$(\forall n \in {}^*\mathbb{N})(\omega_r \leq n \rightarrow r < s_n)$$

being true in ${}^*\mathfrak{R}$. Continuing like last time, since ω_r is a standard real, we get that $\omega_r \leq N$ and thus $r < s_N$ for all $N \in {}^*\mathbb{N}_\infty$. Since this can be done for any real r it follows that s_N is greater than every real i.e. is positive unlimited.

For the 'if'-part of (i), suppose s_N is positive unlimited for all $N \in {}^*\mathbb{N}_\infty$. Fixing a $N \in {}^*\mathbb{N}_\infty$, every hyper natural n, greater than N, must also be unlimited. So by our assumption $(\forall n \in {}^*\mathbb{N})(N \leq n \rightarrow r < s_N)$ is true for each $r \in \mathbb{R}^+$. Binding N with an existential quantification results in the $\mathcal{L}_{{}^*\mathfrak{R}}$-sentence

$$(\exists \omega \in {}^*\mathbb{N})(\forall n \in {}^*\mathbb{N})(\omega \leq n \rightarrow r < s_N)$$

which is true in ${}^*\mathfrak{R}$. By transfer

$$(\exists \omega \in \mathbb{N})(\forall n \in \mathbb{N})(\omega \leq n \rightarrow r < s_N).$$

is true in $\mathfrak{R}$ for any $r \in \mathbb{R}^+$ so (1) is true in ${}^*\mathfrak{R}$ which means s diverges to positive infinity.

The proof of (ii) is analogous. □

With **theorems 5.1** through **5.3** we can see a sequences s asymptotic behavior just by investigating s_N for unlimited N. We demonstrate this with some basic examples

Example. For the real sequence defined as $s_n = \frac{2n^3+n+2}{n^3+2n^2}$ we can find its limit in a similar way to how one would in standard analysis. Consider an unlimited natural N we get $s_N = \frac{2N^3+N+2}{N^3+2N^2} = \frac{2+\frac{1}{N^2}+\frac{2}{N^3}}{1+\frac{2}{N}}$. By **proposition 4.1** we get $2+\frac{1}{N^2}+\frac{2}{N^3} \simeq 2$ and $1+\frac{2}{N} \simeq 1$ which means both are limited. We can therefore substitute them into the numerator and denominator while presering infinite closeness. Thus $\frac{2+\frac{1}{N^2}+\frac{2}{N^3}}{1+\frac{2}{N}} \simeq \frac{2}{1} = 2$ for all $N \in {}^*\mathbb{N}_\infty$ so by **theorem 5.1** s converges to 2.

For a sequence which doesn't converge, consider $t_n = n(-1)^n$. By transfer, every other hypernatural number is even while every other is odd and -1 to the power of an even number is 1, while to the power of an odd number yields -1. That means that $t_N = N$ while $t_{N+1} = -N$, where N is some even hypernatural. By **theorems 5.1** it follows that t doesn't converge to any real number since $t_N \not\simeq t_{N+1}$. Nor does t diverges to $+\infty$ or $-\infty$ by **theorem 5.2** since t_N is positive unlimited while t_{N+1} is negative unlimited.

Theorem 5.3. *A real sequence s*

(i) is bounded from above if and only if s_N is never positive unlimited for any $N \in {}^*\mathbb{N}_\infty$.

(ii) is bounded from below if and only if s_N is never negative unlimited for any $N \in {}^*\mathbb{N}_\infty$.

Proof. For the 'only if'-part of (i) assume s is bounded from above. Then $(\forall n \in \mathbb{N})(s_n \leq m)$ is true in $\mathfrak{R}$ for some real m. By transfer $(\forall n \in {}^*\mathbb{N})({}^*s_n \leq m))$ is true in ${}^*\mathfrak{R}$ so s_n cannot be positive unlimited for any $n \in {}^*\mathbb{N}$, including all $N \in {}^*\mathbb{N}_\infty$.

For the 'if'-part of (i), assume s_N is non positive unlimited, for all $N \in {}^*\mathbb{N}_\infty$. Since s_n is real for all $n \in \mathbb{N}$ it follows that s_n so not positive unlimited for any $n \in {}^*\mathbb{N}$. Fixing any positive unlimited M it follows that $(\forall n \in {}^*\mathbb{N})(n < M)$ is true which implies $(\exists x \in {}^*\mathbb{R})(\forall n \in {}^*\mathbb{N})(n < x)$ is true in ${}^*\mathfrak{R}$. By transfer $(\exists x \in \mathbb{R})(\forall n \in \mathbb{N})(n < x)$ is true in $\mathfrak{R}$, so s is bounded from above.

The proof of (ii) is analogous. □

Corollary 5.4. A real sequence s is bounded, that is bounded from above and below, if and only if s_N is limited for all $N \in {}^*\mathbb{N}_\infty$.

Proof. Follows immediate from **theorem 5.3** since the only numbers not positive or negative unlimited, are limited. □

With these theorems we can derive some classic analysis results.

Theorem 5.5. *Let s, t be real valued sequences that converge to $L, M \in \mathbb{R}$ respectively. If $s_n \leq t_n$ for all $n \in \mathbb{N}$, then $L \leq M$.*

Proof. If $s_n \leq t_n$ for all $n \in \mathbb{N}$, then by transfer it also holds for all $n \in {}^*\mathbb{N}$. By **theorem 5.1** $L = \mathrm{sh}(s_N)$ and $M = \mathrm{sh}(s_N)$ for any fixed $N \in {}^*\mathbb{N}_\infty$, so since $s_N \leq t_N$, it follows from **proposition 4.2** (v) that $L \leq M$. □

Theorem 5.6 (Squeeze theorem)**.** *If $r_n \leq s_n \leq t_n$ for all $n \in \mathbb{N}$ and both r and t converge to $L \in \mathbb{R}$, then s also converges to L.*

Proof. If $r_n \leq s_n \leq t_n$ for all $n \in \mathbb{N}$ then by transfer it also holds for all $n \in {}^*\mathbb{N}$. By **theorem 5.1** $L \simeq r_N \leq s_N \leq t_N \simeq L$ for all $N \in {}^*\mathbb{N}_\infty$ so by **proposition 4.2** (v), $L \leq \mathrm{sh}(s_N) \leq L$ for all $N \in {}^*\mathbb{N}_\infty$ which means s converge to L. □

Theorem 5.7. *A real sequences $s : \mathbb{N} \to \mathbb{R}$ converges if*

(i) s is bounded from above and non-decreasing or

(ii) s is bounded from below and non-increasing.

Proof. For (i), assume s is bounded from above by some $M \in \mathbb{R}$ and is non-decreasing i.e. $s_1 \leq s_2 \leq s_3 \leq \ldots$. Then s is bounded from below by s_1 and bounded-above by M, so by **corollary 5.4** s_N is limited for $N \in {}^*\mathbb{N}_\infty$. Fixing an unlimited natural N, all natural n are less than N so $s_n \leq s_N$ by s being non-decreasing. By **proposition 4.4** (v) it follows that $\mathrm{sh}(s_n) = s_n \leq \mathrm{sh}(s_N)$ for all $n \in \mathbb{N}$, so $\mathrm{sh}(N)$ is a real upper-bound for the image of $\mathbb{N}$ under s, the real set $\{s_n : n \in \mathbb{N}\}$. Let r be a real upper-bound for $\{s_n : n \in \mathbb{N}\}$. Then the $\mathcal{L}_{\mathfrak{R}}$-sentence $(\forall n \in \mathbb{N})(s_n \leq r)$ is true in $\mathfrak{R}$ so, by transfer, $s_N \leq r$ which implies $\mathrm{sh}(s_N) \leq \mathrm{sh}(r) = r$. Thus s_N is the *least* upper-bound of $\{s_n : n \in \mathbb{N}\}$. Since the least upper-bound is unique and this holds for any $N \in {}^*\mathbb{N}_\infty$, every s_N has the same shadow for every $N \in {}^*\mathbb{N}_\infty$ so s converges by **theorem 5.1**.

The proof of (ii) is analogous. □

An important kind of real sequences are so called *Cauchy* sequences, which are sequences where the terms get arbitrarily close to each other, eventually. Formally, a real valued sequence s is Cauchy if the $\mathcal{L}_{\mathfrak{R}}$-sentences

$$(\forall \varepsilon \in \mathbb{R}^+)(\exists \omega \in \mathbb{N})(\forall m, n \in \mathbb{N})(\omega \leq m, n \to |s_m - s_n| < \varepsilon) \tag{4}$$

is true in $\mathfrak{R}$. We can give a hyperreals characterization of Cauchy sequences as follows

Theorem 5.8. *A real valued sequence s is Cauchy if and only if $s_N \simeq s_M$ for all $M, N \in {}^*\mathbb{N}_\infty$.*

Proof. For the 'only if'-part, assume s is Cauchy so (4) is true in $\mathfrak{R}$. Analogous to the proofs of **theorem 5.1** and **5.2** we fix the ε and ω_ε in (4) giving the true $\mathcal{L}_{\mathfrak{R}}$-sentence

$$(\forall m, n \in \mathbb{N})(\omega_\varepsilon \leq m, n \to |s_m - s_n| < \varepsilon).$$

From transfer, it then follows that $|s_M - s_N| < \varepsilon$ for any unlimited M, N. Since this can be carried out for any positive real ε, we get $s_M \simeq s_N$.

For the 'if'-part, assume $s_N \simeq s_M$ for all $N \in {}^*\mathbb{N}_\infty$. Then fixing an unlimited hypernatural N the sentences

$$(\forall m, n \in {}^*\mathbb{N})(N < m, n \to |s_m - s_m| < \varepsilon)$$

is then true where for any $\varepsilon \in \mathbb{R}^+$. Binding N with an existential quantifier yields the $\mathcal{L}_{^*\mathfrak{R}}$-sentences

$$(\exists \omega \in {}^*\mathbb{N})(\forall m, n \in {}^*\mathbb{N})(\omega < m, n \to |s_m - s_m| < \varepsilon)$$

which is true in ${}^*\mathfrak{R}$ By transfer (4) is true in $\mathfrak{R}$ for every positive realε, so s is Cauchy. □

In **theorem 4.3** we showed that the shadow map is equivalent to the completeness of the real numbers by proving the supremum-axiom. Another statement equivalent to the completeness of $\mathbb{R}$ is 'a sequence converges if and only if it is Cauchy'. Although it's equivalent with **theorem 4.3** we prove it here for pedagogical purposes. However we first need the following lemma.

Lemma 5.9. Every Cauchy sequences is bounded

Proof. Assume s is a Cauchy sequence so (4) is true in $\mathfrak{R}$, which implies $(\forall m, n \in \mathbb{N})(\omega_\varepsilon \leq m, n \to |s_m - s_n| < \varepsilon)$ is true in $\mathfrak{R}$ for a fixed $\varepsilon \in \mathbb{R}^+$ and a corresponding ω_ε. By transfer $(\forall m, n \in {}^*\mathbb{N})(\omega_\varepsilon \leq m, n \to |s_m - s_n| < \varepsilon)$ is true in ${}^*\mathfrak{R}$. Fixing m as ω_ε then yields $(\forall n \in {}^*\mathbb{N})(\omega_\varepsilon \leq \omega_\varepsilon, n \to |s_{\omega_\varepsilon} - s_n| < \varepsilon)$ which is equivalent with $(\forall n \in {}^*\mathbb{N})(\omega_\varepsilon \leq n \to -\varepsilon - s_{\omega_\varepsilon} < -s_n < \varepsilon - s_{\omega_\varepsilon})$. Since both s_{ω_ε} and ε are real numbers, it follows that s_n is limited for all $n \in {}^*\mathbb{N}$, including the unlimited ones. Therefore, by **corollary 5.4**, s is bounded. □

Theorem 5.10. *A real valued sequence converges if and only if it is Cauchy.*

Proof. Suppose s converges to some real L, then $s_N \simeq L$ for all $N \in {}^*\mathbb{N}_\infty$ by **theorem 5.1** which implies $s_N \simeq s_M$ for all $M, N \in {}^*\mathbb{N}_\infty$ by $\simeq$ being an equivalence. Ergo s is Cauchy by **theorem 5.6**.

For the converse, suppose s is Cauchy. By **lemma 5.9** s is bounded so by **corollary 5.4**, s_N is limited for all unlimited natural N. Thus $\mathrm{sh}(s_N)$ is defined for all unlimited N and by **theorem 5.8** all extended terms terms have the same shadow. Therefore, by **theorem 5.1**, s converges to $\mathrm{sh}(s_N)$. □

5.2 Series

A real infinite series $\sum_{k=1}^{\infty} a_k$, where a is a real sequences, is said to converge to a real value A if the sequence of partial sums $\left(\sum_{k=1}^{n} a_k\right)$ convergence to A. We also say $\sum_{k=1}^{\infty} a_k$ is diverges if $\left(\sum_{k=1}^{n} a_k\right)$ diverges. For any unlimited hyper naturals N, we define $\sum_{k=1}^{N} a_k$ as the N-th term of the hyperreal extension of the partial sum sequence. If $M \geq N$ we also define $\sum_{k=N}^{M} a_k$ as $\sum_{k=1}^{M} a_k - \sum_{k=1}^{N-1} a_k$.

Remark. Note that $\sum_{k=1}^{N} a_k$ for a unlimited N is NOT a sum, it is just suggestive notation. Hyper-finite sums can be defined for hyper-integers, but that is beyond the scope of this paper. For the interested, the topic is explored in [1] section 12.7.

We apply our previous results about general sequences to partial sum sequences. In particular **theorem 5.1**, **5.6** and **5.7** implies

Proposition 5.11. For any real series $\sum_{k=1}^{\infty}$ the following holds

(i) $\sum_{k=1}^{\infty} a_k = A$ if and only if $\sum_{k=1}^{N} a_k \simeq A \in \mathbb{R}$.

(ii) $\sum_{k=1}^{\infty} a_k$ converges if and only if $\sum_{k=N}^{M} a_k$ is infinitesimal for all unlimited hyper naturals $M \geq N$. $\square$

If $\sum_{k=1}^{\infty} a_k$ converges, then as a special case of (ii) we have $\sum_{N}^{N} a_k = a_N \simeq 0$ for all $N \in {}^*\mathbb{N}_\infty$ which necessitates that the sequence a must converge to 0. Anyone who's read analysis before, knows that the converse doesn't hold, nevertheless we demonstrate with an example.

Example. For any unlimited natural N, by **proposition 4.1**, $\frac{1}{N} \simeq 0$ so the sequences $(\frac{1}{n})$ converges to 0, nevertheless consider $\sum_{k=1}^{\infty} \frac{1}{k}$. Since $(\frac{1}{n})$ is decreasing, by transfer we have $\sum_{k=n}^{m} \frac{1}{k} \geq (m-n)\frac{1}{m}$ for any $m, n \in {}^*\mathbb{N}$ since $m - n$ is the 'number of terms'. In particular, for a unlimited natural N, we get $\sum_{N}^{2N} \frac{1}{k} \geq (2N - N)\frac{1}{2N} = \frac{N}{2N} = \frac{1}{2}$ which is not infinitesimal so. Thus $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges by **proposition 5.11** (ii) .

With **Proposition 5.11** we are able to prove some classical results about series.

Proposition 5.12. If $\sum_{k=1}^{\infty} a_k$ and $\sum_{k=1}^{\infty} b_k$ are convergent series, and c is some real number, then

- $\sum_{k=1}^{\infty}(a_k \pm b_k) = \sum_{k=1}^{\infty} a_k \pm \sum_{k=1}^{\infty} b_k$
- $\sum_{k=1}^{\infty} ca_k = c\sum_{k=1}^{\infty} a_k$

Proof. Let $\sum_{k=1}^{\infty} a_k$ converge to $A \in \mathbb{R}$ and $\sum_{k=1}^{\infty} b_k$ converge to $B \in \mathbb{R}$. Then $\sum_{k=1}^{N} a_k \simeq A$ and $\sum_{k=1}^{N} b_k \simeq B$ for all $N \in {}^*\mathbb{N}_\infty$. The $\mathcal{L}_{\mathfrak{R}}$-sentences $(\forall n \in \mathbb{N})(\sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k = \sum_{k=1}^{n}(a_k + b_k))$ is true in $\mathfrak{R}$ so, by transfer, it holds for all hyper natural n, including the unlimited ones. Thus $\sum_{k=1}^{N}(a_k \pm b_k) = \sum_{k=1}^{N} a_k \pm \sum_{k=1}^{N} b_k \simeq A + B$ for all $N \in {}^*\mathbb{N}_\infty$ so $\sum_{k=1}^{\infty}(a_k \pm b_k) = A + B = \sum_{k=1}^{\infty} a_k \pm \sum_{k=1}^{\infty} b_k$.

The second statement follows from transfer on $(\forall n \in \mathbb{N})(\sum_{k=1}^{n} ca_k = c\sum_{k=1}^{n} a_k)$ and similar reasoning. $\square$

Lemma 5.13. Let $a_n \geq 0$ for all $n \in \mathbb{N}$. Then $\sum_{k=1}^{\infty} a_k$ converges if and only if $\sum_{k=1}^{N} a_k$ is limited for all $N \in {}^*\mathbb{N}_\infty$, which holds if and only if $\sum_{k=1}^{N} a_k$ is limited for some $N \in {}^*\mathbb{N}_\infty$

Proof. If $\sum_{k=1}^{\infty} a_k$ converges, then for all unlimited N there exits some real A such that $\sum_{k=1}^{N} a_k \simeq A$. As a consequence $\sum_{k=1}^{N} a_k$ is limited for all $N \in {}^*\mathbb{N}_\infty$.

If $\sum_{k=1}^{N} a_k$ is limited, for all $N \in {}^*\mathbb{N}_\infty$, then $\sum_{k=1}^{N} a_k$ is limited, for some $M \in {}^*\mathbb{N}_\infty$.

Let M be an unlimited hypernatural such that s_M is limited. Since a is non-negative, we know that $(\forall m, n \in \mathbb{N})(m < n \to \sum_{k=1}^{m} a_k \leq \sum_{k=1}^{n} a_k)$ is true in $\mathfrak{R}$. Applying transfer and by M being unlimited, we get $\sum_{k=1}^{M}$ is an upper-bound for $\{\sum_{k=1}^{m} a_k : m \in \mathbb{N}\}$. Since s_M is limited, we can get that $\text{sh}(s_M)$ is a real upper-bound for $\{\sum_{k=1}^{m} a_k : m \in \mathbb{N}\}$. The real sequence of partial sums is then a non-decreasing sequence bounded from above, so by **theorem 5.6** it converge, so $\sum_{k=1}^{\infty} a_k$ converges. $\square$

Proposition 5.14 (Comparison Test 1)**.** If $0 \leq a_n \leq b_n$ for all $n \in \mathbb{N}$, then

- $\sum_{k=1}^{\infty} a_k$ being divergent implies $\sum_{k=1}^{\infty} b_k$ being divergent.
- $\sum_{k=1}^{\infty} b_k$ being convergent implies $\sum_{k=1}^{\infty} a_k$ being convergent.

Proof. We begin by proving the second statement. If $0 \leq a_n \leq b_n$ for all $n \in \mathbb{N}$, then by induction $0 \leq \sum_{k=1}^{n} a_k \leq \sum_{k=1}^{n} b_k$ for all $n \in \mathbb{N}$ and by transfer this holds for all ${}^*\mathbb{N}$. If $\sum_{k=1}^{\infty} b_k$ converges, then by **lemma 5.13**, $\sum_{k=1}^{N} b_k$ is limited for some $N \in {}^*\mathbb{N}_\infty$ which by $0 \leq \sum_{k=1}^{N} a_k \leq \sum_{k=1}^{N} b_k$, implies $\sum_{k=1}^{N} a_k$ is also limited. Ergo $\sum_{k=1}^{\infty} a_k$ converges by **lemma 5.13**.

The first statement is just the contra-position of the first and therefore equivalent. $\square$

Proposition 5.15 (Comparison Test 2)**.** If a and b are positive real sequences and the sequence $\frac{a}{b}$ converges, then $\sum_{k=1}^{\infty} a_k$ converges if and only if $\sum_{k=1}^{\infty} b_k$ converges.

Proof. Let a and b be positive real sequences and let the sequence $\frac{a}{b}$ converge to some real L. By **theorem 5.1** $\frac{a_n}{b_n} \simeq L$ for all $N \in {}^*\mathbb{N}_\infty$, it follows that $\frac{L}{2} < \frac{a_N}{b_N} < \frac{3L}{2}$ and thus $\frac{L}{2}b_N < a_N < \frac{3L}{2}b_N$ for all $N \in {}^*\mathbb{N}_\infty$ since b is positive. Fixing some unlimited H we then get that $(\forall n \in {}^*\mathbb{N})(H \leq n \to \frac{L}{2}b_n < a_n < \frac{3L}{2}b_n)$. Binding H with an existentialist quantifier we get a true $\mathcal{L}_{{}^*\mathfrak{R}}$-sentences and applying transfer we get that

$$(\exists H \in {}^*\mathbb{N})(\forall n \in {}^*\mathbb{N})\big(H \leq n \to (\frac{L}{2})b_n < a_n < (\frac{3L}{2})b_n\big)$$

is true in ${}^*\mathfrak{R}$. By transfer, this implies that $\frac{L}{2}b_n < a_n < \frac{3L}{2}b_n$ for all natural numbers n, greater then some fixed natural number h. It follows by induction that $\frac{L}{2}\sum_{k=n}^{m} b_k < \sum_{k=n}^{m} a_k < \frac{3L}{2}\sum_{k=n}^{m} b_k$ holds for all m, n such that $h \leq n \leq m$. By transfer

$$(\forall m, n \in {}^*\mathbb{N})\Big(h \leq n \leq m \to \frac{L}{2}\sum_{k=n}^{m} b_k < \sum_{k=n}^{m} a_k < \frac{3L}{2}\sum_{k=n}^{m} b_k\Big)$$

is true in ${}^*\mathfrak{R}$. Since all unlimited hypernaturals are greater then h we get, and

$$\frac{L}{2}\sum_{k=N}^{M} b_k < \sum_{k=N}^{M} a_k < \frac{3L}{2}\sum_{k=N}^{M} b_k$$

for all unlimited M, N such that $N \leq M$.

If $\sum_{k=1}^{\infty} b_k$ converges, then $\frac{L}{2}\sum_{k=N}^{M} b_k$ and $\frac{3L}{2}\sum_{k=N}^{M} b_k$ are infinitesimal for all unlimited M, N such that $N \leq M$ by **proposition 5.11** (ii). The inequality then forces$\sum_{k=N}^{M} a_k$ to be infinitesimal for all unlimited M, N such that $N \leq M$, which means $\sum_{k=1}^{\infty} a_k$ converges.

If $\sum_{k=1}^{\infty} a_k$ converges, then $\sum_{k=N}^{M} a_k$ is infinitesimal for all unlimited hypernatural M, N such that $N \leq M$ by **proposition 5.11** (ii). Since b being positive, the sequences of partial sums is non-decreasing so $0 \leq \sum_{k=n}^{m} b_k = \sum_{k=1}^{m} b_k - \sum_{k=n-1}^{n} b_k$ for all natural numbers $n \leq m$, so by transfer for all hypernatural numbers $M \leq N$. By the above inequality we then get $0 < \frac{L}{2}\sum_{k=N}^{M} b_k < \sum_{k=N}^{M} a_k$ for all unlimited hypernaturals M, N such that $N \leq M$, which implies $\frac{L}{2}\sum_{k=N}^{M} b_k$ being infinitesimal. Ergo $\sum_{k=1}^{\infty} b_k$ converges. □

5.3 Continuity

As a reminder, in standard analysis a real valued function f is continuous at a point in the interval c if, for every opens subset around $f(c)$, there exists an open subset around c which gets mapped into it. This capture's the notion that $f(x)$ should approach $f(c)$ when x approaches c. Formally, continuity is expressed by the $\mathcal{L}_{\mathfrak{R}}$-sentence

$$(\forall \varepsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x \in D_f)(|x - c| < \delta \to |f(x) - f(c)| < \varepsilon) \tag{5}$$

being true in $\mathfrak{R}$. This $\varepsilon\delta$-definition is famously daunting the first time one sees it; non-standard analysis can however, with the use of infinitesimals, capture continuity quite succinctly.

Theorem 5.16. *A function f, defined for a real value c, is continuous at c if and only if $x \simeq c$ implies $f(x) \simeq f(c)$ for all $x \in {}^*D_f$.*

Proof. Suppose f is continuous at c so (5) is true in $\mathfrak{R}$, thus for any positive real ε, there is a corresponding positive real δ_ε such that

$$(\forall x \in D_f)(|x - c| < \delta_\varepsilon \to |f(x) - f(c)| < \varepsilon)$$

is true in $\mathfrak{R}$. By transfer this holds for all of ${}^*D_f$, in particular those x infinitely close to c. Then since $x - c$ is infinitesimal, it follows that it's less then every possible δ_ε thus $|f(x) - f(c)| < \varepsilon$ is true in $\mathfrak{R}$ for any $\varepsilon \in \mathbb{R}^+$, ergo $f(x) \simeq f(c)$.

For the converse, suppose $x \simeq c$ implies $f(x) \simeq f(c)$ for all $x \in {}^*D_f$. Fixing a positive infinitesimal δ, if $|x - c| < \delta$, then $x \simeq c$ so, by assumption, $|x - c| < \delta$ implies $f(x) \simeq f(c)$. Binding δ with an existential quantifier, we get that

$$(\exists \delta \in {}^*\mathbb{R}^+)(\forall x \in {}^*D_f)(|x - c| < \delta \to |f(x) - f(c)| < \varepsilon)$$

is in ${}^*\mathfrak{R}$, for every $\varepsilon \in \mathbb{R}^+$. Applying transfer we infer that f is continuous at c. □

We say that a function f is continuous on a set $X \subseteq \mathbb{R}$ if it is continuous on every point in X, and that f is continuous if it is continuous on its domain. To demonstrate, we investigate the continuity of the cosine function

Example. For any real c and any infinitesimal ε, we have $\cos(c + \varepsilon) - \cos c = \cos c \cos \varepsilon - \sin \varepsilon \sin c - \cos c \simeq \cos c \cdot 1 + 0 \cdot \sin c - \cos c \simeq 0$ by transfer on the addition formula for cosine, and because $\sin \varepsilon \simeq 0, \cos \varepsilon \simeq 1$. So by **theorem 5.16**, cosine is a continuous function.

We now prove the intermediate value theorem and the extreme value theorem - two classic results which are both, in standard analysis, proved by dividing up an interval into smaller sub-interval. The non-standard proof builds on the same idea but considers sub-intervals of infinitesimal length.

Theorem 5.17 (The Intermediate Value Theorem). *If a real function f is continuous on the closed interval $[a, b]$, then for every d strictly in between $f(a)$ and $f(b)$, there exists a $c \in]a, b[$ such that $f(c) = d$.*

Proof. Assume without loss of generality that $f(a) < f(b)$, so $f(a) < d < f(b)$. For every $n \in \mathbb{N}$ we partition $[a, b]$ into n sub-intervals of width $\frac{b-a}{n}$. The partition is defined by its endpoints, we denote the set of endpoints, for n-interval partition as $P_n = \{a + k\frac{b-a}{n} : k = 0, 1, \ldots, n\}$. The set $\{p_k \in P_n : f(p_k) < d\}$ is always finite, having at most n elements, and is never empty since $p_0 = a$ is always in it. Consequentially we can define a sequence s such that $s_n = \max\{p_k \in P_n : f(p_k) < d\}$. By how s is defined, it follows that

$$a \leq s_n < b \text{ and } f(s_n) < d \leq f(s_n + \frac{b-a}{n})$$

for all $n \in \mathbb{N}$. By transfer, the same hold for all $n \in {}^*\mathbb{N}$ for the hyperreal extension of s. Fixing any unlimited natural N, by the first condition s_N is limited so we may define $c = \text{sh}(s_N)$. If follows that $\frac{b-a}{N}$ is infinitesimal, so $s_N + \frac{b-a}{N} \simeq s_N \simeq c$. By continuity $f(s_N + \frac{b-a}{N}) \simeq f(c)$. By the second condition of s, we get that d is between $f(s_N)$ and $f(s_N + \frac{b-a}{N})$, hence $d \simeq f(s_N) \simeq f(c)$. Since both d and $f(c)$ are real, they must be equal by **theorem 4.3**. □

Theorem 5.18 (Extreme Value Theorem). *If a real function f is continuous on the closed interval $[a, b]$, then f assume a maximum and minimum on $[a, b]$ i.e there exists $c, d \in [a, b]$ such that $f(c) \leq f(x) \leq f(d)$ for all $x \in [a, b]$.*

Proof. To obtain a maximum for f in $[a, b]$, consider the partitions $P_n = \{a + k\frac{b-a}{n} : k = 0, 1, \ldots, n\}$ from the previous proof, and let s_n be the least $p_k \in P_n$ such that $f(p_k)$ is maximal. Again s_n is well-defined since for every $n \in \mathbb{N}$, there are a finite amount of endpoints. Consequently

$$a \leq s_n \leq b \text{ and } f(a + k\frac{b-a}{n}) \leq f(s_n)$$

holds for all $n \in \mathbb{N}$ and integers k such that $0 \leq k \leq n$. Applying transfer, the statement holds for all $N \in {}^*\mathbb{N}$ and hyperintegers K such that $0 \leq K \leq N$. The first condition ensures s_N always being limited, so let $d = \text{sh}(s_M)$ for a fixed $M \in {}^*\mathbb{N}_\infty$.

Note that for each $n \in \mathbb{N}$, fixing $r \in [a, b]$, there exists an integer $0 \leq k \leq n$ such that

$$a + k\frac{b-a}{n} \leq r \leq a + (k+1)\frac{b-a}{n}.$$

Then by transfer, for each $N \in {}^*\mathbb{N}$ there exists a hyperinteger $0 \leq K \leq N$ such that

$$a + K\frac{b-a}{N} \leq r \leq a + (K+1)\frac{b-a}{N}.$$

Then since $a + K\frac{b-a}{N} \simeq a + (K+1)\frac{b-a}{N}$ it follows that $r \simeq K\frac{b-a}{N}$. Then for our fixed $M \in {}^*\mathbb{N}_\infty$, and a fixed $r \in [a, b]$ there exists a hyperinteger $0 \leq K_r \leq M$ such that $r \simeq a + K_r\frac{b-a}{M}$. The continuity of f on $[a, b]$, then implies that $f(a + K_r\frac{b-a}{M}) \simeq f(r) \in \mathbb{R}$ and $f(s_M) \simeq f(d) \in \mathbb{R}$. By the second condition of s, we know that $f(a + K\frac{b-a}{M}) \leq f(s_M) \simeq f(d)$ for all hyperintegers K, so by **Proposition 4.4** (iv), $\text{sh}\big(f(a + K_r\frac{b-a}{N})\big) = f(r) \leq \text{sh}\big(f(s_M)\big) = f(d)$. This holds for any $r \in [a, b]$, so it follows that $f(d)$ is the maximum of f on $[a, b]$.

The proof f obtains a minimum is analogous. □

Another advantage of non-standard analysis is that usually hard to understand concept *uniform continuity* is (relatively) more intuitive. In standard analysis, a function f is *uniformly continuous* on X if

$$(\forall \varepsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x, y \in X)(|x - y| < \delta \to |f(x) - f(y)| < \varepsilon) \tag{6}$$

is true in $\mathfrak{R}$. The only difference between this and regular continuity on a set is that the x-quantifier is within the existential quantifier. What this means is that the choice of δ cannot depend on the point we're evaluating continuity at, only the value ε. However, In the hyperreals, uniform continuity can be expressed in the following way.

Theorem 5.19. *A function f is uniformly continuous on a real set X if and only if $x \simeq y$ implies $f(x) \simeq f(y)$ for all $x, y \in {}^*X$.*

Proof. Assume that f is uniformly continuous on $X \subseteq \mathbb{R}$ so (6) is true in $\mathfrak{R}$. Then, fixing some $\varepsilon \in \mathbb{R}^+$, there exists a $\delta_\varepsilon \in \mathbb{R}^+$ such that

$$(\forall x, y \in X)(|x - y| < \delta_\varepsilon \rightarrow |f(x) - f(y)| < \varepsilon).$$

By transfer

$$(\forall x, y \in {}^*X)(|x - y| < \delta_\varepsilon \rightarrow |f(x) - f(y)| < \varepsilon)$$

is true in ${}^*\mathfrak{R}$. If $x, y \in {}^*X$ are infinitely close, then $|x - y|$ is less then every δ_ε, so $|f(x) - f(y)| < \varepsilon$ is true for every $\varepsilon \in \mathbb{R}^+$, i.e. $f(x) \simeq f(y)$.

For the converse, suppose $x \simeq y$ implies $f(x) \simeq f(y)$ for all $x, y \in {}^*X$. Then, fixing a positive infinitesimal δ, the sentence

$$(\forall x, y \in {}^*X)(|x - y| < \delta \rightarrow |f(x) - f(y)| < \varepsilon)$$

is true in for each $\varepsilon \in \mathbb{R}^+$. Binding δ with an existential quantifier and applying transfer we get that

$$(\exists \delta \in \mathbb{R}^+)(\forall x, y \in X)(|x - y| < \delta \rightarrow |f(x) - f(y)| < \varepsilon)$$

is true in ${}^*\mathfrak{R}$ for all $\varepsilon \in \mathbb{R}^+$. Ergo, f is uniformly continuous on X. □

Theorem 5.20. *If a real valued function f is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$.*

Proof. Let f be continuous on $[a, b]$ and consider $x, y \in {}^*[a, b]$ such that $x \simeq y$. Since $x, y \in {}^*[a, b]$ they must be limited. It then follows that $\text{sh}(x) = \text{sh}(y) = c \in \mathbb{R}$. By f being continuous on $[a, b]$ and $x, y \simeq c$ we get $f(x), f(y) \simeq f(c)$ so by $\simeq$ being an equivalence, $f(x) \simeq f(y)$. Thus by **theorem 7.4**, f is uniformly continuous on $[a, b]$. □

Remark. This result doesn't holds for open intervals $]a, b[$ or unbounded intervals of the form $[a, \infty[$, $]-\infty, a]$. In the open case it doesn't holds since $]a, b[$ isn't closed under taking shadow e.g. $a+\varepsilon \in]a, b[$ but $\text{sh}(a+\varepsilon) = a \notin]a, b[$. In the unbounded case, all their elements aren't limited so the shadow of x and y isn't always defined.

We prove an interesting connection between Cauchy sequences and uniformly continuous functions.

Proposition 5.21. If a real valued sequence s is Cauchy and f is uniformly continuous on the co-domain of s, then the sequences $f \circ s$ is also Cauchy.

Proof. If s is Cauchy, then $s_M \simeq s_N$ for all $M, N \in {}^*\mathbb{N}_\infty$ by **theorem 5.8**. By f being uniformly continuous on the co-domain of s, it follows that $f(s_M) \simeq f(s_N)$ for for all $M, N \in {}^*\mathbb{N}_\infty$, hence $f \circ s$ is also Cauchy. □

Regular continuity isn't sufficient since $f(s_M), f(s_N)$ may be non-standard - we demonstrate via example.

Example. The function $f(x) = \frac{1}{x}$ is continuous on $]0, 1]$ since $f(c) - f(c + \varepsilon) = \frac{1}{c} - \frac{1}{c+\varepsilon} = \frac{c+\varepsilon-c}{c^2+c\varepsilon} = \frac{\varepsilon}{c^2+c\varepsilon}$ is infinitesimal for all real $c \in]0, 1]$ and all infinitesimal ε. It is however not uniformly continuous on $]0, 1]$ because, for positive infinitesimals $\varepsilon, \varepsilon^2 \in {}^*]0, 1]$, we have $f(\varepsilon^2) - f(\varepsilon) = \frac{1}{\varepsilon^2} - \frac{1}{\varepsilon} = \frac{1-\varepsilon}{\varepsilon^2}$ which is unlimited despite $\varepsilon \simeq \varepsilon^2$. So by **theorem 5.19**, f isn't uniformly continuous on $]0, 1]$. The sequences s defined as $s_n = \frac{1}{n}$ is Cauchy since it convergence to 0, it also stays entirely in $]0, 1]$. Despite this we see that $(f \circ s)(n) = \frac{1}{\frac{1}{n}} = n$ which divergence to positive infinity, so cannot be Cauchy.

In standard analysis, continuity is usually defined in terms of limits. In non-standard analysis we opted for a more intuitive definition but rest assured that limits can still be defined in the hyperreals.

Theorem 5.22. *Given a real valued function f and a real c*

1. $\lim_{x \to c^-} f(x) = L$ *if and only if* $f(x) \simeq L$ *for all* $x \in {}^*D_f$ *such that* $x \simeq c$ *and* $x < c$.
2. $\lim_{x \to c^+} f(x) = L$ *if and only if* $f(x) \simeq L$ *for all* $x \in {}^*D_f$ *such that* $x \simeq c$ *and* $x > c$.
3. $\lim_{x \to c} f(x) = L$ *if and only if* $f(x) \simeq L$ *for all* $x \in {}^*D_f$ *such that* $x \simeq c$ *and* $x \neq c$.
4. $\lim_{x \to +\infty} f(x) = L$ *if and only if* $f(x) \simeq L$ *for all positive unlimited* $x \in {}^*D_f$.
5. $\lim_{x \to -\infty} f(x) = L$ *if and only if* $f(x) \simeq L$ *for all negative unlimited* $x \in {}^*D_f$.
6. $\lim_{x \to c} f(x) = +\infty$ *if and only if* $f(x)$ *is positive unlimited for all* $x \in {}^*D_f$ *with* $x \simeq c$ *and* $x \neq c$.
7. $\lim_{x \to c} f(x) = -\infty$ *if and only if* $f(x)$ *is negative unlimited for all* $x \in {}^*D_f$ *with* $x \simeq c$ *and* $x \neq c$.

Proof. Part 1,2 and 3 are very similar so we only prove 3 here. For the 'only if'-part, assume $\lim_{x\to c^-} f(x) = L$, which means

$$(\forall \varepsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x \in D_f)(|x - c| < \delta \to |f(x) - L| < \varepsilon)$$

is true in $\mathfrak{R}$. Thus fixing any positive real ε, there is a δ_ε such that

$$(\forall x \in D_f)(|x - c| < \delta_\varepsilon \to |f(x) - L| < \varepsilon)$$

is true in $\mathfrak{R}$. By transfer then, if $x \in {}^*D_f$ is such that $x \simeq c$ and $x \neq c$, then by $|x - c|$ being less then any $\delta_\varepsilon \in \mathbb{R}^+$ we get that $|f(x) - L| < \varepsilon$ is true in ${}^*\mathfrak{R}$, for any $\varepsilon \in \mathbb{R}^+$ - hence $f(x) \simeq L$.

For the 'if'-part assume $x \simeq c$ and $x \neq c$ implies $f(x) \simeq L$, for all $x \in {}^*D_f$. Then fixing a positive infinitesimal δ, we have $|x-c| < \delta \to |f(x)-L| < \varepsilon$ being true for any positive real ε. Binding δ with an existential quantifier we get a true $\mathcal{L}_{{}^*\mathfrak{R}}$ and applying transfer we get that

$$(\exists \delta \in \mathbb{R}^+)(\forall x \in D_f)(|x - c| < \delta \to |f(x) - L| < \varepsilon)$$

is true in $\mathfrak{R}$ for any positive real ε. Ergo $\lim_{x\to c^-} f(x) = L$.

Part 4 and 5 are similar to the proof of **theorem 5.1** while while 6 and 7 are almost identical to the proof of theorem **5.2**. ☐

It follows that that $\lim_{x\to c} f(x) = L$ if and only if $\lim_{x\to c^-} f(x) = \lim_{x\to c^+} f(x) = L$. Also note that this aligns with the standard definition of a function being continuous at a point c if $\lim_{x\to c} f(x) = f(c)$. We now investigate some interesting limits with non-standard methods.

Example.

1. Let $f(x) = \sin\frac{1}{x}$ be the function defined for all non-zero reals, consider $\lim_{x\to c} f(x)$. If $c \neq 0$, then for any $x \simeq c$ we have $\frac{1}{x} \simeq \frac{1}{c}$ so, by sine being continuous we have $\sin\frac{1}{x} \simeq \sin\frac{1}{c}$. Thus $\lim_{x\to c} f(x) = \sin\frac{1}{c}$ for all non-zero c. However if $c = 0$, then for $\lim_{x\to c} f(x) = L$ to exists, $f(\varepsilon)$ must be infinitely close to some real number, for every non-zero infinitesimal ε. Since the inverse of an unlimited number is a non-zero infinitesimal we know that both $\frac{1}{2\pi K}$ and $\frac{1}{2\pi K + \frac{\pi}{2}}$ are infinitesimals where K is an unlimited hyperintegers. By transfer $\sin(\frac{1}{\frac{1}{2\pi K}}) = \sin(2\pi K) = 0$ while $\sin(\frac{1}{\frac{1}{2\pi K+\frac{\pi}{2}}}) = \sin(2\pi K + \frac{\pi}{2}) = 1$ so the limit a at 0 cannot exists.

2. For an opposite example consider the piecewise function

$$g(x) = \begin{cases} x & \text{if } x \text{ is rational,} \\ -x & \text{if } x \text{ is irrational,} \end{cases}$$

defined on the all of $\mathbb{R}$. For all non-zero infinitesimal ε, either $g(\varepsilon) = \varepsilon \simeq 0$ or $g(\varepsilon) = -\varepsilon \simeq 0$, hence $\lim_{x\to 0} g(x) = 0$. As for when c is non-zero, we know every real number is infinitely close to a hyperrational and a hyperirationals by transfer on the true sentence

$$(\forall \varepsilon \in \mathbb{R}^+)(\forall c \in \mathbb{R})(\exists x \in \mathbb{Q})(\exists y \in \overline{\mathbb{Q}})(c \neq x \wedge c \neq y \wedge |c - x| < \varepsilon \wedge |c - y| < \varepsilon).$$

Then for all $c \in {}^*\mathbb{R}$, there exists a hyperrational x, and a hyperirational y, such that $c \simeq x, y$. Moreover $g(x) = x \simeq c$ while $g(y) = -y \simeq -c$, so if c is non-zero, then $g(x) \not\simeq g(y)$ which means the limit $\lim_{x\to c} g(x)$ cannot exists.

5.4 The Derivative

The derivative of a real function f at a real point x, denoted $f'(x)$, represent the rate of change of f in x. Alternatively $f'(x)$ is the slope of the tangent of f passing through x. In standard analysis, $f'(x)$ is defined as the limit

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}. \tag{7}$$

In non-standard analysis however, we can almost use Leibniz' original definition introduced in the **introduction**, but with our new non-standard tools.

Theorem 5.23. *A real valued function f has a derivative at the real point x, denoted $f'(x)$, if and only if f is defined on $\mu(x)$ and*

$$\frac{f(x+\varepsilon)-f(x)}{\varepsilon}\simeq f'(x)$$

for all non-zero infinitesimals ε.

Proof. A real function f has a derivative $f'(x)$ at $x\in\mathbb{R}$ if and only if

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=f'(x)$$

and by **theorem 5.22**, that is equivalent with

$$\frac{f(x+\varepsilon)-f(x)}{\varepsilon}\simeq f'(x)$$

for all non-zero infinitesimal ε. □

So when f is differentiable at x we have

$$f'(x)=\operatorname{sh}\Big(\frac{f(x+\varepsilon)-f(x)}{\varepsilon}\Big).$$

The linearity of the derivative then follows from the linearity of the shadow map. Note the similarity, but also the difference, with Leibniz definition, intruded in the introduction; Let us demonstrate the difference by computing the derivative of $f(x)=x^n$.

Example. Consider the function $f(x)=x^n$, then

$$\frac{f(x+\varepsilon)-f(x)}{\varepsilon}=\frac{(x+\varepsilon)^n-x^n}{\varepsilon}$$

where ε is a non-zero infinitesimal. We know that the binomial theorem applies for infinitesimal by the transfer principle, so $(x+\varepsilon)^n=x^n+nx^{n-1}\varepsilon+\binom{n}{2}x^{n-2}\varepsilon^2\cdots+\varepsilon^n$. The value is never unlimited, so we may substitute it into the fraction we compute

$$\frac{x^n+nx^{n-1}\varepsilon+\binom{n}{2}x^{n-2}\varepsilon^2\cdots+\varepsilon^n-x^n}{\varepsilon}=\frac{nx^{n-1}\varepsilon+\binom{n}{2}x^{n-2}\varepsilon^2\cdots+\varepsilon^n}{\varepsilon}=nx^{n-1}+\cdots+\varepsilon^{n-1}\simeq nx^{n-1},$$

arriving at $f'(x)=nx^{-1}$. In the final step we use the arithmetic rules we proved in **proposition 4.1**. We get around the problem of ε being both non-zero when dividing by it, and zero when removing all its multiples, by the fraction not equaling nx^{n-1}, but being infinitely close to it.

In non-standard, an interesting quantity arises in relation to the derivative, namely

Definition 5.24 (Increment)**.** Given a real number x, a function f defined on $\mu(x)$ and a non-zero infinitesimal Δx, we call the difference $f(x+\Delta x)-f(x)$ the *increment of f at x by Δx*, denoted Δf

The notation is suggestive since Δf stands for the increase in f, when x increments by an infinitesimal amount Δx. The fraction $\frac{\Delta f}{\Delta x}$ is then the slope of the line passing through $(x,f(x))$ and $(x+\Delta x,f(x+\Delta x))$, and when f is differentiable at $x\in\mathbb{R}$, we have

$$\frac{\Delta f}{\Delta x}\simeq f'(x).$$

for non-zero infinitesimal Δx. If f is differentiable at x, it also follows that $\frac{\Delta f}{\Delta x}$ is limited which necessitates $\Delta f=\frac{\Delta f}{\Delta x}\Delta x$ being infinitesimal, so $\Delta f=f(x+\Delta x)-f(x)\simeq 0$, which implies $f(x+\Delta x)\simeq f(x)$, for all infinitesimal Δx. By **theorem 5.16** we then have

Proposition 5.25. If f is differentiable at $x\in\mathbb{R}$, then f is continuous at x. □

Additionally, if $\frac{\Delta f}{\Delta x}\simeq f'(x)$ when f is differentiable at x, we have $\frac{\Delta f}{\Delta x}-f'(x)=\varepsilon\in\mathbb{I}$ and multiplying both sides by Δx we get

$$\Delta f=f'(x)\Delta x+\varepsilon\Delta x \text{ or equivalently } f(x+\Delta x)=f'(x)+\Delta x f(x)+\varepsilon\Delta x.$$

This is result is called the *Incremental Equation* and is interesting because it shows relation between the actual increment of the function Δf, and $f'(x)\Delta x$ which is the increment along the tangent of f at x. Even though they're both infinitesimal, the equation shows that they're difference is infinitesimal, to a 'higher order' than Δx, i.e. $\frac{\Delta f - f'(x)\Delta x}{\Delta x} = \frac{\varepsilon \Delta x}{\Delta x} = \varepsilon$ is still infinitesimal. The second formulation also shows how, for a fixed x, the function $l(\Delta x) = f(x) + f'(x)\Delta x$ linearly approximates $f(x + \Delta x)$ for infinitesimal Δx, with an infinitesimal error compared to Δx.

All the classic rules for the derivative follow

Proposition 5.26. If f and g are function, differentiable at $x \in \mathbb{R}$, then $f+g$, $f \cdot g$ and $\frac{f}{g}$ are also differentiable at $x \in \mathbb{R}$ in the following ways

- $(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x)$,
- $(\frac{f}{g})' = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$ assuming $g(x) \neq 0$.

Proof. Analyzing $\Delta(f \cdot g)$ and using some algebra, we get

$$\begin{aligned}\Delta(f \cdot g) &= f(x + \Delta x)g(x + \Delta x) - f(x)g(x) \\ &= (f(x) + \Delta f)(g(x) + \Delta g) - f(x)g(x) && [f(x + \Delta x) = f(x) + \Delta f] \\ &= f(x)\Delta g + g(x)\Delta f + \Delta f \Delta g.\end{aligned}$$

Diving by Δx then yields

$$\frac{\Delta(f \cdot g)}{\Delta x} = \frac{f(x)\Delta g + g(x)\Delta f + \Delta f \Delta g}{\Delta x} = f(x)\frac{\Delta g}{\Delta x} + \frac{\Delta f}{\Delta x}g(x) + \Delta f \frac{\Delta g}{\Delta x} \simeq f(x)g'(x) + f'(x)g(x).$$

Analyzing $\Delta(\frac{f}{g})$ we get

$$\begin{aligned}\Delta(\frac{f}{g}) &= \frac{f(x + \Delta x)}{g(x + \Delta x)} - \frac{f(x)}{g(x)} \\ &= \frac{f(x + \Delta x)g(x) - f(x)g(x + \Delta x)}{g(x + \Delta x)g(x)} \\ &= \frac{f(x + \Delta x)g(x) - f(x)g(x + \Delta x) + f(x)g(x) - f(x)g(x)}{g(x + \Delta x)g(x)} \\ &= \frac{(f(x + \Delta x) - f(x))g(x) - f(x)(g(x + \Delta x) - g(x))}{g(x + \Delta x)g(x)} \\ &= \frac{\Delta f g(x) - f(x)\Delta g}{g(x + \Delta x)g(x)}.\end{aligned}$$

Dividing by Δx yields

$$\begin{aligned}\frac{\Delta f g(x) - f(x)\Delta g}{\Delta x} \cdot \frac{1}{g(x + \Delta x)g(x)} &= \left(\frac{\Delta f}{\Delta x}g(x) - f(x)\frac{\Delta g}{\Delta x}\right)\frac{1}{g(x + \Delta x)g(x)} \\ &\simeq \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}. && [g(x + \Delta x) \simeq g(x)]\end{aligned}$$

□

Proposition 5.27 (The Chain rule)**.** If f is differentiable at $x \in \mathbb{R}$ and g differentiable at $f(x)$, then $g \circ f$ is differentiable at x with $(g \circ f)'(x) = g'(f(x))f'(x)$.

Proof. By f being differentiable at x it's also continuous at x, so $f(x+\Delta x) \simeq f(x)$ for all non-zero infinitesimal Δx. By g being differentiable at $f(x)$ it is defined for $\mu(f(x))$, so $(g \circ f)(x + \Delta x) = g(f(x + \Delta x))$ is always defined.

Investigating the increment of $g \circ f$ we see

$$\Delta(g \circ f) = g(f(x + \Delta)) - g(f(x)) = g(f(x) + \Delta f) - g(f(x)),$$

or put into words, the increment of $g \circ f$ at x by Δx is equal to the increment of g at $f(x)$ by Δf. Applying the incremental equation then yields

$$\Delta(g \circ f) = g'(f(x))\Delta f + \varepsilon \Delta f$$

for some infinitesimal ε. Dividing both sides by Δx gives

$$\frac{\Delta(g \circ f)}{\Delta x} = g'(f(x))\frac{\Delta f}{\Delta x} + \varepsilon\frac{\Delta f}{\Delta x} \simeq g'(f(x))f'(x)$$

ergo $(g \circ f)'(x) = g'(f(x))f'(x)$. ☐

Theorem 5.28. *If f has a local maximum or minimum at real point x, and is differentiable at x, then $f'(x) = 0$*

Proof. If f has a maximum at x, then by transfer $f(x + \Delta x) \leq f(x)$ for all infinitesimal Δx. Then if ε is a positive infinitesimal and δ is a negative infinitesimal

$$f'(x) \simeq \frac{f(x+\varepsilon) - f(x)}{\varepsilon} \leq 0 \leq \frac{f(x+\delta) - f(x)}{\delta} \simeq f'(x).$$

Since $f'(x)$ is real, this necessitates that $f'(x) = 0$. The case when x is a minimum is analogous. ☐

With the these results together with **theorem 5.17** and **5.18** from **section 5.3** we could prove that for a function f that is continuous on $[a, b]$ and differentiable on $]a, b[$:

- **Rolle's Theorem**: If $f(a) = f(b) = 0$, then there exists a $c \in]a, b[$ such that $f'(c) = 0$,
- **Mean Value Theorem** For some $c \in]a, b[$

$$f'(c) = \frac{f(a) - f(b)}{a - b},$$

- If f' is positive/zero/negative on some sub-interval $]a', b'[$, then f is growing/constant/negative on $[a', b']$.

However these proofs would not include any non-standard reasoning, so we do not go through them here.

5.5 The Riemann Integral

The definite integral $\int_a^b f(x)$ represents the area restricted by the graph of f, the x-axis, $x = a$ and $x = b$, with the slight caveat that area under the x-axis is negative. In standard analysis one partitions $[a, b]$ into a finite number of sub-intervals, approximating the area one each sub-interval by rectangles, and considering the sum of these rectangles areas, so called *Riemann Sums*. The definite integral is then defined as the limit of these sums as the sub-intervals increase and the partitions get finer and finer. with so sums of rectangles of some fixed width.
A well known fun fact is that the notation $\int_a^b f(x)\,dx$ is relic from Leibniz infinitesimal methods [1]. The "$\int$" symbol is an elongated "S" standing for "sum", and what is being summed are the areas of infinitely many rectangles of heights $f(x)$ and whose widths are an infinitesimal width dx. In this section, non-standard analysis allow us to make this intuition precise.

We begin by introducing the standard definition of Riemann sums and the integral

Definition 5.29 (Riemann Sum)**.** Given a function f who's bounded on $[a, b]$ and a partition $P = \{x_0, x_1, \ldots, x_n\} \subset [a, b]$ where $a = x_0 < x_1 < \ldots < x_n = b$. Let $m_i = \inf_{[x_{i-1}, x_i]} f$ and $M_i = \sup_{[x_{i-1}, x_i]} f$ for $i = 1, 2, \ldots, n$, also let $\Delta x_i = x_i - x_{i-1}$. We define the

- Lower Riemann sum, denoted by $L_a^b(f, P)$, as $\sum_{i=1}^n m_i \Delta x_i$,
- Upper Riemann sum, denoted by $U_a^b(f, P)$, as $\sum_{i=1}^n M_i \Delta x_i$

Definition 5.30. We say that a function f bounded on $[a, b]$ is *Riemann integrable* on $[a, b]$ if the greatest lower-bound of f and lowest upper-bound of f on $[a, b]$, are equal. That is, if

$$\sup_P L_a^b(f, P) = \inf_P U_a^b(f, P).$$

In that case we denote their shared value by the integral $\int_a^b f(x)dx$.

In standard analysis one proves that $\int_a^b f(x)dx$ is the integral of f on $[a, b]$ in the following way [1]

Proposition 5.31. A function f who's bounded on $[a, b]$ is Riemann integrable with the integral $\int_a^b f(x)dx$ if and only if

1. $L_a^b(f,P) \leq \int_a^b f(x)dx \leq U_a^b(f,P)$ for ever partition P on $[a,b]$.
2. For every positive real ε, there exists a P such that $U_a^b(f,P) - L_a^b(f,P) < \varepsilon$. □

We want to able to partitions where every sub-interval is a certain width so we define

Definition 5.32. Given an interval $[a,b]$ and a positive real number Δx let n be the least integer such that $a + n\Delta x \geq b$. We then define a partition $P_{\Delta x} = \{x_0, x_1, \ldots, x_n\}$ with $x_k = a + k\Delta x$ for $k = 0, 1, 2, \ldots, n-1$ and $x_n = b$. This gives a partitions $[a,b]$ into $n-1$ sub-intervals $[x_0,x_1], [x_1,x_2], \ldots, [x_{n-2}, x_n - 1]$ of width Δx and a one $[x_{n-1}, x_n]$ of width less then Δx.

Let $L_a^b(f,\Delta x)$ and $U_a^b(f,\Delta x)$ denote $L_a^b(f,P_{\Delta x})$ and $U_a^b(f,P_{\Delta x})$. Fixing f and $[a,b]$, we can view $L_a^b(f,\Delta x)$ and $U_a^b(f,\Delta x)$ as functions from positive reals as domains (if Δx greater then $b-a$ we just get $P_{\Delta x} = \{a,b\}$). It follows $L_a^b(f,\Delta x)$ and $U_a^b(f,\Delta x)$ as functions have hyperreal extension, defined for all positive hyperreals, including positive infinitesimal. We can then consider infinitesimal partitions and prove the following results.

Theorem 5.33. *If f is a continuous function on $[a,b]$, then for any positive infinitesimal Δx*

$$L_a^b(f,\Delta x) \simeq U_a^b(f,\Delta x).$$

Proof. Assume f is continuous on $[a,b]$, it's the bounded by **theorem 5.17**. We begin by considering $U_a^b(f,\Delta x) - L_a^b(f,\Delta x)$ for a real positive Δx. Unfolding the definitions we get

$$U_a^b(f,\Delta x) - L_a^b(f,\Delta x) = \sum_{i=1}^{n} M_i\Delta x_i - \sum_{i=1}^{n} m_i\Delta x_i = \sum_{i=1}^{n}(M_i - m_i)\Delta x_i$$

Since n is finite in this case we may take $M_{\Delta x}$ as the maximum of $M_i - m_i$ for $i = 1, 2, \ldots, n$, i.e. the biggest vertical difference of f on a sub-interval partitioned by $P_{\Delta x}$. This gives an upper-bound

$$\sum_{i=1}^{n}(M_i - m_i)\Delta x_i \leq \sum_{i=1}^{n} M\Delta x_i = M\sum_{i=1}^{n}\Delta x_i = M(x_1 - x_0 + x_2 - x_1 + \ldots + x_n - x_{n-1}) = M(b-a).$$

Since $M_{\Delta x} = M_i - m_i$ and M_i, m_i are supremum and infimum on a closed interval $[x_{i-1}, x_i]$, it follows by the **extreme value theorem** that $c_{\Delta x}, d_{\Delta x} \in [x_{i-1}, x_i]$ such that $f(c_{\Delta x}) = M_i$ and $f(d_{\Delta x}) = m_i$. Since all sub-intervals are at most Δx wide, it follows that $|c_{\Delta x} - d_{\Delta x}| \leq \Delta x$. Putting these results together into a $\mathcal{L}_{\mathfrak{R}}$-sentence we get

$$(\forall \Delta x \in \mathbb{R}^+)(\exists c_{\Delta x}, d_{\Delta x} \in \mathbb{R})\left(|c_{\Delta x} - d_{\Delta x}| \leq \Delta x \wedge U_a^b(f,\Delta x) - L_a^b(f,\Delta x) \leq (f(c_{\Delta x}) - f(d_{\Delta x}))(b-a)\right).$$

Applying transfer and picking a positive infinitesimal Δx gives $c_{\Delta x} \simeq d_{\Delta x}$ by the first conjunct. In addition, since $[a,b]$ is a closed interval, the continuity of f on $[a,b]$ implies uniform continuity on $[a,b]$ by **theorem 6.5**. As a consequence $f(c_{\Delta x}) \simeq f(d_{\Delta x})$, and since $b-a$ is limited, the product $(f(c_{\Delta x}) - f(d_{\Delta x}))(b-a)$ is infinitesimal. This necessitates $U_a^b(f,\Delta x) - L_a^b(f,\Delta x)$ being infinitesimal which gives us $U_a^b(f,\Delta x) \simeq L_a^b(f,\Delta x)$. □

Theorem 5.34. *If f is a non-decreasing or non-increasing function on $[a,b]$, so called monotonic, then for any positive infinitesimal Δx*

$$L_a^b(f,\Delta x) \simeq U_a^b(f,\Delta x).$$

Proof. Assume f is non-decreasing on $[a,b]$ i.e. $x \leq y$ implies $f(x) \leq f(y)$ for all $x, y \in [a,b]$. Again $U_a^b(f,\Delta x) - L_a^b(f,\Delta x) = \sum_{i=1}^{n}(M_i - m_i)\Delta x_i$ and, by f being non-decreasing M_i, m_i always exists as $f(x_{i-1}) = m_i$ and $f(x_i) = M_i$. We also have $\Delta x_i = \Delta x$ for $i = 1, 2, \ldots, n-1$ and $\Delta x_n \leq \Delta x$. Hence $0 \leq \Delta x_i \leq \Delta x$ for $i = 1, 2, \ldots, n$ so we get the upper-bound

$$\begin{aligned}
U_a^b(f,\Delta x) - L_a^b(f,\Delta x) &= \sum_{i=1}^{n}(f(x_i) - f(x_{i-1}))\Delta x_i \\
&\leq \sum_{i=1}^{n}(f(x_i) - f(x_{i-1}))\Delta x \\
&= \Delta x(f(x_1) - f(x_0) - f(x_1) + f(x_2) - \ldots - f(x_{n-1}) + f(x_n)) \\
&= \Delta x(f(b) - f(a)).
\end{aligned}$$

Applying transfer we get $U_a^b(f,\Delta x) - L_a^b(f,\Delta x) \leq \Delta x(f(b) - f(a))$ for all $\Delta x \in {}^*\mathbb{R}^+$. Since $f(b) - f(a)$ is limited we get $U_a^b(f,\Delta x) \simeq L_a^b(f,\Delta x)$ for all positive infinitesimal Δx.

The proof for when f is non-increasing on $[a,b]$ is analogous. □

In the previous section we proved that $U_a^b(f,\Delta x) \simeq L_a^b(f,\Delta x)$ for all positive infinitesimal Δx, when f is continuous or monotonic on $[a,b]$. These result only hold if the lower and upper sums are done on the same infinitesimal partition. This can be further generalized to hold for any infinitesimal partitions.

Proposition 5.35. If $U_a^b(f,\Delta x) \simeq L_a^b(f,\Delta x)$ for all positive infinitesimal Δx, then $L_a^b(f,\Delta y) \simeq L_a^b(f,\Delta z)$ and $U_a^b(f,\Delta y) \simeq U_a^b(f,\Delta z)$ for all positive infinitesimal $\Delta y, \Delta z$, i.e. all upper and lower Riemann integrals of an infinitesimal partition width are infinitely close.

Proof. By refinement of partitions we know that the upper Riemann sum is greater then the lower Riemann sum, regardless of partitions [1]. Thus by transfer, $L_a^b(f,\Delta x) \leq U_a^b(f,\Delta y)$ for all $\Delta x, \Delta y \in {}^*\mathbb{R}^+$. Assume $U_a^b(f,\Delta x) \simeq L_a^b(f,\Delta x)$ for all positive infinitesimal Δx and consider the lower and upper Riemann sums of the positive infinitesimals Δy and Δz. It follows that one of these four cases must be true

$$
\begin{aligned}
&\underline{\textbf{case 1:}} && L_a^b(f,\Delta y) \leq L_a^b(f,\Delta z) \leq U_a^b(f,\Delta y) \leq U_a^b(f,\Delta z),\\
&\underline{\textbf{case 2:}} && L_a^b(f,\Delta y) \leq L_a^b(f,\Delta z) \leq U_a^b(f,\Delta z) \leq U_a^b(f,\Delta y),\\
&\underline{\textbf{case 3:}} && L_a^b(f,\Delta z) \leq L_a^b(f,\Delta y) \leq U_a^b(f,\Delta y) \leq U_a^b(f,\Delta z),\\
&\underline{\textbf{case 4:}} && L_a^b(f,\Delta z) \leq L_a^b(f,\Delta y) \leq U_a^b(f,\Delta z) \leq U_a^b(f,\Delta y).
\end{aligned}
$$

In **case 1**, by our assumption $L_a^b(f,\Delta y) \simeq U_a^b(f,\Delta y)$ and since $L_a^b(f,\Delta z)$ is between them, it must be infinitely close to both. Our assumption also implies $L_a^b(f,\Delta z) \simeq U_a^b(f,\Delta z)$. Thus, by $\simeq$ being an equivalence, all the sums must be infinitely close. Analogous arguments work for the other cases. □

This results then means that if $U_a^b(f,\Delta x) \simeq L_a^b(f,\Delta x)$ for every positive infinitesimal Δx, then every infinitesimal Riemann sums has the same shadow. In this case, the shadow of infinitesimal Riemann sums is the integral of f on $[a,b]$.

Theorem 5.36. *If $U_a^b(f,\Delta x) \simeq L_a^b(f,\Delta x)$ for every positive infinitesimal Δx, then f is integrable on $[a,b]$ and in particular*

$$\int_a^b f(x)dx = \operatorname{sh}(U_a^b(f,\Delta x)) = \operatorname{sh}(L_a^b(f,\Delta x))$$

for any positive infinitesimal Δx.

Proof. We prove this by proving that $\operatorname{sh}(U_a^b(f,\Delta x))$ fulfills both condition in **proposition 5.31**.

Firstly, since all lower sums are less then all upper sums, we have, for any partition P that

$$L_a^b(f,P) \leq U_a^b(f,\Delta x) \simeq \operatorname{sh}(U_a^b(f,\Delta x)) \simeq L_a^b(f,\Delta x) \leq U_a^b(f,P).$$

Since $L_a^b(f,P), U_a^b(f,P)$ and $\operatorname{sh}(U_a^b(f,\Delta x))$ are all real, we may conclude that

$$L_a^b(f,P) \leq \operatorname{sh}(U_a^b(f,\Delta x)) \leq U_a^b(f,P).$$

For the second condition, since $U_a^b(f,\Delta x) \simeq L_a^b(f,\Delta x)$ for all positive infinitesimal Δx, the $\mathcal{L}_{^*\mathfrak{R}}$-sentence

$$(\exists \Delta x \in {}^*\mathbb{R}^+)(U_a^b(f,\Delta x) - L_a^b(f,\Delta x) < \varepsilon)$$

is true in ${}^*\mathfrak{R}$ for every $\varepsilon \in \mathbb{R}^+$. So by transfer there exists a partition $P_{\Delta x}$ such that $U_a^b(f,P_{\Delta x}) - L_a^b(f,P_{\Delta x}) < \varepsilon$ for every positive real ε.

□

Together with **theorem 5.33** and **5.34** we've then proven that if a function is continuous or monotonic on an interval, then it's integrable on that interval.

To prove some classic properties of integrals

Proposition 5.37. If f, g are functions, integrable on $[a,b]$, then

(i) $\int_a^b cf(x)dx = c\int_a^b f(x)dx$,

(ii) $\int_a^b (f(x)+g(x))dx = \int_a^b f(x)dx + \int_a^b g(x)dx$,

(iii) if $a \leq c \leq b$, then $\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx$,

(iv) If $f(x) \leq g(x)$ for all $x \in [a,b]$, then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$,

(v) $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$,

(vi) if $m \leq f(x) \leq M$ for all $x \in [a,b]$, then $m(b-a) \leq \int_a^b f(x)dx \leq M(b-a)$.

Proof. Assume f and g are integrable on $[a,b]$.

For (i) we note that the supremum of cf on $[x_{i-1}, x_i]$ is the supremum of f times c, i.e. cm_i. Consider $U_a^b(cf, \Delta x)$ for a positive real Δx, by definition

$$U_a^b(cf, \Delta x) = \sum_{i=1}^n cm_i \Delta x_i = c \sum_{i=1}^n m_i \Delta x_i = c \cdot U_a^b(f, \Delta x).$$

This holds for any positive real Δx, so by transfer it holds for all positive hyperreal. Thus by **theorem 8.3**

$$\int_a^b cf(x) = \text{sh}(U_a^b(c \cdot f, \Delta x)) = \text{sh}(cU_a^b(f, \Delta x)) = c\,\text{sh}(U_a^b(f, \Delta x)) = c \int_a^b f(x)dx.$$

The proof of (ii) is similar to the proof of 1, except that the supremum of $f+g$ on $[x_{i-1}, x_i]$ is the supremum of f plus the supremum of g on $[x_{i-1}, x_i]$.

For (iii) let $\Delta x = \frac{c-a}{k}$ for some natural number k, and let $P_{\Delta x}$ be the partition on $[a,b]$, $Q_{\Delta x}$ the partition on $[a,c]$ and let $R_{\Delta x}$ be the partition on $[c,b]$. Then $P_{\Delta x}$ consist of $a < a + \frac{c-a}{k} < \cdots < a + k\frac{c-a}{k} = c < \cdots a + (n-1)\frac{c-a}{k} < b$ while $Q_{\Delta x}$ consists of $a < a + \frac{c-a}{k} < \cdots < a + k\frac{c-a}{k} = c$ while $R_{\Delta x}$ consists of $c = a + k\frac{c-a}{k} <= a + (k+1)\frac{c-a}{k} < \cdots < a + (n-1)\frac{c-a}{k} < b$. Thus $P_{\Delta x} = Q_{\Delta x} \cup R_{\Delta x}$, so it follows that

$$U_a^b(f, \Delta x) = \sum_{i=1}^n M_i \Delta x_i = \sum_{i=1}^k M_i \Delta x_i + \sum_{i=k+1}^n M_i \Delta x_i = U_a^c(f, \Delta x) + U_c^b(f, \Delta x)$$

for all $k \in \mathbb{N}$. By transfer it holds for all hypernatural k. If k is unlimited, then consequently $\Delta x = \frac{c-a}{k}$ be infinitesimal and by **theorem 5.36**

$$\int_a^b f(x)dx = \text{sh}(U_a^b(f, \Delta x)) = \text{sh}(U_a^c(f, \Delta x) + U_c^b(f, \Delta x)) = \text{sh}(U_a^c(f, \Delta x)) + \text{sh}(U_c^b(f, \Delta x)) = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

For (iv), assume $f(x) \leq g(x)$ for all $x \in [a,b]$. Then on any interval $[x_{i-1}, x_i]$ the supremum of f, denoted $M_i(f)$ must be less then the supremum of g, denoted $M_i(g)$. Consequently

$$U_a^b(f, \Delta x) = \sum_{i=1}^n M_i(f) \Delta x_i \leq \sum_{i=1}^n M_i(g) \Delta x_i = U_a^b(g, \Delta x)$$

holds for all positive realΔx. by transfer it holds for positive infinitesimal Δx. By **theorem 5.36**

$$\int_a^b f(x)dx = \text{sh}\left(U_a^b(f, \Delta x)\right) \leq \text{sh}\left(U_a^b(g, \Delta x)\right) = \int_a^b g(x)dx.$$

The proof of (v) follows form transfer on the triangle inequality for sums.

For (vi) assume $m \leq f(x) \leq M$ for all $x \in [a,b]$. Then $m \leq M_i \leq M$ for $i = 1, \ldots, n$. Consequently

$$m(b-a) = \sum_{i=1}^n m \Delta x_i \leq \sum_{i=1}^n M_i \Delta x_i = U_a^b(f, \Delta x) \leq \sum_{i=1}^n M \Delta x = (b-a)M$$

holds for every positive Δx. By transfer

$$m(b-a) \leq (U_a^b(f, \Delta x) \leq M(b-a).$$

holds for every positive infinitesimal Δx. Since $m(b-a)$ and $M(b-a)$ are real, it follows that

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a).$$

□

We end this paper by deriving on of the most famous analysis results, namely the Fundamental Theorem of Calculus. The theorem follows immediately from the following.

Theorem 5.38. *Given a function f that's continuous on $[a,b]$, let $F(x) = \int_a^x f(t)dt$. The function F is then differentiable on every $x \in [a,b]$ and $F'(x) = f(x)$.*

There's a very intuitive reason why this holds if one considers the increment of F at x by Δx. Since f is continuous it only change by an infinitesimal amount if when we increment by Δx. The increase in area is then infinitely close to the area a rectangle of width Δx and a height $f(x)$. Thus $\frac{\Delta F}{\Delta x} \simeq \frac{f(x)\Delta x}{\Delta x} \simeq f(x)$. We formally prove this as follows.

Proof. Let $F(x) = \int_a^x f(t)dt$ and consider a real number Δx such that $0 \leq \Delta x \leq b - x$. Then by **proposition 5.37** part (iii)

$$F(x + \Delta x) - F(x) = \int_a^{x+\Delta x} f(t)dt - \int_a^x f(t)dt = \int_x^{x+\Delta x} f(t)dt.$$

Since f is continuous on $[a,b]$ it is continuous on the sub-interval $[x, x + \Delta x]$ and by **the extreme value theorem**, there exists $x_1, x_2 \in [x, x + \Delta x]$ such that $f(x_1)$ is a minimum, and $f(x_2)$. By **proposition 5.37** (vi)

$$f(x_1)(x + \Delta x - x) \leq \int_x^{x+\Delta x} f(t)dt \leq f(x_2)(x + \Delta x - x)$$

so

$$f(x_1)\Delta x \leq F(x + \Delta x) - F(x) \leq f(x_2)\Delta x$$

and since Δx is positive

$$f(x_1) \leq \frac{F(x + \Delta x) - F(x)}{\Delta x} \leq f(x_2).$$

Because this holds for any positive real Δx so by transfer it holds for any positive hyperreal Δx. If Δx is a positive infinitesimal, then since $x \leq x_1, x_2 \leq x + \Delta x$ and f is continuous we have $f(x_1) \simeq f(x_2) \simeq f(x)$. Consequently

$$f(x) \simeq f(x_1) \leq \frac{F(x + \Delta x) - F(x)}{\Delta x} \leq f(x_2) \simeq f(x)$$

for all positive infinitesimal Δx. For negative infinitesimals,

$$F(x + \Delta x) - F(x) = -\left(\int_a^x f(t)dt - \int_a^{x+\Delta x} f(t)dt\right) = -\int_{x+\Delta x}^x f(t)dt = \int_x^{x+\Delta x} f(t)dt.$$

Then by analogous argument

$$\frac{F(x + \Delta x) - F(x)}{\Delta x} \simeq f(x)$$

for all negative infinitesimal Δx and thus for all non-zero infinitesimal Δx, so by **theorem 5.23** $F'(x) = f(x)$. □

Theorem 5.39 (Fundamental Theorem of Calculus)**.** *If G has a continuous derivative g on $[a,b]$, then $\int_a^b g(x)dx = G(b) - G(a)$*

Proof. Let $F(x) = \int_a^x g(t)dt$. By **theorem 5.38** F and G have the same derivative g on $[a,b]$, so they can only differ by some some constant c. It then follows that $G(b) - G(a) = F(b) + c - (F(a) + c) = F(b) - F(a) = \int_a^b f(x)dx - \int_a^a f(x)dx = \int_a^b f(x)dx$. □

Bibliography

[1] Robert Goldblatt. *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis*. 1st. Vol. 188. Graduate Texts in Mathematics. Springer, 1998. ISBN: 978-1-4612-6841-3.

[2] A.B Slomson J.L. Bell. *Models and Ultraproducts: An Introduction*. 2nd Revised. publishing, 1971. ISBN: 0-7204-2054-7.

[3] Abraham Robinson. "The Metaphysics of the Calculus". In: *The Philosophy of Mathematics*. Ed. by Jaakko Hintikka. Oxford Readings in Philosophy. Oxford University Press, 1969, pp. 153–163.

In the proof of **theorem 2.8** it should refer to **proposition 2.4** instead of **proposition 2**.

In the proof of **proposition 2.10**, it should refer to **theorem 2.8** instead of **theorem 2.1**.

In the proof of**theorem 2.11**, it should refer to **proposition 2.7** instead of **proposition 4**

In the first paragraph of **section 3.2**, it should refer to **proposition 2.15** instead of **proposition 2.8**.

In **proposition 4.1**, $\alpha \cdot \beta$ should not be under the indeterminate cases, it should be $\alpha \cdot A$ instead.

The second paragraph of the proof of **theorem 5.3** needs to be rewritten. An improved version reads:
For the 'if'-part of (i), assume s_N is not positive unlimited, for all $N \in {}^*\mathbb{N}_\infty$. Fixing a positive unlimited M, it follows that $(\forall n \in {}^*\mathbb{N})(s_n \leq M)$. Binding M with an existential quantifier yields a $\mathcal{L}_{^*\mathfrak{R}}$-sentences $(\exists M \in {}^*\mathbb{R})(\forall n \in {}^*\mathbb{N})(s_n \leq M)$. Applying transfer means s must be bounded.

In the proof of **theorem 5.34** there should be a clarification that, since f is non-increasing/non-decreasing on $[a, b]$, it is bounded on the interval by $f(a)$ and $f(b)$.

The proofs of part (i) and (ii) of **proposition 5.37** are faulty. For (i) $\sup cf = c \sup f$ is only the case when $c \geq 0$. If $c < 0$, then $\sup cf = c \inf f$ so $U(cf, \Delta x) = cL(f, \Delta x)$ and thus

$$\int_a^b cf(x)dx = \mathrm{sh}(U(cf, \Delta x)) = c\,\mathrm{sh}(L(f, \Delta x)) = c \int_a^b f(x)dx.$$

For (ii) it is wrong that $\sup(f + g) = \sup f + \sup g$, instead $\sup(f + g) \leq \sup f + \sup g$ and $\inf(f + g) \geq \inf f + \inf g$. Consequently

$$L(f, \Delta x) + L(g, \Delta x) \leq L(f + g, \Delta x) \leq U(f + g, \Delta x) \leq U(f, \Delta x) + U(g, \Delta x)$$

and since $L(f, \Delta x) + L(g, \Delta x) \simeq U(f, \Delta x) + U(g, \Delta x)$, it follows that they all have the same shadow. Thus

$$\int_a^b (f(x)+g(x))dx = \mathrm{sh}(U(f+g, \Delta x)) = \mathrm{sh}(U(f, \Delta x)+U(g, \Delta x)) = \int_a^b f(x)dx+\int_a^b g(x)dx.$$