UiO : Department of Mathematics
University of Oslo

Hyperreal Calculus

MAT2000 — Project in Mathematics

Arne Tobias Malkenes Ødegaard
Supervisor: Nikolai Bjørnestøl Hansen

Abstract

This project deals with doing calculus not by using epsilons and deltas, but by using a number system called the hyperreal numbers. The hyperreal numbers is an extension of the normal real numbers with both infinitely small and infinitely large numbers added. We will first show how this system can be created, and then show some basic properties of the hyperreal numbers. Then we will show how one can treat the topics of convergence, continuity, limits and differentiation in this system and we will show that the two approaches give rise to the same definitions and results.

Contents

1 Construction of the hyperreal numbers

1.1 Intuitive construction

We want to construct the hyperreal numbers as sequences of real numbers $\langle r_n \rangle = \langle r_1, r_2, \ldots \rangle$, and the idea is to let sequences where $\lim_{n\to\infty} r_n = 0$ represent infinitely small numbers, or infinitesimals, and let sequences where $\lim_{n\to\infty} r_n = \infty$ represent infinitely large numbers.

However, if we simply let each hyperreal number be defined as a sequence of real numbers, and let addition and multiplication be defined as elementwise addition and multiplication of sequences, we have the problem that this structure is not a field, since

$$\langle 1, 0, 1, 0, \ldots \rangle \odot \langle 0, 1, 0, 1, \ldots \rangle = \langle 0, 0, 0, 0, \ldots \rangle.$$

The way we solve this is by introducing an equivalence relation on the set of real-valued sequences. We want to identify two sequences if the set of indices for which the sequences agree is a *large* subset of $\mathbb{N}$, for a certain technical meaning of large. Let us first discuss some properties we should expect this concept of largeness to have.

- $\mathbb{N}$ itself must be large, since a sequence must be equivalent with itself.
- If a set contains a large set, it should be large itself.
- The empty set $\emptyset$ should not be large.
- We want our relation to be transitive, so if the sequences r and s agree on a large set, and s and t agree on a large set, we want r and t to agree on a large set.

1.2 Ultrafilters

Our model of a large set is a mathematical structure called an ultrafilter.

Definition 1.1 (Ultrafilters)**.** We define an *ultrafilter on* $\mathbb{N}$, $\mathcal{F}$, to be a set of subsets of $\mathbb{N}$ such that:

- If $X \in \mathcal{F}$ and $X \subseteq Y \subseteq \mathbb{N}$, then $Y \in \mathcal{F}$. That is, $\mathcal{F}$ is closed under supersets.
- If $X \in \mathcal{F}$ and $Y \in \mathcal{F}$, then $X \cap Y \in \mathcal{F}$. $\mathcal{F}$ is closed under intersections.
- $\mathbb{N} \in \mathcal{F}$, but $\emptyset \notin \mathcal{F}$.
- For any subset A of $\mathbb{N}$, $\mathcal{F}$ contains exactly one of A and $\mathbb{N} \setminus A$.

We say that an ultrafilter is *free* if it contains no finite subsets of $\mathbb{N}$. Note that a free ultrafilter will contain all cofinite subsets of $\mathbb{N}$ (sets with finite complement) due to the last property of an ultrafilter.

Theorem 1.2. *There exists a free ultrafilter on* $\mathbb{N}$*.*

Proof. See [Kei76, p. 49]. ∎

1.3 Formal construction

Let $\mathcal{F}$ be a fixed free ultrafilter on $\mathbb{N}$. We define a relation $\equiv$ on the set of real-valued sequences $\mathbb{R}^{\mathbb{N}}$ by letting

$$\langle r_n \rangle \equiv \langle s_n \rangle \iff \{n \in \mathbb{N} \mid r_n = s_n\} \in \mathcal{F}.$$

Proposition 1.3 (Equivalence)**.** *The relation $\equiv$ is an equivalence relation on $\mathbb{R}^{\mathbb{N}}$.*

Proof. We check all needed properties of an equivalence relation.

Reflexivity Since the set $\{n \in \mathbb{N} \mid r_n = r_n\} = \mathbb{N}$, and $\mathbb{N} \in \mathcal{F}$, $\equiv$ is reflexive.

Symmetry The sets $\{n \in \mathbb{N} \mid r_n = s_n\}$ and $\{n \in \mathbb{N} \mid s_n = r_n\}$ are the same, so if one belongs to $\mathcal{F}$, so does the other.

Transitivity Assume that $\langle r_n \rangle \equiv \langle s_n \rangle$ and $\langle s_n \rangle \equiv \langle t_n \rangle$. Then both $\{n \in \mathbb{N} \mid r_n = s_n\} \in \mathcal{F}$ and $\{n \in \mathbb{N} \mid s_n = t_n\} \in \mathcal{F}$. Since $\{n \in \mathbb{N} \mid r_n = s_n\} \cap \{n \in \mathbb{N} \mid s_n = t_n\} \subseteq \{n \in \mathbb{N} \mid r_n = t_n\}$, and $\mathcal{F}$ is closed under intersections and supersets, $\{n \in \mathbb{N} \mid r_n = t_n\} \in \mathcal{F}$, and so $\langle r_n \rangle \equiv \langle t_n \rangle$, as desired.

■

Since $\equiv$ is an equivalence relation, we can define the set of hyperreal numbers $^*\mathbb{R}$ as the set of real-valued sequences modulo the equivalence relation $\equiv$. In symbols,

$$^*\mathbb{R} = \{[r] \mid r \in \mathbb{R}^{\mathbb{N}}\} = \mathbb{R}^{\mathbb{N}} / \equiv .$$

We define addition and multiplication of elements in $^*\mathbb{R}$ by doing elementwise addition and multiplication in the related sequences, more formally as

$$[r] + [s] = [\langle r_n \rangle] + [\langle s_n \rangle] = [\langle r_n + s_n \rangle]$$
$$[r] \cdot [s] = [\langle r_n \rangle] \cdot [\langle s_n \rangle] = [\langle r_n \cdot s_n \rangle] .$$

We define the ordering relation $<$ by letting

$$[r] < [s] \iff \{n \in \mathbb{N} \mid r_n < s_n\} \in \mathcal{F}.$$

At this point, let us introduce some notation to make our arguments easier to read. For two sequences $\langle r_n \rangle$ and $\langle s_n \rangle$, we denote the agreement set $\{n \in \mathbb{N} \mid r_n = s_n\}$ by $[\![r = s]\!]$. We can apply the same notation to other relations, so for example we have $[\![r < s]\!] = \{n \in \mathbb{N} \mid r_n < s_n\}$.

Proposition 1.4. *The operations $+$ and $\cdot$ are well-defined, and so is the relation $<$.*

Proof. We first show that $+$ is well-defined. If we have that $\langle r_n \rangle \equiv \langle r'_n \rangle$ and $\langle s_n \rangle \equiv \langle s'_n \rangle$, then $[\![r = r']\!] \in \mathcal{F}$ and $[\![s = s']\!] \in \mathcal{F}$, which means that $[\![r = r']\!] \cap [\![s = s']\!] \in \mathcal{F}$. What we now need to show is that $[\![r + s = r' + s']\!] \in \mathcal{F}$. If, for some $k \in \mathbb{N}$, both $r_k = r'_k$ and $s_k = s'_k$, then $r_k + s_k = r'_k + s'_k$, hence if $k \in [\![r = r']\!] \cap [\![s = s']\!]$, then $k \in [\![r + s = r' + s']\!]$, which shows that $[\![r = r']\!] \cap [\![s = s']\!] \subseteq [\![r + s = r' + s']\!]$. Since $[\![r = r']\!] \cap [\![s = s']\!] \in \mathcal{F}$, so is $[\![r + s = r' + s']\!]$. So if $r \equiv r'$ and $s \equiv s'$, $r + s \equiv r' + s'$, which shows that the operation is well-defined. Showing that $\cdot$ is well-defined is similar.

We will now show that $<$ is well-defined, which means that we need to show that if $\langle r_n \rangle \equiv \langle r'_n \rangle$ and $\langle s_n \rangle \equiv \langle s'_n \rangle$, then if $[\![r < s]\!] \in \mathcal{F}$, then $[\![r' < s']\!] \in \mathcal{F}$. Firstly, assume that $[\![r = r']\!] \in \mathcal{F}$ and that $[\![s = s']\!] \in \mathcal{F}$. Then, we need to prove that if $[\![r < s]\!] \in \mathcal{F}$ then $[\![r' < s']\!] \in \mathcal{F}$. So let us assume that $[\![r < s]\!] \in \mathcal{F}$, and then prove that $[\![r' < s']\!] \in \mathcal{F}$.

By our assumptions, we have that $[\![r = r']\!] \cap [\![s = s']\!] \cap [\![r < s]\!] \in \mathcal{F}$. If $k \in [\![r = r']\!] \cap [\![s = s']\!] \cap [\![r < s]\!]$, then $r_k = r'_k$, $s_k = s'_k$ and $r_k < s_k$, and therefore $r'_k < s'_k$, so $k \in [\![r' < s']\!]$. So, $[\![r = r']\!] \cap [\![s = s']\!] \cap [\![r < s]\!] \subseteq [\![r' < s']\!]$, and since $\mathcal{F}$ is closed under supersets, we conclude that $[\![r' < s']\!] \in \mathcal{F}$, which shows that $<$ is well-defined. ■

1.4 Infinitely small and large numbers

One of the main reasons for constructing the hyperreals is that we want to have access to infinitely large and infinitely small numbers, and now we can prove their existence.

Theorem 1.5. *There exists a number $\varepsilon \in {}^*\mathbb{R}$ such that $0 < \varepsilon < r$ for any positive real number r, and there exists a number $\omega \in {}^*\mathbb{R}$ such that $\omega > r$ for any real number r.*

Proof. First, we need to talk about real numbers in ${}^*\mathbb{R}$. The way to do this is that given a real number $r \in \mathbb{R}$, we can identify this with a hyperreal number ${}^*r \in {}^*\mathbb{R}$ as ${}^*r = \langle r, r, \ldots \rangle$. We will generally omit the $*$-decoration, and simply refer to this number as r.

Now, let us turn to the actual proof. Let $\varepsilon = \left[\langle 1, \frac{1}{2}, \ldots \rangle\right] = \left[\langle \frac{1}{n} \rangle\right]$. For any positive real number r, the set $\{n \in \mathbb{N} \mid \frac{1}{n} > r\}$ must be finite, and therefore $\{n \in \mathbb{N} \mid \frac{1}{n} < r\}$ is cofinite, and hence belongs to our free ultrafilter $\mathcal{F}$. Therefore, we can conclude that $\varepsilon < r$. Also, since $\{n \in \mathbb{N} \mid 0 < \frac{1}{n}\} = \mathbb{N} \in \mathcal{F}$, it must be the case that $0 < \varepsilon$. So the number ε is a hyperreal number which is greater than 0, but smaller than any positive real number.

Let $\omega = [\langle 1, 2, \ldots \rangle] = [\langle n \rangle]$. For any real number r, the set $\{n \in \mathbb{N} \mid r \geq n\}$ is finite, and hence $\{n \in \mathbb{N} \mid r < n\}$ is cofinite, and belongs to $\mathcal{F}$, which means that $\omega > r$. This proves that ω is a hyperreal number greater than any real number. ■

1.5 Enlarging sets

For a given subset A of $\mathbb{R}$ we can define an "enlarged" subset ${}^*A$ of ${}^*\mathbb{R}$ by saying that a hyperreal number r is an element in ${}^*A$ if and only if the set of n such that r_n is an element in A is large. Formally this can be defined as

$$[r] \in {}^*A \iff \{n \in \mathbb{N} \mid r_n \in A\} \in \mathcal{F}.$$

Again, we need to check that this is well-defined. Using the $[\![\ldots]\!]$ notation, let $[\![r \in A]\!] = \{n \in \mathbb{N} \mid r_n \in A\}$. We have that

$$[\![r = r']\!] \cap [\![r \in A]\!] \subseteq [\![r' \in A]\!],$$

so if $r \equiv r'$ and $[\![r \in A]\!] \in \mathcal{F}$, then $[\![r' \in A]\!] \in \mathcal{F}$, which shows that enlargements are well-defined.

An example of this is if $A = \mathbb{N}$ and $\omega = \langle 1, 2, 3, \ldots \rangle$. Then $[\![\omega \in \mathbb{N}]\!] = \mathbb{N} \in \mathcal{F}$, so $\omega \in {}^*\mathbb{N}$. We will refer to the set ${}^*\mathbb{N}$ as the *hypernaturals*. Similarly, if $A = (0, 1)$ and $r = \langle 0.9, 0.99, 0.999, \ldots \rangle$. Then $[\![r \in \mathbb{N}]\!] = \mathbb{N} \in \mathcal{F}$, so $r \in {}^*(0, 1)$.

1.6 Extending functions

An important tool in non-standard analysis is to take a function $f\colon \mathbb{R} \to \mathbb{R}$ and extend it to a function ${}^*f\colon {}^*\mathbb{R} \to {}^*\mathbb{R}$. This is done by applying the function to each element in the sequence representing the given hyperreal number. We define the extension as follows:

$$ {}^*f([\langle r_1, r_2, \ldots \rangle]) = [\langle f(r_1), f(r_2), \ldots \rangle] \,. $$

Again, we need to prove that this is well-defined. First, let $f \circ r$ denote $\langle f(r_1), f(r_2), \ldots \rangle$. In general, $[\![r = r']\!] \subseteq [\![f \circ r = f \circ r']\!]$, and so if $r \equiv r'$, then ${}^*f(r) = f \circ r \equiv f \circ r' = {}^*f(r')$. Hence the function is well-defined.

A function $f\colon A \to \mathbb{R}$ defined on some subset A of $\mathbb{R}$ can also be extended to a function ${}^*f\colon {}^*A \to {}^*\mathbb{R}$, but not in exactly the same way as above. Since r can be in ${}^*A$ without all elements of r being in A, there can be indices i for which $f(r_i)$ is not defined. In order to get around this, we let $f(r_i) = 0$ whenever $r_i \notin A$. More formally, let

$$ s_n = \begin{cases} f(r_n) & \text{if } r_n \in A \\ 0 & \text{otherwise} \end{cases} $$

and define

$$ {}^*f([\langle r_n \rangle]) = [\langle s_n \rangle] \,. $$

Since we have that ${}^*f(r) = f(r)$ whenever $r \in A$, ${}^*f$ extends f. Therefore we will often simply drop the $*$-decoration, and simply refer to the extended function as f as well.

An important subject related to this construction is sequences. A sequence $\langle s_1, s_2, \ldots \rangle$ is simply a function $s\colon \mathbb{N} \to \mathbb{R}$, and so by this construction can be extended to a *hypersequence* $s\colon {}^*\mathbb{N} \to {}^*\mathbb{R}$, which means that the term s_n is defined even when $n \in {}^*\mathbb{N} \setminus \mathbb{N}$.

2 The transfer principle

2.1 Stating the transfer principle

One of the most important tools of non-standard analysis is the transfer principle, a way to show that a certain type of statement is true when talking about the real numbers if and only if a certain related statement is true when talking about the hyperreal numbers. [1]

First, we introduce the set of sentences which the transfer principle applies to. This set is basically the set of all sentences (formulas with no free variables) in a language of first-order logic which consists of a constant for each real number, a function symbol for each real function, and a relation symbol for each

[1] This is a rather cursory introduction to the tranfer principle. For a more in-depth explanation, see [Gol98, pp. 35-47].

relation on the reals. However, instead of using the quantifiers $(\forall x)$ and $(\exists y)$, our sentences use quantifiers of the form $(\forall x \in A)$ and $(\exists y \in B)$ where A and B are subsets of $\mathbb{R}$. Some examples of such sentences are $(\forall n \in \mathbb{N})(\exists m \in \mathbb{N})(m > n)$, $(\exists x \in \mathbb{R})(\forall y \in \mathbb{R})(x + y = y)$ which state respectively that there is no biggest natural number and there is an additive identity for the reals. Let us call such a sentence an $\mathcal{L}$-sentence.

Now, we define the $*$-transform of an $\mathcal{L}$-sentence. We take a sentence φ, and create a related sentence ${}^*\varphi$. An $\mathcal{L}$-sentence φ contains symbols P, f, and r for relations, functions, and constants on $\mathbb{R}$. To create ${}^*\varphi$, we replace P by ${}^*P$ for all relations P, replace f by ${}^*f$ for all functions f, and replace r by ${}^*r$ for all constants r. Some examples of this are:

- The $*$-transform of the sentence $(\forall n \in \mathbb{N})(\exists m \in \mathbb{N})(m > n)$ is $(\forall n \in {}^*\mathbb{N})(\exists m \in {}^*\mathbb{N})(m \mathbin{{}^*{>}} n)$.
- The $*$-transform of $(\forall x \in \mathbb{R})(\sin(x) < 2)$ is $(\forall x \in {}^*\mathbb{R})({}^*\sin(x) \mathbin{{}^*{<}} {}^*2)$.

We will generally follow the conventions that we omit the $*$ for constants, most functions, and simple equalities and inequalities. With these conventions, the above sentences become $(\forall n \in {}^*\mathbb{N})(\exists m \in {}^*\mathbb{N})(m > n)$ and $(\forall x \in {}^*\mathbb{R})(\sin(x) < 2)$.

Now we state the transfer principle, which we will take as true without proof.

Theorem 2.1 (Transfer principle)**.** *An $\mathcal{L}$-sentence φ is true if and only if its $*$-transform ${}^*\varphi$ is true.*

Some remarks are in order. It is worth pointing out that one can go in both directions, that is one can go from $\mathbb{R}$ to ${}^*\mathbb{R}$, and from ${}^*\mathbb{R}$ to $\mathbb{R}$. If one decides to go in this last direction, it is important that the statement is the $*$-transform of an $\mathcal{L}$-sentence, so for example it can contain no hyperreal constants. A way to get around this is by replacing the constant with a variable x, and adding the quantifier $(\exists x \in {}^*A)$ for some $A \subseteq \mathbb{R}$ in front, which is a technique we will use. In many cases, we will not explicitly write down the full sentence, but rather state things like "since $s < n$ for all natural n, by transfer it also also true for any hypernatural n".

2.2 Using the transfer principle

Theorem 2.2. *The structure $\langle {}^*\mathbb{R}, +, \cdot, < \rangle$ is an ordered field with zero and unity.*

Proof. The way we prove this is by using the transfer principle. We take the fact that $\mathbb{R}$ is an ordered field as true. This can be stated by a number of logical sentences. The fact that addition is commutative in $\mathbb{R}$ can be expressed as the sentence $(\forall x, y \in \mathbb{R})(x + y = y + x)$, and so by the transfer principle, we can conclude that $(\forall x, y \in {}^*\mathbb{R})(x + y = y + x)$, and so addition is commutative in ${}^*\mathbb{R}$. We leave out the full details, but this procedure can then be done for all the axioms for ordered fields (since they are all first-order axioms), and so we conclude that $\langle {}^*\mathbb{R}, +, \cdot, < \rangle$ is also an ordered field. ■

Remark. One important property of the standard real numbers is that they are complete, that is any subset of $\mathbb{R}$ which is non-empty and bounded above has a least upper bound. The reason for why this cannot be proven to hold for ${}^*\mathbb{R}$ is that this can only be expressed using second-order logic, since you need to talk about subsets of $\mathbb{R}$, not just elements of $\mathbb{R}$. In fact, ${}^*\mathbb{R}$ is not complete. An example of this is that the open interval of real numbers $(0,1)$ does not have a least upper bound in ${}^*\mathbb{R}$.

Proposition 2.3. *For any two subsets A and B of $\mathbb{R}$, we have that*

- ${}^*(A \cup B) = {}^*A \cup {}^*B$
- ${}^*(A \cap B) = {}^*A \cap {}^*B$
- ${}^*(A \setminus B) = {}^*A \setminus {}^*B$.

Proof. We prove the statement about unions, but the other two statements can be proven similarly. The statement $(\forall x \in \mathbb{R})(x \in (A \cup B) \leftrightarrow x \in A \vee x \in B)$ is true for any two subsets A and B of $\mathbb{R}$, basically by the definition of unions. Using the transfer principle, the statement $(\forall x \in {}^*\mathbb{R})(x \in {}^*(A \cup B) \leftrightarrow x \in {}^*A \vee x \in {}^*B)$ is also true. We also have that for any two subsets X and Y of ${}^*\mathbb{R}$, $(\forall x \in {}^*\mathbb{R})(x \in (X \cup Y) \leftrightarrow x \in X \vee x \in Y)$. Combining these last two statements, letting $X = {}^*A$ and $Y = {}^*B$, we get that $(\forall x \in {}^*\mathbb{R})(x \in {}^*(A \cup B) \leftrightarrow x \in ({}^*A \cup {}^*B))$, which shows that ${}^*(A \cup B) = {}^*A \cup {}^*B$. ■

Remark. It is worth noting that ${}^*\left(\bigcup_{n\in\mathbb{N}} A_n\right)$ does not need to be equal to $\left(\bigcup_{n\in\mathbb{N}} {}^*A_n\right)$. If $A_n = \{n\}$, then ${}^*\left(\bigcup_{n\in\mathbb{N}} A_n\right) = {}^*\mathbb{N}$, but $\left(\bigcup_{n\in\mathbb{N}} {}^*A_n\right) = \mathbb{N}$.

3 Properties of the hyperreals

3.1 Terminology and notation

At this point we introduce some terminology and notation for talking about hyperreal numbers. We say that a hyperreal number b is:

- *limited* if $r < b < s$ for some $r, s \in \mathbb{R}$,
- *positive unlimited* if $r < b$ for all $r \in \mathbb{R}$,
- *negative unlimited* if $b < r$ for all $r \in \mathbb{R}$,
- *unlimited* if it is positive or negative unlimited,
- *positive infinitesimal* if $0 < b < r$ for all positive $r \in \mathbb{R}$,
- *negative infinitesimal* if $r < b < 0$ for all negative $r \in \mathbb{R}$,
- *infinitesimal* if it is positive infinitesimal, negative infinitesimal or 0,
- *appreciable* if it is limited but not infinitesimal.

We will use the terms limited and unlimited, rather than finite and infinite, when referring to individual numbers. Finite and infinite are terms we use for sets only.

For any subset X of ${}^*\mathbb{R}$, we define $X_\infty = \{x \in X \mid x \text{ is unlimited}\}$, $X^+ = \{x \in X \mid x > 0\}$, and $X^- = \{x \in X \mid x < 0\}$. These notations can also be combined, and so X^+_∞ denotes all positive unlimited members of X.

3.2 Arithmetic of hyperreals

When reasoning about hyperreals, it is useful to have certain rules for computing them, for example that the sum of two infinitesimals is itself infinitesimal. Here are some such rules for computing with hyperreal numbers. If ε and δ are infinitesimals, b and c are appreciable, and H and K are unlimited, then:

- $\varepsilon + \delta$ is infinitesimal,
- $b + \varepsilon$ is appreciable,
- $H + \varepsilon$ and $H + b$ are unlimited,
- $b + c$ is limited,
- $-\varepsilon$ is infinitesimal,
- $-b$ is appreciable,
- $-H$ is unlimited,
- $\varepsilon \cdot \delta$ and $\varepsilon \cdot b$ are infinitesimal,
- $b \cdot c$ is appreciable,
- $b \cdot H$ and $H \cdot K$ are unlimited,
- $\frac{1}{\varepsilon}$ is unlimited if $\varepsilon \neq 0$,
- $\frac{1}{b}$ is appreciable,
- $\frac{1}{H}$ is infinitesimal,
- $\frac{\varepsilon}{b}$, $\frac{\varepsilon}{H}$ and $\frac{b}{H}$ are infinitesimal,
- $\frac{b}{c}$ is appreciable,
- $\frac{b}{\varepsilon}$, $\frac{H}{\varepsilon}$ and $\frac{H}{b}$ are unlimited if $\varepsilon \neq 0$.

We do not give a proof for any of these rules, but they can be proven by using the transfer principle, or by reasoning about sequences of reals.

The following expressions do not have such a rule, and can all take on infinitesimal, appreciable, and unlimited values: $\frac{\varepsilon}{\delta}$, $\frac{H}{K}$, $\varepsilon \cdot H$, $H + K$.

3.3 Halos

A hyperreal b is said to be *infinitely close* to a hyperreal c if $b - c$ is infinitesimal, and this is denoted by $b \simeq c$. This defines an equivalence relation on ${}^*\mathbb{R}$, and we define the *halo* of b to be the $\simeq$-equivalence class

$$\operatorname{hal}(b) = \{c \in {}^*\mathbb{R} \mid b \simeq c\}.$$

Said differently, the halo of b is the set of all hyperreals which are infinitely close to b.

Proposition 3.1. *If two real numbers b and c are infinitely close, that is if $b \simeq c$, then $b = c$.*

Proof. Suppose that $b \simeq c$ with b and c real, but that $b \neq c$. Then there is a non-zero real number r such that $b - c = r$. But this contradicts the assumption that $b \simeq c$, since r is not an infinitesimal. ■

Proposition 3.2. *Suppose that b and c are limited, and that $b \simeq b'$ and $c \simeq c'$. Then $b \pm c \simeq b' \pm c'$ and $b \cdot c \simeq b' \cdot c'$. Furthermore, if $c \not\simeq 0$, then $b/c \simeq b'/c'$.*

Proof. From our assumptions, we have that $b - b' = \varepsilon_b$ and $c - c' = \varepsilon_c$, with ε_b and ε_c being infinitesimal. It is also the case that both b' and c' are limited.

We want to show that $b \pm c \simeq b' \pm c'$, and this is done by showing that $(b \pm c) - (b' \pm c')$ is infinitesimal. We have that

$$(b \pm c) - (b' \pm c') = (b - b') \pm (c - c') = \varepsilon_b \pm \varepsilon_c.$$

Since both the sum of and the difference between two infinitesimals is itself infinitesimal by Section 3.2, we have that $(b \pm c) - (b' \pm c')$ is infinitesimal, and hence that $b \pm c \simeq b' \pm c'$.

The case $b \cdot c \simeq b' \cdot c'$ is proven similarly. We have that

$$\begin{aligned} b \cdot c - b' \cdot c' &= b \cdot c - b \cdot c' + b \cdot c' - b' \cdot c' \\ &= b \cdot (c - c') + (b - b') \cdot c' \\ &= b \cdot \varepsilon_c + \varepsilon_b \cdot c' \end{aligned}$$

which is infinitesimal since the product of a limited number with an infinitesimal is infinitesimal and the sum of two infinitesimals is infinitesimal. Hence $b \cdot c \simeq b' \cdot c'$.

For the last case we have that

$$\begin{aligned} \frac{b}{c} - \frac{b'}{c'} &= \frac{b \cdot c' - b' \cdot c}{c \cdot c'} \\ &= \frac{b \cdot c' - b \cdot c + b \cdot c - b' \cdot c}{c \cdot c'} \\ &= \frac{b \cdot (c - c') + c \cdot (b - b')}{c \cdot c'} \\ &= \frac{c \cdot \varepsilon_b - b \cdot \varepsilon_c}{c \cdot c'}. \end{aligned}$$

Now, if $c \not\simeq 0$, the denominator is the product of two appreciable numbers, which is also appreciable. Since the numerator is infinitesimal by a similar argument to the case of products, the quotient is itself infintesimal, and hence $b/c \simeq b'/c'$. ■

Remark. The first part of the prosition, namely that $b \pm c \simeq b' \pm c'$, holds also for unlimited b and c, but the other parts do not. To show this, let H be some positive unlimited number, and let b', c and c' equal H, and let b equal $H + \frac{1}{H}$. Then $b \simeq b'$ and $c \simeq c'$, but

$$b \cdot c - b' \cdot c' = \left(H + \frac{1}{H}\right) \cdot H - H \cdot H = H^2 + 1 - H^2 = 1,$$

which is not infinitesimal, and so $b \cdot c \not\simeq b' \cdot c'$. A similar counterexample can also be produced for b/c.

3.4 Shadows

Theorem 3.3 (Existence of shadows)**.** *Every limited hyperreal b is infinitely close to one and only one real number s. This real number is called the shadow of b, which is denoted by* $\mathrm{sh}(b)$.

Proof. Let $A = \{r \in \mathbb{R} \mid r < b\}$. Since A is a non-empty set which is bounded above, it has a least upper bound of A in $\mathbb{R}$ by the (Dedekind) completeness of $\mathbb{R}$. Call this real number s.

We want to show that $b \simeq s$, and we do this by showing that $|b - s| < \varepsilon$ for all $\varepsilon \in \mathbb{R}^+$. Take any such ε. We show that $|b - s| < \varepsilon$ by showing that $s - \varepsilon < b < s + \varepsilon$. Take the case when $b < s + \varepsilon$. Assume that $s + \varepsilon \leq b$. Then $s < s + \frac{\varepsilon}{2} < s + \varepsilon \leq b$. Since both s and ε are real, so is $s + \frac{\varepsilon}{2}$, and since $s + \frac{\varepsilon}{2} < b$, $s + \frac{\varepsilon}{2} \in A$. But since $s + \frac{\varepsilon}{2} > s$, s is not an upper bound of A. But this is a condradiction, so it must be the case that $b < s + \varepsilon$. Now take the case when $s - \varepsilon < b$. Assume that $b \leq s - \varepsilon$. Then $b \leq s - \varepsilon < s - \frac{\varepsilon}{2} < s$. Since $s - \frac{\varepsilon}{2} \geq b$, $s - \frac{\varepsilon}{2}$ is an upper bound of A, but $s - \frac{\varepsilon}{2} < s$, so s is not the least upper bound of A, which is a contradiction.

We also need to check that there cannot be more than one shadow of b. Assume that there are two reals s and s' which are both infinitely close to b. Thus, by definition, $b \simeq s$ and $b \simeq s'$, and so by transitivity of $\simeq$, $s \simeq s'$. But since both s and s' are real, by Proposition 3.1 we conclude that $s = s'$.[2] ■

Alternative proof. Watch Babylon 5. ■

4 Convergence

4.1 Convergence in hyperreal calculus

The standard way to define convergence in real analysis is that a sequence $\langle s_n \rangle$ converges to the limit $L \in \mathbb{R}$ if for any $\varepsilon \in \mathbb{R}^+$, there exists an $m_\varepsilon \in \mathbb{N}$ such that $|s_n - L| < \varepsilon$ for any $n > m_\varepsilon$. This can be expressed in formal logic by the sentence $(\forall \varepsilon \in \mathbb{R}^+)(\exists m_\varepsilon \in \mathbb{N})(\forall n \in \mathbb{N})(n > m_\varepsilon \to |s_n - L| < \varepsilon)$. The idea that this definition formalizes is that a sequence convergences to a real value L if you get very close to L when you get very far out in the sequence. What we do for non-standard analysis is that we say that a sequence converges to L if it gets infinitely close to L as one gets infinitely far out in the sequence.

The original sequence $\langle s_n \rangle$ is only defined on the naturals, so one can not go infinitely far out, but by using how we defined hypersequences in Section 1.6, we get a new sequence which is defined for all $n \in {}^*\mathbb{N}$, where we can go infinitely far out, and we denote this sequence by $\langle s_n \rangle$ as well.

Theorem 4.1. *A sequence of real numbers $\langle s_n \rangle$ converges to L if and only if $s_n \simeq L$ for all unlimited n.*

Proof. Assume that the sequence $\langle s_n \rangle$ converges to L. We need to show that $s_n \simeq L$ for any unlimited n, and we do this by proving that $|s_n - L| < \varepsilon$ for any positive real ε. So take an $\varepsilon \in \mathbb{R}^+$. By the definition of convergence, there exists a natural number m_ε such that $|s_n - L| < \varepsilon$ whenever $n > m_\varepsilon$. Let k be such a natural number. Then the formal statement

$$(\forall n \in \mathbb{N})(n > k \to |s_n - L| < \varepsilon)$$

must hold. By the transfer principle, it must also be the case that

$$(\forall n \in {}^*\mathbb{N})(n > k \to |s_n - L| < \varepsilon) \tag{1}$$

[2]This proof, along with several other proofs we give in this article, is a modified version of a proof given in [Gol98].

is true. Now, let N be any unlimited number. Since k is limited, we have that $N > k$, and so by (1) can conclude that $|s_N - L| < \varepsilon$. Since this holds for any positive ε, it must be the case that $s_N \simeq L$ is true, which completes the forward direction of the proof.

For the converse, assume that $s_n \simeq L$ for all unlimited n. We want to show that the sequence converges. Take any $\varepsilon \in \mathbb{R}^+$, and fix an unlimited $N \in {}^*\mathbb{N}$. Now, if $n > N$, n must be unlimited, and so $s_n \simeq L$ by our assumption, from which we conclude that $|s_n - L| < \varepsilon$. Formally, this is expressible as $(\forall n \in {}^*\mathbb{N})(n > N \to |s_n - L| < \varepsilon)$. Thus, the sentence

$$(\exists m_\varepsilon \in {}^*\mathbb{N})(\forall n \in {}^*\mathbb{N})(n > m_\varepsilon \to |s_n - L| < \varepsilon)$$

must also be true. By transfer, we can conclude that

$$(\exists m_\varepsilon \in \mathbb{N})(\forall n \in \mathbb{N})(n > m_\varepsilon \to |s_n - L| < \varepsilon)$$

must hold. Since ε was taken to be any positive real, we have that the sentence

$$(\forall \varepsilon \in \mathbb{R}^+)(\exists m_\varepsilon \in \mathbb{N})(\forall n \in \mathbb{N})(n > m_\varepsilon \to |s_n - L| < \varepsilon)$$

must hold. This is indeed the formal statement for stating that the sequence s_n converges, which finishes our proof. ■

4.2 Monotone convergence

A standard theorem about convergence from calculus is the theorem of monotone convergence, which can be stated as

Theorem 4.2. *Let $\langle s_1, s_2, \ldots \rangle$ be a sequence of real numbers which is bounded above and non-decreasing. Then $\langle s_n \rangle$ is convergent.*

The standard proof works by taking the supremum of the set $\{s_n \mid n \in \mathbb{N}\}$, and showing that the sequence converges to this number. The non-standard proof also uses the supremum of that set, but in a very different way.

Proof. Let s_N be an extended term of the sequence, and let b be an upper bound of the sequence.

Since the sequence is non-decreasing, $s_1 \leq s_n$ for any n, and $s_n \leq b$ must also hold for any n since the sequence was bounded above by b. Thus the statement

$$(\forall n \in \mathbb{N})(s_1 \leq n \wedge n \leq b)$$

must be true, and so must its $*$-transform

$$(\forall n \in {}^*\mathbb{N})(s_1 \leq n \wedge n \leq b).$$

Applying this to our extended term s_N, it is clear that s_N is limited and so has a shadow $L = \mathrm{sh}(s_N)$. What we now want to prove is that L is the least upper bound for the set $\{s_n \mid n \in \mathbb{N}\}$. Since a set can only have one least upper bound, this L must be the same for all extended terms, and so all extended terms have the same shadow. Then, for any extended term s_N, $s_N \simeq L$, and then by Theorem 4.1, the sequence must be convergent.

If $m \leq n$, $s_m \leq s_n$ since the sequence is non-decreasing. By transfer, this holds for any $m, n \in {}^*\mathbb{N}$ as well. In particular, if $m \in \mathbb{N}$, and N is the index for our chosen extended term s_N, then $s_m \leq s_N \simeq L$, and hence $s_m \leq L$ since both s_m and L are real. Hence, $L \geq s_i$ for any $i \in \mathbb{N}$, and so L is an upper bound of our set.

Now we show that L is the least upper bound. Let r be any upper bound of our set. Then $(\forall n \in \mathbb{N})(s_n \leq r)$, and so by using transfer, we must have that $s_N \leq r$. Then we have that $L \simeq s_N \leq r$, and then that $L \leq r$, since both L and r are real. So for any upper bound of our set, L is not larger, and so L is the least upper bound, completing our proof. ■

5 Continuity

5.1 Continuity in hyperreal calculus

The standard definition of continuity states that a function f is continuous at c if for any positive real ε, there exists a positive real δ such that $|f(x) - f(c)| < \varepsilon$ whenever $|x - c| < \delta$, which can be expressed by the formal statement $(\forall \varepsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})(|x - c| < \delta \to |f(x) - f(c)| < \varepsilon)$. The intuitive notion in this definition is that $f(x)$ gets arbitrarily close to $f(c)$ when x gets arbitrarily close to c. What our non-standard definition formalizes, is that $f(x)$ is infinitely close to $f(c)$ when x is infinitely close to c.

Theorem 5.1. *A function $f \colon \mathbb{R} \to \mathbb{R}$ is continuous at $c \in \mathbb{R}$ if and only if $f(x) \simeq f(c)$ whenever $x \simeq c$.*

Proof. We start by assuming that f is continuous at c, and also that we have a hyperreal x such that $x \simeq c$. From this, we want to show that $f(x) \simeq f(c)$, and we do this by showing that $|f(x) - f(c)| < \varepsilon$ for all $\varepsilon \in \mathbb{R}^+$.

Take any positive real ε. By the definition of continuity, there exists a δ such that for all real x, $|f(x) - f(c)| < \varepsilon$ whenever $|x - c| < \delta$. Fix such a δ. Then, the statement

$$(\forall x \in \mathbb{R})(|x - c| < \delta \to |f(x) - f(c)| < \varepsilon)$$

must hold, and so by transfer its $*$-transform

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

must also hold. For the x we assumed was infinitesimally close to c, the statement $|x - c| < \delta \to |f(x) - f(c)| < \varepsilon$ is true. But since δ is a positive real and $x \simeq c$, it must be true that $|x - c| < \delta$, and so we can conclude that $|f(x) - f(c)| < \varepsilon$. Since this holds for any $\varepsilon \in \mathbb{R}^+$, it must be true that $f(x) \simeq f(c)$, which is what we needed to show.

For the converse, assume that $f(x) \simeq f(c)$ whenever $x \simeq c$. We want to prove that the formal statement of continuity must be true. First, let ε be any positive real, and let d be any positive infinitesimal. Then, it must be the case that $x \simeq c$ whenever $|x - c| < d$. Then, by assumption, we have that $f(x) \simeq f(c)$, and thus that $|f(x) - f(c)| < \varepsilon$ for any $\varepsilon \in \mathbb{R}^+$. From this we can conclude that if $|x - c| < d$, then $|f(x) - f(c)| < \varepsilon$. This can be expressed formally as

$$(\forall x \in {}^*\mathbb{R})(|x - c| < d \to |f(x) - f(c)| < \varepsilon).$$

Since this is true, the statement

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

must also be true. But this is the $*$-transform of the sentence

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

and so by transfer we can conclude that this last sentence is also true. Since ε was chosen arbitrarily, with no conditions other than it being positive and real, we can conclude that the formal statement of continuity,

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

must be true, which concludes our proof. ■

This theorem only deals with functions which are defined on all of $\mathbb{R}$. In many circumstances it is useful to study functions which are defined only on some subset A of $\mathbb{R}$. The proof of Theorem 5.1 can be easily extended to showing the following theorem.

Theorem 5.2. *The function $f\colon A \rightarrow \mathbb{R}$ is continuous at $c \in A$ if and only if $f(x) \simeq f(c)$ for all $x \in {}^*A$ with $x \simeq c$.*

Note that we here do not require that $f(x) \simeq f(c)$ for all $x, c \in {}^*A$. This turns out to be a stronger condition, and is in fact equivalent with the notion of uniform continuity, which we will discuss later in this section.

5.2 Examples

Here we give some examples of using hyperreal calculus to show that some functions are continuous or discontinuous.

Proposition 5.3. *The function $f(x) = x^2$ is continuous at any $a \in \mathbb{R}$.*

Proof. By Theorem 5.1, it suffices to show that $f(x) \simeq f(a)$ whenever $x \simeq a$. If $x \simeq a$, then $x = a+\varepsilon$ for some infinitesimal ε. Now $f(x) = f(a+\varepsilon) = a^2+2a\varepsilon+\varepsilon^2$. Then $f(x) - f(a) = a^2 + 2a\varepsilon + \varepsilon^2 - a^2 = \varepsilon(2a + \varepsilon)$, which is infinitesimal since the product of a limited number with an infinitesimal is infinitesimal. Hence, whenever $x \simeq a$, $f(x) \simeq f(a)$, so f is a continuous function. ■

Proposition 5.4. *The function f defined by*

$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ 0 & \text{if } x \text{ is irrational} \end{cases}$$

is discontinuous at all $a \in \mathbb{R}$.

Proving this with hyperreal calculus is rather straightforward, but requires establishing some propositions first.

Proposition 5.5. *The extended function ${}^*f$ can be defined as*

$$ {}^*f(x) = \begin{cases} 1 & \text{if } x \in {}^*\mathbb{Q} \\ 0 & \text{if } x \notin {}^*\mathbb{Q}. \end{cases} \tag{2}$$

Proof. By transfer of the true sentences

$$(\forall x \in \mathbb{R})(x \in \mathbb{Q} \to f(x) = 1)$$
$$(\forall x \in \mathbb{R})(x \notin \mathbb{Q} \to f(x) = 0)$$

we can conclude that ${}^*f(x) = 1$ if $x \in {}^*\mathbb{Q}$, and that ${}^*f(x) = 0$ if $x \notin {}^*\mathbb{Q}$, which shows that the definition (2) is a correct definition of ${}^*f$. ∎

Proposition 5.6. *Any halo contains both hyperrationals (members of* ${}^*\mathbb{Q}$*) and hyperirrationals (members of* ${}^*\mathbb{R} \setminus {}^*\mathbb{Q}$*)*

Proof. Since any halo contains some hyperreal number r and the hyperreal number $r+\varepsilon$, where ε is some positive infinitesimal, it also contains all hyperreals between these, the set $X = \{x \in {}^*\mathbb{R} \mid r < x < r + \varepsilon\}$.

Now, since the sentence $(\forall x, y \in \mathbb{R})(\exists z \in \mathbb{Q})(x < y \to x < z \wedge z < y)$ is true, using transfer, and applying the statement to r and $r + \varepsilon$, the statement $(\exists z \in {}^*\mathbb{Q})(r < z \wedge z < r + \varepsilon)$ is true, and so $X \cap {}^*\mathbb{Q} \neq \emptyset$, which means that our given halo contains at least one hyperrational number.

For the other case, since the sentence $(\forall x, y \in \mathbb{R})(\exists z \in (\mathbb{R} \setminus \mathbb{Q}))(x < y \to x < z \wedge z < y)$ is true, using transfer and applying the statement to r and $r + \varepsilon$, the statement $(\exists z \in {}^*(\mathbb{R} \setminus \mathbb{Q}))(r < z \wedge z < r + \varepsilon)$ is true, which means that $X \cap {}^*(\mathbb{R} \setminus \mathbb{Q}) \neq \emptyset$, so our halo contains at least one hyperreal which is a member of ${}^*(\mathbb{R} \setminus \mathbb{Q})$. But by Proposition 2.3, ${}^*(\mathbb{R} \setminus \mathbb{Q}) = {}^*\mathbb{R} \setminus {}^*\mathbb{Q}$, so our halo contains at least one member of ${}^*\mathbb{R} \setminus {}^*\mathbb{Q}$, or a hyperirrational. ∎

Proof of Proposition 5.4. From these two propositions, we can show that f is not continuous in any point. Let c be a rational number. Then $f(c) = 1$. By Proposition 5.6, there is a hyperirrational d in $\mathrm{hal}(c) \setminus {}^*\mathbb{Q}$, with $f(d) = 0$. Since $0 \not\simeq 1$, we have that $c \simeq d$, but $f(c) \not\simeq f(d)$, so f is not continuous in c. Now, let c be an irrational number. Then $f(c) = 0$. By Proposition 5.6, there is a hyperrational $d \in \mathrm{hal}(c) \cap {}^*\mathbb{Q}$, and so $f(d) = 1$. Again we have that $c \simeq d$, but $f(c) \not\simeq f(d)$, so f is not continuous in c. So regardless of whether c is rational or irrational, f is not continuous in c, and therefore f is discontinuous in all points of $\mathbb{R}$. ∎

5.3 Theorems about continuity

Theorem 5.7. *If f and g are continuous at c, then $f + g$, $f - g$ and fg are continuous at c. Furthermore, if $g(c) \neq 0$, then f/g is also continuous at c.*

Proof. Assume that f and g are continuous at c. Hence when $x \simeq c$, we have that $f(x) \simeq f(c)$ and $g(x) \simeq g(c)$, and these values are all limited. It then follows from Proposition 3.2 that

- If $x \simeq c$, then $(f + g)(x) = f(x) + g(x) \simeq f(c) + g(c) = (f + g)(c)$, and so $f + g$ is continuous at c.
- If $x \simeq c$, then $(f - g)(x) = f(x) - g(x) \simeq f(c) - g(c) = (f - g)(c)$, and so $f - g$ is continuous at c.
- If $x \simeq c$, then $(fg)(x) = f(x) \cdot g(x) \simeq f(c) \cdot g(c) = (fg)(c)$, and so fg is continuous at c.

- If $x \simeq c$, then $(f/g)(x) = f(x)/g(x) \simeq f(c)/g(c) = (f/g)(c)$. Note that we require that $g(c) \neq 0$, and so $g(x) \not\simeq 0$, and we can apply Proposition 3.2. Hence f/g is continuous at c.

■

Theorem 5.8. *If f is continuous at c, and g is continuous at $f(c)$, $g \circ f$ is continuous at c.*

Proof. Let $x \simeq c$. Since f is continuous at c, we have that $f(x) \simeq f(c)$. Since g is continuous at $f(c)$, for any number v which is infinitely close to $f(c)$, we have that $g(v) \simeq g(f(c))$. Since $f(x)$ is infinitely close to $f(c)$, we have that $(g \circ f)(x) = g(f(x)) \simeq g(f(c)) = (g \circ f)(c)$, which proves that $g \circ f$ is continuous at c. ■

Theorem 5.9 (The Intermediate Value Theorem)**.** *Let $f: [a,b] \to \mathbb{R}$ be a continuous function. Then for every real number d strictly between $f(a)$ and $f(b)$ there exists a real number $c \in (a,b)$ such that $f(c) = d$.*

Proof. Assume that $f(a) < d < f(b)$. The case where $f(a) > d > f(b)$ is similar. For each $n \in \mathbb{N}$, we partition $[a,b]$ into n subintervals of equal length $\frac{b-a}{n}$. These intervals then have the endpoints $p_k = a + k\frac{b-a}{n}$ for $0 \leq k \leq n$. Now, we let s_n be the greatest endpoint for which $f(p_k) < d$. s_n is then the maximum of the set $\{p_k \mid f(p_k) < d\}$, which exists since the set is finite and non-empty (it contains $p_0 = a$ since $f(a) < d$ by assumption).

Since $f(b) > d$, $p_n = b \notin \{p_k \mid f(p_k) < d\}$. Therefore we have that $a \leq s_n < b$ for all $n \in \mathbb{N}$. By construction of s_n it must be true that $f(s_n) < d \leq f(s_n + \frac{b-a}{n})$ for any $n \in \mathbb{N}$. By transfer, we conclude that both of these statements also hold for any $n \in {}^*\mathbb{N}$.

Now, let N be an unlimited hypernatural. We have that $a \leq s_N < b$, hence s_N is limited and has a shadow $c = \mathrm{sh}(s_N) \in \mathbb{R}$. Now, since N is unlimited, $\frac{b-a}{N}$ is infinitesimal, and so we have that $s_N \simeq c$ and $s_N + \frac{b-a}{N} \simeq c$. Now, by the assumption that f is continuous, and our equivalent formulation of continuity, we have that $f(s_N) \simeq f(c)$ and $f\left(s_N + \frac{b-a}{N}\right) \simeq f(c)$. Therefore, it is the case that

$$f(c) \simeq f(s_N) < d \leq f\left(s_N + \frac{b-a}{N}\right) \simeq f(c).$$

Therefore $f(c) \simeq d$, but since both $f(c)$ and d are real, we can conclude that $f(c) = d$, which completes the proof. ■

5.4 Uniform continuity

The notion of uniform continuity is a strengthening of the ordinary notion of continuity, and can be expressed with the formal sentence $(\forall \varepsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x, y \in A)(|x - y| < \delta \to |f(x) - f(y)| < \varepsilon)$. The big difference here is that for a given ε, the same δ should work for all $x, y \in A$, whereas in the ordinary notion of continuity, δ can depend on x.

Theorem 5.10. *The function $f: A \to \mathbb{R}$ is uniformly continuous on A if and only if $f(x) \simeq f(y)$ whenever $x \simeq y$ for all $x, y \in {}^*A$.*

Proof. This can be proven in a similar manner to the theorem for standard continuity, but then using the formal sentence $(\forall \varepsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x, y \in A)(|x - y| < \delta \to |f(x) - f(y)| < \varepsilon)$. ■

Theorem 5.11. *If f is continuous on $[a, b]$, then f is uniformly continuous on $[a, b]$.*

Proof. Assume that f is continuous. Now, take hyperreals $x, y \in {}^*[a, b]$ with $x \simeq y$. Let $c = \text{sh}(x)$. Then since $a \leq x \leq b$, and $x \simeq c$, then $c \in [a, b]$, and so by assumption f is continuous at c. Since both $c \simeq x$ and $c \simeq y$, we have that $f(c) \simeq f(x)$ and $f(c) \simeq f(y)$ by the continuity of f. By the transitivity of $\simeq$, we conclude that $f(x) \simeq f(y)$, and hence that f is uniformly continuous on $[a, b]$. ■

Remark. This proof does not transfer to more general intervals (for example $(0, 1)$ or $[0, \infty])$ since it is a necessary part of the proof that the shadow of x is contained in the original interval, but for these intervals this is not guaranteed. As an example, let $(0, 1)$ be our interval and let $x = \varepsilon$ be a positive infinitesimal, which is in ${}^*(0, 1)$. Then $c = \text{sh}(x) = 0 \notin (0, 1)$.

Proposition 5.12. *$f(x) = \frac{1}{x}$ is not uniformly continuous on $(0, 1)$.*

Proof. Let H be any positive unlimited hyperreal. Then $H + 1$ is also unlimited. Hence both $\frac{1}{H}$ and $\frac{1}{H+1}$ are positive infinitesimals, and hence we have $\frac{1}{H} \simeq \frac{1}{H+1}$ and $\frac{1}{H}, \frac{1}{H+1} \in {}^*(0, 1)$. However $f\left(\frac{1}{H}\right) = H$ and $f\left(\frac{1}{H+1}\right) = H+1$, but $H \not\simeq H+1$. Therefore we have $x, y \in {}^*(0, 1)$ such that $f(x) \not\simeq f(y)$, so f is not uniformly continuous. ■

6 Limits and derivatives

6.1 Limits in hyperreal calculus

In order to talk about derivatives of functions, we want to be able to talk about limits of functions. In standard analysis, L is the limit of f as x goes to c, written $\lim_{x \to c} f(x) = L$ if for any $\varepsilon \in \mathbb{R}^+$, there exists a $\delta \in \mathbb{R}^+$ such that $|f(x) - L| < \varepsilon$ whenever $|x - c| < \delta$. The intuition behind this definition is that f gets very close to L as x gets very close to c. The definition using non-standard analysis formalizes the intuitive idea that f is infinitely close to L when x is infinitely close to c. Given $c, L \in \mathbb{R}$ and a function f defined on $A \subseteq \mathbb{R}$, we have that

$$\lim_{x \to c} f(x) = L \iff f(x) \simeq L \text{ for all } x \in {}^*A \text{ with } x \simeq c \text{ and } x \neq c.$$

Similarly, one can define different types of limits, both one-sided limits and limits as x tends to ∞. We have that

- $\lim_{x \to c^+} f(x) = L$ iff $f(x) \simeq L$ for all $x \in {}^*A$ with $x \simeq c$ and $x > c$.
- $\lim_{x \to c^-} f(x) = L$ iff $f(x) \simeq L$ for all $x \in {}^*A$ with $x \simeq c$ and $x < c$.
- $\lim_{x \to +\infty} f(x) = L$ iff $f(x) \simeq L$ for all $x \in {}^*A^+_\infty$ (and ${}^*A^+_\infty \neq \emptyset$).

- $\lim_{x\to-\infty} f(x) = L$ iff $f(x) \simeq L$ for all $x \in {}^*A^-_\infty$ (and ${}^*A^-_\infty \neq \emptyset$).

These can be proved in a similar manner to the related theorems for continuity or for convergence, but we will not give the proof here.

6.2 Differentiation in hyperreal calculus

In standard analysis, we say that f is differentiable at x if

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

exists, and if it does, we let $f'(x)$ denote the derivative of f in x and $f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$.

Theorem 6.1. *If f is defined at $x \in \mathbb{R}$, then $L \in \mathbb{R}$ is the derivative of f at x if and only if for every nonzero infinitesimal ε, $f(x+\varepsilon)$ is defined, and*

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

Proof. Let $g(h) = \frac{f(x+h)-f(x)}{h}$. Then the statement that $\lim_{h\to 0} g(h) = L$ is equivalent with f having derivative L at x, and so applying the characterisation of limits from Section 6.1, the theorem follows. ■

This means that when f is differentiable, we can find the derivative as $f'(x) = \mathrm{sh}\left(\frac{f(x+\varepsilon)-f(x)}{\varepsilon}\right)$ for any non-zero infinitesimal ε.

6.3 Examples

Proposition 6.2. *The function $f(x) = x^2$ is differentiable at any $x \in \mathbb{R}$, and $f'(x) = 2x$ for all $x \in \mathbb{R}$.*

Proof. Using the definition, we want to show that $\frac{f(x+\varepsilon)-f(x)}{\varepsilon} \simeq 2x$ for any infinitesimal $\varepsilon \neq 0$ and real x. By straightforward calculations, we have that

$$\begin{aligned}\frac{f(x+\varepsilon)-f(x)}{\varepsilon} &= \frac{(x+\varepsilon)^2-x^2}{\varepsilon}\\ &= \frac{x^2+2x\varepsilon+\varepsilon^2-x^2}{\varepsilon}\\ &= \frac{\varepsilon(2x+\varepsilon)}{\varepsilon}\\ &= 2x+\varepsilon \simeq 2x\end{aligned}$$

Since for any ε, $\frac{f(x+\varepsilon)-f(x)}{\varepsilon} \simeq 2x$, by Theorem 6.1, f is differentiable at all $x \in \mathbb{R}$, and $f'(x) = 2x$, as we wanted to show. ■

Proposition 6.3. *The function $f(x) = |x|$ is not differentiable at $x = 0$.*

Proof. Let ε be some positive infinitesimal. Then

$$\frac{f(x+\varepsilon)-f(x)}{\varepsilon} = \frac{|0+\varepsilon|-|0|}{\varepsilon} = \frac{\varepsilon}{\varepsilon} = 1.$$

However, we also have that

$$\frac{f(x+(-\varepsilon))-f(x)}{\varepsilon}=\frac{|0+(-\varepsilon)|-|0|}{-\varepsilon}=\frac{\varepsilon}{-\varepsilon}=-1.$$

Since $-1 \not\simeq 1$, we have that $\frac{f(x+\varepsilon)-f(x)}{\varepsilon} \not\simeq \frac{f(x+\delta)-f(x)}{\delta}$ for two non-zero infinitesimals ε and $\delta=-\varepsilon$, and so they can not both be infinitely close to the same real number L, which means that f is not differentiable at 0. ∎

6.4 Increments

We introduce some notation to simplify our arguments. Let Δx denote som non-zero infinitesimal, representing a small change or an increment in the value of x. Then we let $\Delta f = f(x+\Delta x) - f(x)$ denote the corresponding increment in the value of f at x. To be explicit, we should write this as $\Delta f(x, \Delta x)$, since this value depends on both those variables, but we will mainly use the more convinient shorthand Δf.

The way we will use this shorthand is to compute $\frac{\Delta f}{\Delta x}$, and if this is always infinitely close to the same real number, then we have that $f'(x) = \mathrm{sh}(\frac{\Delta f}{\Delta x})$. But since $\frac{\Delta f}{\Delta x}$ is just an ordinary fraction of hyperreal numbers, we can compute Δf on its own, something which will be useful.

An important thing to note is that if f is differentiable at x, $\frac{\Delta f}{\Delta x} \simeq f'(x)$, and so $\frac{\Delta f}{\Delta x}$ is limited. Since $\Delta f = \frac{\Delta f}{\Delta x}\Delta x$, we then have that Δf is infinitesimal, and thus $f(x+\Delta x) \simeq f(x)$ for all infinitesimal Δx. This proves that

Theorem 6.4. *If a function $f\colon A \to \mathbb{R}$ is differentiable at x, then f is continuous at x.*

The lemma that follows is needed mainly in our proof of the chain rule.

Lemma 6.5 (Incremental Equation)**.** *If $f'(x)$ exists at real x and Δx is infinitesimal, then there exists an infinitesimal ε, dependent on x and Δx, such that*

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

Proof. Since $f'(x)$ exists, we have that $f'(x) \simeq \frac{\Delta f}{\Delta x}$, and hence that $f'(x) - \frac{\Delta f}{\Delta x} = \varepsilon$ for some infintesimal ε. Multiplying through by Δx and rearranging, we get that $\Delta f = f'(x)\Delta x + \varepsilon \Delta x$, which is what we wanted. ∎

6.5 Theorems about derivatives

Theorem 6.6. *If f and g are differentiable at x, so is $f+g$ and fg, and we have that*

- $(f+g)'(x) = f'(x) + g'(x)$
- $(fg)'(x) = f(x)g'(x) + g(x)f'(x)$.

Proof. We take the case of addition. First we compute $\Delta(f+g)$. We have that

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

and hence that

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

under the assumption that both f and g are differentiable. Since the real value $f'(x) + g'(x)$ is independent of Δx, we conclude, by Theorem 6.1, that $(f+g)'(x) = f'(x) + g'(x)$.

For our proof of the statement regarding multiplication, we need a little trick, namely that $f(x + \Delta x) = f(x) + (f(x + \Delta x) - f(x)) = f(x) + \Delta f$. Then we get that

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

which yields that

$$\begin{aligned}\frac{\Delta(fg)}{\Delta x} &= f(x)\frac{\Delta g}{\Delta x} + g(x)\frac{\Delta f}{\Delta x} + \frac{\Delta f}{\Delta x}\Delta g \\ &\simeq f(x)g'(x) + g(x)f'(x) + 0\end{aligned}$$

where we again use that g and f are differentiable. The last term is 0 since $\frac{\Delta f}{\Delta x}$ is limited and Δg is infinitesimal. Since this last real number is independent of Δx, we conclude, by applying Theorem 6.1, that $(fg)'(x) = f(x)g'(x) + g(x)f'(x)$. ■

Theorem 6.7 (Chain Rule)**.** *If f is differentiable at $x \in \mathbb{R}$, and g is differentiable at $f(x)$, then $g \circ f$ is differentiable at x with derivative $g'(f(x))f'(x)$.*

Proof. For any non-zero infinitesimal Δx, $f(x + \Delta x)$ is defined and $f(x + \Delta x) \simeq f(x)$. Since $g'(f(x))$ exists, g is defined at all points infinitely close to $f(x)$, which means that $(g \circ f)(x + \Delta x) = g(f(x + \Delta x))$ is defined.

Now, we want to express $\Delta(g \circ f)$ in other terms. Again, we use that $f(x + \Delta x) = f(x) + \Delta f$. We get that

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

which shows that $\Delta(g \circ f)$ is also the increment of g at $f(x)$ corresponding to Δf. Using the more explicit notation for increments, we have that

$$\Delta(g \circ f)(x, \Delta x) = \Delta g(f(x), \Delta f).$$

By the incremental equation applied to g, there exists an infinitesimal ε such that

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

and hence that

$$\begin{aligned}\frac{\Delta(g \circ f)}{\Delta x} &= g'(f(x))\frac{\Delta f}{\Delta x} + \varepsilon\frac{\Delta f}{\Delta x} \\ &\simeq g'(f(x))f'(x) + 0\end{aligned}$$

which establishes our claim, namely that $g'(f(x))f'(x)$ is the derivative of $g \circ f$ at x. ■

Theorem 6.8 (Critical Point Theorem)**.** *Let f be defined on some open interval (a, b), and have a maximum or minimum at $x \in (a, b)$. If f is differentiable at x, then $f'(x) = 0$.*

Proof. Let f have a maximum at x. By the transfer principle, we conclude that

$$f(x + \Delta x) \leq f(x) \text{ and thus that } f(x + \Delta x) - f(x) \leq 0$$

for all infinitesimal Δx. Hence for a positive infinitesimal ε and a negative infinitesimal δ, we have that

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

Since $f'(x)$ is real, it must be equal to 0.

The case when f has a minimum is similar. ■

References

[Gol98] Robert Goldblatt. *Lectures on the hyperreals.* An introduction to nonstandard analysis. Springer-Verlag, New York, 1998.

[Kei76] H. Jerome Keisler. *Foundations of infinitesimal calculus.* 1976.