THE HYPERREALS

LARRY SUSANKA

ABSTRACT. In this article we define the hyperreal numbers, an ordered field containing the real numbers as well as infinitesimal numbers. These infinitesimals have magnitude smaller than that of any nonzero real number and have intuitively appealing properties, harkening back to the thoughts of the inventors of analysis.

We use the ultrafilter construction of the hyperreal numbers which employs common properties of sets, rather than the original approach (see A. Robinson *Non-Standard Analysis* [5]) which used model theory.

A few of the properties of the hyperreals are explored and proofs of some results from real topology and calculus are created using hyperreal arithmetic in place of the standard limit techniques.

CONTENTS

Date: June 27, 2018.

1. Historical Remarks and Overview

The historical Euclid-derived conception of a line was as an object possessing "the quality of length without breadth" and which satisfies the various axioms of Euclid's geometric structure. Euclidean constructions involved finitely many points identified on the line by application of certain allowable operations. Arithmetical properties followed from their way of doing things, such as the Archimedean order property: "There is an integer multiple of the smaller of two numbers which exceeds the greater."

Our modern point of view identifies a certain algebraic object, a Dedekind complete ordered field called the real numbers, with infinitely many points on a line. These properties are intended to codify some of our intuition about the concept of "line" in a way that builds on the old view. For instance Dedekind completeness implies the Archimedean order property, though the reverse implication does not hold.

Dedekind completeness, which affirms the existence of a supremum for every bounded set, is a key feature: it guarantees that most of the things we would like to do with real numbers do not cause us to *leave* the real numbers, thereby insulating us from defects of earlier number systems.

Most of us take it for granted that the real numbers can be used to represent *all* the points on the line. However the inventors of calculus had a conception of "vanishingly small quantities" and numbers separated from each other by infinitesimal displacements. They performed arithmetic with their evanescent creations and added up infinitely many of them or calculated infinite multiples of them to produce tangible numbers. They could obtain verifiably correct results by proper use of these methods, though "proper use" was never completely codified. They just *knew* when they were doing it correctly, even if they had problems explaining it to their philosopher critics (such as George Berkeley) who disdained their use, with considerable verve, as intellectually hopeless.

The real number system has no room for "infinitesimals," nor is there a direct way to discuss their reciprocals which must exceed any integer. The concept of limit has replaced both, as introduced—much to the relief of those who were distressed by the widely acknowledged lack of rigor of infinitesimal arithmetic—by Cauchy and others.

In 1960 Abraham Robinson (see [5]) realized that the elements were in place to create a system which *could* represent "infinitesimals" directly and whose properties could be used to prove theorems in our standard world.

Analysis using this augmented number system is called **nonstandard analysis.** The augmented number system itself is called the **hyperreal number system**. Many practitioners of nonstandard analysis believe that it will *become* standard, that the real number system was a stopgap required by limitations in mathematical and logical technique of its nineteenth century creators. At this time it is not universally accepted that this would be beneficial.

- The difficulties involved in the transfer of a nonstandard proof to the standard world can be formidable, which can act as a counterweight to the conceptually simpler proofs one sometimes finds using nonstandard methods.

- The addition of entities to the real number system (infinitesimals and infinite numbers) which can never be measured or sensed but only *imagined* strikes some as esthetically displeasing, given the success of the technology of limits which they are to supplant. But the real numbers are not useful without that technology. If by building a richer structure many technically challenging and important theorems involving limits become trivial, is that not evidence that the augmented structure is the more "natural" one?
- It has been shown that any fact of real analysis which can be proven using the hyperreals is provable using real numbers alone. Critics contend this makes nonstandard methods superfluous. Practitioners claim that, sooner or later, a nonstandard proof will be created whose standard version is too hard for humans to find.
- Some feel that the ultrapower construction of the hyperreal numbers, which seems to be the most accessible realization of this system, fails to have the virtue of specificity. It relies on the existence of free ultrafilters on the natural numbers, $\mathbb{N}$. The axiom of choice implies that these exist but no explicit description of the members of any free ultrafilter is possible.

  This last objection should be taken in perspective. The usual construction of the real numbers themselves is not very closely connected to how we actually *think* of them. Once we know that a construction is possible, that the axioms for a Dedekind complete ordered field are consistent, we are no longer concerned with the specifics of how the real numbers are built. We use the axioms directly to prove theorems involving real numbers and rarely feel that this approach is a barrier to intuition or that we should, instead, work with the nuts and bolts of our construction. Theorems abound in real analysis that rely upon, or are even equivalent to, the axiom of choice. Mathematicians no longer put up much of a fuss about it. Similarly, once we have the properties of the hyperreals we will use these properties, *not the details used to show the properties are consistent,* to prove theorems. Why should the admittedly abstract ultrapower definition of the hyperreals be a problem?

With all these "pros and cons" in mind we embark on the construction of this interesting field and let the reader decide if it is, ultimately, worth the candle. A "fair hearing" on the matter would involve many more applications than we offer here. We follow parts of the discussion in the very readable *Lectures on the Hyperreals* by Robert Goldblatt [1], and that reference is a good next step. Somewhat more advanced topics can be found in the collection of articles *Nonstandard Analysis for the Working Mathematician*, edited by Peter Loeb and Manfred Wolff [3].

2. The Construction

Let $\mathbb{S}$ denote the ring of real valued sequences with the usual pointwise operations. If x is a real number we let s^x denote the constant sequence, $s^x_n = x$ for all n. The function sending x to s^x is a one-to-one ring homomorphism, providing an embedding of the real numbers, $\mathbb{R}$, into $\mathbb{S}$. In the following, wherever it is not too confusing we will not distinguish between $x \in \mathbb{R}$ and the constant function s^x, leaving the reader to derive intent from context.

The ring $\mathbf{S}$ has additive identity 0 and multiplicative identity 1. $\mathbf{S}$ is not a field because if r is any sequence having 0 in its range it can have no multiplicative inverse. There are lots of zero divisors in $\mathbf{S}$.

We need several definitions now. Generally, for any set S, $\mathbb{P}(S)$ denotes the set of all subsets of S. It is called the **power set of S**. Also, a subset of $\mathbb{N}$ will be called **cofinite** if it contains all but finitely many members of $\mathbb{N}$. The symbol $\varnothing$ denotes the empty set. A **partition** of a set S is a decomposition of S into a union of sets, any pair of which have no elements in common.

An **ultrafilter** $\mathbb{H}$ on $\mathbb{N}$ is a family of sets for which:

(i) $\varnothing \notin \mathbb{H} \subset \mathbb{P}(\mathbb{N})$.

(ii) Any intersection of finitely many members of $\mathbb{H}$ is in $\mathbb{H}$.

(iii) $A \subset \mathbb{N},\ B \in \mathbb{H} \Rightarrow A \cup B \in \mathbb{H}$.

(iv) If $V_1, \ldots, V_n$ is any finite partition of $\mathbb{N}$ then $\mathbb{H}$ contains exactly one of the V_i.

If, further,

(v) $\mathbb{H}$ contains every cofinite subset of $\mathbb{N}$.

the ultrafilter is called **free**.

If an ultrafilter on $\mathbb{N}$ contains a finite set then it contains a one-point set, and is nothing more than the family of all subsets of $\mathbb{N}$ containing that point. So if an ultrafilter is *not* free it must be of this type, and is called a **principal ultrafilter**. Our construction below depends on the use of a *free—not a principal—ultrafilter.* The existence of a free ultrafilter containing any given infinite subset of $\mathbb{N}$ is implied by the Axiom of Choice.

We now suppose $\mathbb{H}$ to be a free ultrafilter. **We will not vary $\mathbb{H}$ during the course of the discussions to follow.**

We are going to be using conditions on sequences and sets to define subsets of $\mathbb{N}$. We introduce a convenient shorthand for the usual "set builder" notation. If P is a property that can be true or false for natural numbers we use $[\![P]\!]$ to denote $\{n \in \mathbb{N} \mid P(n) \text{ is true}\}$. This notation will only be employed during a discussion to decide if the set of natural numbers defined by P is in $\mathbb{H}$, or not.

For example, if s, t is a pair of sequences in $\mathbf{S}$ we define three sets of integers

$$[\![s < t]\!], \qquad [\![s = t]\!], \qquad [\![s > t]\!].$$

Since these three sets partition $\mathbb{N}$, exactly one of them is in $\mathbb{H}$, and we declare $s \equiv t$ when $[\![s = t]\!] \in \mathbb{H}$.

2.1. ***Exercise.*** *$\equiv$ is an equivalence relation on $\mathbf{S}$. We denote the equivalence class of any sequence s under this relation by $[s]$.*

Define for each $r \in \mathbf{S}$ the sequence $\tilde{r}$ by

$$\tilde{r}_n = \begin{cases} 0, & \text{if } r_n = 0; \\ (r_n)^{-1}, & \text{if } r_n \neq 0. \end{cases}$$

2.2. ***Exercise.*** *(a) There is at most one constant sequence in any class* $[r]$.

(b) $[0]$ *is an ideal in* $\mathbb{S}$ *so* $\mathbb{S}/[0]$ *is a commutative ring with identity* $[1]$.

(c) Consequently $[r] = r + [0] = \{\, r + t \mid t \in [0]\,\}$ *for all* $r \in \mathbb{S}$.

(d) If $[r] \neq [0]$ *then* $[\tilde{r}][r] = [1]$. *So* $[r]^{-1} = [\tilde{r}]$.

From this exercise, we conclude that ${}^*\mathbb{R}$, defined to be $\mathbb{S}/[0]$, is a field containing an embedded image of $\mathbb{R}$ as a subfield. $[0]$ is a maximal ideal in $\mathbb{S}$. This quotient ring is called the field of **hyperreal numbers**.

We declare $[s] < [t]$ provided $[\![s < t]\!] \in \mathbb{H}$.

Any field with a linear order $<$ is called an ordered field provided

(i) $\quad x + y > 0$ whenever $x, y > 0$

(ii) $\quad xy > 0$ whenever $x, y > 0$

(iii) $\quad x + z > y + z$ whenever $x > y$.

Positive, negative, nonnegative and nonpositive members of an ordered field are defined in the obvious way. The three items in the definition imply that a nonzero field member is positive or negative but not both.

2.3. ***Exercise.*** *(a) The relation given above is a linear order on* ${}^*\mathbb{R}$, *and makes* ${}^*\mathbb{R}$ *into an ordered field. As with any ordered field, we define* $|x|$ *for* $x \in {}^*\mathbb{R}$ *to be* x *or* $-x$, *whichever is nonnegative.*

(b) If x, y *are real then* $x \leq y$ *if and only if* $[x] \leq [y]$. *So the ring isomorphism of* $\mathbb{R}$ *into* ${}^*\mathbb{R}$ *is also an order isomorphism onto its image in* ${}^*\mathbb{R}$.

(c) It is also true that if the sequences s, t *satisfy* $s_n \leq t_n$ *for every* n *then* $[s] \leq [t]$. *The converse, of course, need not hold.*

Because of this last exercise and the essential uniqueness of the real numbers it is common to identify the embedded image of $\mathbb{R}$ in ${}^*\mathbb{R}$ with $\mathbb{R}$ itself. One does not (unless *absolutely necessary*) make, for instance, the two-step transition

$$2 \rightarrow \text{ (constant sequence equal to } 2\ \forall n) \rightarrow [2] = \{\, t \in \mathbb{S} \mid [\![t = 2]\!] \in \mathbb{H}\,\}.$$

We do not distinguish 2 from $[2]$.

Though obviously circular, one does something similar when identifying the rational numbers, $\mathbb{Q}$, with its isomorphic image in $\mathbb{R}$, and $\mathbb{N}$ itself with the corresponding subset of $\mathbb{Q}$. This kind of notational simplification usually does not cause problems.

Now we get to the ideas that prompted the construction. Define the sequence r by $r_n = \frac{1}{n+1}$. For every positive integer k, $\left[\!\!\left[r < \frac{1}{k} \right]\!\!\right] \in \mathbb{H}$. So $0 < [r] < \frac{1}{k}$. We have found a positive hyperreal smaller than (the embedded image of) any real number. This is our first nontrivial infinitesimal number. The sequence $\tilde{r}$ is given by $\tilde{r}_n = n + 1$. So $[r]^{-1} = [\tilde{r}] > k$ for every positive integer k. $[r]^{-1}$ is a hyperreal larger than any real number.

Note that if s is the sequence given by $s_n = \frac{1}{n+2}$ for $n \in \mathbb{N}$ then $[s] < [r]$. So the order structure on infinitesimals will reflect rather delicate information about the *rate* at which "most" of its terms converge to 0, not merely that they do so.

3. Vocabulary

We call a member $x \in {}^*\mathbb{R}$ **limited** if there are members $a, b \in \mathbb{R}$ with $a < x < b$. We will use $\mathbb{L}$ to indicate the limited members of ${}^*\mathbb{R}$. x is called **unlimited** if it not limited. These terms are preferred to "finite" and "infinite," which are reserved for concepts related to cardinality.

If $x, y \in {}^*\mathbb{R}$ and $x < y$ we use $[x, y]_*$ to denote $\{ t \in {}^*\mathbb{R} \mid x \leq t \leq y \}$. This set is called a **closed hyperinterval**. Open and half-open hyperintervals are defined and denoted similarly.

A set $S \subset {}^*\mathbb{R}$ is called **hyperbounded** if there are members x, y of ${}^*\mathbb{R}$ for which S is a subset of the hyperinterval $[x, y]_*$.

Abusing standard vocabulary for ordered sets, S is called **bounded** if x and y can be chosen to be *limited* members of ${}^*\mathbb{R}$. x and y could, in fact, be chosen to be real if S is bounded.

The vocabulary of bounded or hyperbounded above and below can be used.

We call a member $x \in {}^*\mathbb{R}$ **infinitesimal** if $|x| < a$ for every positive $a \in \mathbb{R}$. The only real infinitesimal is, obviously, 0. We will use $\mathbb{I}$ to indicate the infinitesimal members of ${}^*\mathbb{R}$.

x is called **appreciable** if it is limited but not infinitesimal. Hyperreals x and y are said to have **appreciable separation** if $|x - y|$ is appreciable.

We will be working with various subsets S of ${}^*\mathbb{R}$ and adopt the following convention: $S_\infty = S - \mathbb{L} = \{ x \in S \mid x \notin \mathbb{L} \}$. These are the unlimited members of S, if any.

We say two hyperreals x, y are **infinitesimally close** or have **infinitesimal separation** if $|x - y| \in \mathbb{I}$. We use the notation $x \approx y$ to indicate that x and y are infinitesimally close. They have **limited separation** if $|x-y| \in \mathbb{L}$. Otherwise they are said to have **unlimited separation.**

We define the **halo** of x by $\text{halo}(x) = x + \mathbb{I}$. There can be at most one real number in any halo. Whenever $\text{halo}(x) \cap \mathbb{R}$ is nonempty we define the **shadow** of x, denoted $\text{shad}(x)$, to be that unique real number.

The **galaxy** of x is defined to be $\text{gal}(x) = x + \mathbb{L}$. $\text{gal}(x)$ is the set of hyperreal numbers a limited distance away from x. So if x is limited $\text{gal}(x) = \mathbb{L}$.

If n is any fixed positive integer we define ${}^*\mathbb{R}^n$ to be the set of equivalence classes of sequences in $\mathbb{R}^n$ under the equivalence relation $x \equiv y$ exactly when $[\![x = y]\!] \in \mathbb{H}$. We use the notation $[x]$ for the equivalence class of the sequence x under this relation. The case $n = 1$ is discussed at some length above.

If $K \subset \mathbb{R}^n$ we define ${}^*K$, called the **enlargement** of K, to be the set of those classes $[x] \in {}^*\mathbb{R}^n$ for which $[\![x \in K]\!] \in \mathbb{H}$. Note that if $[x] \in {}^*K$ it might be that some values of x are not in K. However, there is always a sequence $y \in [x]$ whose range *is* in K. So we are free to conceive of these classes as corresponding to classes of sequences entirely in K, or at least to select representatives of each class of that type.

This simple idea is important enough to merit special attention. We single it out as a bulleted item below, which we refer to as the **Localization Lemma**. We will refer to an application of the lemma as, simply, **localization.**

- For any $[x] \in {}^*K$, there is always a sequence $y \in [x]$ whose range is in K.

Frequently we will want to compose a sequence $x \in [x] \in {}^*K$ with a function f with domain K. So that $f \circ x$ makes sense, we will choose the representative x to have range in K in that case. This can be done by localization.

By examining the classes of constant sequences in K, with which K itself can be identified, we can consider K to be a subset of ${}^*K$. Obviously, ${}^*\varnothing = \varnothing$. For $A, B \subset \mathbb{R}^n$, $A \subset B$ if and only if ${}^*A \subset {}^*B$.

3.1. ***Exercise.*** *Suppose b is a real number, $b \notin S \subset \mathbb{R}$ and* $\text{halo}(b) \cap {}^*S \neq \varnothing$. *Then there are points in S which are arbitrarily close to b. The converse is also true: if there are points in S other than b which are arbitrarily close to b then* $\text{halo}(b) \cap {}^*S \neq \varnothing$.

We call ${}^*\mathbb{N}$ the set of **hypernatural numbers**, ${}^*\mathbb{Z}$ the ring of **hyperintegers** formed using the set of integers $\mathbb{Z}$ and ${}^*\mathbb{Q}$ the field of **hyperrational numbers**.

If $A \subset \mathbb{R}^n$ and $f\colon A \to \mathbb{R}^k$ we define a function ${}^*f$ from ${}^*A$ to ${}^*\mathbb{R}^k$ as follows. If x is any sequence whose values are all in A define ${}^*f([x]) = [f \circ x]$. By examining constant sequences we see that this definition extends f from A to ${}^*A$.

In particular, a sequence $s\colon \mathbb{N} \to \mathbb{R}^k$ can be extended to a **hypersequence** ${}^*s\colon {}^*\mathbb{N} \to {}^*\mathbb{R}^k$. We call ${}^*s_i$ for $i \in {}^*\mathbb{N}_\infty$ an **extended term** of the sequence s.

It is important to recognize that sequences serve two very distinct purposes in our development. First, they are used as the members of the equivalence classes which *are* the hyperreal numbers. This is where members of the ultrafilter $\mathbb{H}$ enter the picture. Second, they have their ordinary usage from calculus and topology. Hypersequences are derived from sequences of the second kind.

Any function, such as the tangent function, the natural logarithm function or the square root function, whose domain was formerly a subset of $\mathbb{R}$ will be automatically extended to the corresponding enlarged or "starred" subset of ${}^*\mathbb{R}$ by this process. Often the "star" is omitted for common functions and relations. For instance $|x|,\ x < y,\ x^y,\ \sin(x)$ or $\sqrt{x}$ are all used whether x, y are in $\mathbb{R}$ or ${}^*\mathbb{R}$.

4. A Collection of Exercises

4.1. ***Exercise.*** *(a) Every function from $A \subset \mathbb{R}^n$ to $B \subset \mathbb{R}^k$ is a set of ordered pairs of the form $(x, f(x))$ where $x \in A$ and $f(x) \in B$. These ordered pairs correspond in the obvious way to a set W of members of $\mathbb{R}^{n+k}$, and we have a definition for ${}^*W$. In this case, the definition matches the one just given above for ${}^*f$.*

(b) Consider a function f as above. $f\colon A \to B$ is one-to-one or onto B if and only if the corresponding function ${}^*f\colon {}^*A \to {}^*B$ has these properties.

(c) Consider a function f as above. If $W \subset A$ then ${}^*(f(W)) = {}^*f({}^*W)$.

4.2. ***Exercise.*** *The usual order $<$ on $\mathbb{R}$ corresponds to a subset J of $\mathbb{R}^2$ where $(x, y) \in J$ if and only if $x < y$. The set ${}^*J$ corresponds to the linear order we discussed for ${}^*\mathbb{R}$ above.*

4.3. ***Exercise.*** *(a) If $x, y \in \mathbb{R}$ and $x < y$ show for open intervals (x, y) that ${}^*(x, y) = (x, y)_*$, and the same relation holds for half-open and closed intervals.*

(b) Suppose $x, y \in \mathbb{L}$ and $\operatorname{shad}(x) < \operatorname{shad}(y)$. (All limited hyperreals have shadows, as we shall soon see.) Show that $\mathbb{R} \cap [x, y]_*$ is one of four intervals in $\mathbb{R}$ bounded below by $\operatorname{shad}(x)$ and above by $\operatorname{shad}(y)$. It is open or closed on either end depending on whether $\operatorname{shad}(x) < x$, $\operatorname{shad}(y) > y$.

(c) If $S \subset \mathbb{R}$, then $S = \mathbb{R} \cap {}^*S$.

4.4. ***Exercise.*** *Addition of real numbers corresponds to a subset A of $\mathbb{R}^3$ where $(x, y, z) \in A$ if and only if $x + y = z$. The set ${}^*A$ corresponds to the addition we have defined for ${}^*\mathbb{R}$. The same situation holds for multiplication in $\mathbb{R}$ and ${}^*\mathbb{R}$.*

4.5. ***Exercise.*** *Show that if you follow the construction of ${}^*\mathbb{R}$ using a **principal** rather than a **free** ultrafilter then ${}^*\mathbb{R}$ is field and order isomorphic to $\mathbb{R}$.*

4.6. ***Exercise.*** *(a) For any $A, B \subset \mathbb{R}^n$*

$$\begin{aligned} &{}^*(A - B) = {}^*A - {}^*B \quad \textit{and} \quad {}^*(A \cap B) = {}^*A \cap {}^*B \\ \textit{and} \quad &{}^*(A \cup B) = {}^*A \cup {}^*B \quad \textit{and} \quad \mathbb{R}^n \cap {}^*A = A. \end{aligned}$$

(b) Suppose $A \subset \mathbb{R}^n$.
${}^*A - A$ is nonempty if and only if A is infinite.
For $n = 1$, ${}^*A_\infty$ is nonempty if and only if A is unbounded.

(c) No set in ${}^*\mathbb{R}$ that is hyperbounded but not bounded can be the enlargement of a set from $\mathbb{R}$. In other words, if ${}^*S$ is hyperbounded then S is bounded.

(d) Some bounded sets, including intervals, in ${}^*\mathbb{R}$ are not the enlargement of any subset of $\mathbb{R}$. (hint: Consider $\mathbb{I}$.)

(e) If $A \subset {}^*\mathbb{R}$ is finite and A is the enlargement of a set from $\mathbb{R}$ then $A \subset \mathbb{R}$.

(f) No infinite bounded subset of $\mathbb{R}$ is the enlargement of a set from $\mathbb{R}$.

4.7. ***Exercise.*** *Suppose $\delta, \varepsilon \in \mathbb{I} - \{0\}$ (infinitesimal but nonzero) and $a, b \in \mathbb{L} - \mathbb{I}$ (appreciable: limited but not infinitesimal) and $x, y \in {}^*\mathbb{R}_\infty$ (unlimited.) Show:*

$\delta + \varepsilon,\ \delta\varepsilon,\ \delta a,\ \dfrac{a}{x},\ \dfrac{\delta}{a},\ \dfrac{\delta}{x},$ *and* $\sqrt{\delta}$ *are all in* $\mathbb{I}$.

$a + \delta,\ a + b,\ ab,\ \dfrac{a}{b},$ *and* $\sqrt{a}$ *are all in* $\mathbb{L}$.

$x + \delta,\ x + a,\ |x| + |y|,\ xy,\ xa,\ \dfrac{a}{\delta},\ \dfrac{x}{\delta}$ *and* $\sqrt{|x|}$ *are all in* ${}^*\mathbb{R}_\infty$.

If, in addition, $a \in \operatorname{halo}(c)$ and $b \in \operatorname{halo}(d)$ then

$$ab \in \operatorname{halo}(cd),\ a^n \in \operatorname{halo}(c^n)\ \forall n \in \mathbb{Z},\ \frac{a}{b} \in \operatorname{halo}\!\left(\frac{c}{d}\right) \ \textit{and}\ a + b \in \operatorname{halo}(c + d).$$

We conclude that the expected arithmetic relationships are true in ${}^*\mathbb{R}$, that $\mathbb{L}$ is a subring of ${}^*\mathbb{R}$ and that $\mathbb{I}$ is an ideal in the ring $\mathbb{L}$.

4.8. ***Exercise.*** *${}^*\mathbb{R}$ is not Dedekind complete. (hint: $\mathbb{N}$ is bounded above by the member $[t] \in {}^*\mathbb{N}$, where t is the sequence given by $t_n = n$ for all $n \in \mathbb{N}$. But $\mathbb{N}$ can have no least upper bound: if $n \le c$ for all $n \in \mathbb{N}$ then $n \le c - 1$ for all $n \in \mathbb{N}$. As another example consider $\mathbb{I}$. This set is (very) bounded, but has no least upper bound.)*

Every limited hyperreal has a shadow: there is a real number infinitesimally close to each member of $\mathbb{L}$. To see this, select $a \in \mathbb{L}$. Let b denote the infimum of all real numbers exceeding a. If $b-a$ is appreciable there is a real number between a and b, contrary to the definition of b. This fact is important enough to warrant a bullet:

- Every limited hyperreal has a shadow.

4.9. ***Exercise.*** *(a) The set of galaxies and the set of halos each partition* ${}^*\mathbb{R}$*, so the relations "share the same galaxy" and "share the same halo" are both equivalence relations on* ${}^*\mathbb{R}$*.*

(b) The shadow map $\operatorname{shad}\colon \mathbb{L} \to \mathbb{R}$ *is an order preserving ring homomorphism onto* $\mathbb{R}$*. The kernel of this map is* $\mathbb{I}$*. Thus* $\mathbb{R}$ *is field isomorphic to* $\mathbb{L}/\mathbb{I}$*. The natural isomorphism preserves an order on* $\mathbb{L}/\mathbb{I}$*, the order induced by the order on* $\mathbb{L}$*. When restricted to* $\mathbb{R}$*, halo is the inverse isomorphism onto* $\mathbb{L}/\mathbb{I}$*.*

(c) Does $\mathbb{L}$ *contain proper ideals other than* $\{0\}$ *and* $\mathbb{I}$*?*

(d) Suppose $\operatorname{halo}(x) \neq \operatorname{halo}(y)$ *and* $a \in \operatorname{halo}(x)$ *and* $b \in \operatorname{halo}(y)$ *and* $a < b$*. Then every member of* $\operatorname{halo}\left(\frac{x+y}{2}\right)$ *exceeds* a *and is less than* b*. So between every two halos is another halo.*

4.10. ***Exercise.*** *Every halo contains a hyperrational. Also* $\mathbb{I} \cap {}^*\mathbb{Q}$ *is an ideal in* $\mathbb{L} \cap {}^*\mathbb{Q}$*. It follows that* $\mathbb{R}$ *is isomorphic to the quotient ring* $(\mathbb{L} \cap {}^*\mathbb{Q})/(\mathbb{I} \cap {}^*\mathbb{Q})$*.*

4.11. ***Exercise.*** *Is it reasonable to refer to the additive group* ${}^*\mathbb{R}/\,(2\pi{}^*\mathbb{Z})$ *as the "hypercircle?"*

4.12. ***Exercise.*** *Suppose* t *is a bounded sequence in* $\mathbb{R}$*. Prove directly that there is a unique real number* x *so that* $[t] \in \operatorname{halo}(x)$*. (hint: Divide an interval* $A = [a, b]$ *containing* t *into* $B = [a, c]$ *and* $C = [c, b]$ *where* $c = (a + b)/2$*.* $[\![t \in B]\!]$ *or* $[\![t \in C]\!]$ *is in* $\mathbb{H}$*. Produce in this way a nested sequence of nonempty closed sets whose width decreases to* 0*, each of which contains the range of* t *restricted to a member of* $\mathbb{H}$*. Invoke Dedekind completeness to show that the intersection of these sets must contain a real number* x*. Then show this intersection cannot contain two real numbers.)*

4.13. ***Exercise.*** *(a) Show that* $\operatorname{gal}(x) \cap {}^*\mathbb{Z} = x + \mathbb{Z}$ *for each* $x \in {}^*\mathbb{Z}$*. The set of equivalence classes* $\{\, \operatorname{gal}(z) \cap {}^*\mathbb{Z} \mid z \in {}^*\mathbb{Z}\,\}$ *forms a partition of* ${}^*\mathbb{Z}$*. Let* S *be a selection of one element from each member of this partition, where we choose* 0 *from the galaxy* $\mathbb{Z}$*.*

$${}^*\mathbb{Z} = \bigcup_{s \in S} (s + \mathbb{Z}) = \mathbb{Z} \cup \left(\bigcup_{s \,\in\, S - \{0\}} (s + \mathbb{Z}) \right)$$

represents ${}^*\mathbb{Z}$ *as a disjoint union of copies of* $\mathbb{Z}$*, each entirely beyond or entirely before the other.*

(b) Suppose $s + \mathbb{Z}$ *and* $t + \mathbb{Z}$ *are different members of the partition. There is a hyperinteger* n *in* $\operatorname{gal}\left(\frac{s+t}{2}\right)$ *and* $n + \mathbb{Z}$ *cannot intersect either of the two partition members* $s + \mathbb{Z}$ *or* $t + \mathbb{Z}$*. So between every two distinct partition members is another.*

(c) The usual order on the hypernaturals is not a well order.

(d) Suppose $\mathrm{gal}(x) \neq \mathrm{gal}(y)$ *and* $a \in \mathrm{gal}(x)$ *and* $b \in \mathrm{gal}(y)$ *and* $a < b$. *Then every member of* $\mathrm{gal}\left(\frac{x+y}{2}\right)$ *exceeds* a *and is less than* b. *Conclude that between every two distinct galaxies is another.*

4.14. ***Exercise.*** *We suppose that* $A_t,\ t \in S$, *is an indexed set of subsets of* $\mathbb{R}^k$.

$$\text{Then} \quad \bigcup_{t\in S} {}^*A_t \subset {}^*\left(\bigcup_{t\in S} A_t\right) \quad \text{and} \quad \bigcap_{t\in S} {}^*A_t \supset {}^*\left(\bigcap_{t\in S} A_t\right).$$

Show that equality rather than containment is guaranteed if S *is finite. Show by counterexample that equality is not guaranteed if* S *is countably infinite.*

5. Transfer

In this section we will discuss a variety of situations in which a statement about real numbers implies one about hyperreals, and conversely. We discuss generalities in the exercise below, and a number of particular instances to illustrate the method after that. Limits to the procedure are discussed.

5.1. ***Exercise.*** *In this exercise we suppose* K *is any subset of* $\mathbb{R}^{n+k}$ *for positive integers* n *and* k. *We identify* $\mathbb{R}^{n+k}$ *with* $\mathbb{R}^n \times \mathbb{R}^k$ *when that is convenient.*

K *might correspond to the set of* $(x, f(x))$ *where* $f\colon A \to B$ *and* $A \subset \mathbb{R}^n$ *and* $B \subset \mathbb{R}^k$. *Also,* K *might correspond to an order relation on* $\mathbb{R}$, *consisting of ordered pairs* $(x,y) \in \mathbb{R}^2$ *where* $(x,y) \in K$ *exactly when* $x < y$. $K \subset \mathbb{R}^3$ *could correspond to a binary operation where* $(x,y,z) \in K$ *exactly when* $x + y = z$. *Examples of constructions represented by sets of this kind abound.*

(a) If $A \subset \mathbb{R}^n$ *and* $B \subset \mathbb{R}^k$. *Consider the two statements*

$$(\forall (x,y) \in K)(x \in A \Rightarrow y \in B) \quad \text{and} \quad (\forall (x,y) \in {}^*K)(x \in {}^*A \Rightarrow y \in {}^*B).$$

Show that one is true exactly when the other is true.

(b) If $A \subset \mathbb{R}^n$ *and* $B \subset \mathbb{R}^k$. *Consider the two statements*

$$(\exists (x,y) \in K)(x \in A \text{ and } y \in B) \quad \text{and} \quad (\exists (x,y) \in {}^*K)(x \in {}^*A \text{ and } y \in {}^*B).$$

Show that one is true exactly when the other is true.

(c) Results similar to (a) and (b) hold for sentences obtained by negating the sentences from (a) or (b), or concatenating a finite number of such sentences using the connectives "and" or "or."

(d) Simple set membership situations can be created by letting $K \subset \mathbb{R}^2$ *be the set of diagonal pairs* (x,x) *for* $x \in \mathbb{R}$. *Then* (a) *becomes*

$$(\forall x \in A)(x \in B) \quad \text{and} \quad (\forall x \in {}^*A)(x \in {}^*B)$$

while (b) reduces to

$$(\exists x \in A)(x \in B) \quad \text{and} \quad (\exists x \in {}^*A)(x \in {}^*B).$$

The last exercise involves examples of what are called **transfer principles**: conditions under which a certain type of statement involving ${}^*\mathbb{R}$ and a related condition involving $\mathbb{R}$ are true or false simultaneously.

The principles above posit situations in which transfer is *guaranteed* to be valid. Our conditions require *the sentence in the "hyper universe" to be the "star transfer" of a sentence involving finitely many bound variables and membership in finitely* *many constant sets from* $\mathbb{R}^n$ *for various* $n \geq 1$. *Each quantification symbol is associated with explicit membership in a constant set.*

We perform the transfer from the "standard world" by taking any constant member of $\mathbb{R}^k$ to the hyperreal constant with which it is identified and enlarging all constant sets, including functions (which are identified with sets of ordered pairs.) All logical, quantification and grouping symbols are left unchanged.

In particular, we do *not* have transfer of sentences involving the power set on ${}^*\mathbb{R}$, nor can we transfer statements between the "hyper universe" and the "real universe" that involve the countable union or intersection of sets, even if each set involved *is* transferable.

Statements that involve constants from ${}^*\mathbb{R} - \mathbb{R}$ cannot be transferred to the "real universe" directly.

However, the hyperreal constant has a purpose in the statement, and that purpose can often be satisfied by prefacing the statement with

$$(\exists z \in {}^*\mathbb{R})(\text{property of the constant used in the statement})$$

and using z in place of the constant. *This* statement might be transferable.

Even with all these restrictions and provisos, there are a large number of useful facts or theorems susceptible to transfer by our criteria if properly phrased. More advanced treatments create more comprehensive transfer theorems, which facilitate discussion of topics such as measure theory, probability, functional analysis and so forth.

We now illustrate the method of transfer with a number of examples.

The Archimedean order property does not hold in ${}^*\mathbb{R}$. That is because if x is an ordinary positive integer and y is positive and unlimited, no integer multiple of x can exceed y.

However the statement

$$(\forall x \in (0,\infty))(\forall y \in (x,\infty))(\exists n \in \mathbb{N})(nx \in (y,\infty))$$

is a statement of the Archimedean order property in $\mathbb{R}$ that is given as a combination of statements of the form (a) and (b) of the last exercise. This property is true, and is equivalent to

$$(\forall x \in (0,\infty)_*)(\forall y \in (x,\infty)_*)(\exists n \in {}^*\mathbb{N})(nx \in (y,\infty)_*)$$

So if you allow hyperinteger multiples, a version of this basic and important fact is true in ${}^*\mathbb{R}$.

This is typical of transfer in action. Learning to phrase a property of interest in language that allows transfer is the issue.

Here is another example.

$$(\forall x \in \mathbb{R})(\exists n \in \mathbb{Z})(\ x \in (n, n+1]\ ).$$

This is the statement that every real number is in an interval bounded by consecutive integers. We transfer and obtain the same fact involving hyperintegers about hyperreals.

$$(\forall x \in {}^*\mathbb{R})(\exists n \in {}^*\mathbb{Z})(\ x \in (n, n+1]_*\ ).$$

It is true that a subset of real numbers is bounded exactly when its enlargement is hyperbounded, a fact we saw above. A transfer argument can demonstrate this.

$$(\exists x, y \in \mathbb{R})(\forall t \in S)(t \in [x, y]) \iff (\exists x, y \in {}^*\mathbb{R})(\forall t \in {}^*S)(t \in [x, y]_*).$$

Note that the phrase $(\forall t \in S)(t \in [x, y])$ has the same meaning as $S \subset [x, y]$ so the subset relation for constant sets does transfer. Using it shortens the last transfer to the more readable

$$(\exists x, y \in \mathbb{R})(S \subset [x, y]) \iff (\exists x, y \in {}^*\mathbb{R})({}^*S \subset [x, y]_*).$$

Consider the condition $(\forall x, y \in \mathbb{R})(\exists z \in \mathbb{R})(x + y < z)$. The condition involves addition and an inequality. It (and similar statements) can be transferred. Let $S \subset \mathbb{R}^3$ denote the addition operation: $(x, y, z) \in S$ exactly when $x + y = z$. Let $T \subset \mathbb{R}^2$ denote the order on $\mathbb{R}$: $(a, b) \in T$ exactly when $a < b$. So

$$\begin{aligned}&(\forall x, y \in \mathbb{R})(\exists z \in \mathbb{R})(x + y < z)\\ &\qquad\qquad \iff (\forall x, y \in \mathbb{R})(\exists w, z \in \mathbb{R})(\ (x, y, w) \in S\ )(\ (w, z) \in T\ ).\end{aligned}$$

So the equivalent statement about sums and order in ${}^*\mathbb{R}$ follows by transfer.

The statement

$$(\forall x, y \in \mathbb{R})(\ \forall S \in \mathbb{P}([x, y])\ )(\exists z \in \mathbb{R})(\forall w \in \mathbb{R})(S \subset [x, w] \Rightarrow [x, z] \subset [x, w])$$

encapsulates Dedekind completeness in $\mathbb{R}$. But this is not in the form of a sentence that transfers. The power set does not transfer: it is not a subset of $\mathbb{R}^k$ for any k. If you settle on a constant S and constant x, y with $S \subset [x, y]$ you get a sentence that does transfer.

$$(\exists z \in \mathbb{R})(\forall w \in \mathbb{R})(S \subset [x, w] \Rightarrow [x, z] \subset [x, w])$$

becomes

$$(\exists z \in {}^*\mathbb{R})(\forall w \in {}^*\mathbb{R})({}^*S \subset [x, w]_* \Rightarrow [x, z]_* \subset [x, w]_*)$$

which references the enlarged set ${}^*S$ contained in the constant interval $[x, y]_*$ with real endpoints.

This is an extremely limited form of completeness. It merely says that the enlargement of a set known to be bounded in $\mathbb{R}$ has a least upper bound in ${}^*\mathbb{R}$. It says nothing about hyperbounded but not bounded sets or even limited sets which are not the enlargement of some bounded set of real numbers. These subsets of ${}^*\mathbb{R}$ cannot be mentioned in a sentence transferred from the real world. And, in fact, the least upper bound property need not be true for such sets.

Finally, we give as an example of transfer the following rather important result.

Suppose s is a sequence and $A \subset \mathbb{R}$. If the values of s are in A for all but finitely many integers then ${}^*s_k \in {}^*A$ for all unlimited k, essentially by definition of ${}^*s$.

On the other hand:

- If ${}^*s_k \in {}^*A$ for all sufficiently large unlimited k then $s_k \in A$ for all but finitely many integers k, which in turn implies that ${}^*s_k \in {}^*A$ for *all* unlimited k.

To see this we transfer the statement

$$(\exists n \in {}^*\mathbb{N})(\forall k \in {}^*\mathbb{N})(\ k \geq n \Rightarrow {}^*s_k \in {}^*A\ )$$

to the standard world as

$$(\exists n \in \mathbb{N})(\forall k \in \mathbb{N})(\ k \geq n \Rightarrow s_k \in A\ ).$$

This last asserts that $s_k \in A$ for all but finitely many k.

5.2. ***Exercise.*** *Prove the triangle inequality for hyperreal numbers using the transfer principle.*

5.3. ***Exercise.*** *In this exercise we consider the set P of prime integers and the **hyperprimes**, ${}^*P$.*

(a) $(\forall n \in \mathbb{N})(n > 1)(\exists p \in P)(\exists m \in \mathbb{N})(n = pm)$*. Use this true statement to conclude that every hyperinteger exceeding* 1 *has a hyperprime factor.*

(b) $(\forall p \in P)(\forall m, n \in \mathbb{N})(\ p = mn \Rightarrow (p = m \text{ or } p = n)\ )$*. This implies that the only positive hyperinteger factors of a hyperprime are itself and* 1*.*

(c) Suppose n is a positive hyperinteger and p a hyperprime. There exists $k \in {}^*\mathbb{N}$ for which p^k is a factor of n but p^{k+1} is not.

(d) If m and n are positive integers with all the same prime power factors then $m = n$. Create an equivalent statement in transferable form to deduce the same fact for hyperintegers and hyperprime powers. (hint: Try the contrapositive.)

5.4. ***Exercise.*** *In this exercise we consider the real sequence s given by $s_n = n!$ and its associated hypersequence. Once again, let P be the set of prime integers.*

(a) If k is a hyperinteger, ${}^*s_k$ is a hyperinteger divisible by every hyperinteger not exceeding k.

(b) Transfer $(\exists x, k \in {}^*\mathbb{N})(1 < x \leq k)(x \text{ divides } 1 + {}^*s_k)$. Use the transferred statement (which is false) to conclude that for any positive hypernatural k there is a hypernatural (namely $1 + {}^*s_k$) not divisible by any hypernatural exceeding 1 *and less than or equal to k.*

*(c) Let S denote any countably infinite subset of P. Let t_n denote the $(n+1)$st member of S in the natural order for each $n \in \mathbb{N}$. Define $w \colon \mathbb{N} \to S$ to be the sequence with w_n equal to the product $t_0 t_1 \cdots t_n$. Show that ${}^*w_k$ is divisible by every member of S for any unlimited hypernatural k and it is **not** divisible by any member of P not in S. Use this to show that the cardinality of the hypernaturals is at least as big as the cardinality of the countably infinite subsets of P. Conclude that the hypernaturals have cardinality at least as big as that of the continuum.*

*(d) On the other hand, the cardinality of the set of real valued sequences is also the cardinality of the continuum. The hyperreals are a partition of this set, so the cardinality of the hyperreals cannot exceed that of the continuum. Coupled with part (c), we find that **the cardinality of the hypernaturals is the same as that of the hyperreals themselves, both equal to the cardinality of the continuum**, in contrast to the situation for the integers and the real numbers.*

(e) Every halo has cardinality of the continuum.

(f) Suppose A is any infinite set of real numbers. Must ${}^*A$ have cardinality of the continuum?

5.5. ***Exercise.*** *From the geometry of the unit circle and the definition of the angle parameter,*

$$\sin(0) = 0 \qquad \sin\left(\frac{\pi}{2}\right) = \cos(0) = 1 \ \textit{and}$$

for all real t, $\cos^2(t) + \sin^2(t) = 1$ *and*

$\cos(t) = \cos(-t)$ *and* $\sin(t) = -\sin(t)$ *and*

for real t, s $\cos(t+s) = \cos(t)\cos(s) - \sin(t)\sin(s)$ *and*

$\sin(s+t) = \sin(s)\cos(t) + \sin(t)\cos(s)$ *and*

for all $t \in (0,1)$ $\quad 0 < \cos(t) < \dfrac{\sin(t)}{t} < 1.$

So these facts all hold for hyperreal t and s or (in the last line) t in the hyperinterval $(0,1)_*$*. Show that if t is positive and infinitesimal*

$$\frac{\sin(t)}{t} \in \text{halo}(1) \quad \textit{and} \quad \frac{1-\cos(t)}{t^2} \in \text{halo}\left(\frac{1}{2}\right).$$

6. The Rearrangement and Hypertail Lemmas

- Suppose $0 \le \varepsilon \in \mathbb{I}$ and $\varepsilon = [t]$. For every positive $\delta \in {}^*\mathbb{R}$ there is a sequence s **whose range is contained in the range of t** and for which $0 < [s] < \delta$.

To see this, let $\frac{1}{2}\delta = [w]$ for a sequence w. Define the set $B = [\![0 < w < t]\!]$. This could be a finite set, or even empty. However unless $B \in \mathbb{H}$, and hence infinite, the result we want is trivial since we can choose $s = t$.

So we may presume that δ is infinitesimal and $B \in \mathbb{H}$. Since $[t] \in \mathbb{I}$, for each $b \in B$ the set $A_b = [\![t < w_b]\!] \in \mathbb{H}$, and in particular is nonempty. Let a_b be a selection of a member of A_b for each $b \in B$.

Now define the sequence s by

$$s_k = \begin{cases} t_k, & \text{if } k \notin B; \\ t_{a_k}, & \text{if } k \in B. \end{cases}$$

The idea here is that if δ and $\varepsilon = [t]$ are any two nonnegative infinitesimals, and if δ is not actually 0, you can create a nonnegative infinitesimal less than δ using a sequence whose only values are selected from the values of t, reorganized so as to get small quickly enough on a member of $\mathbb{H}$.

A very similar result holds if both δ and ε are nonpositive.

We will have several occasions to refer to this result, which we call the **Rearrangement Lemma**.

Moving on to the second result of this section, suppose s is a sequence, $k = [t]$ is an unlimited hypernatural and ${}^*s_k = [s \circ t] \in \mathbb{I}$. Suppose also that $j = [u]$ is a *second* unlimited hypernatural.

k is unlimited so $t(A)$ is an unbounded set of integers for any $A \in \mathbb{H}$.

To say that ${}^*s_k \in \mathbb{I}$ means that for any real $\varepsilon > 0$ the set $[\![-\varepsilon < s \circ t < \varepsilon]\!] \in \mathbb{H}$.

For each $n \in \mathbb{N}$ let $B_n = \left[\!\!\left[-\frac{1}{n+1} < s \circ t < \frac{1}{n+1}\right]\!\!\right]$.

For each $n \in \mathbb{N}$, since $B_n \in \mathbb{H}$ the set $t(B_n)$ is unbounded and so has a member q_n exceeding u_n.

So $d = [q] > j$ and $s_d \in \mathbb{I}$.

Conversely, if ${}^*s_k \notin \mathbb{I}$ for unlimited k then there is a real $\varepsilon > 0$ for which the set $B = [\![\ |s \circ t| > \varepsilon\ ]\!] \in \mathbb{H}$, and is therefore infinite. So for each n there is a member q_n of B exceeding u_n. Once again, $d = [q] > j$ and $s_d \notin \mathbb{I}$.

Our conclusion is the **Hypertail Lemma**.

Suppose s is a real valued sequence.

- If ${}^*s_k \in \mathbb{I}$ *for one* $k \in {}^*\mathbb{N}_\infty$ then ${}^*s_k \in \mathbb{I}$ for an arbitrarily large $k \in {}^*\mathbb{N}_\infty$.
- If ${}^*s_k \notin \mathbb{I}$ *for one* $k \in {}^*\mathbb{N}_\infty$ then ${}^*s_k \notin \mathbb{I}$ for an arbitrarily large $k \in {}^*\mathbb{N}_\infty$.

6.1. ***Exercise.*** *(a) Modify the Hypertail Lemma to show that if* $x \in \mathbb{R}$ *and* s *is a real sequence and* ${}^*s_k \in \mathrm{halo}(x)$ *for some* $k \in {}^*\mathbb{N}_\infty$ *then* ${}^*s_k \in \mathrm{halo}(x)$ *for arbitrarily large* $k \in {}^*\mathbb{N}_\infty$*. Also, if* ${}^*s_k \notin \mathrm{halo}(x)$ *for some* $k \in {}^*\mathbb{N}_\infty$ *then* ${}^*s_k \notin \mathrm{halo}(x)$ *for arbitrarily large* $k \in {}^*\mathbb{N}_\infty$*.*

(b) Is this result true (either part) if $x \in {}^*\mathbb{R} - \mathbb{R}$*?*

(c) Why can't transfer be used to prove the Hypertail Lemma immediately?

7. Open, Closed and Boundary For Subsets of $\mathbb{R}$

First we mention a few basic definitions from the topology of the real numbers.

A subset S of $\mathbb{R}$ is called **open** if it is a **neighborhood** of each of its points: that is, for each $x \in S$ there is a real $\varepsilon > 0$ for which $(x - \varepsilon, x + \varepsilon) \subset S$. A set S is called **closed** if its complement $\mathbb{R} - S = \mathbf{S^c}$ is open.

A point $x \in \mathbb{R}$ is said to be a **boundary point of S** when every neighborhood of x contains points in both S and S^c. The set of boundary points of S is denoted $\partial\mathbf{S}$.

Note that $\partial S = \partial(S^c)$.

Also, for any S, the set $S \cup \partial S$ is the smallest closed set containing S. This set is denoted $\overline{\mathbf{S}}$ and called the **closure of S**.

The **interior of S** is the biggest open set inside S and denoted $\mathbf{S^o}$. It is $S - \partial S = S - \overline{S^c}$. Also, $\overline{S} = S^o \cup \partial S$.

$\mathbb{R}$ and $\varnothing$ are open and closed. Finite intersections of open sets are open, and any union of open sets is open. On the other hand, finite unions of closed sets are closed while any intersection of closed sets is closed.

For $a, b \in \mathbb{R}$ and $a < b$ the interval $[a, b]$ is closed, (a, b) is open but $[a, b)$ is neither open nor closed.

If S is a neighborhood of x, we have

$$(\exists \varepsilon \in \mathbb{R})(\varepsilon > 0)(\ (x - \varepsilon, x + \varepsilon) \subset S\ ).$$

This transfers to

$$(\exists \varepsilon \in {}^*\mathbb{R})(\varepsilon > 0)(\ (x - \varepsilon, x + \varepsilon)_* \subset {}^*S\ ).$$

Now suppose that for some nonzero infinitesimal ε and $x \in S$ the hyperreal $x+\varepsilon$ fails to be in ${}^*S$. So $x+\varepsilon$ must be in the complement ${}^*S^c$. By localization we can choose a sequence t so that $\varepsilon = [t]$ and all values of the sequence $x+t$ are in S^c. By the Rearrangement Lemma we can reorganize some of the values of t to create a new sequence s for which $[x+s] \in {}^*S^c$ and so that $[x+s]$ is arbitrarily close to x.

This means that if there is *any* member of halo(x) not in ${}^*S$ then there is *no* interval $(x-\varepsilon, x+\varepsilon)_* \subset {}^*S$ for *any* positive infinitesimal ε.

In particular, if S is a neighborhood of x the entire halo of x, not just some infinitesimal interval around x, lies entirely in ${}^*S$.

We note that if S is open and $x \in {}^*S$ it is quite possible for halo$(x) \not\subset {}^*S$. If $S = (0,1)$ then ${}^*S$ contains a point $x \in$ halo(1) but halo$(x) =$ halo$(1) \not\subset (0,1)_*$.

- For $S \subset \mathbb{R}$, S is open exactly when halo$(x) \subset {}^*S$ for all $x \in S$.
- For $S \subset \mathbb{R}$, S is closed exactly when halo$(x) \cap {}^*S = \varnothing$ for all $x \in S^c$.
- For $S \subset \mathbb{R}$, $x \in \partial S$ if and only if halo$(x) \cap {}^*S \neq \varnothing \neq$ halo$(x) \cap {}^*S^c$.

8. The Hyperreal Approach to Real Convergent Sequences

The key technical concept used to create proofs in real analysis is the idea of limit—most fundamentally, the limit of a sequence of real numbers. We explore the relationship between standard and nonstandard versions of these tools.

We suppose throughout that $s \colon \mathbb{N} \to \mathbb{R}$ is a sequence with hypersequence extension ${}^*s \colon {}^*\mathbb{N} \to {}^*\mathbb{R}$.

It is customary, and a habit we will adopt, to use the same symbol for a function and its nonstandard extension. Thus $f(c)$ or s_k will, henceforth, denote ${}^*f(c)$ or ${}^*s_k$ if c or k are hyperreal or hypernatural but not real. This will rarely cause confusion, but when it might the ${}^*f$ or ${}^*s$ notation will be resurrected.

In particular, we will use s_k for an extended term ${}^*s_k$ of a sequence s, relying on the fact that k is unlimited to indicate the intended meaning.

x is a **cluster point of s** precisely if

$$(\forall \varepsilon \in \mathbb{R})(\varepsilon > 0)(\forall n \in \mathbb{N})(\exists k \in \mathbb{N})(\ k > n\ )(\ s_k \in (x-\varepsilon, x+\varepsilon)\ ).$$

The transfer of this statement is: x is a cluster point of the sequence s exactly when

$$(\forall \varepsilon \in {}^*\mathbb{R})(\varepsilon > 0)(\forall n \in {}^*\mathbb{N})(\exists k \in {}^*\mathbb{N})(\ k > n\ )(\ s_k \in (x-\varepsilon, x+\varepsilon)_*\ ).$$

Since in this last statement ε can be infinitesimal, an implication of this is that there are extended terms with arbitrarily large subscript in the halo of x whenever x is a cluster point of s.

Suppose k is **any** unlimited hyperinteger and $s_k \in$ halo(x) for real x. By the Hypertail Lemma we conclude that $s_k \in$ halo(x) for arbitrarily large unlimited k, and so x is a cluster point of s.

Recall the **Bolzano-Weierstrass Theorem**, which states that a bounded sequence has a cluster point. Any extended term s_k of a bounded sequence is limited, so its shadow $x =$ shad(s_k) is defined. So $s_k \in$ halo(x) and the theorem is proved.

- x is a cluster point of s exactly when $(\exists k \in {}^*\mathbb{N}_\infty)(s_k \in \text{halo}(x))$.
- If s has a bounded extended term, then the shadow of that term is a cluster point of s.
- **Bolzano-Weierstrass Theorem** Any bounded sequence has at least one cluster point.

The sequence s **converges to a real number x** precisely if

$$(\forall \varepsilon \in \mathbb{R})(\varepsilon > 0)(\exists n \in \mathbb{N})(\forall k \in \mathbb{N})(\ k > n \Rightarrow s_k \in (x - \varepsilon, x + \varepsilon)\ ).$$

In other words, each ε-neighborhood of x contains a "tail" of the sequence s.

The transfer of this statement is: s converges to a real number x precisely if

$$(\forall \varepsilon \in {}^*\mathbb{R})(\varepsilon > 0)(\exists n \in {}^*\mathbb{N})(\forall k \in {}^*\mathbb{N})(\ k > n \Rightarrow s_k \in (x - \varepsilon, x + \varepsilon)_*\ ).$$

In this last statement ε can be infinitesimal, so one implication is that for every **sufficiently large** hyperinteger k, $s_k \in \text{halo}(x)$. But if this is to be true, the Hypertail Lemma then requires $s_k \in \text{halo}(x)$ for **every** unlimited k.

Note that since any halo contains at most one real number, a real sequence can converge to at most one real number, called the **limit of the sequence**. The notation $\lim_{\mathbf{i}\to\infty} \mathbf{s_i}$ is sometimes used for the number x when the sequence s converges to x.

- A sequence s converges to x exactly when $s_k \in \text{halo}(x)$ for every $k \in {}^*\mathbb{N}_\infty$.
- A sequence has at most one limit.

Here is a consequence. Suppose s is a sequence and $s_n \geq s_{n+1}$ for all n. If s is not bounded it is easy to see that every extended term is unlimited. So we presume s to be bounded. If s_k and s_j are two extended terms then each has a shadow. Suppose $k < j$. By the Hypertail Lemma there is a hyperinteger d with $j < d$ and $s_d \in \text{halo}(\text{shad}(s_k))$. But $s_k \leq s_j \leq s_d$ so s_j is in $\text{halo}(\text{shad}(s_k))$ as well. We conclude that extended terms all have the same shadow and every nondecreasing bounded sequence converges to the common shadow of any extended term. The nonincreasing sequences are handled similarly, yielding:

- Every bounded monotone real sequence converges.

As an application, suppose $x \in (0, 1)$. So the sequence $s_n = x^n$ is decreasing and bounded and so has a limit, obviously less than $x < 1$. This limit is 0, because if n is an unlimited hypernatural, $x^{2n} - x^n = x^n(x^n - 1) \in \mathbb{I}$, and the only way that can happen is if $x^n \in \mathbb{I}$.

Another traditional topic in the study of sequences is **Cauchy sequences.**

A sequence s is called Cauchy if

$$(\forall \varepsilon \in (0, \infty))(\exists n \in \mathbb{N})(\forall k, j \in \mathbb{N})(\ k \geq j \geq n \Rightarrow |s_k - s_j| < \varepsilon\ ).$$

By an easy argument, we see that all Cauchy sequences are bounded, so all the extended terms of the sequence are bounded and therefore have shadows.

The triangle inequality also implies that all convergent sequences are Cauchy.

Another application of the triangle inequality allows the Cauchy condition to be rephrased as

$$(\forall \varepsilon \in (0,\infty))(\exists n \in \mathbb{N})(\forall k \in \mathbb{N})(\ k \geq n \Rightarrow |s_k - s_n| < \varepsilon\ ).$$

For each extended term s_n let x_n be $\text{shad}(s_n)$. By transfer we have

$$(\exists n \in {}^*\mathbb{N}_\infty)(\forall k \in {}^*\mathbb{N}_\infty)(\ k \geq n \Rightarrow s_k \in \text{halo}(x_n)\ ).$$

But if $s_k \notin \text{halo}(x_n)$ for *any* unlimited k then the Hypertail Lemma says $s_k \notin \text{halo}(x_n)$ for *arbitrarily large* unlimited k. So in fact if s is Cauchy we must have $s_k \in \text{halo}(x_n)$ for *all* unlimited k.

Since there can be no more than one real number in any halo, we have just shown that Cauchy sequences converge.

If a sequence s is not a bounded sequence, it is easy to find unlimited hypernaturals k, n with $s_k \notin \text{halo}(s_n)$. We have:

- s is Cauchy if and only if $(\forall k, n \in {}^*\mathbb{N}_\infty)(s_k \in \text{halo}(s_n))$.
- A Cauchy sequence converges to the shadow of any extended term.

8.1. ***Exercise.*** *Formulate and prove convergence results for the sum, product and ratio of pairs of convergent sequences using nonstandard methods.*

9. Series

Suppose given a real sequence $a\colon \mathbb{N} \to \mathbb{R}$. Define the **sequence of partial sums** of a by $s_n = \sum_{i=0}^{n} a_i$ for $n \in \mathbb{N}$. Any sequence formed in this way is called a **series**, and the limit, if it exists, is denoted $\sum_{i=0}^{\infty} a_i$. We will discover criteria for convergence of series using nonstandard methods, rather than the usual (real) standard techniques.

First let's consider a specific series, formed from $a_n = x^n$ for $x \in (-1, 1)$.

$$s_n = 1 + x + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

We have seen that $|x^n|$ converges to 0 so if s_n is an extended term

$$\frac{1 - x^{n+1}}{1 - x} \in \text{halo}\left(\frac{1}{1-x}\right) \text{ and so } \sum_{i=0}^{\infty} x^n = \frac{1}{1-x}.$$

We should take a moment to reflect on this result. The proof that x^n converges to 0 and so $\sum_{i=0}^{\infty} x^n = \frac{1}{1-x}$ takes a few lines of hyperreal arithmetic and is essentially complete, as stated.

The same proof using real methods is by no means arduous and the outline would proceed in a similar fashion. However to actually create a complete proof, a pair of perfunctory "ε, N" arguments would need to be created, introducing intermediary entities, estimates and algebra that probably (for most students at least) make the result seem *less obvious than it should be.*

You don't get something for nothing. The hyperreals are harder to build than the real numbers and have some important features which are, possibly, less than obvious. But the hyperreals contain the mechanism of limit-taking as part of their arithmetic and order structure. With these hyperreals in hand, taking limits corresponds to basic arithmetic, rather than messy extra arguments needed to rigorize the obvious.

Getting back to our generic series $s_n = \sum_{i=1}^n a_i$, we define for any hypernaturals n, m the hyperreal $\sum_{i=m}^n a_i$ to be $s_n - s_m$. Recalling facts from the last section regarding convergent sequences,

- s converges to the real number x exactly when $s_n \in \text{ halo}(x) \ \forall n \in {}^*\mathbb{N}_\infty$.
- s converges exactly when $\displaystyle\sum_{i=m}^{n} a_i \in \mathbb{I} \quad \forall m, n \in {}^*\mathbb{N}_\infty$.

$a_{n+1} = s_{n+1} - s_n$ for all n, so $a_n = \sum_{i=n}^{n+1} a_i$ for unlimited n. If s converges this must be in $\mathbb{I}$ for all unlimited n which implies that a converges to 0.

- If s converges then a converges to 0.

Suppose a and b are nonnegative sequences. Suppose also that $a_n \le b_n$ for all $n \in \mathbb{N}$. For every pair of integers n, m with $n \ge m$ we have

$$0 \le \sum_{i=m}^{n} a_i \le \sum_{i=m}^{n} b_i$$

so the same relationship holds for unlimited n, m. But if $\sum_{i=m}^n b_i \in \mathbb{I}$ then $\sum_{i=m}^n a_i \in \mathbb{I}$ too, while if $\sum_{i=m}^n a_i \notin \mathbb{I}$ then $\sum_{i=m}^n b_i \notin \mathbb{I}$.

We have, therefore, the basic **comparison theorem** for series convergence.

- If $0 \le a \le b$ and $\sum_{i=0}^\infty b_i$ exists, then so does $\sum_{i=0}^\infty a_i$.
- Similarly if $\sum_{i=0}^\infty a_i$ fails to exist, then neither does $\sum_{i=0}^\infty b_i$.

Any series s given by $s_n = \sum_{i=0}^n a_i$ is called **absolutely convergent** if $\sum_{i=0}^\infty |a_i|$ exists.

For integers m and n with $n \ge m$ we see that

$$0 \le \sum_{i=m}^{n} |a_i| - \sum_{i=m}^{n} a_i \le 2 \sum_{i=m}^{n} |a_i|$$

so the same relationship holds for unlimited n and m.

If s is absolutely convergent $\sum_{i=m}^n |a_i|$ is infinitesimal for unlimited n and m which, by the inequalities above, implies that $\sum_{i=m}^n a_i \in \mathbb{I}$ as well. We have proven:

- Any absolutely convergent series converges.

Suppose given a positive sequence a for which the sequence $r_n = \frac{a_{n+1}}{a_n}$ converges to a number x.

If $x < 1$ there is a number y with $x < y < 1$. Only finitely many values of r can exceed y, so let L be an integer with $r_j \le y$ when $j \ge L$.

For any integer k with $k \ge L$ we have

$$a_k \le y a_{k-1} \le y^2 a_{k-2} \le \cdots \le y^{k-L} a_L.$$

So we find that for integers $n > m \geq L$

$$\sum_{i=m}^{n} a_i \leq \sum_{i=m}^{n} y^{i-L} a_L = y^{m-L} a_L \sum_{i=m}^{n} y^{i-m}$$
$$= y^{m-L} a_L \sum_{i=0}^{n-m} y^i = y^{m-L} a_L \frac{1 - y^{n-m+1}}{1-y}.$$

The same relationship holds for unlimited n and m, which obviously exceed L. So we see for these n and m that y^{m-L} and therefore $\sum_{i=m}^{n} a_i$ are in $\mathbb{I}$ and the series converges.

On the other hand, if x exceeds 1 an identical argument with inequalities reversed shows that $\sum_{i=m}^{n} a_i \notin \mathbb{I}$ and the series does not converge. We have proven a version of **the ratio test**:

- If a is a positive sequence and the sequence r given by $r_n = \frac{a_{n+1}}{a_n}$ converges to a number x then the series formed from a converges if $x < 1$ and does not converge if $x > 1$.

We finish with a result called the **limit comparison test**.

Suppose a and b are two sequences with positive terms and the sequence p defined by $p_n = \frac{a_n}{b_n}$ converges to x. So there is a positive integer L and number $\varepsilon > 0$ so that if $n \geq L$ then $a_n \leq (x+\varepsilon)b_n$. Since L is finite and the terms of a and b are never 0, we can choose ε large enough so that $a_n \leq (x+\varepsilon)b_n$ for *all* n.

We see now that for any integers n, m with $n \geq m$ we have

$$\sum_{i=m}^{n} a_i \leq (x+\varepsilon) \sum_{i=m}^{n} b_i$$

so the same result holds for unlimited n and m. If $\sum_{i=0}^{\infty} b_i$ exists then $\sum_{i=m}^{n} b_i$ is infinitesimal for unlimited m, n so $\sum_{i=m}^{n} a_i$ is infinitesimal for unlimited n, m too. We conclude:

- If a and b are two sequences with positive terms and the sequence p defined by $p_n = \frac{a_n}{b_n}$ converges to x and if $\sum_{i=0}^{\infty} b_i$ exists then so too will $\sum_{i=0}^{\infty} a_i$.

We could obviously go further, but it seems the point has been made: a nonstandard presentation of series *including proofs* goes about as smoothly as the typical standard discussion of this material *without proofs.*

10. More on Limits

Now suppose $f \colon S \to \mathbb{R}$ and that $(c-\delta, c+\delta)$ contains a member of S *other than* c for every real $\delta > 0$. We do not presume that $c \in S$.

Transfer of this condition implies that there are members of ${}^*S$ other than c (if c is in fact in S) of arbitrarily small infinitesimal separation from c.

We say that the **limit of f as x approaches c is the real number L** provided that for every $\varepsilon > 0$ there is a $\delta > 0$ so that $f(x) \in (L-\varepsilon, L+\varepsilon)$ whenever $x \in S$ and $0 < |x - c| < \delta$.

Notation $\lim_{\mathbf{x}\to\mathbf{c}} \mathbf{f}(\mathbf{x})$ is used for L when the limit of f as x approaches c is L.

$\lim_{x\to c} f(x) = L$ exactly when

$$(\forall\varepsilon \in \mathbb{R})(\varepsilon > 0)(\exists\delta \in \mathbb{R})(\delta > 0)$$
$$(\ x \in S \cap ((c-\delta, c) \cup (c, c+\delta)) \Rightarrow f(x) \in (L-\varepsilon, L+\varepsilon)\ ).$$

The transfer of this statement is

$$(\forall\varepsilon \in {}^*\mathbb{R})(\varepsilon > 0)(\exists\delta \in {}^*\mathbb{R})(\delta > 0)$$
$$(\ x \in {}^*S \cap ((c-\delta, c)_* \cup (c, c+\delta)_*) \Rightarrow f(x) \in (L-\varepsilon, L+\varepsilon)_*\ )$$

In this last condition ε could be infinitesimal, so we have the implication that if L is the limit of f at c there is a hyperreal $\delta > 0$ for which $f(y) \in \text{halo}(L)$ whenever $y \in {}^*S \cap (c-\delta, c+\delta)_*$.

Suppose there is some $y = c + [t] \in {}^*S \cap \text{halo}(c)$, and for which $f(y) \notin \text{halo}(L)$. That means there is a real $\varepsilon > 0$ with $f(y) \notin (L-2\varepsilon, L+2\varepsilon)_*$.

Therefore the set $B = [\![\ f(c+t) \notin (L-\varepsilon, L+\varepsilon)\ ]\!] \in \mathbb{H}$. Applying the Rearrangement Lemma to t we find that there are arbitrarily small infinitesimals $[s]$ with $z = c + [s] \in {}^*S \cap \text{halo}(c)$ but $f(z) \notin (L-\varepsilon, L+\varepsilon)_*$.

So in fact if the "limit implication" found above fails for *any* infinitesimal δ it fails for *every* infinitesimal δ. So:

- $\displaystyle\lim_{x\to c} f(x) = L$ exactly when $(\ \forall x \in {}^*S \cap \text{halo}(c)\ )\ (\ x \neq c \ \Rightarrow\ f(x) \in \text{halo}(L)\ )$.

Since halos of different real numbers are disjoint, this implies that limits are unique when they exist.

10.1. ***Exercise.*** *Formulate definitions for the one sided limits: the* ***limit from below*** $\lim_{\mathbf{x}\nearrow\mathbf{c}} \mathbf{f}(\mathbf{x})$ *and the* ***limit from above*** $\lim_{\mathbf{x}\searrow\mathbf{c}} \mathbf{f}(\mathbf{x})$.

Suppose ${}^*S$ contains positive unlimited members. This will happen exactly when S contains arbitrarily large positive members.

If L is a real number, we say f has the **limit L at infinity** provided that, for each $\varepsilon > 0$ there is an integer N so that $|f(x) - L| < \varepsilon$ whenever $x \in S$ and $x > N$.

The notation $\lim_{\mathbf{x}\to\infty} \mathbf{f}(\mathbf{x})$ is used for L when the limit of f as x approaches infinity is L. Once again, these limits are unique when they exist.

As before, there is a nonstandard version of this condition:

- $\displaystyle\lim_{x\to\infty} f(x) = L$ exactly when $(\ \forall x \in {}^*S_\infty\ )\ (\ x > 0 \ \Rightarrow\ f(x) \in \text{halo}(L)\ )$.

10.2. ***Exercise.*** *(a) Prove that the condition of the bullet above is equivalent to the earlier definition of the limit at infinity.*

(b) If $S = \mathbb{N}$, the function f is a sequence. Our definition of "limit at infinity" is the same as our earlier definition of "limit of a sequence" in that case.

(c) Formulate a definition for the ***limit at minus infinity,*** $\lim_{\mathbf{x}\to-\infty} \mathbf{f}(\mathbf{x})$.

Sometimes a limit can fail to exist because the function values under consideration gyrate wildly through numbers great and small. But other times the failure is easier to describe: the so-called **infinite limits.**

As before, we presume that $(c-\delta, c+\delta)$ contains a member of S *other than* c for every $\delta > 0$.

We write $\lim_{\mathbf{x}\to\mathbf{c}} \mathbf{f}(\mathbf{x}) = \infty$ to indicate that for each integer N there is a $\delta > 0$ so that $f(x) > N$ whenever $x \in S$ and $0 < |x - c| < \delta$.

The nonstandard translation is:

- $\lim\limits_{x\to c} f(x) = \infty$ exactly when ($x \in {}^*S \cap \text{halo}(c) \Rightarrow f(x)$ is positive and unlimited.)

Now suppose ${}^*S$ contains positive unlimited members.

We write $\lim_{\mathbf{x}\to\infty} \mathbf{f}(\mathbf{x}) = \infty$ to indicate that for each integer N there is an integer K so that $f(x) > N$ whenever $x \in S$ and $x > K$.

The nonstandard translation is:

- $\lim\limits_{x\to\infty} f(x) = \infty$ exactly when

$$(x \in {}^*S_\infty \text{ and } x > 0) \Rightarrow (f(x) \text{ is positive and unlimited.})$$

10.3. ***Exercise.*** *(a) Prove that the standard and nonstandard versions of these "infinite limits" definitions actually are equivalent.*

(b) Modify the definitions to create definitions for "negatively infinite limits."

We reiterate that we do not use the locution "the limit exists and is infinite" or variants of that phrase. Only real number limits are said to exist. These infinite limits are simply descriptions of *how* a limit fails to exist.

11. Continuity and Uniform Continuity

Suppose $S \subset \mathbb{R}$. A member c of S is called an **isolated point of S** when there is an open interval containing c but no other member of S.

If a point c of S is *not* isolated, every open interval containing c contains a member of S other than c. This is equivalent to saying that halo(c) contains points of ${}^*S$ other than c.

A function $f\colon S \to \mathbb{R}$ is called **continuous at c** $\in$ **S** exactly when either c is an isolated point of S or $\lim_{x\to c} f(x) = f(c)$.

f is called **continuous** if it is continuous at c for every $c \in S$.

f is continuous at $c \in S$ exactly when $(x \in {}^*S \cap \text{halo}(c) \Rightarrow f(x) \in \text{halo}(f(c)))$.

An analogous result equating continuity at x with a condition on sequences converging to x is also listed below.

We see that if $S \subset \mathbb{R}$ and $f\colon S \to \mathbb{R}$ then the following are equivalent:

- f is continuous at $x \in S$
- $f({}^*S \cap \text{halo}(x)) \subset \text{halo}(f(x))$.
- $f(s_k) \in \text{halo}(f(x))$ for $x \in S$ and

  for every sequence s in S converging to $x \in S$ and $k \in {}^*\mathbb{N}_\infty$.

We present a fact from the the usual discussion of continuity to illustrate the nonstandard methods in action.

Suppose both f and g are real valued and continuous at a point x in the interior of their common domains. Suppose $g(x) \neq 0$. We will show that the function h defined by $h(x) = \frac{f(x)}{g(x)}$ is continuous at x.

Suppose $y \in \text{halo}(x)$. So $g(y) = g(x) + j$ and $f(y) = f(x) + k$ for infinitesimals j and k. Note that $g(y)$ is in the halo of $g(x)$, so is itself appreciable. We have:

$$\begin{aligned} h(y) - h(x) &= \frac{f(y)}{g(y)} - \frac{f(x)}{g(x)} = \frac{f(y)g(x) - g(y)f(x)}{g(y)g(x)} \\ &= \frac{kg(x) - jf(x)}{g(y)g(x)} = k\frac{g(x)}{g(y)g(x)} - j\frac{f(x)}{g(y)g(x)} \in \mathbb{I}. \end{aligned}$$

Since y was a generic member of $\text{halo}(x)$ we have finished the proof of continuity.

If you recall, the standard proof uses three linked $\varepsilon-\delta$ invocations, using estimates on the magnitudes of f and g near x. It is by no means hard as these things go, but it seems pretty clear that our argument is conceptually direct and cleans things up nicely.

11.1. ***Exercise.*** *Formulate and prove continuity results for the sum and product and composition of pairs of continuous functions using nonstandard methods.*

We will now examine uniform continuity.

A function f as above is called **uniformly continuous** if

$$\begin{aligned} &(\forall \varepsilon \in \mathbb{R})(\varepsilon > 0)(\exists \delta \in \mathbb{R})(\delta > 0)(\forall x \in S) \\ &\qquad (\ y \in S \cap (x - \delta, x + \delta) \Rightarrow f(y) \in (f(x) - \varepsilon, f(x) + \varepsilon)\ ). \end{aligned}$$

This statement transfers to

$$\begin{aligned} &(\forall \varepsilon \in {}^*\mathbb{R})(\varepsilon > 0)(\exists \delta \in {}^*\mathbb{R})(\delta > 0)(\forall x \in {}^*S) \\ &\qquad (\ y \in {}^*S \cap (x - \delta, x + \delta)_* \Rightarrow f(y) \in (f(x) - \varepsilon, f(x) + \varepsilon)_*\ ). \end{aligned}$$

In words: For each positive hyperreal ε there is a positive hyperreal δ so that whenever x, y are in the domain of f and $|x - y| < \delta$ then $|f(x) - f(y)| < \varepsilon$.

This is distinct in two ways from the statement about simple continuity. First, for each ε the same δ must "work" for all x. Second, both x and y values come from ${}^*S$. For continuity, x was taken from S, and only y was from ${}^*S$.

Our transferred condition for uniform continuity implies that there is an infinitesimal $\delta > 0$ so that $f(x) \in \text{halo}(f(y))$ whenever $x, y \in {}^*S$ and $|x - y| < \delta$.

Suppose there is *any* infinitesimal $\delta = [j]$ with both $x = [t]$ and $x + \delta$ in ${}^*S$ but for which $f(x + \delta) \notin \text{halo}(f(x))$. So there must be a real $\varepsilon > 0$ for which

$$[\![\ |f(t + j) - f(t)| > \varepsilon\ ]\!] \in \mathbb{H}.$$

Since members of $\mathbb{H}$ are infinite, this set contains arbitrarily large integers. Since δ is infinitesimal, this set contains integers n for which $t_n + j_n$ is arbitrarily close to t_n. So f cannot be uniformly continuous.

We conclude:

- If $S \subset \mathbb{R}$ and $f \colon S \to \mathbb{R}$ then f is uniformly continuous on S if and only if $\forall x \in {}^*S,\ f({}^*S \cap \text{halo}(x)) \subset \text{halo}(f(x))$.

We use this criterion to show the sine function to be uniformly continuous on $\mathbb{R}$.

Suppose x and $y = x + t$ are two members of halo$(x) \in {}^*\mathbb{R}$. From our earlier calculations, $sin(t) = w \in \mathbb{I}$ and $1 - cos(t) = u \in \mathbb{I}$. So

$$\begin{aligned}\sin(y) &= \sin(x+t) = \sin(x)\cos(t) + \sin(t)\cos(x)\\ &= \sin(x)(1-u) + w\ \cos(x) = \sin(x) + (w\ \cos(x) - u\ \sin(x)) \in \text{halo}(\sin(x)).\end{aligned}$$

The function f given by $f(x) = \frac{1}{x}$ and defined on $\mathbb{R} - \{0\}$, on the other hand, is continuous but *not* uniformly continuous.

For real $x \neq 0$ and nonzero infinitesimal ε

$$\frac{1}{x+\varepsilon} - \frac{1}{x} = \frac{x - (x+\varepsilon)}{x(x+\varepsilon)} = \varepsilon \frac{-1}{x(x+\varepsilon)} \in \mathbb{I}.$$

However if x is allowed to get infinitesimally close to 0, this difference can be made appreciable no matter how small ε might be.

The function g given by $g(x) = x^2$ and defined on $\mathbb{R}$ is continuous and is *not* uniformly continuous.

Suppose $x \in \mathbb{R}$ and $\varepsilon \neq 0$ is in $\mathbb{I}$.

$$(x+\varepsilon)^2 - x^2 = 2x\varepsilon + \varepsilon^2 \in \mathbb{I}.$$

But if x is unlimited, this need not be infinitesimal. There is no uniform infinitesimal bound on the difference, independent of unlimited x.

We finish the section with proofs of a couple of standard results from real topology, proved using nonstandard methods.

First, suppose $f\colon [a,b] \to \mathbb{R}$ is continuous. Pick $x, y \in [a,b]_*$. Suppose $y \in$ halo(x). Both x and y are bounded and so possess shadows. They are in the same halo so they have the same shadow. $[a,b]$ is closed so that shadow must be in $[a,b]$. Let $z =$ shad(x). By continuity both $f(x)$ and $f(y)$ must be in halo$(f(z))$ and so $f(y)$ is in halo$(f(x))$. Our conclusion:

- Any continuous function defined on a closed and bounded interval is uniformly continuous.

Next let's consider an important result called the **Intermediate Value Theorem**.

Suppose f is continuous on the interval $[a,b]$ and d is a real number between $f(a)$ and $f(b)$. For specificity suppose $f(a) < d < f(b)$ and we leave other cases to the reader.

For each $n \in \mathbb{N}$ consider the numbers $f\left(a + \frac{k(b-a)}{n+1}\right)$ for $0 \leq k \leq n+1$.

There is an integer k_n with

$$f\left(a + \frac{k_n(b-a)}{n+1}\right) > d \quad \text{but}$$

$$f\left(a + \frac{k(b-a)}{n+1}\right) \leq d \text{ for } 0 \leq k < k_n.$$

Letting t_n denote $a + \frac{(k_n-1)(b-a)}{n+1}$ and $\triangle_n = \frac{b-a}{n+1}$ we have

$$f(t_n) \leq d < f(t_n + \triangle_n) \quad \forall n \in \mathbb{N}.$$

By transfer, the same inequality holds for $^*f$ and extended terms of the sequences.

Both t and $t{+}\vartriangle$ are bounded sequences so their extended terms have shadows.

If k is unlimited, the extended terms t_k and $t_k{+}\vartriangle_k$ are infinitesimally separated so they share a shadow c. The interval $[a, b]$ is closed so $c \in [a, b]$.

By uniform continuity $f(c)$ and $f(t_k)$ and $f(t_k{+}\vartriangle_k)$ are in the same halo.

$$f(t_k) \leq d \leq f(t_k{+}\vartriangle_k)$$

so that halo must be halo(d). A halo can have at most one real member, so $f(c) = d$.

- Any function defined and continuous on an interval $[a, b]$ attains all real values between $f(a)$ and $f(b)$ for some real number in $[a, b]$.

It now follows by standard arguments that a continuous function f defined on a closed interval $[a, b]$ actually achieves both a maximum and minimum value at points in that interval, a fact known as the **Extreme Value Theorem**.

To see this, use uniform continuity to show that $f([a, b])$ is a bounded set, and so that set has an infimum i and supremum s. If there is no $c \in [a, b]$ at which $f(c) = s$ then $g(x) = \frac{1}{f(x)-s}$ is continuous on $[a, b]$. But the range of g is unbounded, a contradiction. So there is a point c at which $f(c) = s$. The case for the infimum is argued similarly.

- Any function defined and continuous on an interval $[a, b]$ attains both a maximum and a minimum value at members of $[a, b]$.

Let's consider any function that is defined and continuous and on an interval $[a, b]$. It is implied from the remarks above that f attains both a maximum and minimum value on $[a, b]$ and also does not "skip" any values between these maximum and minimum values, so the range of f is a closed interval $[r, s]$. This is also true for f restricted to any closed subinterval of $[a, b]$.

Now let's pose the additional condition that f is one-to-one. Under this assumption we claim that $f(a)$ is one of r or s while $f(b)$ is the other.

To see this, we first consider r. Suppose $f(c) = r$ and c is neither a nor b. Then $f(a) > r$ and $f(b) > r$. So pick any t satisfying both $f(a) > t > r$ and $f(b) > t > r$. The intermediate value theorem applied to f restricted to $[a, c]$ implies there is a $d \in [a, c]$ for which $f(d) = t$. Similarly, for f restricted to $[c, b]$ there is an $e \in [c, b]$ for which $f(e) = t$. This violates our condition that f is one-to-one.

The case for s is handled similarly. Our conclusion:

- The range of any function defined and continuous on an interval is itself an interval. If the domain is the closed interval $[a, b]$ then the range is a closed interval $[r, s]$. If f is one-to-one, $f(a)$ is one of the endpoints of the range, r or s, while $f(b)$ is the other endpoint.

11.2. ***Exercise.*** *This last remark has useful consequences. Suppose* $f: [a, b] \to [r, s]$ *is continuous, one-to-one and onto* $[r, s]$.

(a) For open subinterval (d, e) *of* $[a, b]$, $f(\,(a, b)\,)$ *is an open subinterval of* $[r, s]$.

(b) If A *is any open subset of* $[a, b]$ *then* $f(A)$ *is an open subset of* $[r, s]$.

(c) The function inverse to f is continuous.

12. Derivatives

A function $f\colon (a,b) \to \mathbb{R}$ is called **differentiable at c** $\in (a,b)$ if there is a real number denoted $\mathbf{f'(c)}$ for which

$$\frac{f(c+h)-f(c)}{h} \in \text{halo}(f'(c)) \quad \text{for all nonzero } h \in \mathbb{I}.$$

This can be rephrased as follows:

$$(\forall h \in \mathbb{I})(h \neq 0)(\exists j_h \in \mathbb{I})(\ f(c+h) - f(c) = (\ f'(c) + j_h\ )h\ ).$$

It follows instantly that differentiability implies continuity.

If f is differentiable at c, $\text{shad}\left(\frac{f(c+h)-f(c)}{h}\right) = f'(c)$ for all infinitesimal $h \neq 0$.

Even more, suppose $f'(c) > 0$. The definition of differentiability given above implies that

$$(\exists \varepsilon \in (0,\infty)_*)(\forall h \in (0,\varepsilon)_*)(\ f(c+h) > f(c) > f(c-h)\ ).$$

This transfers to the statement

$$(\exists \varepsilon \in (0,\infty))(\forall h \in (0,\varepsilon))(\ f(c+h) > f(c) > f(c-h)\ ).$$

A similar result holds for a negative derivative with inequalities reversed. We have:

- A function differentiable at a point c with positive derivative is strictly increasing on some real interval around c.
- A function differentiable at a point c with negative derivative is strictly decreasing on some real interval around c.

f is called **differentiable on** $(\mathbf{a},\mathbf{b})$ if it is differentiable at every $c \in (a,b)$. The function f' is called the **derivative of f**.

Let's give some specific examples of derivatives.

The sine function is differentiable and its derivative is the cosine function. That is because for each $x \in \mathbb{R}$ and nonzero infinitesimal h

$$\begin{aligned}\frac{\sin(x+h)-\sin(x)}{h} &= \frac{\sin(x)\cos(h)+\cos(x)\sin(h)-\sin(x)}{h}\\ &= \cos(x)\frac{\sin(h)}{h} + \sin(x)\frac{\cos(h)-1}{h}.\end{aligned}$$

We saw earlier that $\frac{\cos(h)-1}{h^2}$ is appreciable so

$$\sin(x)\frac{\cos(h)-1}{h^2}h = \sin(x)\frac{\cos(h)-1}{h}$$

is infinitesimal. We conclude that

$$\frac{\sin(x+h)-\sin(x)}{h} \in \text{halo}(cos(x)).$$

The cosine function is also differentiable.

$$\begin{aligned}\frac{\cos(x+h)-\cos(x)}{h} &= \frac{\cos(x)\cos(h)-\sin(x)\sin(h)-\cos(x)}{h}\\ &= \cos(x)\frac{\cos(h)-1}{h} - \sin(x)\frac{\sin(x)}{h}.\end{aligned}$$

As in the previous calculation, the first term in the last line is infinitesimal, while the second is in the halo of $-\sin(x)$.

Suppose f and g are generic functions, differentiable at point c in the interior of their common domain. For infinitesimal nonzero h, there are infinitesimals k and w with

$$\frac{f(c+h)-f(c)}{h} = f'(c)+k \quad \text{and} \quad \frac{g(c+h)-g(c)}{h} = g'(c)+w.$$

Since g is continuous at c there is infinitesimal z with $g(c+h) = g(c)+z$. So

$$\begin{aligned}&\frac{f(c+h)g(c+h)-f(c)g(c)}{h}\\ &\quad= \frac{f(c+h)g(c+h)-f(c)g(c+h)+f(c)g(c+h)-f(c)g(c)}{h}\\ &\quad= g(c+h)\frac{f(c+h)-f(c)}{h} + f(c)\frac{g(c+h)-g(c)}{h}\\ &\quad= (g(c)+z)(f'(c)+k)+f(c)(g'(c)+w) \in \text{halo}(\, f'(c)g(c)+f(c)g'(c)\,).\end{aligned}$$

This is the **product rule** from standard calculus.

12.1. ***Exercise.*** *(a) Formulate and prove differentiation results for the sum and ratio (this is the **quotient rule**) of pairs of differentiable functions using nonstandard methods.*

(b) Show (by induction and, in the negative case by the quotient rule) that the derivative of the function f given by $f(x) = x^n$ is $f'(x) = nx^{n-1}$.

(c) Conclude that polynomials and rational functions (wherever defined) are differentiable.

- **(product rule)** If f and g are differentiable at c then so is fg and $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$.
- **(quotient rule)** If f and g are differentiable at c and $g(c) \neq 0$ then $\frac{f}{g}$ is differentiable at c and $\left(\frac{f}{g}\right)'(c) = \frac{f'(c)g(c)-f(c)g'(c)}{(g(c))^2}$.
- **(sum rule)** If f and g are differentiable at c then so is $f+g$ and $(f+g)'(c) = f'(c)+g'(c)$.

Now suppose that f and g are two functions and g is differentiable at c and f is differentiable at the number $g(c)$. So if h is a nonzero infinitesimal, continuity of g at c implies that $g(c+h)-g(c)$ is infinitesimal. So if $g(c+h)-g(c)$ is nonzero too we have

$$\begin{aligned}\frac{f(g(c+h)) - f(c)}{h}& \\ &= \frac{f(g(c+h)) - f(g(c))}{g(c+h) - g(c)} \cdot \frac{g(c+h) - g(c)}{h} \\ &\in \text{halo}(\, f'(g(c))g'(c)\,).\end{aligned}$$

If $g(c+h) - g(c)$ is actually 0 then $0 = \frac{g(c+h)-g(c)}{h} \in \text{halo}(g'(c))$ and we find $\frac{f(g(c+h))-f(g(c))}{h} \in \text{halo}(\, f'(g(c))g'(c)\,) = 0$ in this case too. We have just proven the **chain rule**.

- **(chain rule)** If g is differentiable at c and f is differentiable at $g(c)$ then $f \circ g$ is differentiable at c and $(f \circ g)'(c) = f'(g(c))\, g'(c)$.

An easy and useful application of this is to function pairs inverse to each other. Suppose f and g are inverse functions (to each other) and $f(c) = d$ and $f'(c)$ exists and is nonzero. So f is either strictly increasing or strictly decreasing in some small neighborhood of c, which implies that g is defined and either strictly increasing or strictly decreasing on some small neighborhood of d. Since f and g represent the same relationship—repackaged—it *should* be possible to determine everything about g, including derivatives, by examining f. The chain rule applied to $f \circ g$ immediately implies that $g'(d)$ exists and is $\frac{1}{f'(c)}$.

However if $f'(c) = 0$, the fact that $\frac{f(g(c+h))-f(g(c))}{g(c+h)-g(c)} \cdot \frac{g(c+h)-g(c)}{h} \in \text{halo}(1)$ for nonzero infinitesimal h means that $\frac{g(c+h)-g(c)}{h}$ is unlimited: that is, g is not differentiable at d.

13. Results Related to the Mean Value Theorem

Now suppose f is differentiable on (a,b) and continuous on $[a,b]$. We know from the Extreme Value Theorem that f attains a maximum at some $c \in [a,b]$. We suppose further that $c \in (a,b)$.

Let $(c-\varepsilon, c+\varepsilon)$ denote an interval contained in (a,b), for some real $\varepsilon > 0$. If $h \in (0,\varepsilon)$ then $f(c-h) - f(c) \le 0$ and $f(c+h) - f(c) \le 0$. Since this inequality is true for every sufficiently small positive h it must be true for all positive infinitesimal h by transfer and rearrangement. So for infinitesimal positive h

$$\frac{f(c+h) - f(c)}{h} \le 0 \le \frac{f(c-h) - f(c)}{-h}.$$

But the shadow of both outer terms is the real number $f'(c)$ so that shadow must be 0.

Similarly, $f'(c) = 0$ if c is the location of a minimum of f and $c \in (a,b)$.

- If $f\colon [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b) and if $c \in (a,b)$ is a location of either the maximum or the minimum value of f on $[a,b]$ then $f'(c) = 0$.

There are numerous useful consequences of this result, proven by standard means.

For instance suppose a function f is continuous on an interval $[a, b]$ and differentiable on (a, b).

Let h be the function defined by $h(x) = (f(b)-f(x))(b-a)-(f(b)-f(a))(b-x)$. Since $h(a) = h(b) = 0$ there must be a point c in the interval (a, b) at which h achieves either a maximum or a minimum value. But then $h'(c) = 0$ which yields $-f'(c)(b-a) + (f(b) - f(a)) = 0$. We have the **Mean Value Theorem**:

- If $f\colon [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) and then there is a point $c \in (a, b)$ at which
$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

This implies immediately another useful result. If f as above is neither strictly increasing nor strictly decreasing there must be numbers c, d and e with $a \leq c < d < e \leq b$ but for which $f(d)$ is *not* strictly between $f(c)$ and $f(e)$. Since f is continuous, the intermediate value theorem implies there are numbers x, y with $x \neq y$ and $c \leq x < y \leq e$ and $f(x) = f(y)$. The mean value theorem implies that there is a point between x and y at which the derivative of f is 0. So:

- If $f\colon [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) and if f' is never 0 on (a, b) then f is either strictly increasing on $[a, b]$ or strictly decreasing on $[a, b]$.

Now suppose $f\colon [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) and z is a number strictly between $f'(c)$ and $f'(d)$ for $c < d$ with $c, d \in (a, b)$.

The function h given by $h(x) = f(x) - zx$ is differentiable and $h'(c)$ and $h'(d)$ are nonzero and of opposite sign. So h is strictly increasing on some small neighborhood of one of c or d, and strictly decreasing on some small neighborhood of the other. So by the Intermediate Value Theorem (four cases) there are unequal points $x, y \in [c, d]$ with $h(x) = h(y)$ and by the mean value theorem there is a point $w \in (x, y)$ with $h'(w) = 0$. This means $f'(w) = z$. Our conclusion:

- If $f\colon [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) and if z is between $f'(c)$ and $f'(d)$ then there is a point w between c and d at which $f'(w) = z$.

The result above says that whatever discontinuities a derivative might have, they are *not* the kind that cause it to skip intermediate values.

- If $f\colon [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b) and if f' is never 0 on (a, b) then f' is either always positive or always negative.

Suppose f and g are two functions continuous on $[a, b]$ and differentiable on (a, b) and suppose $g'(x)$ is never 0 for $x \in (a, b)$. In particular, the mean value theorem implies that $g(a) \neq g(b)$.

Consider the function K given by
$$K(x) = f(x) - f(a) - (g(x) - g(a))\frac{f(b) - f(a)}{g(b) - g(a)}.$$

Note $K(b) = K(a) = 0$ so there is a point $c \in (a,b)$ with $K'(c) = 0$. This yields the **Cauchy Mean Value Theorem**:

- If $f, g \colon [a,b] \to \mathbb{R}$ are continuous on $[a,b]$ and differentiable on (a,b) and if g' is never 0 on (a,b) then $g(a) \neq g(b)$ and there exists $c \in (a,b)$ with
$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}.$$

As a final result in this section we prove several versions of **l'Hôpital's Rule**.

Suppose f, f', g and g' are defined in some interval (a,b) and $\lim_{x\nearrow b} \frac{f'(x)}{g'(x)} = L$. So g' must be nonzero on some interval bounded above by b, and the Mean Value Theorem then implies that g itself must be nonzero and either strictly increasing or strictly decreasing in some interval (c,b). If $L \neq 0$ then, in addition and by the same reasoning, there must be some interval of the same form upon which both f and f' are also nonzero, and if $L \neq 0$ we choose (c,b) to satisfy these additional conditions as well.

Case One: $L = \infty$ and $\lim_{x\nearrow b} f(x) = \lim_{x\nearrow b} g(x) = 0$.

For each real number K, there is a number s with $b > s > c$ for which $\frac{f'(x)}{g'(x)} > K$ whenever $b > x \geq s$, and so by the Cauchy Mean Value Theorem, $s < t < b$ implies that
$$\frac{f(s)-f(t)}{g(s)-g(t)} > K$$
and this same inequality holds for the enlargements of f and g to $(c,b)_*$.

If $b = \infty$ (replace "$x \nearrow b$" with "$x \to \infty$") we can choose unlimited t in which case both $f(t)$ and $g(t)$ are in $\mathbf{I}$. So $\frac{f(s)}{g(s)} > K$. Since this can be done for each K, we have $\lim_{x\to\infty} \frac{f(x)}{g(x)} = \infty$.

If $b < \infty$ we choose t infinitesimally close to b and the same result follows: $\lim_{x\nearrow b} \frac{f(x)}{g(x)} = \infty$.

Case Two: $0 \leq L < \infty$ and $\lim_{x\nearrow b} f(x) = \lim_{x\nearrow b} g(x) = 0$.

For any real numbers K and M with $K < L < M$, there is a number s with $b > s > c$ for which $K < \frac{f'(x)}{g'(x)} < M$ whenever $b > x \geq s$, and so by the Cauchy Mean Value Theorem, $s < t < b$ implies that
$$K < \frac{f(s)-f(t)}{g(s)-g(t)} < M$$
and these same inequalities hold, as before, for the enlargements of f and g to $(c,b)_*$.

If $b = \infty$ we again choose unlimited t, for which both $f(t)$ and $g(t)$ are in $\mathbf{I}$. So $K < \frac{f(s)}{g(s)} < M$. Since this can be done for each K and M subject only to the condition $K < L < M$, we have $\frac{f(s)}{g(s)} \in \text{halo}(L)$ whenever s is positive and unlimited. We conclude that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = L$.

If $b < \infty$ we choose $t < b$ and $t \in \text{halo}(b)$. The same result follows, and we find $\lim_{x\nearrow b} \frac{f(x)}{g(x)} = L$.

Case Three: $L = \infty$ and $\lim_{x\nearrow b} g(x) = \infty$.

For each positive integer K, there is a number s with $b > s > c$ for which $\frac{f'(x)}{g'(x)} > K$ whenever $b > x \geq s$.

By the Cauchy Mean Value Theorem, $s < t < b$ implies that

$$\frac{f(t)-f(s)}{g(t)-g(s)} = \frac{f(t)}{g(t)}\,\frac{\left(1-\frac{f(s)}{f(t)}\right)}{\left(1-\frac{g(s)}{g(t)}\right)} > K$$

and this same inequality holds for the enlargements of f and g to $(c,b)_*$.

If $b = \infty$ choose unlimited t and if $b < \infty$ choose t infinitesimally close to b. Then $\frac{g(s)}{g(t)} \in \mathbb{I}$ so

$$\frac{f(t)}{g(t)}\left(1-\frac{f(s)}{f(t)}\right) > K + i \quad \text{for some infinitesimal } i.$$

This implies several things. First, $f(t) - f(s) > (K+i)g(t)$ and $g(t)$ is unlimited but $f(s)$ is not, so $f(t)$ must be unlimited. This in turn implies $\frac{f(s)}{f(t)}$ is infinitesimal and we conclude that

$$\frac{f(t)}{g(t)} > K + j \quad \text{for some infinitesimal } j.$$

This is true for any positive integer K so $\frac{f(t)}{g(t)}$ is unlimited and $\lim_{x\nearrow b} \frac{f(x)}{g(x)} = \infty$.

Case Four: $0 < L < \infty$ and $\lim_{x\nearrow b} g(x) = \infty$.

For any real numbers K and M with $0 < K < L < M$, there is a number s with $b > s > c$ for which $K < \frac{f'(x)}{g'(x)} < M$ whenever $b > x \geq s$, and so by the Cauchy Mean Value Theorem, $s < t < b$ implies that

$$0 < K < \frac{f(s)-f(t)}{g(s)-g(t)} = \frac{f(t)}{g(t)}\,\frac{\left(1-\frac{f(s)}{f(t)}\right)}{\left(1-\frac{g(s)}{g(t)}\right)} < M$$

and these same inequalities hold, as before, for the enlargements of f and g to $(c,b)_*$.

If $b = \infty$ we choose unlimited t and if $b < \infty$ choose $t < b$ and $t \in \text{halo}(b)$. So $\frac{g(s)}{g(t)}$ is infinitesimal and once again

$$0 < g(t)(K+i) < f(s) - f(t) < g(t)(M+i) \quad \text{for some infinitesimal } i$$

which implies that $f(t)$ is unlimited. So $\frac{f(s)}{f(t)}$ is *also* infinitesimal and

$$K + j < \frac{f(t)}{g(t)} < M + l \quad \text{for certain infinitesimals } j, l.$$

This implies that $\frac{f(t)}{g(t)} \in \text{halo}(L)$ for these t and so $\lim_{x\nearrow b} \frac{f(x)}{g(x)} = L$.

Case Five: $0 = L$ and $\lim_{x\nearrow b} g(x) = \infty$.

If f' is nonzero in some interval (c,b) the argument of Case Four carries through unchanged by choosing $K = 0$, yielding the conclusion $\lim_{x\nearrow b} \frac{f(x)}{g(x)} = 0$.

This argument does *not* work if f', and consequently f itself, can be zero somewhere in any interval of the form (c,b). We modify it as follows.

For any real numbers K and M with $K<0<M$, there is a number s with $b>s>c$ for which $K<\frac{f'(x)}{g'(x)}<M$ whenever $b>x\geq s$, and so by the Cauchy Mean Value Theorem, $s<t<b$ implies that

$$K<\frac{f(s)-f(t)}{g(s)-g(t)}=\frac{1}{g(t)}\,\frac{f(t)-f(s)}{\left(1-\frac{g(s)}{g(t)}\right)}<M$$

and these same inequalities hold, as before, for the enlargements of f and g to $(c,b)_*$.

If $b=\infty$ we choose unlimited t and if $b<\infty$ choose $t<b$ and $t\in\text{halo}(b)$. Then

$$K+i<\frac{f(s)-f(t)}{g(t)}=\frac{f(s)}{g(t)}-\frac{f(t)}{g(t)}<M+i \quad \text{for some infinitesimal } i.$$

Since $\frac{f(s)}{g(t)}$ is infinitesimal, so must be $\frac{f(t)}{g(t)}$ for each of these t, and we find that $\lim_{x\nearrow b}\frac{f(x)}{g(x)}=L$.

13.1. ***Exercise.*** *Make minor modifications of the given statements and proofs found above to create similar limit theorems corresponding to the cases listed below.*

(a) Consider all five cases with limits switched to "$\lim_{x\searrow a}$*."*

(b) Consider the first four cases (now eight cases, after you do part (a)) with the sign of L *switched.*

(c) Consider the last three cases (now grown to ten cases after you do parts (a) and (b)) with the limit of g *switched to* $-\infty$*.*

Our combined conclusion is **l'Hôpital's Rule**:

> Suppose $f,g\colon(a,b)\to\mathbb{R}$ are differentiable on (a,b). In the following statements, the cases of $a=-\infty,\ b=\infty$ or $L=\pm\infty$ are not excluded.

- Suppose $\lim_{x\nearrow b}f(x)=\lim_{x\nearrow b}g(x)=0$ or $\lim_{x\nearrow b}g(x)=\pm\infty$.

  $\lim_{x\nearrow b}\frac{f'(x)}{g'(x)}=L$ implies $\lim_{x\nearrow b}\frac{f(x)}{g(x)}=L$.
- Suppose $\lim_{x\searrow a}f(x)=\lim_{x\searrow a}g(x)=0$ or $\lim_{x\searrow a}g(x)=\pm\infty$.

  $\lim_{x\searrow a}\frac{f'(x)}{g'(x)}=L$ implies $\lim_{x\searrow a}\frac{f(x)}{g(x)}=L$.

14. Riemann Integral Preliminaries

The standard definition of the Riemann integral is too complicated for us to transfer with our transfer methods. So some initial effort must be devoted to standard proofs of a few theorems, until we arrive at situations we recognize to be equivalent to Riemann integrability which we *can* transfer.

Those readers who are concentrating on nonstandard methods only should scan the exercises and the bulleted items and skip to the next section.

Suppose f is any function **bounded** on the interval $[a,b]$.

A **partition** P of the interval $[a, b]$ is a finite selection of members of $[a, b]$ containing both of the endpoints a and b. If the partition contains $n+1$ distinct elements we will enumerate them as $x_0, x_2, \ldots, x_n$ where $x_i < x_{i+1}$ for $i = 0, \ldots, n$.

The **mesh** of a partition, mesh(P), is the maximum value of $\triangle \mathbf{x_i} = x_i - x_{i-1}$ for $i = 1, \ldots, n$.

A function $q\colon \{1, \ldots, n\} \to [a, b]$ is said to be **subordinate** to partition P if P has $n+1$ distinct members and $q_i \in [x_{i-1}, x_i]$ for $i = 1, \ldots, n$.

For each $i = 0, \ldots, n-1$ the set $f([x_i, x_{i+1}])$ is bounded and so has a supremum $M_i(P)$ and infimum $m_i(P)$. If q is subordinate to P, $m_i(P) \le f(q_i) \le M_i(P)$ for each i.

We define

$$\mathcal{L}(P, f) = \sum_{i=1}^{n} m_i(P) \triangle x_i$$

$$\text{and} \quad \mathcal{R}(P, q, f) = \sum_{i=1}^{n} f(q_i) \triangle x_i \quad \text{(subordinate } q\text{)}$$

$$\text{and} \quad \mathcal{U}(P, f) = \sum_{i=1}^{n} M_i(P) \triangle x_i .$$

These are called **lower, Riemann and upper sums**, respectively.

Note that if M is the supremum of the set $f([a, b])$ and m is the infimum,

$$m(b-a) \le \mathcal{L}(P, f) \le \mathcal{R}(P, q, f) \le \mathcal{U}(P, f) \le M(b-a).$$

The point is that these three sums are uniformly bounded, independently from choice of P or q.

A partition S is said to **refine** (or to be a **refinement** of) the partition P if $P \subset S$. Any two partitions P and T have a common refinement: namely, $P \cup T$.

It is a fact that if S is a refinement of P,

$$\mathcal{L}(P, f) \le \mathcal{L}(S, f) \le \mathcal{U}(S, f) \le \mathcal{U}(P, f)$$

which is proven by examining any interval created by P and noting it is subdivided into (possibly) smaller subdivisions by S.

f is called **Riemann integrable** if there is a real number, denoted $\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{f}\, \mathbf{dx}$ and called **the (Riemann) integral of f**, so that for all $\varepsilon > 0$ there is a $\delta > 0$ for which mesh$(P) < \delta$ implies

$$\left| \int_a^b f\, dx - \mathcal{R}(P, q, f) \right| < \varepsilon \qquad \text{for all } q \text{ subordinate to } P.$$

Given the inequalities above this will happen exactly when for every $\varepsilon > 0$ there is a $\delta > 0$ for which mesh$(P) < \delta$ implies

$$\mathcal{U}(P, f) - \mathcal{L}(P, f) < \varepsilon.$$

It is extremely common and convenient, particularly when f is given by a formula, to write $\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{f}(\mathbf{x})\, \mathbf{dx}$ for $\int_a^b f\, dx$. An example of this usage for the function

f given by $f(x) = x^2 - \sin(x)$ is to write $\int_a^b x^2 - \sin(x)\, dx$ in place of $\int_a^b f\, dx$ or, equivalently, $\int_a^b y^2 - \sin(y)\, dy$ if the variable x is in use elsewhere.

It is also common to define, when $a < b$, the number $\int_{\mathbf{b}}^{\mathbf{a}} \mathbf{f\, dx}$ to be $-\int_a^b f\, dx$ and $\int_{\mathbf{a}}^{\mathbf{a}} \mathbf{f\, dx} = 0$.

14.1. ***Exercise.*** *(a) Suppose $c \in (a,b)$. Any partition P of $[a,b]$ can be used to create partitions R of $[a,c]$ and S of $[c,b]$ and so that neither* $\text{mesh}(R)$ *nor* $\text{mesh}(S)$ *exceeds* $\text{mesh}(P)$*. Conversely, given any partitions R of $[a,c]$ and S of $[c,b]$, the set $R \cup S$ is a partition of $[a,b]$ whose mesh is the greater of* $\text{mesh}(R)$ *and* $\text{mesh}(S)$*. Conclude that $\int_a^b f\, dx$ exists exactly when both $\int_a^c f\, dx$ and $\int_c^b f\, dx$ exist, and in that event $\int_a^b f\, dx = \int_a^c f\, dx + \int_c^b f\, dx$.*

(b) Show that if $\int_a^b f\, dx$ and $\int_a^b g\, dx$ exist and c is a real constant then $\int_a^b f + cg\, dx$ exists and $\int_a^b f + cg\, dx = \int_a^b f\, dx + c\int_a^b g\, dx$.

Note that if the function f happens to be continuous, for any real $\varepsilon > 0$ there is a $\delta > 0$ so that $\text{mesh}(P) \in (0, \delta)$ implies $0 \le M_i - m_i < \varepsilon$. But then

$$0 \le \boldsymbol{\mathcal{U}}(P,f) - \boldsymbol{\mathcal{L}}(P,f) < \sum_{i=0}^{n} \varepsilon \triangle x_i = \varepsilon(b-a).$$

This implies f is Riemann integrable.

A function $f \colon [a,b] \to \mathbb{R}$ is called **piecewise continuous** on $[a,b]$ if there are only a finite number of points in $[a,b]$ at which f fails to be continuous and each discontinuity is a “jump” discontinuity: that is, limits from the left exist at each point of $(a,b]$ and limits from the right exist at each point of $[a,b)$. Since there are at most a finite number of discontinuities, a piecewise continuous function defined on a closed and bounded interval is bounded.

14.2. ***Exercise.*** *Show that a piecewise continuous function $f \colon [a,b] \to \mathbb{R}$ is Riemann integrable. (hint: Break up $[a,b]$ into a finite number of intervals, including tiny intervals containing each point at which f fails to be continuous.)*

- If $f \colon [a,b] \to \mathbb{R}$ is piecewise continuous then f is Riemann integrable.

In the event that a function *is* Riemann integrable any sequence of partitions whose mesh converges to 0 can be used to calculate its integral as a limit.

Specifically, if f is integrable and for each $i \in \mathbb{N}$ the set P_i is a partition of $[a,b]$, and if $\text{mesh}(P_i)$ converges to 0 then

$$\lim_{i\to\infty} \boldsymbol{\mathcal{U}}(P_i,f) = \lim_{i\to\infty} \boldsymbol{\mathcal{L}}(P_i,f) = \int_a^b f\, dx.$$

On the other hand, it is also true (but considerably harder to prove) that if a sequence of partitions P_i of $[a,b]$ exists for which $\lim_{i\to\infty} \boldsymbol{\mathcal{U}}(P_i,f) = \lim_{i\to\infty} \boldsymbol{\mathcal{L}}(P_i,f)$ then the common limit is $\int_a^b f\, dx$.

We outline a proof of this fact below. It is this condition, involving a sequence converging to the integral, which will be used in the transfer process for a nonstandard definition of Riemann integrability.

Suppose $\lim_{i\to\infty} \boldsymbol{\mathcal{U}}(P_i, f) = \lim_{i\to\infty} \boldsymbol{\mathcal{L}}(P_i, f)$ and $\varepsilon > 0$. One of the partitions used to form this limit, which we denote R rather than P_i, satisfies

$$\boldsymbol{\mathcal{U}}(R, f) - \boldsymbol{\mathcal{L}}(R, f) < \frac{\varepsilon}{2}.$$

Suppose R has n distinct members $r_0, \ldots, r_n$. Let t be the least among the numbers $\triangle r_i$ for $i = 1, \ldots n$. Let δ be the lesser of t and $\frac{\varepsilon}{4n(M-m)}$, where we assume to avoid triviality that $M - m \neq 0$: that is, f is not the zero function.

We now suppose that S is *any* partition of $[a, b]$ whose mesh is less than δ. We will argue that

$$\boldsymbol{\mathcal{U}}(S, f) - \boldsymbol{\mathcal{L}}(S, f) < \varepsilon$$

and since ε was arbitrary, and so was S subject only to the restriction on its mesh, we can conclude that f is integrable.

Suppose S has distinct members $s_0, \ldots, s_k$. Some of the intervals $[s_i, s_{i+1}]$ contain one of members of R, but none of these intervals contain two members of R. That is because $\text{mesh}(S) < t$.

So the intervals $[s_i, s_{i+1}]$ can be broken into two groups: those that contain one member of R and those that contain none. In the latter case, $[s_i, s_{i+1}]$ is contained in the interior of one of the partition intervals from R.

Let J denote those subscripts i for which $[s_i, s_{i+1}]$ contains exactly one member of R. There can be no more than $2n$ members of J. Let K denote the rest of the possible subscripts of members of S.

For members of K, the supremum $M_i(S)$ of the value of f on $[s_i, s_{i+1}]$ cannot exceed the supremum of f on the interval of the partition R within which it resides. The infimum $m_i(S)$ of f on $[s_i, s_{i+1}]$ is not less than the infimum of f on the interval of the partition R within which it resides. It follows that

$$\sum_{i\in K} M_i(S)\triangle s_i - \sum_{i\in K} m_i(S)\triangle s_i \leq \boldsymbol{\mathcal{U}}(P, f) - \boldsymbol{\mathcal{L}}(P, f) < \frac{\varepsilon}{2}.$$

On the other hand, for those i in J, $M_i(S) - m_i(S) \leq M - m$ and there are at most $2n$ such terms, and no $\triangle s_i$ exceeds $\frac{\varepsilon}{4n(M-m)}$. So

$$\sum_{i\in J} M_i(S)\triangle s_i - \sum_{i\in J} m_i(S)\triangle s_i \leq 2n(M - m)\frac{\varepsilon}{4n(M - m)} < \frac{\varepsilon}{2}.$$

Since

$$\sum_{i=1}^{k} M_i(S)\triangle s_i - \sum_{i\in J} m_i(S)\triangle s_i = \sum_{i\in J}(M_i(S) - m_i(S))\triangle s_i + \sum_{i\in K}(M_i(S) - m_i(S))\triangle s_i$$

the result is proven.

- Suppose $f\colon [a, b] \to \mathbb{R}$ is bounded and P_i for $i \in \mathbb{N}$ is a sequence of partitions of $[a, b]$ with $\lim_{i\to\infty} \text{mesh}(P_i) = 0$.

  f is Riemann integrable exactly when

  $$\lim_{i\to\infty} \boldsymbol{\mathcal{U}}(P_i, f) = \lim_{i\to\infty} \boldsymbol{\mathcal{L}}(P_i, f),$$

  and in that case the common limit is $\int_a^b f\, dx$.

15. The Infinitesimal Approach to Integration

In the last section we found that continuous (or piecewise continuous) functions on a closed interval are Riemann integrable, and any sequence of partitions with mesh converging to 0 can be used to calculate the value of the integral of any integrable function.

We use the notation of the last section, and select a fixed sequence of partitions.

For each $n \geq 1$ there is an integer k so that $a + \frac{k-1}{n} < b$ but $a + \frac{k}{n} \geq b$. We define P_n to be the partition with members $x_0(n), \ldots, x_k(n)$ where $x_i(n) = a + \frac{i}{n}$ for $i = 0, \ldots, k-1$ and $x_k(n) = b$.

Note that the mesh of this partition is $\frac{1}{n}$ or $b - a$, whichever is smaller.

If f is a real valued function defined on $[a, b]$, the sequences given by

$$\mathcal{U}_n(f) = \sum_{i=1}^{k} M_i(P_n) \triangle x_i(n) \quad \text{and} \quad \mathcal{L}_n(f) = \sum_{i=1}^{k} m_i(P_n) \triangle x_i(n)$$

both converge to a common limit exactly when f is Riemann integrable. So f is integrable exactly when there is a real number L for which

$$\mathcal{U}_n(f) \in \text{halo}(L) \quad \text{and} \quad \mathcal{L}_n(f) \in \text{halo}(L)$$

for any unlimited hyperinteger n. In that case, $L = \int_a^b f\,dx$.

From this point we confine consideration to integrals of piecewise continuous functions defined on a closed interval.

Define, for each $x \in (a, b]$, the function $A(x) = \int_a^x f\,dx$, and define $A(a) = 0$.

We are going to show that A is continuous on $[a, b]$ and differentiable at each point x of continuity of f in (a, b), and its derivative is $f(x)$ at these points. The usual integration theorems follow.

Suppose x and $x + h$ are in $[a, b]$ for $h \in (0, b - x)$. So:

$$A(x + h) - A(x) = \int_x^{x+h} f\,dx.$$

Since f is bounded above by M and below by m, there are real numbers $m_h \in [m, M]$, the infimum of the values of f on $[x, x+h]$ and $M_h \in [m, M]$, the supremum of the values of f on $[x, x + h]$. So for all real $h \in (0, b - x)$

$$h\, m_h \leq A(x + h) - A(x) \leq h\, M_h.$$

By transfer, this statement holds for positive infinitesimal h as well. The transferred statement is

$$(\forall h \in (0, b - x)_*)(\exists m_h \in [m, M]_*)(\exists M_h \in [m, M]_*) \left(m_h \leq \frac{A(x + h) - A(x)}{h} \leq M_h \right).$$

This implies that A is continuous on $[a, b)$.

If x is a point of continuity of f and if h is infinitesimal the outer two numbers are in $\text{halo}(f(x))$ so $\text{shad}\left(\frac{A(x+h)-A(x)}{h}\right) = f(x)$.

A similar argument holds for negative infinitesimal h. We conclude:

- If $f\colon [a,b] \to \mathbb{R}$ is piecewise continuous and A is the function defined by $A(x) = \int_a^x f\,dx$ then A is continuous on $[a,b]$ and differentiable at each point of continuity of f in (a,b). If $x \in (a,b)$ is a point of continuity of f then $A'(x) = f(x)$.

Corollaries are the Fundamental Theorem of Calculus, Integration by Substitution and Integration by Parts:

- If $f\colon [a,b] \to \mathbb{R}$ is piecewise continuous and G is any function continuous on $[a,b]$ and for which $G'(x)$ exists and equals $f(x)$ for all points x of continuity of f in (a,b) then

(Fundamental Theorem of Calculus) $$\int_a^b f\,dx = G(b) - G(a).$$

- If $f \circ g$ is piecewise continuous on $[a,b]$ and g is continuous on $[a,b]$ and g is appropriately differentiable on (a,b) then

(Integration by Substitution) $$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(x)\,dx.$$

- If u and v are continuous on $[a,b]$ and appropriately differentiable on (a,b) then

(Integration by Parts)

$$u(b)v(b) - u(a)v(a) = \int_a^b u(x)v'(x)\,dx + \int_a^b v(x)u'(x)\,dx.$$

15.1. ***Exercise.*** *(a) Prove the Fundamental Theorem of Calculus in the generality specified above.*

(b) Justify Integration by Substitution, and explain what is meant by "appropriately differentiable" in the statement of that result.

(c) Justify Integration by Parts, and explain what is meant by "appropriately differentiable" in the statement of that result.

16. An Example of Euler, Revisited

16.1. ***Exercise.*** *(a) Define the function* $\ln\colon (0,\infty) \to \mathbb{R}$ *by*

$$\ln(x) = \int_1^x \frac{1}{t}\,dt.$$

Show that this function obeys the standard functional patterns of a logarithm function and find the derivative of $\ln$.

(b) Since the function defined above has positive derivative and is defined on an interval it is one-to-one and has a differentiable inverse $\exp\colon \mathbb{R} \to (0,\infty)$.

Show that $\exp' = \exp$ *and that* $\exp$ *obeys the standard functional patterns for an exponential function.*

For any positive real x and sufficiently large integer n,

$$\frac{x}{n} \geq \int_1^{1+\frac{x}{n}} \frac{1}{t}\,dt = \ln\Big(1+\frac{x}{n}\Big) \geq \frac{1}{1+\frac{x}{n}} \cdot \frac{x}{n}.$$

So for any sufficiently large integer n

$$x \geq \ln\Big(\Big(1+\frac{x}{n}\Big)^n\Big) \geq \frac{1}{1+\frac{x}{n}} \cdot x.$$

So for any unlimited hyperinteger n

$$\ln\Big(\Big(1+\frac{x}{n}\Big)^n\Big) \in \text{halo}(x).$$

We conclude that for positive x

$$\exp(x) = \lim_{n\to\infty}\Big(1+\frac{x}{n}\Big)^n.$$

The same is true for negative x by a similar argument. So $\exp(x)$ is the shadow of $\left(1+\frac{x}{n}\right)^n$ for any unlimited hypernatural n.

Expanding $\left(1+\frac{x}{n}\right)^n$ for positive integer n yields

$$\begin{aligned}\Big(1+\frac{x}{n}\Big)^n =& \sum_{k=0}^{n}\frac{n!}{(n-k)!\,k!}\cdot\frac{x^k}{n^k} = \sum_{k=0}^{n}\frac{n!}{(n-k)!\,n^k}\cdot\frac{x^k}{k!}\\ =&1+\frac{n}{n-1}\cdot\frac{x}{1}+\frac{n(n-1)}{n^2}\cdot\frac{x^2}{2!}+\frac{n(n-1)(n-2)}{n^3}\cdot\frac{x^3}{3!}\\ &+\cdots+\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\frac{x^k}{k!}+\cdots\end{aligned}$$

Examining the term in the expansion corresponding to a specific integer power k of x, we see that for unlimited n

$$\frac{n(n-1)\cdots(n-k+1)}{n^k}\cdot\frac{x^k}{k!} \in \text{halo}\Big(\frac{x^k}{k!}\Big),$$

which leads us (and led Euler) to speculate that

$$\exp(x) = \sum_{k=0}^{\infty}\frac{x^k}{k!}.$$

Let s denote the sequence $s_n = \sum_{i=0}^{n}\frac{x^i}{i!}$ of partial sums for this series. By the ratio test, the series given by s converges absolutely, so it converges to some number L. In particular, for any $\varepsilon > 0$ there is an integer N so that $n \geq N$ implies $|L - s_n| \leq \sum_{i=n}^{\infty}\frac{|x|^i}{i!} < \varepsilon$.

Let b denote the sequence given by $b_n = \left(1+\frac{x}{n}\right)^n$. For each n define $a_{n,k}$ to be the sum of terms in the expansion of b_n corresponding to powers of x not exceeding k and let $r_{n,k} = b_n - a_{n,k}$. $r_{n,k}$ is the sum (if any) of higher powers of x in the expansion of b_n.

Note that for natural numbers n, k with $0 \leq k < n$

$$\left| \frac{n(n-1)\cdots(n-k+1)}{n^k} \cdot \frac{x^k}{k!} \right| \leq \frac{|x|^k}{k!} \quad \text{and so} \quad |r_{n,k}| \leq \sum_{i=k+1}^{\infty} \frac{|x|^i}{i!}.$$

Also, for each integer k and $\varepsilon > 0$ there is an integer N for which $n \geq N$ implies $|s_k - a_{n,k}| < \varepsilon$. This means that for each unlimited k there is unlimited $n > k$ for which $|s_k - a_{n,k}|$ is infinitesimal.

So for hypernaturals n, k with $0 \leq k < n$

$$\begin{aligned} |L - \exp(x)| =& |L - s_k + s_k - a_{n,k} - r_{n,k} + b_n - \exp(x)| \\ \leq& |L - s_k| + |s_k - a_{n,k}| + |r_{n,k}| + |b_n - \exp(x)|. \end{aligned}$$

If k is unlimited so is n, which means that the first and fourth terms in the last line are infinitesimal. This forces the third term to be infinitesimal too, since it cannot exceed the first term, regardless of n. Now choose $n > k$ to make the second term infinitesimal.

We have proven:

- For any real x,

$$\exp(x) = \lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

References

[1] Goldblatt, R., *Lectures on the Hyperreals*. Springer-Verlag, New York, NY, 1998
[2] Henle, J. and Kleinberg, E., *Infinitesimal Calculus*. Dover Publications, Mineola, NY, 2003
[3] Loeb, P. and Wolff, M., *Nonstandard Aalysis for the Working Mathematician*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000
[4] Roberts, A. M., *Nonstandard Analysis*. Dover Publications, Mineola, NY, 2003
[5] Robinson, A., *Non-Standard Analysis (Rev. Ed.)*. Princeton University Press, Princeton, NJ, 1996

Index