Infinitesimal analysis

2000 *Mathematics Subject Classification.* Primary 26E35; 97A20, 97C30

Abstract. Infinitesimal analysis. Course notes for 88-503. See
https://u.math.biu.ac.il/~katzmik/88-503.html

Contents

Part 1

Basic Infinitesimal Analysis

CHAPTER 1

Infinitesimal calculus

1. Course site https://u.math.biu.ac.il/~katzmik/88-503.html
2. Final exam 90%, homework 10%.
3. Textbook: Goldblatt, *Lectures on the Hyperreals* [**6**].

1.1. From natural numbers to real numbers

Leibniz was the co-inventor of the calculus; see https://u.math.biu.ac.il/~katzmik/leibniz.html

Leibniz used infinitesimals dx, dy, etc., to develop notions such as the differential quotient

$$\frac{dy}{dx}$$

and integral

$$\int f(x)\,dx.$$

In non-infinitesimal approaches to the calculus, one starts with the natural numbers $\mathbb{N}$ and develops the sequence of successive extensions

$$\mathbb{N} \hookrightarrow \mathbb{Z} \hookrightarrow \mathbb{Q} \hookrightarrow \mathbb{R}. \tag{1.1.1}$$

All of these number systems are equipped with an order relation $<$, satisfying the following trichotomy:

For any a, b, either $a < b$ or $a = b$ or $a > b$.

Each successive extension as in (1.1.1) enables us to express additional mathematical facts and phenomena:

(1) $\mathbb{Z}$ enables us to subtract any pair of numbers, something that was not possible in $\mathbb{N}$.
(2) $\mathbb{Q}$ enables us to speak of ratios of arbitrary numbers (so long as the denominator is nonzero), something that was not possible in $\mathbb{Z}$.
(3) $\mathbb{R}$ enables us to speak of the length of the diagonal of the unit square and the length of the unit circle, something that was not possible in $\mathbb{Q}$.

Here $\mathbb{R}$ is the unique complete Archimedean ordered field. In Section 1.2, we will explain why one needs to go beyond $\mathbb{R}$.

1.2. A new ordered extension, microscopes, and telescopes

In infinitesimal calculus, one exploits a further extension to an ordered field denoted ${}^*\mathbb{R}$ called the *hyperreal numbers*:

$$\mathbb{N} \hookrightarrow \mathbb{Z} \hookrightarrow \mathbb{Q} \hookrightarrow \mathbb{R} \hookrightarrow {}^*\mathbb{R}.$$

These enable more convenient definitions of the key notions of the calculus, such as derivative, integral, continuity, limit, etc.

Definition 1.2.1. A hyperreal number ϵ is called *infinitesimal* if

$$-a < \epsilon < a$$

for every positive real number a.

Definition 1.2.2 (Finite and infinite numbers). We define finite and infinite numbers.

(1) If $\epsilon > 0$ is infinitesimal, then $H = \frac{1}{\epsilon}$ is *positive infinite*.[1]
(2) If $\epsilon < 0$ is infinitesimal, then $H = \frac{1}{\epsilon}$ is *negative infinite*.
(3) Hyperreal numbers which are not infinite numbers are called *finite*.

We view infinitesimals with a Keisler microscope (see Figure 1.2). We view infinite numbers with a Keisler telescope (see Figure 1.2).

Definition 1.2.3 (Relation of infinite proximity). Numbers r and s are called *infinitely close* if the difference $r - s$ is infinitesimal.

1.3. Extension Principle

Let us now formulate the extension principle.

Theorem 1.3.1 (Extension Principle). *This is in three parts.*

(1) *The real numbes are properly contained in the hyperreal numbers, and the order relation $<$ for the real numbers extends to the order relation for the hyperreal numbers.*
(2) *There is a hyperreal number that is greater than zero but less than any positive real number.*
(3) *For every real function f of one variable or more variables, we are given a corresponding hyperreal function ${}^*f$ of the same number of variables. Such an ${}^*f$ is called the* natural extension *of f.*

Let us give a more detailed analysis of infinitesimals.

[1] In the long run, it turns out to be counterproductive to use the term "infinite." It is often replaced by *unlimited*, especially in the context of axiomatic approaches to infinitesimal analysis; see Part 2.

DEFINITION 1.3.2. A hyperreal number b is said to be

- a *positive infinitesimal* if b is positive but less than every positive real number.
- a *negative infinitesimal* if b is negative but greater than every negative real number.
- *infinitesimal* if b is either positive infinitesimal, negative infinitesimal, or 0.

EXAMPLE 1.3.3 (Example of application of Extension Principle). Since $+$ is a real function of two variables, its natural extension $^*+$ is a hyperreal function of two variables. Similar remarks apply to the product operation.

We will usually drop the asterisks on functions when this does not lead to confusion.

1.4. Introduction to the Transfer Principle

This section contains a preliminary discussion of the Transfer Principle. More detailed presentations appear in Sections 3.6 and 4.11.

THEOREM 1.4.1 (Transfer Principle). *Every* real statement *that holds for one or more particular real functions, hold for the hyperreal extensions of these functions.*

EXAMPLE 1.4.2. [Examples of real statements] Here are seven examples.

(1) closure law for addition: for any x and y, $x + y$ is defined.
(2) commutative law for addition: $x + y = y + x$.
(3) a rule for order: if $0 < x < y$ then $0 < \frac{1}{y} < \frac{1}{x}$.
(4) division by zero is not allowed: $\frac{x}{0}$ is undefined.
(5) an algebraic identity: $(x - y)^2 = x^2 - 2xy + y^2$.
(6) a trigonometric identity: $\sin^2 x + \cos^2 x = 1$.
(7) a rule for logarithms: if $x > 0$ and $y > 0$ then $\log_{10}(xy) = \log_{10} x + \log_{10} y$.

REMARK 1.4.3. The kind of statements the transfer principle applies to will be treated in more detail in Chapter 5.2.

One can use the transfer principle to define hyperreal functions as follows.

EXAMPLE 1.4.4 (Using transfer to define functions). Here are three examples.

(1) The *square root* function is defined by the real statement

$$y = \sqrt{x} \text{ if and only if } y^2 = x \text{ and } y \geq 0.$$

By transfer, the square root is defined for all nonnegative hyperreal x.

(2) the *absolute value* function is defined by the real statement

$$y = |x| \text{ if and only if } y = \sqrt{x^2}.$$

(3) the common log is defined by the real statement

$$y = \log_{10} x \text{ if and only if } 10^y = x.$$

DEFINITION 1.4.5. A hyperreal number b is said to be

(a) *finite* if b is between two real numbers.
(b) *positive infinite* if b is greater than every real number.
(c) *negative infinite* if b is less than every real number.

We will use the transfer principle to prove the following proposition.

PROPOSITION 1.4.6. *If ϵ is a positive infinitesimal, then $\frac{1}{\epsilon}$ is positive infinite.*

PROOF. Let r be any positive real. Since ϵ is infinitesimal, we have $0 < \epsilon < \frac{1}{r}$. Applying the transfer principle as in Example 1.4.2 item (3), we obtain $0 < r < \frac{1}{\epsilon}$. This is true for each positive real number r. It follows that $\frac{1}{\epsilon}$ is positive infinite. $\square$

1.5. Three orders of magnitude for hyperreal numbers

DEFINITION 1.5.1. A hyperreal number is *appreciable*[2] if it is finite but not infinitesimal.

REMARK 1.5.2. We have defined three orders of magnitude for hyperreal numbers: infinitesimal, appreciable, infinite.

THEOREM 1.5.3. *This is in four parts.*

(1) *Every hyperreal number which is between two infinitesimals, is infinitesimal.*
(2) *Every hyperreal number which is between two finite hyperreal numbers, is finite.*
(3) *Every hyperreal number which is greater than some positive infinite number, is positive infinite.*
(4) *Every hyperreal number which is less than some negative infinite number, is negative infinite.*

[2] Mashma'uti

DEFINITION 1.5.4. Two hyperreal numbers b and c are said to be infinite close to each other: written

$$b \approx c$$

if their difference $b - c$ is infinitesimal. The relation $\approx$ is called the relation of infinite proximity.

If b and c are real and $b \approx c$ then $b = c$.

THEOREM 1.5.5. (i) *$a \approx a$;*
(ii) *if $a \approx b$ then $b \approx a$;*
(iii) *if $a \approx b$ and $b \approx c$ then $a \approx c$.*

THEOREM 1.5.6. *Assume $a \approx b$. Then*
(1) *If a is infinitesimal then so is b.*
(2) *If a is appreciable then so is b.*
(3) *If a is infinite then so is b.*

1.6. Standard part principle, shadow

THEOREM 1.6.1 (Standard Part Principle). *Every finite hyperreal number is infinitely close to exactly one real number.*

PROOF. (Optional for those familiar with Dedekind cuts) Let b be a finite hyperreal. Using the order relation of ${}^*\mathbb{R}$, the number b defines a Dedekind cut on the rationals $\mathbb{Q} \subseteq {}^*\mathbb{R}$. Let r be the real number coresponding to such a Dedekind cut. Then r is the standard part of b. □

DEFINITION 1.6.2 (Shadow). Let b be a finite hyperreal number. The standard part, or shadow,

$$\mathbf{sh}(b)$$

of b is the real number which is infinitely close to b.

THEOREM 1.6.3. *Eight rules for working with standard part:*
(1) $\mathbf{sh}(-a) = -\,\mathbf{sh}(a)$.
(2) $\mathbf{sh}(a+b) = \mathbf{sh}(a) + \mathbf{sh}(b)$.
(3) $\mathbf{sh}(a-b) = \mathbf{sh}(a) - \mathbf{sh}(b)$.
(4) $\mathbf{sh}(ab) = \mathbf{sh}(a)\,\mathbf{sh}(b)$.
(5) *If $\mathbf{sh}(b) \neq 0$ then $\mathbf{sh}\,\frac{a}{b} = \frac{\mathbf{sh}\,a}{\mathbf{sh}\,b}$.*
(6) $\mathbf{sh}(a^n) = (\mathbf{sh}(a))^n$.
(7) *If $a \geq 0$ then $\mathbf{sh}(\sqrt[n]{a}) = \sqrt[n]{\mathbf{sh}(a)}$.*
(8) *If $a \leq b$ then $\mathbf{sh}(a) \leq \mathbf{sh}(b)$.*

PROOF OF ITEM (4). We write $a = r + \epsilon$ and $b = s + \delta$. Then $ab = rs + r\epsilon + s\epsilon + \epsilon\delta \approx rs$. Hence $\mathbf{sh}(ab) = rs = \mathbf{sh}(a)\,\mathbf{sh}(b)$. □

1.7. Infinitesimal increments, slope

Definition 1.7.1. We will use Δx, Δy for infinitesimal increments.[3]

Definition 1.7.2. Let f be a function and a a real number. A real number s is said to be the *slope* of f at a if $s = \mathbf{sh}\left(\frac{f(a+\Delta x)-f(a)}{\Delta x}\right)$ for every nonzero infinitesimal Δx.

Definition 1.7.3. Let f be a real function of one real variable. The derivative of f is the new function f' whose value at x is the slope of f at x. In symbols,

$$f'(x) = \mathbf{sh}\left(\frac{f(x+\Delta x)-f(x)}{\Delta x}\right).$$

1.8. Dependent and independent variables

Definition 1.8.1. In equation $y = f(x)$, we say that y is the *dependent variable* and x is the *independent variable*.

When $y = f(x)$, we introduce a new independent variable Δx and a new dependent variable Δy, by equation $\Delta y = f(x+\Delta x) - f(x)$. Then Δy is called the y-increment. The derivative can be expressed as $\mathbf{sh}\left(\frac{\Delta y}{\Delta x}\right)$.

Example 1.8.2. When calculating the slope of $y = x^2$ at a point c, after a series of algebraic manipulations we obtain that $\frac{\Delta y}{\Delta x} = 2c + \Delta x$. The slope is then obtained by discarding the remaining term Δx. Then whenever Δx is infinitesimal, we obtain $\mathbf{sh}(2x + \Delta x) = 2x$.

Example 1.8.3. Find $f'(x)$ given $f(x) = \sqrt{x}$ in domain $x \geq 0$.
Case 1: $x < 0$ since $\sqrt{x}$ is undefined, $f'(x)$ does not exist.
Case 2: $x = 0$. When $\Delta x < 0$, Δy is undefined. Hence $f'(0)$ does not exist.
Case 3. $x > 0$. If $y = \sqrt{x}$ then we obtain a ratio

$$\begin{aligned}\frac{\Delta y}{\Delta x} &= \frac{\sqrt{x+\Delta x}-\sqrt{x}}{\Delta x} = \frac{(\sqrt{x+\Delta x}-\sqrt{x})(\sqrt{x+\Delta x}+\sqrt{x})}{\Delta x(\sqrt{x+\Delta x}+\sqrt{x})}\\ &= \frac{1}{\sqrt{x+\Delta x}+\sqrt{x}}.\end{aligned}$$

[3] Hefresh

Applying the standard part, we obtain, using the rules of Theorem 1.6.3:

$$\begin{aligned}
\mathbf{sh}\left(\frac{\Delta y}{\Delta x}\right) &= \mathbf{sh}\left(\frac{1}{\sqrt{x+\Delta x}+\sqrt{x}}\right) \\
&= \frac{1}{\mathbf{sh}(\sqrt{x+\Delta x}+\sqrt{x})} \\
&= \frac{1}{\mathbf{sh}\sqrt{x+\Delta x}+\mathbf{sh}\sqrt{x}} = \frac{1}{2\sqrt{x}}.
\end{aligned}$$

Thus when $x > 0$, $f'(x) = \frac{1}{2\sqrt{x}}$. The domain of f' is the set $\{x > 0\}$.

THEOREM 1.8.4 (Increment Theorem). *Let $y = f(x)$. Suppose $f'(x)$ exists at a certain point x, and Δx is infinitesimal. Then Δy is infinitesimal, and there exists an infinitesimal ϵ which depends on x and Δx such that*

$$\Delta y = f'(x)\Delta x + \epsilon\Delta x.$$

PROOF. If $\Delta x = 0$ then $\Delta y = 0$ and we set $\epsilon = 0$. If $\Delta x \neq 0$, then we obtain a relation of infinite proximity

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

Hence for some infinitesimal ϵ we obtain $\frac{\Delta y}{\Delta x} = f'(x) + \epsilon$, or equivalently $\Delta y = f'(x)\Delta x + \epsilon\,\Delta x$. □

1.9. Differentials *dx, dy*

Independent and dependent variables were introduced in Section 1.8. We now introduce a new dependent variable dy.

DEFINITION 1.9.1. The *differential* of y, denoted dy, is the dependent variable defined by

$$dy = f'(x)\Delta x.$$

To keep the notation uniform, we denote Δx by dx. We summarize the notation introduced so far.

DEFINITION 1.9.2. Let $y = f(x)$. The differential of x is the independent variable $dx = \Delta x$. The differential of y is the dependent variable $dy = f'(x)dx$.

When $dx \neq 0$, one can write $f'(x) = \frac{dy}{dx}$. The increment theorem can then be reformulated as follows.

COROLLARY 1.9.3 (Reformulation of the Increment Theorem). *Here is a short form of the theorem:*

$$\Delta y = dy + \epsilon\,dx.$$

The "d" notation can also be applied to *terms* as follows. Consider the term $\tau(x)$ given by a specific function $\tau(x) = f(x)$. Then we will write

$$d(\tau(x)) = f'(x)dx.$$

EXAMPLE 1.9.4. Some examples of d applied to terms:
(1) $d(x^3) = 3x^2 dx$.
(2) $d(\ln x) = \frac{dx}{x}$.

THEOREM 1.9.5 (Sum Rule). *Suppose u and v depend on an independent variable x. Then for any value of x where du and dv exist,*

$$\frac{d(u+v)}{dx} = \frac{du}{dx} + \frac{dv}{dx},$$

or equivalently $d(u+v) = du + dv$.

PROOF. Let $y = u + v$, let $\Delta x \neq 0$ be infinitesimal, and compute the corresponding Δy:

$$\Delta y = (u + \Delta u) + (v + \Delta v) - y = \Delta u + \Delta v.$$

Dividing by Δx, we obtain $\frac{\Delta y}{\Delta x} = \frac{\Delta u + \Delta v}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x}$. Applying standard part to the equality $\frac{\Delta y}{\Delta x} = \frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x}$ we obtain

$$\mathbf{sh}\left(\frac{\Delta y}{\Delta x}\right) = \mathbf{sh}\left(\frac{\Delta u}{\Delta x} + \frac{\Delta v}{\Delta x}\right) = \mathbf{sh}\left(\frac{\Delta u}{\Delta x}\right) + \mathbf{sh}\left(\frac{\Delta v}{\Delta x}\right).$$

It follows that $\frac{dy}{dx} = \frac{du}{dx} + \frac{dv}{dx}$, as required. □

1.10. Leibniz rule

THEOREM 1.10.1 (Leibniz Rule). *Suppose u and v depend on x. Then for any value of x where du and dv exist,*

$$\frac{d(uv)}{dx} = u\,\frac{du}{dx} + v\,\frac{du}{dx},$$

or equivalently $d(uv) = u\,dv + v\,du$.

PROOF. Let $y = uv$. Then

$$\begin{aligned}\mathbf{sh}\left(\frac{\Delta y}{\Delta x}\right) &= \mathbf{sh}\left(u\frac{\Delta v}{\Delta x} + v\frac{\Delta u}{\Delta x} + \Delta u\frac{\Delta v}{\Delta x}\right)\\ &= u\,\mathbf{sh}\left(\frac{\Delta v}{\Delta x}\right) + v\,\mathbf{sh}\left(\frac{\Delta u}{\Delta x}\right) + 0\cdot\mathbf{sh}\left(\frac{\Delta v}{\Delta x}\right)\end{aligned}$$

as required. □

A similar argument with standard part proves the quotient rule.

1.11. Inverse function rule

DEFINITION 1.11.1. Two functions f and g are called *inverse functions* if the two equations

$$y = f(x), \quad x = g(y) \tag{1.11.1}$$

have the same graphs in the (x, y)-plane.

THEOREM 1.11.2 (Inverse Function Rule). *Suppose f and g are inverse functions in the sense of* (1.11.1). *If both derivatives $f'(x)$ and $g'(y)$ are nonzero then*

$$f'(x) = \frac{1}{g'(y)};$$

equivalently, $\frac{dy}{dx} = \frac{1}{dx/dy}$.

PROOF. Let Δx be a nonzero infinitesimal. Let Δy be the corresponding change in y. Then Δy is also infinitesimal because $f'(x)$ exists and $\Delta y = \Delta x(f'(x) + \epsilon)$. Since $f'(x)$ is nonzero, $f'(x)$ is appreciable and therefore so is $f'(x) + \epsilon$. Hence Δy is nonzero.

By the rule of standard parts,

$$f'(x)g'(y) = \mathbf{sh}\left(\frac{\Delta y}{\Delta x}\right) \cdot \mathbf{sh}\left(\frac{\Delta x}{\Delta y}\right) = \mathbf{sh}\left(\frac{\Delta y}{\Delta x} \cdot \frac{\Delta x}{\Delta y}\right) = \mathbf{sh}(1) = 1,$$

as required. □

See further on infinitesimal analysis in Chapter 6. In the next Chapter, we will present a construction of hyperreal fields.

CHAPTER 2

The ultrapower construction of the hyperreals

Before discussing the construction of the hyperreals, we will present a motivational discussion of comparing sequences in Section 2.2. First we review a construction of the real numbers via Cauchy sequences of rational numbers.

2.1. Equivalence classes of Cauchy sequences

A sequence $r = \langle r_1, r_2, r_3, \ldots \rangle$ is called *Cauchy* if for every $\epsilon > 0$ there exists an index n such that if $m > n$ then $|x_m - x_n| < \epsilon$. Cauchy sequences $r = \langle r_1, r_2, r_3, \ldots \rangle$ and $s = \langle s_1, s_2, s_3, \ldots \rangle$ are said to be Cantor-equivalent if $r_n - s_n$ tends to 0 as n tends to infinity, and then one writes $r \sim s$. Now let $CS(\mathbb{Q})$ be the space of all Cauchy sequences of rational numbers. Since the relation $\sim$ is an equivalence relation, we can form the quotient space $CS(\mathbb{Q})/\sim$. A standard result in real analysis is that this quotient space is isomorphic to the field of real numbers $\mathbb{R}$.

2.2. What is a large set?

Let $r = \langle r_1, r_2, r_3, \ldots \rangle$ and $s = \langle s_1, s_2, s_3, \ldots \rangle$ be real-valued sequences. We are going to say that r and s are equivalent if they agree at a "large" number of ranks of the index, i.e., if their *agreement set*

$$E_{rs} = \{n : r_n = s_n\}$$

is large in some sense that is to be determined. Whatever "large" means, there are some properties we will want such a notion to have:

(1) $\mathbb{N} = \{1, 2, 3, \ldots\}$ must be large, in order to ensure that any sequence will be equivalent to itself.

(2) Equivalence is to be a transitive relation, so if E_{rs} and E_{st} are large, then E_{rt} must be large. Since

$$E_{rs} \cap E_{st} \subseteq E_{rt},$$

this suggests the following requirement:

If A and B are large sets, and $A \cap B \subseteq C$, then C is large.

In particular, this entails that if A and B are large, then so is their intersection $A \cap B$, while if A is large, then so is any of its supersets $C \supseteq A$.

(3) The empty set $\varnothing$ is *not* large; otherwise by the previous requirement all subsets of $\mathbb{N}$ would be large, and so all sequences would be equivalent.

There are natural situations in which all three requirements are fulfilled.

EXAMPLE 2.2.1. One such situation is when a set $A \subseteq \mathbb{N}$ is declared to be large if it is *cofinite*, i.e., its complement $\mathbb{N} - A$ is finite. This means that A contains "almost all" members of $\mathbb{N}$. Although this is a plausible notion of largeness, it is not adequate to our goals.

The number system we are constructing is to be linearly ordered, and a natural way to achieve this, in terms of our general approach, is to take the equivalence class of a sequence r to be less than that of s if the set

$$L_{rs} = \{n : r_n < s_n\}$$

is large. But consider the sequences

$$r = \langle 1, 0, 1, 0, 1, 0, \ldots \rangle$$

and

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

Their agreement set E_{rs} is empty, so they determine distinct equivalence classes, one of which should be less than the other. But L_{rs} (the even numbers) is the complement of L_{sr} (the odds), so both are infinite and neither is cofinite.

It emerges that our definition of largeness should require the following condition:

For any subset A of $\mathbb{N}$, one of A and $\mathbb{N} - A$ is large.

The other requirements imply that A and $\mathbb{N} - A$ cannot both be large, or else $A \cap (\mathbb{N} - A) = \emptyset$ would be. Thus the large sets are precisely the complements of the ones that are not large. Either the even numbers form a large set or the odd ones do, but they cannot both do so, so which is it to be?

Can there in fact be such a notion of largeness, and if so, how do we show it? The answer is provided in terms of the notion of a *filter* studied in Section 2.3.

2.3. Filters and ultrafilters

The properties discussed in Section 2.2 motivate the following definition. Let I be a nonempty set. The power set of I is the set

$$\mathcal{P}(I) = \{A : A \subseteq I\}$$

of all subsets of I.

Definition 2.3.1. A *filter* on I is a nonempty collection $F \subseteq \mathcal{P}(I)$ of subsets of I satisfying the following two axioms:

(1) Intersections: if $A, B \in F$, then $A \cap B \in F$.
(2) Supersets: if $A \in F$ and $A \subseteq B \subseteq I$, then $B \in F$.

Thus to show $B \in F$, it suffices to show

$$A_1 \cap \ldots \cap A_n \subseteq B,$$

for some n and some $A_1, \ldots, A_n \in F$.

Proposition 2.3.2. *A filter F contains the empty set $\varnothing$ if and only if $F = \mathcal{P}(I)$.*

Proof. If F contains the empty set, by the superset property F must contain every subset of I.

If F is the power set $\mathcal{P}(I)$, then it contains all subsets of I and in particular the empty set. $\square$

Definition 2.3.3. A filter F is *proper* if $\varnothing \notin F$.

Every filter contains I, and in fact the collection $\{I\}$ is the smallest filter on I. Recall that if $A \subseteq I$ then $A^c = I - A$.

Definition 2.3.4. An *ultrafilter* is a proper filter that satisfies

for any $A \subseteq I$, either $A \in F$ or $A^c \in F$.

Ultrafilters will play a key role in the construction of fields of hyperreals; see Section 2.9.

2.4. Examples of filters

Definition 2.4.1 (Principal ultrafilter). Let $i \in I$. Then

$$F^i = \{A \subseteq I : i \in A\}$$

is an ultrafilter, called the *principal ultrafilter* generated by i.

Proposition 2.4.2. *If the set I is finite, then every ultrafilter on I is of the form F^i for some $i \in I$, and so is principal.*

PROOF. If I is finite then its power set is finite. By the intersection property, the intersection of all members of F is a member $A \in F$. If A has more than one element, then A is not maximal. Thus A must contain a single element $a \in I$. It follows that $F = F^i$. $\square$

DEFINITION 2.4.3 (Fréchet filter). The filter

$$F^{co} = \{A \subseteq I : I - A \text{ is finite}\}$$

is the cofinite, or Fréchet, filter on I.

The filter F^{co} is proper iff I is infinite.

PROPOSITION 2.4.4. *If I is infinite, then F^{co} is not an ultrafilter.*

PROOF. To fix ideas, we assume that I includes $\mathbb{N}$. Let $A \subseteq \mathbb{N} \subseteq I$ be the set of all even natural numbers. Then neither A nor A^c is a member of F^{co}. $\square$

DEFINITION 2.4.5 (Union of filters). Suppose $\{F_x : x \in X\}$ is a collection of filters on I that is linearly ordered by set inclusion, in the sense that either $F_x \subseteq F_y$ or $F_y \subseteq F_x$ for any $x, y \in X$. Then the union

$$\bigcup_{x \in X} F_x = \{A \subseteq I : \ (\exists x \in X)\, A \in F_x\}$$

is a filter on I.

2.5. Facts about filters

We list some useful facts about filters.

PROPOSITION 2.5.1. *If a collection $F \subseteq P(I)$ satisfies the superset axiom, then $F \neq \varnothing$ iff $I \in F$. Hence $I \subseteq F$ for any filter F.*

PROOF. By definition, a filter is a nonempty collection. $\square$

PROPOSITION 2.5.2. *An ultrafilter F satisfies*

$$A \cap B \in F \quad \textit{iff} \quad A \in F \textit{ and } B \in F$$

and

$$A \cup B \in F \quad \textit{iff} \quad A \in F \textit{ or } B \in F$$

and

$$A^c \in F \quad \textit{iff} \quad A \notin F.$$

PROPOSITION 2.5.3. *Let F be an ultrafilter and $\{A_1, \ldots, A_n\}$ a finite collection of pairwise disjoint ($A_i \cap A_j = \varnothing$ if $i \neq j$) sets such that*

$$A_1 \cup \cdots \cup A_n \in F.$$

Then $A_i \in F$ for exactly one i such that $1 \leq i \leq n$.

PROOF. Let's give the proof in the case $n = 2$; the general case is similar. Suppose $A_1 \cup A_2 \in F$. Then one of them must be in F for otherwise the union will also not be in F. To fix ideas, suppose $A_1 \in F$. Then its complement A_1^c is not in F. But $A_2 \subseteq A_1^c$ by hypothesis. Therefore $A_2 \notin F$ as required. □

COROLLARY 2.5.4. *If an ultrafilter contains a finite set, then it contains a one-element set and is principal. Hence a nonprincipal ultrafilter must contain all cofinite sets.*

This is a crucial property used in the construction of infinitesimals and infinitely large numbers.

COROLLARY 2.5.5 (Maximality). *F is an ultrafilter on I iff it is a* maximal *proper filter on I, i.e., a proper filter that cannot be extended to a larger proper filter on I.*

2.6. The ring of real-valued sequences

Let $\mathbb{N} = \{1, 2, \ldots\}$.

DEFINITION 2.6.1. $\mathbb{R}^{\mathbb{N}}$ is the set of all sequences of real numbers.

A typical member of $\mathbb{R}^{\mathbb{N}}$ has the form $r = \langle r_1, r_2, r_3, \ldots \rangle$, which may be denoted more briefly as

$$\langle r_n : n \in \mathbb{N} \rangle$$

or just

$$\langle r_n \rangle.$$

DEFINITION 2.6.2 (Arithmetic operations). For $r = \langle r_n \rangle$ and $s = \langle s_n \rangle$, we set

$$r \oplus s = \langle r_n + s_n : n \in \mathbb{N} \rangle,$$

and

$$r \odot s = \langle r_n \cdot s_n : n \in \mathbb{N} \rangle.$$

PROPOSITION 2.6.3. $(\mathbb{R}^{\mathbb{N}}, \oplus, \odot)$ *is a commutative ring with zero* $\mathbf{0} = \langle 0, 0, 0, \ldots \rangle$ *and unity*

$$\mathbf{1} = \langle 1, 1, \ldots \rangle,$$

and additive inverse (reciprocal) given by

$$-r = \langle -r_n : n \in \mathbb{N} \rangle.$$

PROPOSITION 2.6.4. *The ring $\mathbb{R}^{\mathbb{N}}$ is not a field.*

PROOF. Consider the product of the sequences

$$\langle 1,0,1,0,1,\ldots\rangle \odot \langle 0,1,0,1,0,\ldots\rangle = \mathbf{0}.$$

Then one of the two sequences on the left of this equation are nonzero elements of $\mathbb{R}^{\mathbb{N}}$ with a zero product; hence neither can have a multiplicative inverse. Indeed, no sequence that has at least one zero term can have such an inverse in $\mathbb{R}^{\mathbb{N}}$. □

2.7. Equivalence modulo an ultrafilter

Let F be a fixed nonprincipal ultrafilter on the set $I = \mathbb{N}$. Such an F will be used to construct a quotient ring of $\mathbb{R}^{\mathbb{N}}$.[1]

DEFINITION 2.7.1. We define a relation $\equiv$ on $\mathbb{R}^{\mathbb{N}}$ by setting

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

When this relation holds it may be said that the two sequences agree on a *large* set, or agree *almost everywhere modulo* F, or agree at almost all n.

PROPOSITION 2.7.2. *The relation* $\equiv$ *has the following properties.*

(1) $\equiv$ *is an equivalence relation on* $\mathbb{R}^{\mathbb{N}}$.
(2) *if* $r\equiv r'$ *and* $s\equiv s'$, *then*

$$r\oplus s \equiv r'\oplus s' \text{ and } r\odot s\equiv r'\odot s'.$$

(3) *We have a pair of inequivalent sequences:*

$$\left\langle 1, \tfrac{1}{2}, \tfrac{1}{3}, \ldots\right\rangle \not\equiv \langle 0,0,0,\ldots\rangle.$$

2.8. A suggestive logical notation

It is suggestive to denote the agreement set $\{n\in\mathbb{N} : r_n = s_n\}$ by

$$[[r = s]],$$

rather than E_{rs} as in Section 2.2. Thus

$$r\equiv s \ \text{ iff } \ [[r=s]]\in F.$$

This idea can be applied to other logical assertions, such as inequalities, by defining

$$\begin{aligned}
[[r < s]] &= \{n\in\mathbb{N} : r_n < s_n\},\\
[[r > s]] &= \{n\in\mathbb{N} : r_n > s_n\},\\
[[r \le s]] &= \{n\in\mathbb{N} : r_n \le s_n\}.
\end{aligned}$$

[1] The use of F in the construction of such a quotient would seem to suggest that infinitesimal analysis depends on ultrafilters in an essential way. It turns out that this is not the case; see Chapter 10.

2.9. The ultrapower construction; definition of ${}^*\mathbb{R}$

DEFINITION 2.9.1. The equivalence class of a sequence $r \in \mathbb{R}^{\mathbb{N}}$ under the relation $\equiv$ will be denoted by $[r]$.

Thus

$$[r] = \{s \in \mathbb{R}^{\mathbb{N}} : r \equiv s\}.$$

DEFINITION 2.9.2 (Defining ${}^*\mathbb{R}$). The quotient set (set of equivalence classes) of $\mathbb{R}^{\mathbb{N}}$ by $\equiv$ is

$${}^*\mathbb{R} = \left\{[r] : r \in \mathbb{R}^{\mathbb{N}}\right\}.$$

DEFINITION 2.9.3 (Operations). We define operations on ${}^*\mathbb{R}$ as follows:

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

and

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

and

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

By properties given in Section 2.7, these notions are well-defined, which means that they are independent of the equivalence class representatives chosen to define them.

DEFINITION 2.9.4 (Simplified notation). A simpler notation is to write $[r_n]$ for the equivalence class

$$[\langle r_n : n \in \mathbb{N} \rangle]$$

of the sequence whose n^{th} term is r_n.

The definitions of addition and multiplication then take the simple form

$$[r_n] + [s_n] = [r_n + s_n]$$

and

$$[r_n] \cdot [s_n] = [r_n \cdot s_n].$$

2.10. ${}^*\mathbb{R}$ as an ordered field

THEOREM 2.10.1. *The ring ${}^*\mathbb{R}$ equipped with relations $+, \cdot, <$ is an ordered field with zero $[\mathbf{0}]$ and unity $[\mathbf{1}]$.*

PROOF. As a quotient ring of $\mathbb{R}^{\mathbb{N}}$, the ring ${}^*\mathbb{R}$ is a commutative ring with zero $[0]$ and unity $[1]$, and additive inverses given by

$$-[\langle r_n : n \in \mathbb{N} \rangle] = [\langle -r_n : n \in \mathbb{N} \rangle].$$

Let us show that it has multiplicative inverses. Suppose $[r] \neq [0]$ so that $r \not\equiv 0$, i.e., $\{n \in \mathbb{N} : r_n = 0\} \notin F$.

Since F is an ultrafilter, the complementary set

$$J = \{n \in \mathbb{N} : r_n \neq 0\} \in F$$

(J is a member of F). Define a sequence s by setting

$$s_n = \begin{cases} \frac{1}{r_n} & \text{if } n \in J \\ 0 & \text{otherwise} \end{cases}$$

Then the set $[[r \odot s = 1]]$ is equal to J, so $[[r \odot s = 1]] \in F$, giving $r \odot s \equiv 1$ and hence

$$[r] \cdot [s] = [r \odot s] = [1]$$

in ${}^*\mathbb{R}$. This means that the element $[s]$ is the multiplicative inverse $[r]^{-1}$ of $[r]$.

Let us show that the ordering $<$ on ${}^*\mathbb{R}$ is linear. Observe that $\mathbb{N}$ is the disjoint union of the three sets

$$[[r < s]], \quad [[r = s]], \quad [[s < r]].$$

By Proposition 2.5.3, exactly one of the three belongs to F. Therefore exactly one of the three relations

$$[r] < [s], \quad [r] = [s], \quad [s] < [r]$$

is true.

Similarly, one shows that the set $\{[r] : [\mathbf{0}] < [r]\}$ of positive elements in ${}^*\mathbb{R}$ is closed under addition and multiplication. $\square$

2.11. Including the reals in the hyperreals

We can identify a real number $r \in \mathbb{R}$ with the constant sequence $\mathbf{r} = \langle r, r, r, \ldots \rangle$ and hence assign to it the element

$$[\mathbf{r}] = [\langle r, r, r, \ldots \rangle].$$

THEOREM 2.11.1. *The map $r \mapsto [\mathbf{r}]$ is an order-preserving field isomorphism from $\mathbb{R}$ to ${}^*\mathbb{R}$.*

COROLLARY 2.11.2. *${}^*\mathbb{R}$ is an ordered field extension of $\mathbb{R}$.*

2.12. Infinitesimals and infinite numbers

DEFINITION 2.12.1. A number $\alpha \in {}^*\mathbb{R}$ is *infinitesimal* if it is smaller than every positive real number and bigger than its negative:

$$\forall r \in \mathbb{R}\ (r > 0 \implies -r < \alpha < r).$$

EXAMPLE 2.12.2. Let $\varepsilon = \langle 1, \frac{1}{2}, \frac{1}{3}, \ldots \rangle = \langle \frac{1}{n} : n \in \mathbb{N} \rangle$. Then

$$[[0 < c]] = \{n \in \mathbb{N} : 0 < \tfrac{1}{n}\} = \mathbb{N} \in F.$$

Thus $[\mathbf{0}] < [\varepsilon]$ in ${}^*\mathbb{R}$. But if r is any positive real number, then the set

$$[[c < r]] = \{n \in \mathbb{N} : \frac{1}{n} < r\}$$

is cofinite because the sequence ε converges to 0 in $\mathbb{R}$. Now, since F is nonprincipal, it contains all cofinite sets by Proposition 2.5.4. Therefore $[[\varepsilon < r]] \in F$ and thus $[\varepsilon] < [\mathbf{r}]$ in ${}^*\mathbb{R}$. It follows that $[\varepsilon]$ is a positive infinitesimal.

DEFINITION 2.12.3. A positive number $H \in {}^*\mathbb{R}$ is *infinite* if it is bigger than every real number:

$$\forall r \in \mathbb{R} \quad H > r.$$

EXAMPLE 2.12.4. Let $\omega = \langle 1, 2, 3, \ldots \rangle$. Then for any $r \in \mathbb{R}$, the set

$$[[r < \omega]] = \{n \in \mathbb{N} : r < n\} \tag{2.12.1}$$

is cofinite as there are only finitely many integers less than r. Therefore the set (2.12.1) belongs to F, showing that $[\mathbf{r}] < [\omega]$ in ${}^*\mathbb{R}$.

Thus $[\omega]$ is "infinitely large" compared to all real numbers. In fact $\varepsilon \cdot \omega = 1$, so $[\omega] = [\varepsilon]^{-1}$ and $[\varepsilon] = [\omega]^{-1}$.

CHAPTER 3

Enlarging sets and functions; Transfer Principle

3.1. Enlargements of Sets

In Chapter 2, we enlarged the real line $\mathbb{R}$ to a hyperreal line ${}^*\mathbb{R}$ by means of the ultrapower construction. Recall that

(1) $\mathbb{R}^{\mathbb{N}}$ is the space of real-valued sequences;
(2) the field ${}^*\mathbb{R}$ is a the set of equivalence classes in $\mathbb{R}^{\mathbb{N}}$ of the equivalence relation $\equiv$;
(3) sequences $r = \langle r_n \rangle$ and $s = \langle s_n \rangle$ are equivalent, $r \equiv s$, if and only if they coincide of a "large" set of indices, meaning that

$$\{n \in \mathbb{N} : r_n = s_n\} \in F, \tag{3.1.1}$$

where
(4) F is a fixed nonprincipal ultrafilter on $\mathbb{N}$;
(5) we will use the $[[\ldots]]$ notation:

$$[[r \in A]] = \{n \in \mathbb{N} : r_n \in A\}.$$

Remark 3.1.1. Sometimes we use the term "equal for almost all n (modulo F)" to refer to the situation as in (3.1.1).

Recall that the equivalence class of a sequence r is denoted $[r] \in {}^*\mathbb{R}$. A subset $A \subseteq \mathbb{R}$ can also be "enlarged" to a subset

$${}^*\!A \subseteq {}^*\mathbb{R}.$$

What are the elements of ${}^*\!A$?

Definition 3.1.2. For each $r \in \mathbb{R}^{\mathbb{N}}$, put

$$[r] \in {}^*\!A \ \text{ iff } \ \{n \in \mathbb{N} : r_n \in A\} \in F.$$

Thus we are declaring, by the almost-all criterion, that the hyperreal $[\langle r_n \rangle]$ is in ${}^*\!A$ if and only if r_n is in A for almost all n.

Proposition 3.1.3. *The enlargement of A is well-defined.*

Proof. Let r' be another sequence from the same class. Then

$$[[r = r']] \cap [[r \in A]] \subseteq [[r' \in A]].$$

It follows by a defining property of the ultrafilter (intersections) that

$$r \equiv r' \text{ and } [[r \in A]] \in F \text{ implies } [[r' \in A]] \in F.$$

According to Definition 3.1.2, we have

$$[r] \in {}^*\!A \text{ iff } [[r \in A]] \in F,$$

as required. □

REMARK 3.1.4. If $s \in A$, then $[[\mathbf{s} \in A]] = \mathbb{N} \in F$ (where $\mathbf{s} = \langle s, s, \ldots \rangle$ as usual), so $[\mathbf{s}] \in {}^*\!A$.

Identifying s with $[\mathbf{s}]$, we may regard ${}^*\!A$ as a superset of A:

$$A \subseteq {}^*\!A.$$

Elements of the complement ${}^*\!A - A$ may be thought of as new "nonstandard", or "ideal", members of A that live in ${}^*\mathbb{R}$.

EXAMPLE 3.1.5. Let $A = \mathbb{N}$, and $\omega = \langle 1, 2, 3, \ldots \rangle$ as in Section 2.12. Then $[[\omega \in \mathbb{N}]] = \mathbb{N} \in F$, so $[\omega] \in {}^*\mathbb{N}$. The hyperreal $[\omega]$ is a "nonstandard natural number".

THEOREM 3.1.6. *For any infinite subset A of $\mathbb{R}$, the set ${}^*\!A$ has nonstandard members.*

PROOF. If $A \subseteq \mathbb{R}$ is infinite, then there is a sequence r of elements of A whose terms are all distinct. Then $[[r \in A]] = \mathbb{N} \in F$, so $[r] \in {}^*\!A$. But for each element $s \in A$, the set

$$\{n : r_n = s\}$$

is either $\varnothing$ or a singleton, neither of which can belong to F (finite sets are always negligible). Therefore $[r] \neq [\mathbf{s}]$. Hence $[r] \in {}^*\!A - A$. □

The converse of this theorem is also true (see Section 3.9) so the property of having nonstandard members exactly characterizes the infinite sets.

3.2. Extending functions

A function $f : \mathbb{R} \to \mathbb{R}$ extends to ${}^*\!f : {}^*\mathbb{R} \to {}^*\mathbb{R}$ as follows. First, for each sequence $r \in \mathbb{R}^{\mathbb{N}}$, let $f \circ r$ be the sequence $\langle f(r_1), f(r_2), \ldots \rangle$.

DEFINITION 3.2.1. We set

$${}^*\!f([r]) = [f \circ r].$$

In other words,

$$^*f([\langle r_1, r_2, \ldots\rangle]) = [\langle f(r_1), f(r_2), \ldots\rangle], \tag{3.2.1}$$

or in the simplified notation,

$$^*f([r_n]) = [f(r_n)].$$

PROPOSITION 3.2.2. *The function $^*f$ of* (3.2.1) *is well-defined.*

PROOF. In general, we have

$$[[r = r']] \subseteq [[f \circ r = f \circ r']].$$

In particular, we obtain

$$r \equiv r' \text{ implies } f \circ r \equiv f \circ r',$$

proving the proposition. □

Observe that $^*f$ obeys the almost-all criterion:

$$^*f([r]) = [s] \text{ iff } [[f \circ r = s]] \in F,$$

which occurs if and only if

$$\{n \in \mathbb{N} : f(r_n) = s_n\} \in F$$

which occurs if and only if

$$f(r_n) = s_n \text{ for almost all } n.$$

EXAMPLE 3.2.3. The sine function is extended to all of $^*\mathbb{R}$ by

$$^*\sin([r]) = [\langle \sin(r_1), \sin(r_2), \ldots\rangle] = [\sin(r_n)].$$

3.3. Partial Functions and Hypersequences

Let $f : A \to \mathbb{R}$ be a function whose domain A is a subset of $\mathbb{R}$ (e.g., $f(x) = \tan x$). Then f extends to a function $^*f : {^*A} \to {^*\mathbb{R}}$ whose domain is the enlargement of A, i.e., $\mathrm{dom}(^*f) = {^*(\mathrm{dom} f)}$.

To define this extension, take $r \in \mathbb{R}^{\mathbb{N}}$ with $[r] \in {^*A}$, so that

$$[[r \in A]] = \{n \in \mathbb{N} : r_n \in A\} \in F.$$

Let

$$s_n = \begin{cases} f(r_n) & \text{if } n \in [[r \in A]] \\ 0 & \text{if } n \notin [[r \in A]] \end{cases}$$

(it is enough to define s_n for almost all n). Then put

$$^*f([r]) = [s].$$

Remark 3.3.1. Essentially, we have defined

$$^*f([r_n]) = [f(r_n)]$$

as in Section 3.2, but with a modification to account for the complication that $f(r_n)$ may be undefined for some n. The construction works because $f(r_n)$ exists for almost all n modulo F.

It is readily shown that if $r \in A$, then $^*f([\mathbf{r}]) = {}^*(f(r))$, or identifying $[\mathbf{r}]$ with r as before, we have $^*f(r) = f(r)$, so $^*f$ extends f.

Remark 3.3.2 (Dropping the star). We will often drop the * symbol and just use f for the extended function, as well. It is a particularly natural practice for the more common mathematical functions.

For instance, the function $\sin x$ is now defined for all hyperreals $x \in {}^*\mathbb{R}$.

An important case of this construction concerns sequences. A real-valued sequence is just a function $s : \mathbb{N} \to \mathbb{R}$, and so the construction extends this to a hypersequence $s : {}^*\mathbb{N} \to {}^*\mathbb{R}$.

Corollary 3.3.3. *The n-th term s_n of the sequence is defined even when $n \in {}^*\mathbb{N} - \mathbb{N}$.*

3.4. Enlarging Relations

Let P be a k-ary relation on $\mathbb{R}$. Thus P is a set of k-tuples, namely a subset of $\mathbb{R}^k$.

Example 3.4.1. An example of a binary relation ($k = 2$) is the order relation $<$ on $\mathbb{R}$. We have $[[r < s]] = \{n \in \mathbb{N} : r_n < s_n\}$.

More generally, for given sequences $r^1, \ldots, r^k \in \mathbb{R}^{\mathbb{N}}$, define

$$[[P(r^1, \ldots, r^k)]] = \{n \in \mathbb{N} : P(r^1_n, \ldots, r^k_n)\}.$$

Just as $<$ extends to a relation on $^*\mathbb{R}$, any relation P can be enlarged to a k-ary relation $^*P$ on $^*\mathbb{R}$, i.e., a subset of $(^*\mathbb{R})^k$.

For this we use the notation $^*P([r^1], \ldots, [r^k])$ to mean that the k-tuple $([r^1], \ldots, [r^k])$ belongs to $^*P$. The definition is:

$$^*P([r^1], \ldots, [r^k]) \text{ iff } [[P(r^1 \ldots, r^k)]] \in F$$

which occurs if and only if

$$P(r^1_n, \ldots, r^k_n) \text{ for almost all } n \text{ (modulo } F)$$

As always with a definition involving equivalence classes named by particular elements, it must be shown that the notion is well-defined. In this case we can prove

$$[[r^1 = s^1]] \cap \cdots \cap [[r^k = s^k]] \cap [[P(r^1 \ldots, r^k)]] \subseteq [[P(s^1, \ldots, s^k)]],$$

so that if

$$\big(r^1 \equiv s^1 \text{ and } \dots \text{ and } r^k \equiv s^k \text{ and } [[P(r^1,\dots,r^k)]] \in F\big),$$

then $[[P(s^1,\dots,s^k)]] \in F$.

Definition 3.4.2. Let r^j be a real number. We will use the notation ${}^*r^j$ for the equivalence class $[\mathbf{r}^j] = [\langle r^j, r^j, \dots\rangle]$.

When $r^1,\dots,r^k$ are real numbers,

$$P(r^1,\dots,r^k) \text{ iff } {}^*P({}^*r^1,\dots,{}^*r^k),$$

showing that ${}^*P$ is an extension of P.

3.5. Relations encompass sets and functions

Our definition of the k-ary relation ${}^*P$ encompasses the work on extensions of sets and functions.

Example 3.5.1. A subset A of $\mathbb{R}$ is just a unary relation ($k = 1$), so the definition of ${}^*A$ is a special case of that of ${}^*P$.

Example 3.5.2. If P is any of the relations $=$, $<$, $>$, $\leq$ on ${}^*\mathbb{R}$, then ${}^*P$ is the corresponding relation that we defined on ${}^*\mathbb{R}$. Indeed, given sequences r and s, we have

$$[r] = [s] \text{ iff } [[r = s]] \in F,$$

$$[r] < [s] \text{ iff } [[r < s]] \in F,$$

and so on.

Example 3.5.3. An m-ary function $f : \mathbb{R}^m \to \mathbb{R}$ can be identified with its $(m+1)$-ary graph

$$\text{Graph } f = \{\langle r^1 \dots, r^m, s\rangle : f(r^1,\dots,r^m) = s\}.$$

Then the extension of Graph f to ${}^*\mathbb{R}$ is just the graph of the extension

$${}^*f : {}^*\mathbb{R}^m \to {}^*\mathbb{R}$$

of f i.e.,

$${}^*(\text{Graph } f) = \text{Graph}({}^*f).$$

Moreover, Graph f is defined even when f is a partial function (see Section 3.3), and so the case of partial functions is covered as well.

3.6. Introduction to the transfer principle

QUESTION 3.6.1. What properties are preserved in passing from $\mathbb{R}$ to ${}^*\mathbb{R}$?

We have already seen a number of examples: e.g., properties of an ordered field; see Theorem 2.10.1.

We will now consider some more examples in order to illustrate the powerful logical transfer principle that underlies them.

To formulate this principle we will need to develop a precise language in which to describe transferable properties. Ultimately this will allow us to abandon the ultrapower description of ${}^*\mathbb{R}$ and ultrafilter calculations.

REMARK 3.6.2. Similarly, the Dedekind completeness principle allows us to abandon the view of real numbers as cuts or equivalence classes of Cauchy sequence of rationals.

Later we will see that the strength of nonstandard analysis lies in the ability to transfer properties back from ${}^*\mathbb{R}$ to $\mathbb{R}$, providing a new technique for exploring real analysis. We will provide several examples of transforming statements in Sections 3.7 through 3.12.

3.7. Transforming Statements: the Archimedean Principle

The statement

$$\forall x\, \exists m\, (x < m \text{ and } m \in \mathbb{N})$$

is true when the variable x ranges over $\mathbb{R}$.

However, the statement is no longer true when x ranges over ${}^*\mathbb{R}$. For example, the formula fails for the hyperreal $x = [\langle 1, 2, 3, \ldots \rangle]$.

But if $\mathbb{N}$ is replaced by its "*-transform" ${}^*\mathbb{N}$, the result is the statement

$$\forall x\, \exists m\, (x < m \text{ and } m \in {}^*\mathbb{N}),$$

which is true when x ranges over all of ${}^*\mathbb{R}$.

This example shows that in order to determine the truth value of a sentence,[1] we need to specify what values a quantified variable is allowed to take. We can achieve this by using bounded quantifiers,[2] a notational device that displays the range of quantification explicitly. Thus the first sentence can be conveniently written as

$$\forall x \in \mathbb{R}\, \exists m \in \mathbb{N}\, (x < m),$$

[1] Pasuk

[2] kamatim chasumim

which is a true statement. Its *-transform

$$\forall x \in {}^*\mathbb{R}\ \exists m \in {}^*\mathbb{N}\ (x < m)$$

is also true. On the other hand, the statement

$$\forall x \in {}^*\mathbb{R}\ \exists m \in \mathbb{N}\ (x < m)$$

is false.

3.8. Density of the Rationals

The density of the rationals is expressed by the true statement

$$\forall x, y \in \mathbb{R}\,(x < y \text{ implies } \exists q \in \mathbb{Q}\,(x < q < y)).$$

The *-transform

$$\forall x, y \in {}^*\mathbb{R}\,(x < y \text{ implies } \exists q \in {}^*\mathbb{Q}\,(x < q < y))$$

is also true. In particular, it is true when x and y are infinitely close. Then the statement asserts the existence of a hyperrational between x and y and therefore necessarily also infinitely close to both x and y.

3.9. Finite sets

Let $A = \{r_1, \dots, r_k\}$ be a finite subset of $\mathbb{R}$. Then the statement

$$\forall x \in A\,(x = r_1 \text{ or } x = r_2 \text{ or } \cdots \text{ or } x = r_k)$$

is true, and so is its *-transform

$$\forall x \in {}^*\!A\,(x = {}^*r_1 \text{ or } x = {}^*r_2 \text{ or } \cdots \text{ or } x = {}^*r_k).$$

Since we identify r_i with ${}^*r_i$ in viewing $\mathbb{R}$ as a subset of ${}^*\mathbb{R}$, this implies that ${}^*\!A = A$. We therefore obtain the following proposition.

Proposition 3.9.1. *Finite sets of standard numbers admit no non-standard elements.*

Question 3.9.2. Why does this argument not work for infinite sets (see Theorem 3.1.6) ?

The answer is that there is no corresponding formula that one could apply transfer to.

3.10. Finitary set operations

We continue our analysis of the *-transform. If $A, B \subseteq \mathbb{R}$, then the statement

$$\forall x \in \mathbb{R}\,(x \in A \cup B \text{ iff } x \in A \text{ or } x \in B)$$

transforms to the true statement

$$\forall x \in {}^*\mathbb{R}\,(x \in {}^*(A \cup B) \text{ iff } x \in {}^*\!A \text{ or } x \in {}^*B),$$

which shows that ${}^*(A \cup B) = {}^*\!A \cup {}^*B$.

QUESTION 3.10.1. Question: why does the argument not work for unions of infinitely many sets?

3.11. Discreteness of natural numbers

If $n \in \mathbb{N}$, then the statement

$$\forall x \in \mathbb{N}\,(n \leq x \leq n+1 \text{ implies } x = n \text{ or } x = n+1)$$

transforms to

$$\forall x \in {}^*\mathbb{N}\,(\,{}^*\!n \leq x \leq {}^*(n+1) \text{ implies } x = {}^*\!n \text{ or } x = {}^*(n+1)),$$

which again is true. Since $n = {}^*n$ and likewise ${}^*(n+1) = n+1$, this shows that there are no nonstandard members of ${}^*\mathbb{N}$ occurring between any standard natural numbers. Also, there are no members of ${}^*\mathbb{N}$ smaller than 1, i.e.,

$$\forall x \in {}^*\mathbb{N}\,(x \geq 1);$$

hence any member of ${}^*\mathbb{N} - \mathbb{N}$ must be greater than all members of $\mathbb{N}$, and so is infinite (see Section 2.12).

3.12. Unbounded sets of real numbers

Consider an infinite $H \in {}^*\mathbb{N}$. Then we can deduce the Archimedean principle in the following way. If r is any real number, then $r < H$, since H is infinite. It follows that the statement

$$\exists n \in {}^*\mathbb{N}(r < n)$$

is true. This is the *-transform of the statement

$$\exists n \in \mathbb{N}(r < n),$$

and as we shall see, a statement must be true if its *-transform is. This shows that there is a positive integer greater than r.

More generally, this argument can be used to show the following.

PROPOSITION 3.12.1. *If the enlargement* $^*A$ *of a set* A *of reals contains an infinite member, then* A *itself must be unbounded in* $\mathbb{R}$*, in the sense that for any real* r *there is a member of* A *that is greater than* r*.*

In brief: if $^*A$ has an infinite nonstandard member, then A has arbitrarily large standard members.

REMARK 3.12.2. The *-transform of a statement arises by attaching the "*" prefix to symbols that name particular entities, but not attaching it to variable symbols. The precise definition of *-transform will be presented later.

CHAPTER 4

Relational structures, *-transform, transfer

4.1. Relational Structures

The examples given in Chapter 3 used a semiformal logical symbolism to express statements that were asserted to be true or false of the structures $\mathbb{R}$ and ${}^*\mathbb{R}$. This symbolism will now be explicitly described.

Remark 4.1.1. Our first task is to distinguish clearly between a *relational structure* and the *language* it uses.

Definition 4.1.2. A *relational structure*[1] is a system of the form

$$\mathcal{S} = \langle K, Rel_{\mathcal{S}}, Fun_{\mathcal{S}} \rangle,$$

where K is a nonempty set, $Rel_{\mathcal{S}}$ is a collection of finitary relations on K, and $Fun_{\mathcal{S}}$ is a collection of finitary functions[2] on K (possibly including partial functions).

For instance, associated with any set K is the *full* structure

$$\langle K, Rel_K, Fun_K \rangle,$$

based on K, where Rel_K consists of *all* the finitary relations on K, and Fun_K consists of all the finitary functions on K. Since sets are unary relations, a full structure includes all subsets of K in Rel_K.

Definition 4.1.3. The full structure based on $\mathbb{R}$ will be denoted by $\mathcal{R}$.

Associated with the full structure is the structure

$${}^*\mathcal{R} = \left\langle {}^*\mathbb{R}, \{{}^*P : P \in Rel_{\mathbb{R}}\}, \{{}^*f : f \in Fun_{\mathbb{R}}\} \right\rangle.$$

Thus ${}^*\mathcal{R}$ consists of the extensions ${}^*P$ and ${}^*f$ of all relations and functions on $\mathbb{R}$, as defined in Chapter 3.

Remark 4.1.4. The structure ${}^*\mathcal{R}$ is *not* a full structure, since there are relations on ${}^*\mathbb{R}$ that are not of the form ${}^*P$ for any $P \in Rel_{\mathbb{R}}$.

[1] mivne yachasim

[2] I.e., functions of finitely many variables

4.2. The Language of a Relational Structure

Associated with each relational structure $\mathcal{S}$ is a language $\mathcal{L}_{\mathcal{S}}$ based on the following alphabet:

- Logical Connectives:[3]
  - $\wedge$ and
  - $\vee$ or
  - $\neg$ not (negation)
  - $\rightarrow$ implies
  - $\leftrightarrow$ if and only if
- Quantifier Symbols:
  - $\forall$ for all
  - $\exists$ there exists
- Parentheses: (,), [,]
- Variables: A countable collection of symbols, for which we use letters like x, y, z, x_1, x', etc.

A richer language associated with a universe is developed in [**6**, p. 166, Section 13.7].

4.3. Terms of the language

A *term*[4] of $\mathcal{L}_{\mathcal{S}}$ is a string of symbols defined inductively by the following rules:

- Each variable is an $\mathcal{L}_{\mathcal{S}}$-term.
- Each element s of K is an $\mathcal{L}_{\mathcal{S}}$ term, called a *constant.*
- If $f \in Fun_{\mathcal{S}}$ is an m-ary function, and $\tau_1, \ldots, \tau_m$ are $\mathcal{L}_{\mathcal{S}}$-terms, then $f(\tau_1, \ldots, \tau_m)$ is an $\mathcal{L}_{\mathcal{S}}$-term.

4.4. What Does a Term Name?

Definition 4.4.1. A *closed term* is one that has no variables and therefore is made up of constants and function symbols.

Such a term is intended to name a particular element of the structure $\mathcal{S}$. But there are many opportunities in mathematics to write down symbolic expressions that have no meaning because the element they purport to name does not exist, as in $\tan(\frac{\pi}{2})$.

A closed term is undefined if it does not name anything. Here are the rules that determine when, and what, a closed term names:

- The constant s names itself.

[3] chiburim logi'im

[4] shem etzem

- If $\tau_1, \ldots, \tau_m$ name the elements $s_1, \ldots, s_m$, respectively, and the m-tuple $(s_1, \ldots, s_m)$ is in the domain of a function f, then $f(\tau_1, \ldots, \tau_m)$ names the element $f(s_1, \ldots, s_m)$.
- $f(\tau_1, \ldots, \tau_m)$ is undefined if one of $\tau_1, \ldots, \tau_m$ is undefined, or if they are all defined but name an m-tuple that is not in the domain of f.

4.5. Atomic formulae of the language

The *atomic formulae* of the language $\mathcal{L}_\mathcal{S}$ are obtained by introducing the *relations* available in the relational structure. They are strings of the form

$$P(\tau_1, \ldots, \tau_k)$$

where $P \in \mathit{Rel}_\mathcal{S}$ is k-ary, and the τ_i are $\mathcal{L}_\mathcal{S}$-terms. Such strings assert basic relationships between elements of K and serve as the building blocks for more complex expressions.

We also use conventional notation for atomic formulae where appropriate. For binary relations ($k = 2$) there is the usual infix notation: $P(\tau_1, \tau_2)$ is written

$$\tau_1 = \tau_2$$

when P is the identity relation $\{(a, b) \in K \times K : a = b\}$, and as

$$\tau_1 < \tau_2$$

when $P = \{(a, b) : a < b\}$. Similarly for the relations $\tau_1 > \tau_2$, $\tau_1 \leq \tau_2$, $\tau_1 \geq \tau_2$.

When $k = 1$ we have unary atomic formulae of the form $P(\tau)$, with P being a subset of K. Such a formula expresses membership of P and so will usually be written in the form $\tau \in P$.

4.6. Formulae

Formulas are built out of atomic formulas by introducing logical connectives and quantifiers as follows.

- Each atomic $\mathcal{L}_\mathcal{S}$-formula is an $\mathcal{L}_\mathcal{S}$-formula.
- If ϕ and ψ are $\mathcal{L}_\mathcal{S}$ -formulae, then so are $\phi \wedge \psi$, $\phi \vee \psi$, $\neg\phi$, $\phi \to \psi$, $\phi \leftrightarrow \psi$.
- If ϕ is an $\mathcal{L}_\mathcal{S}$-formula, x is any variable symbol, and $P \in \mathit{Rel}_\mathcal{S}$ is unary, i.e., P is a subset of K, then

  $$(\forall x \in P)\phi, \ (\exists x \in P)\phi$$

  are $\mathcal{L}_\mathcal{S}$-formulae. Here P is the *bound* of the quantifier in question.

A formula is said to be *defined* if and only if all of its closed terms are defined.

Parentheses will be inserted or deleted in formulae where convenient to aid legibility. Various abbreviations and informalities will be used, such as writing

$$x \leq y \leq z$$

for the formula $(x \leq y) \wedge (y \leq z)$, or collapsing a string of similar quantifiers with the same bound like

$$(\forall x \in P)(\forall y \in P)(\forall z \in P)$$

to the form $(\forall x, y, z \in P)$.

4.7. Sentences

A sentence[5] is a particular type of formula.

An occurrence of the variable x within a formula ψ is called *bound*[6] if it is located within a formula of the form $(\forall x \in P)\phi$ or $(\exists x \in P)\phi$ that is part of ψ. An occurrence that is not bound is *free*. Thus in

$$(x < 1) \wedge (\forall x \in \mathbb{N})(x > y),$$

the first occurrence of x is free, while the others are bound, and the only occurrence of y is free.

If a formula contains free variables, then it has no particular meaning until we assign some values to those free variables. Thus the above formula makes a true assertion if $x = y = 0$, but if $x = 2$, then it cannot be true whatever the value of y is.

Definition 4.7.1. A *sentence* is a formula in which all variables are bound.

The role of each symbol in a sentence is determined. There are no free variables that need to be assigned a value, and if the closed terms of the sentence are all defined then it has a fixed meaning and makes a definite assertion. A defined sentence is either true or false.

Definition 4.7.2. An *atomic sentence* in the language $\mathcal{L}_S$ is an atomic formula $P(\tau_1, \ldots, \tau_k)$ that is a sentence.

This means that the terms $\tau_1, \ldots, \tau_k$ are all closed, i.e., the formula has no free variables.

[5] pasuk

[6] kashur or mekumat

4.8. Truth and Quantification

Suppose that there is only one variable, say x, that has any free occurrence in a certain formula ϕ. Then we write $\phi(s)$ for the sentence that is obtained by substituting the constant s in place of all free occurrences of x in ϕ. For example, if ϕ is the formula

$$\tan(-x) = -\tan(x),$$

then $\phi(\pi/2)$ is the (undefined) atomic sentence

$$\tan(-\pi/2) = -\tan(\pi/2).$$

Now consider the truth of a defined sentence of the form $(\forall x \in P)\phi$. Here only the variable x can have any free occurrence in ϕ, so we can form sentences of the type $\phi(s)$. Intuitively, $(\forall x \in P)\phi$ asserts that whatever ϕ "says about x" is true of each member of P, provided that this is defined, and so it asserts that the sentence $\phi(s)$ is true for every element s of P for which it is defined. Thus

> $(\forall x \in P)\phi$ is true if and only if for all s in P, if the sentence $\phi(s)$ is defined, then it is true.

For example, the following sentence is true:

$$(\forall x \in \mathbb{R})[\tan(-x) = -\tan(x)].$$

The corresponding analysis of the existential quantifier is $(\exists x \in P)\phi$ is true if and only if there is some $s \in P$ for which $\phi(s)$ is (defined and) true.

The standard meanings of the symbolic connectives $\wedge$, $\vee$, $\neg$, $\to$, $\leftrightarrow$ are given by the rules:

- $\phi \wedge \psi$ is true if and only if ϕ is true and ψ is true.
- $\phi \vee \psi$ is true if and only if ϕ is true or ψ is true.
- $\neg\phi$ is true if and only if ϕ is not true (i.e., is false).
- $\phi \to \psi$ is true if and only if the truth of ϕ implies that of ψ (i.e., either ϕ is false or else ψ is true).
- $\phi \leftrightarrow \psi$ is true if and only if $\phi \to \psi$ and $\psi \to \phi$ are true (i.e., ϕ and ψ are either both true or both false).

Remark 4.8.1. These rules reduce the calculation of the truth value of a sentence to the determination of the truth value of atomic sentences.

For atomic sentences we have the following proposition.

Proposition 4.8.2. *$P(\tau_1, \dots, \tau_k)$ is true if and only if the closed terms $\tau_1, \dots, \tau_k$ are all defined and the k-tuple of elements they name belongs to P.*

This exact formulation of the syntax of mathematical statements, with an associated account of their truth conditions, makes the theory of infinitesimals possible. We are able to distinguish exactly which properties are transferable between $\mathbb{R}$ and ${}^*\mathbb{R}$ because we can give an explicit description of the sentences that express such properties.

4.9. *-Transforms

A formula in the language $\mathcal{L}_{\mathcal{R}}$ of the real-number structure $\mathcal{R}$ has symbols P, f for relations and functions of $\mathcal{R}$. It can be turned into a formula of the language $\mathcal{L}_{{}^*\mathcal{R}}$ of the hyperreal structure ${}^*\mathcal{R}$ by replacing P by ${}^*P$, and f by ${}^*f$. Any constant r naming a real number is left as is, since we identify r in $\mathcal{R}$ with ${}^*r$ in ${}^*\mathcal{R}$.

DEFINITION 4.9.1. The *-transform ${}^*\tau$ of an $\mathcal{L}_{\mathcal{R}}$-term τ is obtained by replacing each function symbol f occurring in τ by ${}^*f$, leaving the variables and constants of τ alone.

More formally, we can give the definition by induction on the formation of τ, using the following rules:

- If τ is a variable or an $\mathcal{L}_{\mathcal{R}}$-constant , then ${}^*\tau$ is just τ.
- If τ is $f(\tau_1, \ldots, \tau_m)$, then ${}^*\tau$ is ${}^*f({}^*\tau_1, \ldots, {}^*\tau_m)$.

The *-transform ${}^*\phi$ of an $\mathcal{L}_{\mathcal{R}}$-formula ϕ is obtained as follows:

- replace each term τ occurring in ϕ by ${}^*\tau$;
- replace the relation symbol P of any atomic formula occurring in ϕ by ${}^*P$; and
- replace the "bound" P of any quantifier $(\forall x \in P)$ or $(\exists x \in P)$ occurring in ϕ by ${}^*P$.

We tend to drop the * symbol when referring to the transforms of some of the more well-known relations like $=, \neq, <, \geq$, etc., and well-known mathematical functions like $\sin, \cos, \log, e^x$, etc. For instance,

$$^*(\pi < f(x+1)) = (\pi < {}^*f(x+1)),$$

$$^*(\sin e^x \in \mathbb{Q}) = (\sin e^x \in {}^*\mathbb{Q}),$$

and so on. Even further, it would do no harm to drop the * symbol in referring to the extension ${}^*f$ of any function f. If this practice is adopted systematically, then the transform ${}^*\tau$ of each term τ will just be τ itself. Then atomic formulae like

$$\tau_1 = \tau_2,$$

etc. that express basic equalities and inequalities will be left alone under *-transformation, while a membership formula $\tau \in P$ becomes

$\tau \in {}^*P$. With all these conventions in place, the general procedure for "adding the stars" reduces simply to replacing

$$P(\tau_1, \ldots \tau_k) \text{ by } {}^*P(\tau_1, \ldots, \tau_k),$$

$$\forall x \in P \text{ by } \forall x \in {}^*P,$$

$$\exists x \in P \text{ by } \exists x \in {}^*P.$$

To summarise all of this in words; the essence of *-transformation is to

(1) replace the bound P of any quantifier by its enlargement ${}^*P$; and
(2) replace relations appearing in atomic formulae by their enlargements, but only in the (unary) case of a membership formula ($\tau \in P$), or for relations of arity greater than one other than the common relations $=, \neq <, \geq$, etc.

4.10. Preliminaries to the Transfer Principle

The notion of an $\mathcal{L}_{\mathcal{R}}$ sentence and its *-transform enables a formalisation of the notion of an appropriately formulated statement as discussed in Chapter 3.

Hence it provides a first answer to the question as to which properties are subject to transfer between $\mathbb{R}$ and ${}^*\mathbb{R}$: any property expressible by an $\mathcal{L}_{\mathcal{R}}$-sentence is transferable. Formally, the transfer principle is stated as follows:

> A defined[7] $\mathcal{L}_{\mathcal{R}}$-sentence ϕ is true if and only if ${}^*\phi$ is true.

As a first illustration of this, beyond the examples given earlier, consider the following.

THEOREM 4.10.1. ${}^*\mathbb{R}$ *is an ordered field.*

PROOF. The fact that $\mathbb{R}$ is an ordered field can be expressed in a finite number of $\mathcal{L}_{\mathcal{R}}$-sentences, like

$$(\forall x, y \in \mathbb{R})(x + y = y + x),$$

$$(\forall x \in \mathbb{R})(x \cdot 1 = x),$$

$$(\forall x, y \in \mathbb{R})(x < y \ \vee \ x = y \ \vee \ y < x),$$

and so on. By transfer we conclude that the *-transforms of these sentences are true, showing that ${}^*\mathbb{R}$ is an ordered field.

[7] chuki

In particular, to show that multiplicative inverses exist in ${}^*\mathbb{R}$, instead of making an ultrapower construction of the inverses as in the proof given earlier we simply observe that it is true that

$$(\forall x \in \mathbb{R})[x \neq 0 \rightarrow (\exists y \in \mathbb{R}) x \cdot y = 1]$$

and conclude by transfer that

$$(\forall x \in {}^*\mathbb{R})[x \neq 0 \rightarrow (\exists y \in {}^*\mathbb{R}) x \cdot y = 1],$$

completing the proof. □

For another example, consider the extension of closed intervals.

Example 4.10.2. Consider the closed interval

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$$

in the real line defined by points $a, b \in \mathbb{R}$. Then it is true that

$$(\forall x \in \mathbb{R})(x \in [a, b] \leftrightarrow a \leq x \leq b),$$

so by transfer we see that the enlargement of $[a, b]$ is the hyperreal interval defined by a and b:

$${}^*[a, b] = \{x \in {}^*\mathbb{R} : a \leq x \leq b\}.$$

Similarly, we can transfer to ${}^*\mathbb{R}$ many familiar facts about standard mathematical functions. Thus the following are true:

$$(\forall x \in {}^*\mathbb{R}) \sin(\pi - x) = \sin x,$$

$$(\forall x \in {}^*\mathbb{R}) \cosh x + \sinh x = e^x,$$

$$(\forall x, y \in {}^*\mathbb{R}^+) \log xy = \log x + \log y.$$

4.11. Transfer principle

All of the examples of Section 4.10 involve taking a quantified $\mathcal{L}_{\mathcal{R}}$-sentence of the form $(\forall x, y, \ldots \in \mathbb{R})\phi$ and transforming it to an $\mathcal{L}_{{}^*\mathcal{R}}$-sentence $(\forall x, y, \ldots \in {}^*\mathbb{R})\, {}^*\phi$. They are instances of the following general principle.

> Upward (Universal) Transfer: if a property holds for all real numbers, then it holds for all hyperreal numbers.

The meaning of the expression "property" is specified in terms of the formal language $\mathcal{L}_{\mathcal{R}}$. To use nonstandard analysis we need to develop the ability to show that a given property can be expressed in a transferable form.

Dual to upward transfer is

> Downward (Existential) Transfer: if there exists a hyperreal number satisfying a certain property, then there exists a real number with this property.

EXAMPLE 4.11.1. Take a real-valued sequence $s : \mathbb{N} \to \mathbb{R}$ for which we can show (by some means) that the extended hypersequence ${}^*s : {}^*\mathbb{N} \to {}^*\mathbb{R}$ never takes infinitely large values. Then downward transfer can be used to conclude that the original sequence must be bounded in $\mathbb{R}$. To see this, let ω be a member of ${}^*\mathbb{N} - \mathbb{N}$. By hypothesis it is true that

$$(\forall n \in {}^*\mathbb{N})(\,|\,{}^*s(n)| < \omega). \tag{4.11.1}$$

The sentence (4.11.1) is not the $*$-transform of an $\mathcal{L}_{\mathcal{R}}$-sentence, because it contains the constant ω. But the constant can be removed by introducing an existentially quantified variable. Namely, we observe that the sentence implies

$$(\exists y \in {}^*\mathbb{R})(\forall n \in {}^*\mathbb{N})(|\,{}^*s(n)| < y). \tag{4.11.2}$$

Downward transfer applied to (4.11.2) yields

$$(\exists y \in \mathbb{R})(\forall n \in \mathbb{N})(|s(n)| < y).$$

Put informally, from the existence of a hyperreal bound on ${}^*s$ we infer the existence of a real bound on s. Typically, in order to show that a real number of a certain type exists, it may be easier to show that a hyperreal of this type exists and then apply downward transfer.

In the next chapter, we will provide some details on justifying transfer.

CHAPTER 5

Transfer, Łoś, and arithmetic of hyperreals

5.1. Justfying transfer

In the context of the ultrapower construction of the ordered field ${}^*\mathbb{R}$, we used the following notation:

- $[r]$ is the equivalence class of a sequence $r \in \mathbb{R}^{\mathbb{N}}$;
- $[[r < s]]$ is $\{n \in \mathbb{N} : r_n < s_n\}$, etc.
- $\mathcal{L}_{\mathcal{R}}$ is the language of the full relational system of the real numbers.

We repeatedly used the criterion that a particular property was to hold of hyperreals $[r], [s], \ldots$ if and only if

> the corresponding property held of the real numbers $r_n, s_n, \ldots$ for almost all n.

In fact, this almost-all criterion works for any property expressible by an $\mathcal{L}_{\mathcal{R}}$-formula. That is ultimately why the transfer principle holds.

To spell this out some further technical notation is needed.

Definition 5.1.1. For a formula ϕ we write

$$\phi(x_1, \ldots, x_p)$$

to indicate that the list $x_1, \ldots, x_p$ includes all the variables that occur *free* in the formula ϕ.

Then

$$\phi(s_1, \ldots, s_p)$$

is the sentence obtained by replacing each free occurrence of x_i in ϕ by the constant s_i.

Example 5.1.2. If $\phi(x_1, x_2)$ is the formula

$$(\exists y \in \mathbb{Q})(x_1^2 + x_2^2 < y),$$

then $\phi(\pi, \sqrt{2})$ is the sentence

$$(\exists y \in \mathbb{Q})(\pi^2 + (\sqrt{2})^2 < y).$$

Now let $\phi(x_1, \ldots, x_p)$ be a formula of $\mathcal{L}_{\mathcal{R}}$, and let $r^1, \ldots, r^p \in \mathbb{R}^{\mathbb{N}}$. We set

$$[[\phi(r^1, \ldots, r^p)]] = \{n \in \mathbb{N} : \phi(r^1_n, \ldots, r^p_n) \text{ is true } \}.$$

This extends the definitions of $[[r = s]]$, $[[r < s]]$, etc. to $\mathcal{L}_{\mathcal{R}}$-formulae in general.

Example 5.1.3. Consider the following typical statements:

$$\begin{gathered}
[r] = [s] \text{ iff } [[r = s]] \in \mathcal{F}, \\
[r] < [s] \text{ iff } [[r < s]] \in \mathcal{F}, \\
[r] \in {}^*\!A \text{ iff } [[r \in A]] \in \mathcal{F} \\
{}^*P([r^1], \ldots, [r^k]) \text{ iff } [[P(r^1, \ldots, r^k)]] \in \mathcal{F}.
\end{gathered}$$

Such statements are cases of the following fundamental result.

> For any $\mathcal{L}_{\mathcal{R}}$-formula $\phi(x_1, \ldots, x_p)$ and any sequences $r^1, \ldots, r^p \in \mathbb{R}^{\mathbb{N}}$, the sentence ${}^*\phi([r^1), \ldots, [r^p])$ is true if and only if $\phi(r^1_n, \ldots, r^p_n)$ is true for almost all $n \in \mathbb{N}$.

In other words,

$${}^*\phi([r^1], \ldots, [r^p]) \text{ is true iff } [[\phi(r^1, \ldots, r^p)]] \in \mathcal{F}.$$

This result is known as Łoś's theorem, after the Polish mathematician who first proved it in the early 1950s. It includes transfer as a special case, because if ϕ is a sentence, then it has no free variables, so that $\phi(s_1, \ldots, s_p)$ is just the sentence ϕ and likewise for ${}^*\phi$. Hence $[[\phi(r^1, \ldots, r^p)]]$ is $\mathbb{N}$ if ϕ is true and $\emptyset$ otherwise, independently of the sequences r^j. Since $\emptyset \notin \mathcal{F}$, Los's theorem in this case simply says

$${}^*\phi \text{ is true iff } \phi \text{ is true},$$

which is the transfer principle.

A proof of Łoś's theorem would proceed by induction on the formation of the formula ϕ, considering first atomic formulae and then dealing with inductive cases for the logical connectives and quantifiers. We will not enter into those details here, but rely on the examples already discussed to lend plausibility to the assertion of Łoś's theorem, and hence to transfer.

5.2. Extending Transfer

We defined general relational structures $\mathcal{S}$ and their languages $\mathcal{L}_{\mathcal{S}}$, but applied these ideas only to the language $\mathcal{L}_{\mathcal{R}}$ in describing the transfer principle.

In fact, it is possible to use the ultrapower construction to build an "enlargement" of *any* structure $\mathcal{S}$ and obtain a transfer principle for it. For instance, by replacing $\mathbb{R}$ by $\mathbb{C}$ this would give us a way of embarking on the nonstandard study of complex analysis.

REMARK 5.2.1. The language $\mathcal{L}_{\mathcal{R}}$ is limited by the fact that its quantifiable variables can range only over elements of $\mathbb{R}$ or elements of a given set $A \subseteq \mathbb{R}$. They cannot range over more complicated entities.

Thus they cannot range over subsets of $\mathbb{R}$ (i.e., over elements of the power set $\mathcal{P}(\mathbb{R})$), sequences, real-valued functions, etc.

EXAMPLE 5.2.2. Consider the Dedekind completeness principle,

> Every subset of $\mathbb{R}$ that is nonempty and bounded above has a least upper bound.

This principle cannot be formulated in $\mathcal{L}_{\mathcal{R}}$ because the language does not allow quantifiers of the type

$$\forall x \in \mathcal{P}(\mathbb{R})$$

that apply to a variable (x) whose range of values is the set of all subsets of $\mathbb{R}$.

REMARK 5.2.3. Later on, a language will be introduced that does have such "higher-order" quantifiers and for which an appropriate transfer principle exists.

We will soon see that $\mathcal{L}_{\mathcal{R}}$ is powerful enough to develop a great deal of the standard theory of $\mathbb{R}$, including the convergence of sequences and series, differential and integral calculus, and the basic topology of the real line. Indeed, for the next half-dozen chapters we will not relate to the ultrapower construction and explore all these topics using only the fact that ${}^*\mathbb{R}$ is an ordered field with the following properties:

- it has $\mathbb{R}$ as a subfield;
- it includes infinite numbers $H \in {}^*\mathbb{N} - \mathbb{N}$, hence infinitesimals (such as $\frac{1}{H}$), and
- satisfies the transfer principle.

5.3. Hyperreals

Members of ${}^*\mathbb{R}$ are called hyperreal numbers.
Members of $\mathbb{R}$ are real and sometimes called standard.
${}^*\mathbb{Q}$ consists of hyperrationals.
${}^*\mathbb{Z}$ consists of hyperintegers.
${}^*\mathbb{N}$ consists of hypernaturals.

PROPOSITION 5.3.1. ${}^*\mathbb{Q}$ *consists precisely of quotients* m/n *of hyperintegers* $m, n \in {}^*\mathbb{Z}$.

PROOF. Apply transfer to the sentence

$$\forall x \in \mathbb{R}\, [x \in \mathbb{Q} \leftrightarrow \exists y, z \in \mathbb{Z}(z \neq 0 \wedge x = y/z)]$$

to obtain the desired conclusion. □

5.4. Infinite, Infinitesimal, and Appreciable Numbers

We now examine the basic arithmetical and algebraic structure of ${}^*\mathbb{R}$, particularly in its relation to the structure of $\mathbb{R}$.

A hyperreal number b is:

- finite if $r < b < s$ for some $r, s \in \mathbb{R}$;
- positive infinite if $r < b$ for all $r \in \mathbb{R}$;
- negative infinite if $b < r$ for all $r \in \mathbb{R}$;
- infinite if it is positive or negative infinite;
- positive infinitesimal if $0 < b < r$ for all positive $r \in \mathbb{R}$;
- negative infinitesimal if $r < b < 0$ for all negative $r \in \mathbb{R}$.
- infinitesimal if it is positive infinitesimal, negative infinitesimal, or 0.
- appreciable[1] if it is finite but not infinitesimal, i.e., $r < |b| < s$ for some $r, s \in \mathbb{R}^+$.

Thus all real numbers, and all infinitesimals, are finite. The only infinitesimal real is 0: all other reals are appreciable. An appreciable number is one that is neither infinitely small nor infinitely big. Observe that b is

- finite iff $|b| < n$ for some $n \in \mathbb{N}$;
- infinite iff $|b| > n$ for all $n \in \mathbb{N}$;
- infinitesimal iff $|b| < \frac{1}{n}$ for all $n \in \mathbb{N}$;
- appreciable iff $\frac{1}{n} < |b| < n$ for some $n \in \mathbb{N}$.

We introduce the following notation.

(1) The set ${}^*\mathbb{N} - \mathbb{N}$ of infinite hypernaturals is denoted ${}^*\mathbb{N}_\infty$;
(2) ${}^*\mathbb{R}^+_\infty$ denotes the set of positive infinite hyperreals;
(3) ${}^*\mathbb{R}^-_\infty$ the set of negative infinite numbers.

This notation may be adapted to an arbitary subset $X \subseteq {}^*\mathbb{R}$, setting $X_\infty = \{x \in X : X \text{ is infinite }\}$, and similarly $X^+ = \{x \in X : x > 0\}$, etc.

DEFINITION 5.4.1. We use $\mathbb{L}$ for the set of all finite numbers, and $\mathbb{I}$ for the set of infinitesimals.

[1]mashmauti

5.5. Arithmetic of Hyperreals

Let ϵ, δ be infinitesimal, b, c appreciable, and H, K infinite. Then sums have the following properties:

- $\epsilon + \delta$ is infinitesimal;
- $b + \epsilon$ is appreciable;
- $b + c$ is finite (possibly infinitesimal);
- $H + \epsilon$ and $H + b$ are infinite.

Opposites:

- $-\epsilon$ is infinitesimal;
- $-b$ is appreciable;
- $-H$ is infinite.

Products:

- $\epsilon \cdot \delta$ and $\epsilon \cdot b$ are infinitesimal;
- $b \cdot c$ is appreciable;
- $b \cdot H$ and $H \cdot K$ are infinite.

Reciprocals:

- $\frac{1}{\epsilon}$ is infinite if $\epsilon \neq 0$;
- $\frac{1}{b}$ is appreciable;
- $\frac{1}{H}$ is infinitesimal.

Quotients:

- $\frac{\epsilon}{b}$, $\frac{\epsilon}{H}$, $\frac{b}{H}$ are infinitesimal;
- $\frac{b}{c}$ is appreciable (if $c \neq 0$);
- $\frac{b}{\epsilon}$, $\frac{H}{\epsilon}$, and $\frac{H}{b}$ are infinite ($\epsilon, b \neq 0$).

Roots:

- If $\epsilon > 0$, $\sqrt[n]{\epsilon}$ is infinitesimal;
- If $b > 0$, $\sqrt[n]{b}$ is appreciable;
- If $H > 0$, $\sqrt[n]{H}$ is infinite.

Indeterminate Forms:[2]

$$\frac{\epsilon}{\delta}, \frac{H}{K}, \epsilon \cdot H, H + K.$$

The following corollary follows from these rules.

Corollary 5.5.1. *The set $\mathbb{L}$ of finite numbers and the set $\mathbb{I}$ of infinitesimals are each a subring of ${}^*\mathbb{R}$.*

Also, the infinitesimals form an ideal in the ring of finite numbers. What then is the associated quotient ring $\mathbb{L}/\mathbb{I}$? This will be explained in Theorem 5.8.3.

[2] tzura bilti mugderet

With regard to nth roots, for fixed $n \in \mathbb{N}$ the function $x \mapsto \sqrt[n]{x}$ is defined for all positive reals, so extends to a function defined for all positive hyperreals.

PROPOSITION 5.5.2. *Every positive hyperreal has a hyperreal nth root for all $n \in {}^*\mathbb{N}$.*

PROOF. To find nth roots for infinite n, consider the statement

$$(\forall n \in \mathbb{N})(\forall x \in \mathbb{R}^+)(\exists y \in \mathbb{R})(y^n = x).$$

The statement asserts that any positive real has a real nth root for all $n \in \mathbb{N}$. Its transform asserts that every hyperreal has a hyperreal nth root for all $n \in {}^*\mathbb{N}$. □

5.6. On the Use of "Finite" and "Infinite"

A set is regarded as being finite if it has n elements for some $n \in \mathbb{N}$, and therefore is in bijective correspondence with the set

$$\{1, 2, \dots, n\} = \{k \in \mathbb{N} : k \leq n\}.$$

If H is an infinite hypernatural, then the collection

$$\{1, 2, \dots, H\} = \{k \in {}^*\mathbb{N} : k \leq H\}$$

is set-theoretically infinite but, by transfer, has many properties enjoyed by finite sets.[3] Collections of this type are called *hyperfinite*, and will be examined fully later. They are fundamental to the methodology of hyperreal analysis.

5.7. Halos, Galaxies, and Real Comparisons

A hyperreal number b is infinitely close to a hyperreal number c (in symbols: $b \approx c$), if $b - c$ is infinitesimal. This defines an equivalence relation on ${}^*\mathbb{R}$.

DEFINITION 5.7.1. The *halo* of b is the $\approx$-equivalence class

$$hal(b) = \{c \in {}^*\mathbb{R} : b \approx c\}.$$

Hyperreals b, c are of finite distance apart (in symbols: $b \sim c$) if $b - c$ is finite.

DEFINITION 5.7.2. The *galaxy* of b is the $\sim$-equivalence class

$$gal(b) = \{c \in {}^*\mathbb{R} : b \sim c\}.$$

Thus, b is infinitesimal iff $b \approx 0$, and finite iff $b \sim 0$.

[3]In the context of axiomatic nonstandard analysis, it is more appropriate to use terms *limited* and *unlimited* in place of finite and infinite; see Section 10.4.

COROLLARY 5.7.3. *We have* $hal(0) = \mathbb{I}$*, the set of infinitesimals, while* $gal(0) = \mathbb{L}$*, the set of finite hyperreals.*

Abraham Robinson called $hal(b)$ the "monad" of b in [**14**].

REMARK 5.7.4 (Real Comparisons). We discuss the comparison of the sizes of two real numbers b, c. If $b > c$, then the halos of the two numbers are disjoint, with everything in $hal(b)$ greater than everything in $hal(c)$. Thus to show that $b \leq c$ it is enough to show that something in $hal(b)$ is less than or equal to something in $hal(c)$. In particular, this will hold if there is some x with either $b \approx x \leq c$ or $b \leq x \approx c$.

5.8. Shadows

THEOREM 5.8.1. *Every finite hyperreal* b *is infinitely close to exactly one real number, called the shadow of* b*, denoted by* $sh(b)$*.*

PROOF. Let $A = \{r \in \mathbb{R} : r < b\}$. Since b is finite, there exist real r, s with $r < b < s$, so the set A is nonempty and bounded above in $\mathbb{R}$ by s. By the completeness of $\mathbb{R}$, it follows that A has a least upper bound $c \in \mathbb{R}$.

Let us show $b \approx c$. We take any positive real $\epsilon \in \mathbb{R}$. Since c is an upper bound of A, we cannot have $c + \epsilon \in A$; hence $b \leq c + \epsilon$. Also, if $b \leq c - \epsilon$, then $c - \epsilon$ would be an upper bound of A, contrary to the fact that c is the smallest such upper bound. Hence $b \not\leq c - \epsilon$. Altogether then, $c - \epsilon < b \leq c + \epsilon$, so $|b - c| \leq \epsilon$. Since this holds for all positive real ϵ, we obtain that b is infinitely close to c.

Finally, for uniqueness, if $b \approx c' \in \mathbb{R}$, then as $b \approx c$, we obtain $c \approx c'$, and so $c = c'$, since both are real. $\square$

THEOREM 5.8.2. *If* b *and* c *are finite and* $n \in \mathbb{N}$*, then*

(1) $sh(b \pm c) = sh(b) \pm sh(c)$,
(2) $sh(b \cdot c) = sh(b) \cdot sh(c)$,
(3) $sh(\frac{b}{c}) = \frac{sh(b)}{sh(c)}$ *if* $sh(c) \neq 0$ *(i. e., if* c *is appreciable),*
(4) $sh(b^n) = sh(b)^n$,
(5) $sh(|b|) = |sh(b)|$,
(6) $sh(\sqrt[n]{b}) = \sqrt[n]{sh(b)}$ *if* $b \geq 0$,
(7) *if* $b \leq c$ *then* $sh(b) \leq sh(c)$.

We see from these last facts that the shadow map $sh : b \mapsto sh(b)$ is an order-preserving homomorphism from the ring $\mathbb{L}$ of finite numbers onto $\mathbb{R}$. The kernel of this homomorphism is the set $\{b \in \mathbb{L} : sh(b) = 0\}$ of infinitesimals, and the cosets of the kernel are the halos $hal(b)$ for finite b. We therefore obtain the following theorem.

THEOREM 5.8.3. *The quotient ring* $\mathbb{L}/\mathbb{I}$ *is isomorphic to the real number field* $\mathbb{R}$ *by the correspondence* $hal(b) \mapsto sh(b)$*. Hence* $\mathbb{I}$ *is a maximal ideal of the ring* $\mathbb{L}$*.*

The shadow $sh(b)$ is often called the standard part of b.

5.9. Shadows and Completeness

The existence of shadows of finite numbers follows from the Dedekind completeness of $\mathbb{R}$. In fact, their existence turns out to be an alternative formulation of completeness, as the next result shows.

THEOREM 5.9.1. *The assertion "every finite hyperreal number is infinitely close to to a real number" implies the completeness of* $\mathbb{R}$*.*

PROOF. Let $s : \mathbb{N} \to \mathbb{R}$ be a Cauchy sequence. Recall that this means that its terms get arbitrarily close to each other as we move along the sequence.

In particular, there exists a $k \in \mathbb{N}$ such that all terms of s beyond s_k are within a distance of 1 of each other, i.e., the sentence

$$\forall m, n \in \mathbb{N}\ (m, n \geq k \to |s_m - s_n| < 1)$$

is true. By transfer, the *-transform of this sentence is also true, and applies to the extended hypersequence

$$\langle s_n : n \in {}^*\mathbb{N} \rangle$$

as defined above. In particular, if we take H to be an infinite member of ${}^*\mathbb{N}$, then $k, H \geq k$, so

$$|s_k - s_H| < 1.$$

It follows that s_H is finite. By the assertion quoted in the statement of the theorem, it follows that $s_H \approx L$ for a suitable $L \in \mathbb{R}$. We will show that the original sequence s converges to the real number L.

If ϵ is any positive real number, then since s is Cauchy, there exists $j_\epsilon \in \mathbb{N}$ such that beyond s_{j_ϵ} all terms are within ϵ of each other:

$$\forall m, n \in \mathbb{N}\ (m, n \geq j_\epsilon \to |s_m - s_n| < \epsilon). \tag{5.9.1}$$

But now we can show that beyond s_{j_ϵ} all terms are within ϵ of L. The essential reason is that all such terms are within ϵ of s_H, which is itself infinitely close to L. For if $m \in \mathbb{N}$ with $m \geq j_\epsilon$, we have $m, H \geq j_\epsilon$, so by transfer of the sentence (5.9.1), we get that s_m is within ϵ of s_H:

$$|s_m - s_H| < \epsilon.$$

Since s_H is infinitely close to L, this forces s_m to be within ϵ of L. Indeed,

$$|s_m - L| \le |s_m - s_H| + |s_H - L| < \epsilon + \text{infinitesimal}.$$

Since $s_m - L$ and ϵ are real, it follows that $|s_m - L| \le c$.

This establishes that all the terms $s_j, s_{j+1}, s_{j+2}, \ldots$ are within ϵ of L, which is enough to prove that the sequence s converges to the real number L. Thus we have demonstrated that every real Cauchy sequence is convergent in $\mathbb{R}$, a property that is equivalent to Dedekind completeness. $\square$

5.10. The Hypernaturals

We now develop a more detailed description of ${}^*\mathbb{N}$. First, by transfer, ${}^*\mathbb{N}$ is seen to be closed under addition and multiplication. Next observe that the only finite hypernaturals are the members of $\mathbb{N}$. For if $k \in {}^*\mathbb{N}$ is finite, then $k \le n$ for some $n \in \mathbb{N}$. But then by transfer of the sentence

$$\forall x \in \mathbb{N}\, (x \le n \to x = 1 \vee x = 2 \vee \cdots \vee x = n)$$

it follows that

$$k \in \{1, 2, \ldots, n\}, \text{ so } k \in \mathbb{N}.$$

Thus all members of ${}^*\mathbb{N} - \mathbb{N}$ are infinite, and hence greater than all members of $\mathbb{N}$.

Definition 5.10.1. Fixing $K \in {}^*\mathbb{N} - \mathbb{N}$, put

$$\gamma(K) = \{K\} \cup \{K \pm n : n \in \mathbb{N}\}.$$

Then all members of $\gamma(K)$ are infinite, and together form a "copy of $\mathbb{Z}$" under the ordering $<$. Moreover, it may be seen that

$$\gamma(K) = \{H \in {}^*\mathbb{N} : K \sim H\} = gal(K) \cap {}^*\mathbb{N},$$

the restriction to ${}^*\mathbb{N}$ of the galaxy of K. The set $\gamma(K)$ will be called a ${}^*\mathbb{N}$-galaxy. We can also view $\mathbb{N}$ itself as a ${}^*\mathbb{N}$-galaxy, since $\mathbb{N} = gal(1) \cap {}^*\mathbb{N}$.

Definition 5.10.2. We define $\gamma(K) = \mathbb{N}$ when $K \in \mathbb{N}$.

Then in general,

$$\gamma(K) = \gamma(H) \text{ iff } K \sim H,$$

and the ${}^*\mathbb{N}$-galaxies may be ordered by setting

$$\gamma(K) < \gamma(H) \text{ iff } K < H$$

whenever $K \not\sim H$ (i.e., whenever $|K - H|$ is infinite) .

There is no greatest ${}^*\mathbb{N}$-galaxy, since $\gamma(K) < \gamma(2K)$. Also, there is no smallest infinite one: since one of K and $K+1$ is even (by transfer) and $\gamma(K) = \gamma(K+1)$, we can assume that K is even and note that $K/2 \in {}^*\mathbb{N}$.

Then $\gamma(K/2) < \gamma(K)$ and $K/2$ infinite when K is. Finally, between any two ${}^*\mathbb{N}$-galaxies there is a third, for if $\gamma(K) < \gamma(H)$, with K, H both even, then

$$\gamma(K) < \gamma((H+K)/2) < \gamma(H).$$

COROLLARY 5.10.3. *The ordering $<$ of ${}^*\mathbb{N}$ consists of $\mathbb{N}$ followed by a densely ordered set of ${}^*\mathbb{N}$-galaxies (copies of $\mathbb{Z}$) with no first or last such galaxy.*

5.11. Convergence of Sequences

A real-valued sequence $\langle s_n : n \in \mathbb{N}\rangle$ is a function $s : \mathbb{N} \to \mathbb{R}$, and so extends to a hypersequence $s : {}^*\mathbb{N} \to {}^*\mathbb{R}$. Hence the term s_n becomes defined for infinite hypernaturals $n \in {}^*\mathbb{N}_\infty$ (a fact that was already used earlier), and in this case we say that s_n is an extended term of the sequence.

DEFINITION 5.11.1. The collection

$$\{s_n : n \in {}^*\mathbb{N}_\infty\}$$

of extended terms is the *extended tail* of s.

In real analysis, $\langle s_n : n \in \mathbb{N}\rangle$ converges to the limit $L \in \mathbb{R}$ when each open interval $(L-\epsilon, L+\epsilon)$ around L in $\mathbb{R}$ contains some standard tail of the sequence, i.e., contains all the terms from some point on (with this point depending on ϵ) . Formally, this is expressed by the statement

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

which is intended to capture the idea that we can approximate L as closely as we like by moving far enough along the sequence. It turns out that this is equivalent to the requirement that if we go "infinitely far" along the sequence, then we become infinitely close to L:

THEOREM 5.11.2. *A real-valued sequence $\langle s_n : n \in \mathbb{N}\rangle$ converges to a number $L \in \mathbb{R}$ if and only if $s_n \approx L$ for all infinite n.*

PROOF. Suppose $\langle s_n : n \in \mathbb{N}\rangle$ converges to L. Fix an $H \in {}^*\mathbb{N}_\infty$. In order to show that $s_H \approx L$ we have to show that $|s_H - L| < c$ for any positive real c. But given such an c, the standard convergence condition implies that there is an $m_c \in \mathbb{N}$ such that the standard tail beyond s_{m_c} is within c of L:

$$(\forall n \in \mathbb{N})(n > m_c \to |s_n - L| < c).$$

Then by (upward) transfer this holds for the extended tail as well:

$$(\forall n \in {}^*\mathbb{N})(n > m_c \to |s_n - L| < c).$$

But in fact, $H > m_c$ because H is infinite and m_c is finite, and so this last sentence implies $|s_H - L| < c$ for each $c > 0$, as required.

For the converse, suppose $s_n \approx L$ for all infinite n. We have to show that any given interval $(L-c, L+c)$ in $\mathbb{R}$ contains some standard tail of the sequence. The essence of the argument is to invoke the fact that the extended tail is infinitely close to L, hence contained in ${}^*(L-c, L+c)$, and then apply transfer.

To spell this out, fix an infinite $H \in {}^*\mathbb{N}_\infty$. Then for any $n \in {}^*\mathbb{N}$, if $n > H$, it follows that n is also infinite, so $s_n \approx L$ and therefore $|s_n - L| < c$. This shows that

$$(\forall n \in {}^*\mathbb{N})(n > H \to |s_n - L| < c).$$

Hence the sentence

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

is true. By downward transfer, we obtain

$$(\exists z \in \mathbb{N})(\forall n \in \mathbb{N})(n > z \to |s_n - L| < c),$$

giving the desired conclusion. □

Thus convergence to L amounts to the requirement that the extended tail of the sequence is contained in the halo of L. In this characterisation the role of the standard tails is taken over by the extended tail, while the standard open neighbourhoods $(L-c, L+c)$ are replaced by the "infinitesimal neighbourhood" $hal(L)$.

CHAPTER 6

Sequences, series, continuity

6.1. Limits

We saw in Theorem 5.11.2 that a sequence $s \in \mathbb{R}^{\mathbb{N}}$ converges to a number $L \in \mathbb{R}$ if and only if $s_H \approx L$ for all infinite $H \in {}^*\mathbb{N} - \mathbb{N}$.

Corollary 6.1.1. *A real-valued sequence has at most one limit.*

Proof. If $\langle s_n \rangle$ converges to both L and M in $\mathbb{R}$ then taking an infinite n, we have $s_n \approx L$ as well as $s_n \approx M$. Thus $L \approx M$ and since both are real, we obtain $L = M$. $\square$

6.2. Boundedness and Divergence

Recall that the *extended terms* of a sequence $\langle s_n \rangle$ are the terms at rank $n \in {}^*\mathbb{N} - \mathbb{N}$.

Theorem 6.2.1. *A real-valued sequence $\langle s_n \rangle$ is bounded in $\mathbb{R}$ if and only if its extended terms are all finite.*

Proof. To say that $\langle s_n \colon n \in \mathbb{N} \rangle$ is bounded in $\mathbb{R}$ means that it is contained within some real interval $[-b, b]$, or equivalently that its absolute values $|s_n|$ have some real upper bound b:

$$(\forall n \in \mathbb{N}) \; |s_n| < b.$$

Then by upward transfer the extended sequence is contained in ${}^*[-b, b]$, i.e., $|s_n| < b$ for all $n \in {}^*\mathbb{N}$; hence the term s_n is finite in general.

For the converse, suppose s_n is finite for all infinite $n \in {}^*\mathbb{N}_\infty$.

Then it is finite for all $n \in {}^*\mathbb{N}$.

Choose a positive infinite hyperreal $r \in {}^*\mathbb{R}^+_\infty$. Then the entire extended sequence lies in the interval $\{x \in {}^*\mathbb{R} \colon -r < x < r\}$ and we can therefore apply transfer.

Namely, we have $|s_n| < r$ for all $n \in {}^*\mathbb{N}$, so the sentence

$$(\exists y \in {}^*\mathbb{R})(\forall n \in {}^*\mathbb{N}) \; |s_n| < y$$

is true. But then by downward transfer it follows that there is some real number that is an upper bound to $|s_n|$ for all $n \in \mathbb{N}$. $\square$

DEFINITION 6.2.2. We say that $\langle s_n \rangle$ *diverges to infinity* if for each real r there is some $n \in \mathbb{N}$ such that all terms of the standard tail $s_n, s_{n+1}, s_{n+2}, \ldots$ are greater than r. Correspondingly, $\langle s_n \rangle$ diverges to minus infinity if $\langle -s_n \rangle$ diverges to infinity.

COROLLARY 6.2.3. *A real-valued sequence*

(1) *diverges to infinity if and only if all of its extended terms are positive infinite; and*
(2) *diverges to minus infinity if and only if all of its extended terms are negative infinite.*

6.3. Cauchy sequences

The traditional definition of a Cauchy sequence $\langle s_n \rangle$ is one that satisfies

$$\lim_{m,n\to\infty} |s_n - s_m| = 0,$$

meaning that the terms get arbitrarily close to each other as we move along the sequence. Formally this is rendered by the sentence

$$(\forall \epsilon \in \mathbb{R}^+)(\exists j \in \mathbb{N})(\forall m, n \in \mathbb{N})\,(m, n \geq j \to |s_m - s_n| < \epsilon).$$

COROLLARY 6.3.1. *A real-valued sequence $\langle s_n \rangle$ is Cauchy in $\mathbb{R}$ if and only if all its extended terms are infinitely close to each other, i.e., iff $s_m \approx s_n$ for all $m, n \in {}^*\mathbb{N}_\infty$.*

We will give a proof via infinitesimals of the Cauchy convergence criterion.

THEOREM 6.3.2 (Cauchy's Convergence Criterion). *A real-valued sequence converges in $\mathbb{R}$ if and only if it is Cauchy.*

PROOF. Suppose $\langle s_n : n \in \mathbb{N} \rangle$ is Cauchy. Then it is bounded. Take an infinite number $m \in {}^*\mathbb{N}_\infty$. Then s_m is finite and so it has a shadow $L \in \mathbb{R}$. But all extended terms of the sequence are infinitely close to each other, hence are infinitely close to s_m, and therefore are infinitely close to L as $s_m \approx L$. This shows that the extended tail of the sequence is contained in the halo of L, implying that $\langle s_n \rangle$ converges to $L \in \mathbb{R}$.

Conversely, suppose the sequence converges to L. By the hyperreal criterion for convergence, all of its extended terms are in $hal(L)$. By the triangle inequality, all the extended terms are infinitely close to each other. By Corollary 6.3.1, the sequence is Cauchy. $\square$

REMARK 6.3.3. The assertion that Cauchy sequences converge is often taken as an "axiom" for the real number system, and is equivalent to the Dedekind completeness assertion that sets that are bounded above have least upper bounds in $\mathbb{R}$. We used Dedekind completeness to prove the existence of shadows. The existence of shadows in turn implies convergence of Cauchy sequences.

6.4. Bolzano–Weierstrass theorem

We will give a direct proof of the Bolzano–Weierstrass theorem using infinitesimals.

THEOREM 6.4.1 (Bolzano–Weierstrass). *Every bounded sequence of real numbers has a cluster point*[1] *in* $\mathbb{R}$.

Recall the following.

DEFINITION 6.4.2. A real number L is a cluster point of the real-valued sequence $\langle s_n : n \in \mathbb{N}\rangle$ if each open interval $(L-e, L+e)$ in $\mathbb{R}$ contains infinitely many terms of the sequence.

This is expressed by the sentence

$$(\forall c \in \mathbb{R}^+)(\forall m \in \mathbb{N})(\exists n \in \mathbb{N})\ (n > m \wedge |s_n - L| < e). \tag{6.4.1}$$

From this it can be shown that the original sequence has a subsequence converging to L. Cluster points are also known as limit points of the sequence.

THEOREM 6.4.3. *A number $L \in \mathbb{R}$ is a cluster point of the real-valued $\langle s_n : n \in \mathbb{N}\rangle$ if and only if the sequence has an extended term infinitely close to L, i.e., iff $s_H \approx L$ for some infinite H.*

PROOF. Assume that (6.4.1) holds. Let e be a positive infinitesimal and $m \in {}^*\mathbb{N}_\infty$. Then by transfer of (6.4.1), there is some $n \in {}^*\mathbb{N}$ with $n > m$, and hence n is infinite, and

$$|s_n - L| < e \approx 0.$$

Thus s_n is an extended term infinitely close to L. (Indeed, the argument shows that any interval of infinitesimal width around L contains terms arbitrarily far along the extended tail.)

Conversely, suppose there is an infinite H with $s_H \approx L$. To prove the sentence (6.4.1), take any positive $\epsilon \in \mathbb{R}$ and $m \in \mathbb{N}$. Then $H > m$ and $|s_n - L| < \epsilon$. This shows that

$$(\exists n \in {}^*\mathbb{N})(n > m \ \wedge\ |s_n - L| < \epsilon).$$

By downward transfer, $|s_n - L| < \epsilon$ for some standard integer $n \in \mathbb{N}$, with $n > m$. $\square$

[1]nekudat hitztabrut

This characterisation shows that a shadow of an extended term is a cluster point of a real sequence, and indeed that the cluster points are precisely the shadows of those extended terms that have them, i.e., are finite. But if the sequence is bounded, then all of its extended terms are finite and so have shadows that must be cluster points.

Proof of Bolzano–Weierstrass theorem. Consider an extended term s_H of the sequence. Take its shadow $L = sh(s_H)$. Then L is a cluster point of the sequence. □

6.5. Series

A real infinite series $\sum_1^\infty a_i$ is convergent if and only if the sequence $s = \langle s_n : n \in \mathbb{N}\rangle$ of partial sums

$$s_n = a_1 + \cdots + a_n$$

is convergent. We will write $\sum_1^n a_i$ for s_n , and $\sum_m^n a_i$ for $s_n - s_{m-1}$ when $n \geq m$.

Extending s to a hypersequence $\langle s_n : n \in {}^*\mathbb{N}\rangle$, we obtain that s_n and s_{m-1} are defined for all hyperintegers n, m, so the expressions $\sum_1^n a_i$ and $\sum_m^n a_i$ become meaningful for all $n, m \in {}^*\mathbb{N}$, and may be thought of as hyperfinite sums when n is infinite. Applying our results on convergence of sequences to the sequence of partial sums, we have:

- $\sum_1^\infty a_i = L$ in $\mathbb{R}$ iff $\sum_1^n a_i \approx L$ for all infinite n;
- the series $\sum_1^\infty a_i$ converges in $\mathbb{R}$ iff $\sum_m^n a_i \approx 0$ for all infinite m, n with $m \leq n$.

The second of these is given by the Cauchy convergence criterion (Theorem 6.3.2), since $\sum_m^n a_i \approx 0$ iff $s_n \approx s_{m-1}$ for infinite m, n. Taking the case $m = n$ here, we get that if the series $\sum_1^\infty a_i = L$ converges, then $a_n \approx 0$ whenever n is infinite.

Corollary 6.5.1. *if $\sum_1^\infty a_i$ converges, then $\lim_{i\to\infty} a_i = 0$.*

Corollary 6.5.2. *For a convergent real series we have*

$$\sum_1^\infty a_i = sh \sum_1^n a_i = L$$

for any infinite n.

6.6. Continuous functions

Let f be an $\mathbb{R}$-valued function defined on an open interval (a, b) of $\mathbb{R}$. In passing to ${}^*\mathbb{R}$, we may regard f as being defined for all hyperreal x between a and b, since ${}^*(a, b) = \{x \in {}^*\mathbb{R} : a < x < b\}$.

Informally, we describe the assertion

f is continuous at a point c in the interval (a, b)

as meaning that $f(x)$ stays "close to" $f(c)$ whenever x is "close to" c. The way Cauchy put it in 1821 was that

> the function $f(x)$ is continuous with respect to x between the given bounds if between these bounds an infinitely small increase in the variable always produces an infinitely small increase in the function.

From the enlarged perspective of ${}^*\mathbb{R}$, this account can be made precise as follows.

THEOREM 6.6.1. *f is continuous at the real point c if and only if $f(x) \approx f(c)$ for all $x \in {}^*\mathbb{R}$ such that $x \approx c$, i. e., iff*

$$f(hal(c)) \subseteq hal(f(c)).$$

PROOF. The standard definition is that f is continuous at c iff for each open interval $(f(c) - \epsilon, f(c) + \epsilon)$ around $f(c)$ in $\mathbb{R}$ there is a corresponding open interval $(c - \delta, c + \delta)$ around c that is mapped into $(f(c) - c, f(c) + c)$ by f. Since $a < c < b$, the number δ can be chosen small enough so that the interval $(c - \delta, c + \delta)$ is contained within (a, b), ensuring that f is indeed defined at all points that are within δ of c.

Continuity at c is thus formally expressed by the sentence

$$(\forall \epsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)(\forall x \in \mathbb{R})\,(|x - c| < \delta \to |f(x) - f(c)| < \epsilon). \quad (6.6.1)$$

Now suppose $x \approx c$ implies $f(x) \approx f(c)$. To show that (6.6.1) holds, let ϵ be a positive real number. Then we have to find a real δ small enough to fulfill (6.6.1). First we show that this can be achieved if "small enough" is replaced by "infinitely small", and then apply transfer.

Let d is any positive infinitesimal.

Then for any $x \in {}^*\mathbb{R}$, if $|x - c| < d$, we have $x \approx c$, hence $f(x) \approx f(c)$ by assumption, so $|f(x) - f(c)| < \epsilon$, as ϵ is real. Replacing d by an existentially quantified variable, this shows that the sentence

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

is true. By downward transfer we then infer

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

proving (6.6.1).

Conversely, assume that (6.6.1) holds. Let ϵ be any positive real. Then by (6.6.1) there is a positive $\delta \in \mathbb{R}$ such that the sentence

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

is true, and hence by upward transfer we have

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

But now if $x \approx c$ in ${}^*\mathbb{R}$, then $|x - c| < \delta$, and so by this last sentence

$$|f(x) - f(c)| < \epsilon.$$

Since this holds for arbitrary $\epsilon \in \mathbb{R}^+$, it follows that $f(x) \approx f(c)$.

In other words, the halo $hal(c)$ is mapped by f into the interval $(f(c) - \epsilon, f(c) + \epsilon)$ for any positive real ϵ, and hence is mapped into the halo $hal(f(c))$. □

An inspection of the first part of this proof reveals that in order to establish the standard criterion for continuity at a point c it suffices to know that $f(x) \approx f(c)$ for all x that are within some positive infinitesimal distance d of c. Thus we have this stronger conclusion.

Corollary 6.6.2. *The following are equivalent.*

1. f *is continuous at* $c \in \mathbb{R}$*;*
2. $f(x) \approx f(c)$ *whenever* $x \approx c$*.*
3. *There is some positive* $d \approx 0$ *such that* $f(x) \approx f(c)$ *whenever* $|x - c| < d$*.*

6.7. Continuity in A

If A is a subset of the domain of a function f, then f is continuous on the set A if it is continuous at all points c that belong to A. Sometimes we would like A to be something other than an open interval (a, b), such as a halfopen or closed interval $(a, b]$, $[a, b)$, or $[a, b]$, or a union of such sets. In this case the definition of continuity is modified to specify that for each positive ϵ there is a corresponding δ such that $f(x)$ belongs to $(f(c) - \epsilon, f(c) + \epsilon)$ whenever x is a point of A that belongs to

$$(c - \delta, c + \delta).$$

In other words, the bounded quantification of x in sentence (6.6.1) is restricted to the set A, and we say that f is continuous at all points $c \in A$ if

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

Definition 6.7.1. The formula $\Phi(\epsilon, f, A)$ with free variables ϵ, f, A is the formula

$$(\forall c \in A)(\exists \delta \in \mathbb{R}^+)(\forall x \in A)\ (|x - c| < \delta \to |f(x) - f(c)| < \epsilon). \quad (6.7.1)$$

Then continuity of f on A is expressed by the formula

$$(\forall \epsilon \in \mathbb{R}^+)\, \Phi(\epsilon, f, A).$$

Such a reformulation will be useful in analyzing uniform continuity in Section 7.1.

Reworking the proofs of the above theorem and corollary, we obtain the following hyperreal characterisation of continuity on A.

THEOREM 6.7.2. *The following are equivalent.*

(1) *f is continuous at c in A.*
(2) *$f(x) \approx f(c)$ for all $x \in {}^*\!A$ with $x \approx c$.*
(3) *There is some positive $d \approx 0$ such that $f(x) \approx f(c)$ for all $x \in {}^*\!A$ with $|x - c| < d$.*

It would be natural at this point to ask whether continuity of f on A entails that the condition $f(hal(c)) \subseteq hal(f(c))$ must hold for all points $c \in {}^*\!A$ and not just the real ones. It turns out that this need not be so: it is a stronger requirement, which, remarkably, is equivalent to the standard notion of uniform continuity. We take this up in Section 7.1.

6.8. Continuity of the sine function

To illustrate the use of Theorem 6.6.1, let c be real and $x \approx c$. Then $x = c + \epsilon$ for an infinitesimal ϵ, and

$$\begin{aligned}
\sin x - \sin c &= \sin(c + \epsilon) - \sin c \\
&= \sin c \cos \epsilon + \cos c \sin \epsilon - \sin c \\
&= \sin c(\cos \epsilon - 1) + \cos c \sin \epsilon \\
&= \text{an infinitesimal},
\end{aligned}$$

since $\cos c \approx 1$ and $\sin \epsilon \approx 0$ while $\sin c$ and $\cos c$ are real. Hence $\sin x \approx \sin c$. This proves that the sine function is continuous at all $c \in \mathbb{R}$.

Note that in this proof we used the addition formula

$$\sin(c + \epsilon) = \sin c \cos \epsilon + \cos c \sin \epsilon.$$

This holds for all real numbers, and hence by transfer it holds for all hyperreals.

6.9. The Intermediate Value Theorem

This fundamental result of standard real analysis states the following.

Theorem 6.9.1. *If the real function f is continuous on the closed interval $[a, b]$ in $\mathbb{R}$, then for every real number d strictly between $f(a)$ and $f(b)$ there exists a real $c \in (a, b)$ such that $f(c) = d$.*

There is an intuitively appealing proof of this using infinitesimals. The basic idea is to partition the interval $[a, b]$ into subintervals of equal infinitesimal width, and locate a subinterval whose end points have f-values on either side of d. Then c will be the common shadow of these end points. In this way we "pin down" the point at which the f-values pass through d.

Proof. We deal with the case $f(a) < f(b)$, so that $f(a) < d < f(b)$. First, for each finite $n \in \mathbb{N}$, partition $[a, b]$ into n equal subintervals of width $(b - a)/n$. Thus these intervals have end points

$$p_k = a + \frac{k(b-a)}{n}$$

for $0 \leq k \leq n$. Then let s_n be the greatest partition point whose f-value is less than d. Indeed, the set

$$\{p_k : f(p_k) < d\}$$

is finite and nonempty (it contains $p_0 = a$ but not $p_n = b$). Hence s_n exists as the maximum of this set, and is given by some p_k with $k < n$.

Now, for all $n \in \mathbb{N}$ we have

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

and so by transfer, these conditions hold for all $n \in {}^*\mathbb{N}$. To obtain an infinitesimal-width partition, choose an infinite hypernatural H. Then s_H is finite, as $a \leq s_H < b$, so has a shadow $c = sh(s_H) \in \mathbb{R}$.

Here by transfer, s_H is a number of the form $a + K(b - a)/H$ for some $K \in {}^*\mathbb{N}$). But $(b-a)/H$ is infinitesimal, so s_H and $s_H + (b-a)/H$ are both infinitely close to c. Since f is continuous at c and c is real, it follows by Theorem 6.6.1 that $f(s_H)$ and $f(s_H + (b - a)/H)$ are both infinitely close to $f(c)$. But

$$f(s_H) < d \leq f(s_H + (b-a)/H),$$

so d is also infinitely close to $f(c)$. Since $f(c)$ and d are both real, they must then be equal. □

6.10. The Extreme Value Theorem

Theorem 6.10.1. *If the real function f is continuous on the closed interval $[a, b]$ in $\mathbb{R}$, then f attains an absolute maximum and an absolute minimum on $[a, b]$, i.e., there exist real $c, d \in [a, b]$ such that*

$$f(c) \leq f(x) \leq f(d) \text{ for all } x \in [a, b].$$

PROOF. To obtain the asserted maximum we construct an infinitesimal width partition of $[a, b]$, and show that there is a particular partition point whose f-value is as big as any of the others. Then d will be the shadow of this particular partition point. As with the intermediate value theorem, the construction is first approximated by finite partitions with subintervals of standard width $\frac{1}{n}$. In these cases there is always a partition point with maximum f-value. Then transfer is applied.

For each finite $n \in \mathbb{N}$, we partition $[a, b]$ into n equal subintervals, with end points $a + k(b - a)/n$ for $0 \leq k \leq n$. Then let $s_n \in [a, b]$ be a partition point at which f takes its largest value. In other words, for all integers k such that $0 \leq k \leq n$,

$$a \leq s_n \leq b \text{ and } f(a + k(b - a)/n) \leq f(s_n) \tag{6.10.1}$$

By transfer, (6.10.1) holds for all $n \in {}^*\mathbb{N}$ and all hyperintegers k such that

$$0 \leq k \leq n.$$

Similarly to the intermediate value theorem, choose an infinite hypernatural N and put $d = sh(s_N) \in \mathbb{R}$. Then by continuity

$$f(s_N) \approx f(d). \tag{6.10.2}$$

DEFINITION 6.10.2. The *infinitesimal-width partition* P is

$$P = \{a + k(b - a)/N : k \in {}^*\mathbb{N} \text{ and } 0 \leq k \leq N\}$$

The partition P has the important property that it provides infinitely close approximations to all real numbers between a and b. The halo of each $x \in [a, b]$ contains points from this partition. To show this, observe that if x is an arbitrary real number in $[a, b]$, then for each $n \in \mathbb{N}$ there exists an integer $k < n$ with

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

Hence by transfer there exists a hyperinteger $K < N$ such that x lies in the interval

$$\left[a + \frac{K(b - a)}{N}, a + \frac{(K + 1)(b - a)}{N}\right]$$

of infinitesimal width $(b-a)/N$. Therefore $x \approx a+K(b-a)/N$, so x is indeed infinitely close to a member of P. It follows by continuity of f at x that

$$f(x) \approx f\left(\frac{a + K(b - a)}{N}\right) \tag{6.10.3}$$

But the values of f on P are dominated by $f(s_N)$, as (6.10.1) holds for all $n \in {}^*\mathbb{N}$, so

$$f(a + K(b-a)/N) \leq f(s_N) \tag{6.10.4}$$

Putting (6.10.2), (6.10.3), and (6.10.4) together gives

$$f(x) \approx f\left(a + \frac{K(b-a)}{N}\right) \leq f(s_N) \approx f(d),$$

which implies $f(x) \leq f(d)$, since $f(x)$ and $f(d)$ are real. Thus f attains its maximum value at d. The proof that f attains a minimum is similar. $\square$

CHAPTER 7

Uniform continuity and convergence, derivatives

7.1. Uniform Continuity, LSEQ operator

Let $A \subseteq \mathbb{R}$ and $f : A \to \mathbb{R}$. Traditionally, the uniform continuity of f on A is defined as follows.

DEFINITION 7.1.1. The function f is *uniformly continuous* on A if the following sentence is true:

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

This definition should be compared to the formal sentence just prior to Theorem 6.7.2. Essentially, this says that for a given ε, the same δ for the continuity condition works at all points of A. More precisely, consider the formula $\Phi(\epsilon, f, A)$ of (6.7.1):

$$(\forall c \in A)(\exists \delta \in \mathbb{R}^+)(\forall x \in A)\,(|x-c| < \delta \to |f(x) - f(c)| < \epsilon). \tag{7.1.1}$$

Continuity of f in A is expressed by

$$(\forall \epsilon \in \mathbb{R}^+)\,\Phi(\epsilon, f, A). \tag{7.1.2}$$

DEFINITION 7.1.2 (LSEQ). Let Ψ be a formula containing one occurrence of the existence quantifier. The transformation of Ψ, denoted LSEQ(Ψ) (for "leftward shift of the existence quantifier") consists in shifting the existence quantifier in Ψ all the way to the left of the formula.

EXAMPLE 7.1.3. Applied to formula Φ of (7.1.1), LSEQ produces formula LSEQ(Φ) given by

$$(\exists \delta \in \mathbb{R}^+)(\forall c \in A)(\forall x \in A)\,(|x-c| < \delta \to |f(x) - f(c)| < \epsilon).$$

Then uniform continuity of f on A can be presented as follows:

$$(\forall \varepsilon \in \mathbb{R}^+)\,\mathrm{LSEQ}(\Phi(\epsilon, f, A)).$$

This formula strengthens the condition of continuity (7.1.2).

The LSEQ operator may be useful in clarifying the axiom of Idealisation.[1]

THEOREM 7.1.4. *A function f is uniformly continuous on A if and only if $x \approx y$ implies $f(x) \approx f(y)$ for all hyperreals $x, y \in {}^*\!A$.*

Proof. Exercise.[2]

REMARK 7.1.5. This theorem displays the distinction between uniform and ordinary continuity in a more intuitive and readily comprehensible way than the traditional definitions do. For by Theorem 7.1.4, f is continuous on $A \subseteq \mathbb{R}$ iff $x \approx y$ implies $f(x) \approx f(y)$ for $x, y \in {}^*\!A$ with y standard. Thus uniform continuity amounts to preservation of the "infinite closeness" relation $\approx$ at all hyperreal points in the enlargement ${}^*\!A$ of A, while continuity only requires preservation of this relation at the real points.

THEOREM 7.1.6. *If the real function f is continuous on the closed interval $[a, b] \subseteq \mathbb{R}$, then f is uniformly continuous on $[a, b]$.*

PROOF. Take hyperreals $x, y \in {}^*[a, b]$ with $x \approx y$. Let $c = sh(x)$. Then since $a \le x \le b$ and $x \approx c$, we have $c \in [a, b]$. By hypothesis, f is continuous at c. Applying Theorem 6.6.1, we get $f(x) \approx f(c)$ and $f(y) \approx f(c)$, whence $f(x) \approx f(y)$. Hence f is uniformly continuous by Theorem 7.1.4. $\square$

7.2. Permanence principles

One of the distinctive features of nonstandard analysis is the presence of so-called permanence principles, which assert that

> certain functions must exist, or be defined, on a larger domain than that which is originally used to define them.

For instance, any real function $f : A \to \mathbb{R}$ automatically extends to the enlargement ${}^*\!A$ of its real domain A. In discussing continuity of a real function f at a real point c, we may want (the extension

[1] Consider the following axiom of Countable Idealisation (CI) in an axiomatic set theory such as IST, BST, or SPOT (here CI is somewhat analogous to saturation; see Section 9.5). Let ϕ be an $\in$-formula with arbitrary parameters. Then

$$\forall^{\mathbf{st}} n \in \mathbb{N}\ \exists x\ \forall m \in \mathbb{N}\ (m \le n \to \phi(m, x)) \ \leftrightarrow\ \exists x\ \forall^{\mathbf{st}} n \in \mathbb{N}\ \phi(n, x). \tag{7.1.3}$$

Now let Ψ be the formula $\forall^{\mathbf{st}} n \in \mathbb{N}\ \exists x\ \forall m \in \mathbb{N}\ (m \le n \to \phi(m, x))$. The axiom of Countable Idealisation can then be reformulated as the implication $\Phi \to \mathrm{LSEQ}(\Psi)$. Here $\mathrm{LSEQ}(\Psi)$ is of course equivalent to the right-hand side of (7.1.3).

[2] This is proved in Section 7.1 of `https://u.math.biu.ac.il/~katzmik/tidg.pdf`

of) f to be defined at points infinitely close to c. For this it suffices that f be defined on some real neighbourhood $(c-\varepsilon, c+\varepsilon)$ in $\mathbb{R}$, for then the domain of the extension of f includes the enlarged interval ${}^*(c-\varepsilon, c+\varepsilon)$, which contains the halo $hal(c)$ of c. But the converse of this is also true:

PROPOSITION 7.2.1. *If the extension of f is defined on $hal(c)$, then f must be defined on some real interval of the form $(c-\varepsilon, c+\varepsilon)$, and hence on ${}^*(c-\varepsilon, c+\varepsilon)$.*

In fact, it can be shown that for this last conclusion it suffices that f be defined on some hyperreal interval $(c-d, c+d)$ of infinitesimal radius d. This is our first example of a permanence statement that is sometimes called Cauchy's principle. It asserts that if a property holds for all points within some infinitesimal distance of c, then it must actually hold for all points within some real (hence appreciable) distance of c.

PROOF OF PROPOSITION 7.2.1. At present we can show this for the transforms of properties expressible in the formal language $\mathcal{L}_{\mathcal{R}}$. Let $\phi(x)$ be a formula of this language for which there is some positive $d \approx 0$ such that

$$^*\!\phi(x) \text{ is true for all hyperreal } x \text{ with } c-d < x < c+d.$$

Then the sentence

$$(\exists y \in {}^*\mathbb{R}^+)(\forall x \in {}^*\mathbb{R})\ (|x-c| < y \rightarrow {}^*\!\phi)$$

is seen to be true by interpreting y as d. But then by downward transfer there is some real $\varepsilon > 0$ such that

$$(\forall x \in \mathbb{R})(|x-c| < \varepsilon \rightarrow \phi),$$

so that ϕ is true throughout $(c-\varepsilon, c+\varepsilon)$ in $\mathbb{R}$. Hence by upward transfer back to ${}^*\mathbb{R}$,

$$(\forall x \in {}^*\mathbb{R})(|x-c| < \epsilon \rightarrow {}^*\!\phi),$$

showing that

$$^*\!\phi(x) \text{ is true for all hyperreal } x \text{ with } c-\varepsilon < x < c+\varepsilon.$$

This completes the proof of the proposition. $\square$

REMARK 7.2.2. In this argument c is a real number. Later it will be shown that permanence works for any hyperreal number in place of c, and applies to a broader class of properties than those expressible in the language L_R.

7.3. Pointwise and uniform convergence of $\langle f_n \rangle$

Let $\langle f_n : n \in \mathbb{N} \rangle$ be a sequence of functions $f_n : A \to \mathbb{R}$ defined on some subset $A \subseteq \mathbb{R}$.

Definition 7.3.1. The sequence is said to converge pointwise to the function $f : A \to \mathbb{R}$ if for each $x \in A$ the $\mathbb{R}$-valued sequence $\langle f_n(x) : n \in \mathbb{N} \rangle$ converges to the number $f(x)$.

Symbolically, this asserts that

$$(\forall x \in A) \lim_{n\to\infty} f_n(x) = f(x),$$

which is rendered in full by the sentence

$$(\forall x \in A)(\forall \varepsilon \in \mathbb{R}^+)(\exists m \in \mathbb{N})(\forall n \in \mathbb{N})\ (n > m \to |f_n(x) - f(x)| < \varepsilon).$$

In this statement, the integer m that is asserted to exist depends on the choice of $x \in A$ as well as on ε. The stronger property is that of uniform convergence.

Definition 7.3.2. We say that $\langle f_n : n \in \mathbb{N} \rangle$ converges uniformly to the function f if m depends only on ε in the sense that for a given ε, the same m works for all $x \in A$:

$$(\forall \varepsilon \in \mathbb{R}^+)(\exists m \in \mathbb{N})(\forall x \in A)(\forall n \in \mathbb{N})(n > m \to |f_n(x) - f(x)| < \varepsilon).$$

Now, we know how to extend a sequence of numbers to a hypersequence but at this point we would like to do the same for a sequence of functions. For $n \in \mathbb{N}$, the function f_n extends to a function with domain ${}^*\!A$, but we would like to define $f_n : {}^*\!A \to {}^*\mathbb{R}$ also for infinite rank n. To achieve this we first identify the original sequence $(f_n : n \in \mathbb{N})$ of functions with the single function

$$F : \mathbb{N} \times A \to \mathbb{R}$$

defined by setting $F(n, x) = f_n(x)$ for all $n \in \mathbb{N}$ and $x \in A$. This function F has an extension

$${}^*F : {}^*\mathbb{N} \times {}^*\!A \to {}^*\mathbb{R},$$

which can then be used to define $f_n : {}^*\!A \to {}^*\mathbb{R}$ by setting $f_n(x) = {}^*F(n, x)$.

Thus we now have a hypersequence of functions $(f_n : n \in {}^*\mathbb{N})$ as required.

Lemma 7.3.3. *For each standard integer $n \in \mathbb{N}$, the new construction of f_n just reproduces the extension of the original function f_n.*

PROOF. Let n be a standard integer. We apply upward transfer to

$$(\forall x \in A)\ (f_n(x) = F(n, x)),$$

proving the lemma. □

Moreover, for each $x \in A$, the real-number sequence $s = \langle f_n(x) : n \in \mathbb{N}\rangle$ has as its extension the hypersequence $\langle f_n(x) : n \in {}^*\mathbb{N}\rangle$. This follows by transfer of

$$(\forall n \in \mathbb{N})(s(n) = F(n, x)).$$

In view of the characterisation of converging number sequences given above we can thus immediately infer the following result.

THEOREM 7.3.4 (Characterisation of pointwise convergence). *The sequence $(f_n : n \in \mathbb{N})$ of real-valued functions defined on $A \subseteq \mathbb{R}$ converges pointwise to the function $f : A \to \mathbb{R}$ if and only if for each $x \in A$ and each infinite $n \in {}^*\mathbb{N}$, one has $f_n(x) \approx f(x)$.*

On the other hand, we have the following characterisation of uniform convergence.

THEOREM 7.3.5. *A sequence $(f_n : n \in \mathbb{N})$ converges uniformly to the function $f : A \to \mathbb{R}$ if and only if for each $x \in {}^*A$ and each infinite $n \in {}^*\mathbb{N}$, $f_n(x) \approx f(x)$.*

Proof. Exercise.

The ideas underlying this characterisation are well illustrated by the following example.

EXAMPLE 7.3.6. Consider the behaviour of the sequence $(f_n : n \in \mathbb{N})$ given by $f_n(x) = x^n$ on $A = [0, 1]$. This converges pointwise to the function f that is constantly zero on $[0, 1)$ and has $f(1) = 1$:

$$f(x) = \begin{cases} 0 \text{ if } 0 \le x < 1 \\ 1 \text{ if } x = 1 \end{cases}$$

Thus when $x < 1$, the sequence $(x^n : n \in \mathbb{N})$ converges to 0, but as x moves towards 1 the rate of convergence slows down, in the sense that for a fixed $\varepsilon \in \mathbb{R}^+$, as x approaches 1 we have to move further and further along the sequence of powers of x before reaching a point where the terms are less than ε. Ultimately, when x becomes infinitely close to 1 (but still less than 1), it takes “infinitely long” for x^n to become infinitely close to 0. Indeed, by transferring the statement of pointwise convergence and taking ε to be a positive infinitesimal, it follows that there will be some $M \in {}^*\mathbb{N}$ such that for $n > M$ we have $x^n < \varepsilon$ and hence $x^n \approx 0$ for this fixed x infinitely close to 1. Note that since $x < 1$ we have $f(x) = 0$.

Now, this M will be infinite, because when n is finite, $x \approx 1$ implies $x^n \approx 1$. Hence

> the set $\{x^n : n \in \mathbb{N}\}$ is contained entirely within the halo of 1.

But there is a permanence principle[3] that concludes from this that

> there is some infinite H such that $\{x^n : n \leq H\}$ is contained in the halo of 1

(cf. Robinson's sequential lemma to be treated below). In particular, $x^H \not\approx 0$, i.e.,

$$f_H(x) \not\approx f(x),$$

showing that the condition of Theorem 7.3.5 is violated, and therefore that the original standard sequence $(f_n : n \in \mathbb{N})$ is not uniformly convergent to f.

7.4. Continuity of a Uniform Limit

A sequence $\langle f_n : n \in \mathbb{N}\rangle$ of continuous functions can converge pointwise to a discontinuous function. We have just discussed the standard example: take $f_n(x) = x^n$ on $A = [0, 1]$. Under the hypothesis of uniform convergence this phenomenon cannot occur. We now present a hyperreal approach to this classical result.

Theorem 7.4.1. *If the functions $\langle f_n : n \in \mathbb{N}\rangle$ are all continuous on $A \subseteq \mathbb{R}$, and the sequence converges uniformly to the function $f : A \to \mathbb{R}$, then f is continuous on A.*

Proof. Let c belong to A. To prove that f is continuous at c, we invoke Theorem 6.7.2 (hyperreal characterisation of continuity). If $x \in {}^*\!A$ with $x \approx c$, we need to prove $f(x) \approx f(c)$, i.e., $|f(x) - f(c)| < \varepsilon$ for any positive real ε. The key to this is to analyse the inequality

$$|f(x) - f(c)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(c)| + |f_n(c) - f(c)|. \quad (7.4.1)$$

On the right-hand side of the inequality, the middle term $|f_n(x) - f_n(c)|$ will be infinitesimal for any $n \in \mathbb{N}$ because $x \approx c$ and f_n is continuous at c.

We will show that by taking a large enough n, the first and last terms on the right can be made small enough that the sum of the three terms is less than ε.

[3]The property of lying in a halo is not expressible in the language L_R that we have developed so far.

To see how this works in detail, for a given $\varepsilon \in \mathbb{R}^+$ we apply the definition of uniform convergence as follows. We apply the definition to the number $\varepsilon/4$ to get that there is some integer $m \in \mathbb{N}$ such that

$$n > m \text{ implies } |f_n(x) - f(x)| < \varepsilon/4$$

for all $n \in \mathbb{N}$ and all $x \in A$, and hence for all $n \in {}^*\mathbb{N}$ and all $x \in {}^*A$ by upward transfer.

Now fix n as a standard integer, say by setting $n = m + 1$. Then for any $x \in {}^*A$ with $x \approx c$ it follows, since $x, c \in {}^*A$, that

$$|f_n(x) - f(x)|,\ |f_n(c) - f(c)| < \varepsilon/4,$$

and so in (7.4.1) we get

$$|f(x) - f(c)| < \varepsilon/4 + \text{ infinitesimal } + \varepsilon/4 < \varepsilon$$

as required. □

REMARK 7.4.2. This proof is a combination of standard and nonstandard arguments: it uses the hyperreal characterisation of continuity of f_n and f, but the standard definition of uniform convergence of $(f_n : n \in \mathbb{N})$ rather than the characterisation given by Theorem 7.3.5.

7.5. The derivative

We come now to an examination – from the modern infinitesimal perspective – of the cornerstone concept of the calculus.

The derivative of a function f at a real number x is the real number $f'(x)$ that represents the rate of change of the function as it varies near x. Alternatively, it is the slope of the tangent to the graph of f at x. Formally it is defined as the number

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

THEOREM 7.5.1. *If f is defined at $x \in \mathbb{R}$, then the real number $L \in \mathbb{R}$ is the derivative of f at x if and only if for every nonzero infinitesimal ε, $f(x+\varepsilon)$ is defined and*

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

PROOF. Let $g(h) = \frac{f(x+h)-f(x)}{h}$ and apply the characterisation of the expression

$$\text{“}\lim_{h \to 0} g(h) = L\text{”} \tag{7.5.1}$$

given earlier.[4] □

[4]Specify where @@.

REMARK 7.5.2 (Shadow). When f is differentiable (i.e., has a derivative) at x, we have

$$f'(x) = sh\left(\frac{f(x+\varepsilon) - f(x)}{\varepsilon}\right)$$

for all infinitesimal $\varepsilon \neq 0$.

If (7.5.1) holds only for all positive infinitesimal ε, then L is the right-hand derivative of f at x, defined classically as

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

Similarly, if (7.5.1) holds for all negative $c \approx 0$, then L is the left-hand derivative given by the limit as $h \to 0^-$.

7.6. Increments and Differentials

Let Δx denote an arbitrary nonzero infinitesimal representing a change or increment in the value of variable x. The corresponding increment in the value of the function $y = f(x)$ at x is

$$\Delta y = f(x + \Delta x) - f(x).$$

To be explicit we should denote this increment by $\Delta y(x, \Delta x)$, since its value depends both on the value of x and the choice of the infinitesimal Δx. This notation will be exploited in (7.8.1). The more abbreviated notation is, however, convenient and suggestive. If f is differentiable at $x \in \mathbb{R}$, Theorem 7.5.1 implies that

$$\frac{\Delta y}{\Delta x} \approx f'(x),$$

so the quotient $\frac{\Delta y}{\Delta x}$ is finite. Hence as

$$\Delta y = \frac{\Delta y}{\Delta x} \Delta x,$$

it follows that the increment Δy in f is infinitesimal. Thus $f(x+\Delta x) \approx f(x)$ for all infinitesimal Δx, and this proves the following result.

THEOREM 7.6.1. *If f is differentiable at $x \in \mathbb{R}$, then f is continuous at x.*

DEFINITION 7.6.2. The *differential* of f at x corresponding to Δx is defined to be $dy = f'(x)\Delta x$.

Thus whereas Δy represents the increment of the "y-coordinate" along the graph of f at x, the differential dy represents the increment along the tangent line to this graph at x. Writing dx for Δx, the definition of dy yields

$$\frac{dy}{dx} = f'(x).$$

Now, since $f'(x)$ is finite and Δx is infinitesimal, it follows that dy is infinitesimal. Hence dy and Δy are infinitely close to each other. In fact, their difference is infinitely smaller than Δx, for if

$$\varepsilon = \frac{\Delta y}{\Delta x} - f'(x),$$

then ε is infinitesimal, because $\frac{\Delta y}{\Delta x} \approx f'(x)$, and

$$\Delta y - dy = \Delta y - f'(x)\Delta x = \varepsilon \Delta x,$$

which is also infinitesimal (being a product of infinitesimals). But

$$\frac{\Delta y - dy}{\Delta x} = \frac{\varepsilon \Delta x}{\Delta x} \approx 0,$$

and in this sense $\Delta y - dy$ is infinitesimal compared to Δx. These relationships are summarised in the following theorem.

THEOREM 7.6.3. *[Incremental Equation] If $f'(x)$ exists at real x and $\Delta x = dx$ is infinitesimal, then Δy and dy are infinitesimal, and there is an infinitesimal ε, dependent on x and Δx, such that*

$$\Delta y = f'(x)\Delta x + \varepsilon \Delta x = dy + \varepsilon dx,$$

and so

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

This last equation elucidates the role of the derivative function f' as the best linear approximation to the function f at x. For the graph of the linear function

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

gives the tangent to f at x when the origin is translated to the point $(x, 0)$, and $l(\Delta x)$ differs from $f(x + \Delta x)$ by the amount $\varepsilon \Delta x$, which we saw above is itself infinitely smaller than Δx when Δx is infinitesimal, and in that sense is "negligible".

7.7. Rules for Derivatives

THEOREM 7.7.1. *If f and g are differentiable at $x \in \mathbb{R}$, then so are the functions $f + g$, fg, and f/g, provided that $g(x) \neq 0$. Moreover,*

(1) $(f + g)'(x) = f'(x) + g'(x)$,
(2) $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$,
(3) $(f/g)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$.

Proof. We prove Leibniz's rule (2), and leave the others as exercises. If $\Delta x \neq 0$ is infinitesimal, then, by Theorem 7.5.1, $f(x+\Delta x)$ and $g(x+\Delta x)$ are both defined, and hence so is

$$(fg)(x+\Delta x) = f(x+\Delta x)\, g(x+\Delta x).$$

Then the increment of fg at x corresponding to Δx is

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

It follows that

$$\frac{\Delta(fg)}{\Delta x} = \frac{\Delta f}{\Delta x}g(x) + f(x)\frac{\Delta g}{\Delta x} + \Delta f\frac{\Delta g}{\Delta x} \approx f'(x)g(x) + f(x)g'(x) + 0,$$

since $\frac{\Delta f}{\Delta x} \approx f'(x)$, $\frac{\Delta g}{\Delta x} \approx g'(x)$, $\Delta f \approx 0$ and all quantities involved are finite. Hence by Theorem 7.5.1, the expression $f'(x)g(x) + f(x)g'(x)$ is the derivative of fg at x. □

7.8. Chain Rule

Theorem 7.8.1. *If f is differentiable at $x \in \mathbb{R}$, and g is differentiable at $f(x)$, then $g \circ f$ is differentiable at x with derivative $g'(f(x))f'(x)$.*

Proof. Let Δx be a nonzero infinitesimal. Then $f(x+\Delta x)$ is defined and $f(x+\Delta x) \approx f(x)$. But g is defined at all points infinitely close to $f(x)$, since $g'(f(x))$ exists, so $(g\circ f)(x+\Delta x) = g(f(x+\Delta x))$ is defined. □

Now let

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

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

be the increments of f and $g \circ f$ at x corresponding to Δx. Then Δf is infinitesimal, and

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

which shows, crucially, that

$\Delta(g\circ f)$ is also the increment of g at $f(x)$ corresponding to Δf.

In the full incremental notation, this reads

$$\Delta(g\circ f)(x,\Delta x) = \Delta g(f(x),\Delta f). \tag{7.8.1}$$

By the incremental equation (Theorem 7.6.3) for g, it then follows that there exists an infinitesimal ε such that

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

Hence

$$\frac{\Delta(g \circ f)}{\Delta x} = g'(f(x))\frac{\Delta f}{\Delta x} + \varepsilon\frac{\Delta f}{\Delta x} \approx g'(f(x))f'(x) + 0,$$

establishing that $g'(f(x))f'(x)$ is the derivative of $g \circ f$ at x.

7.9. Critical Point Theorem

THEOREM 7.9.1. *Let f have a maximum or a minimum at x on some real interval (a, b). If f is differentiable at x, then $f'(x) = 0$.*

PROOF. Suppose f has a maximum at x. By transfer,

$$f(x + \Delta x) \leq f(x)$$

for all infinitesimal Δx. Hence if ε is positive infinitesimal and δ is negative infinitesimal,

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

and so as $f'(x)$ is real, it must be equal to 0. The case of f having a minimum at x is similar. □

Using the critical point and extreme value theorems, the following results can be successively derived about a function f that is continuous on $[a, b] \subseteq \mathbb{R}$ and differentiable on (a, b). The proofs do not require any further reasoning about infinitesimals or limits.

THEOREM 7.9.2 (Rolle's Theorem). *If $f(a) = f(b) = 0$, then $f'(x) = 0$ for some $x \in (a, b)$.*

THEOREM 7.9.3 (Mean Value Theorem). *For some $x \in (a, b)$, $f'(x) = \frac{f(b)-f(a)}{b-a}$.*

THEOREM 7.9.4. *If f' is zero/positive/negative on (a, b), then f is constant/increasing/decreasing on $[a, b]$.*

7.10. Inverse function theorem

The material in this section is optional.

LEMMA 7.10.1. *The inverse $x = g(y)$ of a continuous strictly monotone function $f(x)$ is continuous at y.*

PROOF. Using the intermediate value theorem and monotonicity of f it can be shown that g is defined on some real open interval around y. Let Δy be a nonzero infinitesimal. Now, if $g(y+\Delta y)$ were not infinitely close to $g(y)$, then there would be a real number r on the x-axis strictly between them. But then, by monotonicity of would be a real number on the y-axis strictly between $y + \Delta y$ and y. Since y is real, this would mean that $y + \Delta y$ and y were an appreciable distance apart, which is not so. Hence

$$\Delta x = g(y + \Delta y) - g(y)$$

is infinitesimal and is nonzero. This establishes that g is continuous at y. □

THEOREM 7.10.2. *Let f be continuous and strictly monotone (increasing or decreasing) on (a, b), and suppose g is the inverse function of f. If f is differentiable at x in (a, b), with $f'(x) \neq 0$, then g is differentiable at $y = f(x)$, with $g'(y) = \frac{1}{f'(x)}$.*

PROOF. The result $g'(f(x)) = 1/f'(x)$ would follow easily by the chain rule applied to the equation $g(f(x)) = x$ if we knew that g was differentiable at $f(x)$. But that is what we have to prove!

We need to show that

$$\frac{g(y + \Delta y) - g(y)}{\Delta y} \approx \frac{1}{f'(x)}.$$

Observe that Δx is, by definition, the increment $\Delta g(y, \Delta y)$ of g at y corresponding to Δy. Since $g(y) = x$, the last equation gives

$$g(y + \Delta y) = x + \Delta x,$$

so

$$f(x + \Delta x) = f(g(y + \Delta y)) = y + \Delta y.$$

Hence

$$\Delta y = f(x + \Delta x) - f(x) = \Delta f,$$

the increment of f at x corresponding to Δx. Altogether we have

$$\frac{\Delta f(x, \Delta x)}{\Delta x} = \frac{\Delta y}{\Delta x}$$

and

$$\frac{\Delta g(y, \Delta y)}{\Delta y} = \frac{\Delta x}{\Delta y} = \frac{\Delta x}{\Delta f}.$$

Put more briefly, we have shown that

$$\frac{\Delta g}{\Delta y} = \frac{1}{\Delta f/\Delta x}.$$

To derive from this the conclusion $g'(y) = \frac{1}{f'(x)}$, we invoke the hypothesis that $f'(x) \neq 0$ (which is essential: consider what happens at $x = 0$

when $f(x) = x^3$). Since $sh(\Delta f/\Delta x) = f'(x)$, it follows that $\Delta f/\Delta x$ is appreciable. But then

$$sh\left(\frac{\Delta x}{\Delta f}\right) = sh\left(\frac{1}{\Delta f/\Delta x}\right) = \frac{1}{f'(x)}.$$

Therefore,

$$\frac{\Delta g}{\Delta y} = \frac{\Delta x}{\Delta y} \approx \frac{1}{f'(x)}.$$

Because Δy is an arbitrary nonzero infinitesimal, this establishes that the real number $1/f'(x)$ is the derivative of g at y, as required. $\square$

CHAPTER 8

Internal sets, external sets, and transfer

8.1. Internal sets

In the construction of ${}^*\mathbb{R}$ as an ultrapower in Section 2.9, each sequence of points $r = \langle r_n : n \in \mathbb{N} \rangle$ in $\mathbb{R}$ gives rise to the single point $[r]$ of ${}^*\mathbb{R}$, which we also denote by the more informative symbol $[r_n]$. Equality of ${}^*\mathbb{R}$-points is given by

$$[r_n] = [s_n] \text{ iff } \{n \in \mathbb{N} : r_n = s_n\} \in F.$$

This description works for other kinds of entities than points. We will show that

- a sequence of subsets of $\mathbb{R}$ determines a single subset of ${}^*\mathbb{R}$;
- a sequence of functions on $\mathbb{R}$ determines a single function on ${}^*\mathbb{R}$.

Definition 8.1.1. An *internal set* in ${}^*\mathbb{R}$ is given by the following construction. Given a sequence $\langle A_n : n \in \mathbb{N} \rangle$ of subsets $A_n \subseteq \mathbb{R}$, define a subset $[A_n] \subseteq {}^*\mathbb{R}$ by specifying, for each $[r_n] \in {}^*\mathbb{R}$,

$$[r_n] \in [A_n] \text{ iff } \{n \in \mathbb{N} : r_n \in A_n\} \in F.$$

It must be checked that this is a well-defined notion that does not depend on how points are named, which means that if $[r_n] = [s_n]$ then

$$\{n \in \mathbb{N} : r_n \in A_n\} \in F \text{ iff } \{n \in \mathbb{N} : s_n \in A_n\} \in F.$$

This is a slight extension of the argument given in Section 2.9.

Remark 8.1.2. A first application will be an internal well-order principle for ${}^*\mathbb{N}$; see Section 8.4.

Here are some examples.

Example 8.1.3. If $\langle A_n \rangle$ is a constant sequence with $A_n = A \subseteq \mathbb{R}$ for all $n \in \mathbb{N}$, then the internal set $[A_n]$ is just the enlargement ${}^*\!A$ of A defined in Section 3.1. Hence we may also denote ${}^*\!A$ as $[A]$.

Corollary 8.1.4. *The enlargement of any subset of* $\mathbb{R}$ *is an internal subset of* ${}^*\mathbb{R}$.

In particular, we see that ${}^*\mathbb{N}$, ${}^*\mathbb{Z}$, and ${}^*\mathbb{Q}$ and ${}^*\mathbb{R}$ itself are all internal, as is any finite subset $A \subseteq \mathbb{R}$, since in that case $A = {}^*\mathbb{R}$.

EXAMPLE 8.1.5. More generally, any finite set $X = \{[r_n^1], \dots, [r_n^k]\}$ of hyperreals is internal, for then $X = [A_n]$, where $A_n = \{r_n^1, \dots, r_n^k\}$.

EXAMPLE 8.1.6. If $a < b$ in ${}^*\mathbb{R}$, then the hyperreal open interval

$$(a, b) = \{x \in {}^*\mathbb{R} : a < x < b\}$$

is internal. Indeed, if $a = [a_n]$ and $b = [b_n]$, then (a, b) is the internal set defined by the sequence $\langle (a_n, b_n) : n \in \mathbb{N} \rangle$ of real intervals $(a_n, b_n) \subseteq \mathbb{R}$. This follows because

$$[a_n] < [r_n] < [b_n] \text{ iff } \{n \in \mathbb{N} : a_n < r_n < b_n\} \in F.$$

Similarly, the hyperreal intervals $(a, b]$, $[a, b)$, $[a, b]$, and $\{x \in {}^*\mathbb{R} : a < x\}$ are internal.

REMARK 8.1.7. If a is positive infinite, then each of these intervals is disjoint from $\mathbb{R}$, so none of them can be the enlargement ${}^*\!A$ of a set $A \subseteq \mathbb{R}$, since ${}^*\!A$ always includes the (real) members of A.

EXAMPLE 8.1.8. If $H \in {}^*\mathbb{N}$, then the set

$$\{k \in {}^*\mathbb{N} : k \le H\} = \{1, 2, \dots, H\}$$

is internal. If $H = [H_n]$, then this is the internal set $[A_n]$, where

$$A_n = \{k \in \mathbb{N} : k \le H_n\} = \{1, 2, \dots, H_n\}$$

(since $H \in {}^*\mathbb{N}$, we have $\{n : H_n \in \mathbb{N}\} \in F$, so we may as well assume $H_n \in \mathbb{N}$ for all $n \in \mathbb{N}$).

EXAMPLE 8.1.9. If $H = [H_n] \in {}^*\mathbb{N}$, then the set

$$\left\{\frac{k}{H} : k \in {}^*\mathbb{N} \cup \{0\} \text{ and } k \le H\right\} = \left\{0, \frac{1}{H}, \frac{2}{H}, \dots \frac{H-1}{H}, 1\right\}$$

is the internal set $[A_n]$, where

$$A_n = \left\{0, \frac{1}{H_n}, \frac{2}{H_n}, \dots, \frac{H_n - 1}{H_n}, 1\right\}$$

These last two examples illustrate the notion of hyperfinite set, which will be studied in Section 13.8.

8.2. Algebra of internal sets

THEOREM 8.2.1. *Internal sets have the following properties:*

(1) *The collection of internal sets is closed under the standard finite set operations* $\cap$, $\cup$, *and* $-$, *with*

$$[A_n] \cap [B_n] = [A_n \cap B_n],$$

$$[A_n] \cup [B_n] = [A_n \cup B_n],$$

$$[A_n] - [B_n] = [A_n - B_n].$$

(2) $[A_n] \subseteq [B_n]$ *iff* $\{n \in \mathbb{N} : A_n \subseteq B_n\} \in F$.
(3) $[A_n] = [B_n]$ *iff* $\{n \in \mathbb{N} : A_n = B_n\} \in F$.
(4) $[A_n] \neq [B_n]$ *iff* $\{n \in \mathbb{N} : A_n \neq B_n\} \in F$.

PROOF. (1) Exercise.

(2) If $[A_n] \not\subseteq [B_n]$, then there is some hyperreal $[r_n] \in [A_n] - [B_n]$, so by (1) we have

$$I = \{n \in \mathbb{N} : r_n \in A_n - B_n\} \in F.$$

Define J by

$$J = \{n \in \mathbb{N} : A_n \subseteq B_n\}.$$

Then $I \subseteq J^c$, so $J^c \in F$ and hence $J \notin F$.

Conversely, if $J \notin F$, then $J^c \in F$, so choosing $r_n \in A_n - B_n$ for each $n \in J^c$ and r_n arbitrary for $n \in J$, the argument reverses to give a point $[r_n] \in [A_n] - [B_n]$.

(3) This follows from (2) and closure properties of F.[1]

(4) Exercise. $\square$

Part (3) above is important for what it says about the sequence $\langle A_n : n \in \mathbb{N}\rangle$ that determines a certain internal set. We can replace this sequence by another $\langle B_n : n \in \mathbb{N}\rangle$ without changing the resulting internal set, provided that $A_n = B_n$ for F-almost all n. Thus we are free to alter A_n arbitrarily when n is outside a set that belongs to F. For instance, if $[A_n]$ is nonempty, then as $\emptyset = [\emptyset]$, we can assume that $A_n \neq \emptyset$ for every $n \in \mathbb{N}$ while if $[A_n]$ is a subset of ${}^*\mathbb{N}$, then as ${}^*\mathbb{N} = [\mathbb{N}]$, we can assume that $A_n \subseteq \mathbb{N}$ for every n. Moreover, we can combine finitely many such conditions, using the closure of F under finite intersections.

COROLLARY 8.2.2. *If* $[A_n]$ *is a nonempty subset of* ${}^*\mathbb{N}$, *we can assume that* $\emptyset \neq A_n \subseteq \mathbb{N}$ *for every* $n \in \mathbb{N}$.

---

[1]Note that the result is not a matter of the definition of $[A_n]$ via F, since equality of $[A_n]$ and $[B_n]$ is defined independently of F to mean “having the same members”.

8.3. Subsets of Internal Sets

The fact that the intersection of two internal sets is internal allows us to prove the following result.

THEOREM 8.3.1. *If a set A of real numbers is internal, then so is every subset of A.*

PROOF. Let $X \subseteq A$. Then ${}^*X$ is internal, so if A is internal, then so is $A \cap {}^*X$. But since $A \subseteq \mathbb{R}$,

$$A \cap {}^*X = A \cap {}^*X \cap \mathbb{R} = A \cap X$$

so $X = A \cap X$ is internal. □

This result will be used in Section 9.2 to show that actually the only internal subsets of $\mathbb{R}$ are the finite ones.

8.4. Internal Least Number Principle

A characteristic feature of $\mathbb{N}$ is that each of its nonempty subsets has a least member.[2] The same is not true, however, for ${}^*\mathbb{N}$. Namely the set ${}^*\mathbb{N} - \mathbb{N}$ of infinite hypernaturals has no least member, for if H is infinite, then so is $H - 1$. But we do have the following result.

THEOREM 8.4.1. *Any nonempty internal subset of ${}^*\mathbb{N}$ has a least member.*

PROOF. Let $[A_n]$ be a nonempty internal subset of ${}^*\mathbb{N}$. Then by Corollary 8.2.2, we can assume that A_n is nonempty for each $n \in \mathbb{N}$, and so A_n has a least member r_n . This defines a point $[r_n] \in {}^*\mathbb{R}$ with

$$\{n \in \mathbb{N} : r_n \in A_n\} = \mathbb{N} \in F,$$

so $[r_n] \in [A_n]$. Moreover, if $[s_n] \in [A_n]$, then

$$\{n \in \mathbb{N} : s_n \in A_n\} \in F \text{ and } \{n \in \mathbb{N} : s_n \in A_n\} \subseteq \{n \in \mathbb{N} : r_n \leq s_n\},$$

leading to the conclusion $[r_n] \leq [s_n]$ in ${}^*\mathbb{R}$. Hence $[A_n]$ indeed has a least member, namely the hyperreal number $[r_n]$ determined by the sequence of least members of the sets A_n. Writing "min X" for the least element of a set X, this construction can be expressed concisely by the equality

$$\min[A_n] = [\min A_n].$$

□

[2]Indeed this holds for any subset of $\mathbb{Z}$ that has a lower bound.

8.5. Internal induction

The least number principle for $\mathbb{N}$ is equivalent to the following principle of induction.

> A subset of $\mathbb{N}$ that contains 1 and is closed under the successor function $n \mapsto n+1$ must be equal to $\mathbb{N}$.

The corresponding assertion about subsets of ${}^*\mathbb{N}$ is not in general true, and can only be derived for internal sets, as follows.

THEOREM 8.5.1 (Internal Induction). *If X is an internal subset of ${}^*\mathbb{N}$ that contains 1 and is closed under the successor function $n \mapsto n+1$, then $X = {}^*\mathbb{N}$.*

PROOF. Let $Y = {}^*\mathbb{N} - X$. Then Y is internal so if it is nonempty, it has a least element n. Then $n \neq 1$, as $1 \in X$, so $n - 1 \in {}^*\mathbb{N}$. But now $n - 1 \notin Y$, as n is least in Y, so $n - 1 \in X$, and therefore $n = (n-1)+1$ is in X by closure under successor. This contradiction forces us to conclude that $Y = \emptyset$, and so $X = {}^*\mathbb{N}$. $\square$

8.6. The Overflow Principle

The set $\mathbb{N}$ cannot be internal, or else by internal induction it would be equal to ${}^*\mathbb{N}$. Thus if an internal set X contains all members of $\mathbb{N}$, then since X cannot be equal to $\mathbb{N}$, it must "overflow" into ${}^*\mathbb{N}-\mathbb{N}$. This explains the term *overflow principle*.[3] Indeed, we will see that X must contain the initial segment of ${}^*\mathbb{N}$ up to some infinite hypernatural. In fact, a slightly stronger statement than this can be demonstrated by assuming only that X contains "almost all" members of $\mathbb{N}$, as follows.

THEOREM 8.6.1. *Let X be an internal subset of ${}^*\mathbb{N}$ and $k \in \mathbb{N}$. If $n \in X$ for all $n \in \mathbb{N}$ with $k \leq n$, then there is an infinite $K \in {}^*\mathbb{N}$ with $n \in X$ for all $n \in {}^*\mathbb{N}$ with $k \leq n \leq K$.*

PROOF. If all infinite hypernaturals are in X, then any infinite $K \in {}^*\mathbb{N}$ will do. Otherwise there are infinite hypernaturals not in X. If we can show that there is a *least* such infinite number H, then all infinite numbers smaller than H will be in X, giving the desired result. To spell this out: if ${}^*\mathbb{N} - X$ has infinite members, then these must be greater than k, and so the set

$$Y = \{n \in {}^*\mathbb{N} : k < n \in {}^*\mathbb{N} - X\}$$

is nonempty. But Y is internal, by the algebra of internal sets, since it is equal to

$$({}^*\mathbb{N} - \{1, \dots, k\}) \cap ({}^*\mathbb{N} - X).$$

[3] Hatzafa

Hence Y has a least element H by the internal least number principle. Then H is a hypernatural that is greater than k but not in X, so it must be the case that $H \notin \mathbb{N}$, because of our hypothesis that all finite $n \geq k$ are in X. Thus H is infinite. Then $K = H - 1$ is infinite and meets the requirements of the theorem: H is the least hypernatural greater than k that is not in X, so every $n \in {}^*\mathbb{N}$ with $k \leq n \leq H - 1$ does belong to X. □

8.7. Internal order-completeness

The principle of order-completeness, attributed to Dedekind, asserts that every nonempty subset of $\mathbb{R}$ with an upper bound in $\mathbb{R}$ must have a least upper bound in $\mathbb{R}$. The corresponding statement about ${}^*\mathbb{R}$ is false.

EXAMPLE 8.7.1. $\mathbb{R}$ itself is a nonempty subset of ${}^*\mathbb{R}$ that is bounded but has no least upper bound. This is because the upper bounds of $\mathbb{R}$ in ${}^*\mathbb{R}$ are precisely the positive infinite numbers, and there is no least positive infinite number.

Just as for the least number principle, order-completeness is preserved in passing from $\mathbb{R}$ to ${}^*\mathbb{R}$ for internal sets:

THEOREM 8.7.2. *If a nonempty internal subset of ${}^*\mathbb{R}$ is bounded above/ below, then it has a least upper/ greatest lower bound in ${}^*\mathbb{R}$.*

PROOF. We treat the case of upper bounds. In effect, the point of the proof is to show that the least upper bound of a bounded internal set $[A_n]$ is the hyperreal number determined by the sequence of least upper bounds of the A_n's:

$$lub[A_n] = [lub\, A_n].$$

More precisely, it is enough to require that F-almost all A_n's have least upper bounds to make this work.

Suppose that a nonempty internal set $[A_n]$ has an upper bound $[r_n]$. Write $A_n \leq x$ to mean that x is an upper bound of A_n in $\mathbb{R}$, and put

$$J = \{n \in \mathbb{N} : A_n \leq r_n\}.$$

We want $J \in F$. If not, then $J^c \in F$. But if $n \in J^c$, there exists some a_n with $r_n < a_n \in A_n$. This leads to the conclusion $[r_n] < [a_n] \in [A_n]$, contradicting the fact that $[r_n]$ is an upper bound of $[A_n]$. It follows that $J \in F$. Since $[A_n] \neq \emptyset$, this then implies

$$J' = \{n \in \mathbb{N} : \emptyset \neq A_n \leq r_n\} \in F.$$

Now, if $n \in J'$, then A_n is a nonempty subset of $\mathbb{R}$ bounded above (by r_n), and so by the order-completeness of $\mathbb{R}$, A_n has a least upper bound $s_n \in \mathbb{R}$. Then if $[b_n] \in [A_n]$,

$$\{n \in \mathbb{N} : b_n \in A_n\} \cap J' \subseteq \{n \in \mathbb{N} : b_n \leq s_n\},$$

leading to $[b_n] \leq [s_n]$, and showing that $[s_n]$ is an upper bound of $[A_n]$. Finally, if $[t_n]$ is any other upper bound of $[A_n]$, then $\{n : A_n \leq t_n\} \in F$ by the same argument as for $[r_n]$, and

$$\{n \in \mathbb{N} : A_n \leq t_n\} \cap J' \subseteq \{n \in \mathbb{N} : s_n \leq t_n\}.$$

so we get $[s_n] \leq [t_n]$. This shows that $[s_n]$ is indeed the least upper bound of $[A_n]$ in ${}^*\mathbb{R}$. $\square$

Internal completeness is discussed further in Section 15.5.

8.8. External sets

Definition 8.8.1. A subset of ${}^*\mathbb{R}$ is external if it is not internal.

Many of the properties that are special to the structure of ${}^*\mathbb{R}$ define external sets.

Example 8.8.2 (Infinite hypernaturals). Since ${}^*\mathbb{N} - \mathbb{N}$ has no least member, the internal least number principle implies that it cannot be internal.

Example 8.8.3. [Finite Hypernaturals] If $\mathbb{N}$ were internal, then so too would be ${}^*\mathbb{N} - \mathbb{N}$, which we have just seen to be false. Alternatively, by the internal induction principle, if $\mathbb{N}$ were internal, it would be equal to ${}^*\mathbb{N}$.

Example 8.8.4 (Real numbers). $\mathbb{R}$ is external, for if it were internal, then so too would be $\mathbb{R} \cap {}^*\mathbb{N} = \mathbb{N}$.

Alternatively, as noted earlier $\mathbb{R}$ is bounded but has no least upper bound in ${}^*\mathbb{R}$, so must fail to be internal by the internal order-completeness property.

The fact that $\mathbb{N}$ is external will be used in the next section to show that all infinite subsets of $\mathbb{R}$ are external.

Example 8.8.5 (Finite hyperreals). The set $\mathbb{L}$ of finite numbers is external for the same reason $\mathbb{R}$ is: it is bounded above by all members of ${}^*\mathbb{R}^+_\infty$, but has no least upper bound. Since

$$\mathbb{L} = \bigcap \{(-b, b) : b \text{ is infinite }\},$$

it follows that the intersection of an infinite family of internal sets can fail to be internal.

Observe that if X is an internal set that includes $\mathbb{L}$, then $X \neq \mathbb{L}$, and so X must contain infinite members. In fact, by considering lower and upper bounds of ${}^*\mathbb{R}^+_\infty - X$ and ${}^*\mathbb{R}^-_\infty - X$, respectively, we can show that if X is an internal set with $\mathbb{L} \subseteq X$, then $[-b, b] \subseteq X$ for some infinite b.

Example 8.8.6 (Infinitesimals). The set $\mathbb{I} = hal(0)$ of infinitesimals is bounded above (by any positive real), so if it were internal, it would have a least upper bound $b \in {}^*\mathbb{R}$. Such a b would have to be positive but less than every positive real, forcing $b \approx 0$. But then $b < 2b \in \mathbb{I}$, so b cannot be an upper bound of $\mathbb{I}$ after all.

By similar reasoning, any halo $hal(r)$ is seen to be an external set, as are its "left and right halves" $\{x > r : x \approx r\}$ and $\{x < r : x \approx r\}$.

Corollary 8.8.7. *If X is any internal subset of $\mathbb{I}$, then the least upper bound and greatest lower bound of X must be infinitesimal, and so $X \subseteq [-\varepsilon, \varepsilon]$ for some $\varepsilon \approx 0$.*

CHAPTER 9

Defining internal sets, Saturation

A central application of analysis with infinitesimals is a hyperfinite construction of the Lebesgue measure via counting measures. An important principle used in the construction is the saturation principle that will be dealt with in Sections 9.5 and 13.1.

An elegant application of saturation is a short proof of Cantor's theorem on nested sequences of compact sets; see Section 9.7.

9.1. Geometric example of internal set definition

To develop an ultrafiter-free definition of internal sets, we will develop a method of defining them via formulas not referring to the ultrapower construction. We first consider the following geometric example that will help motivate the general definition.

EXAMPLE 9.1.1 (r-neighborhoods). Let $A \subseteq \mathbb{R}^m$ be a subset, and let $r > 0$. The r-neighborhood $U_r A$ of A in $\mathbb{R}^m$ is the set

$$U_r A = \{x \in \mathbb{R}^m : (\exists a \in A)\,(|x-a| < r)\}.$$

Now consider the formula $\phi(x, r, A)$ specified as

$$(\exists a \in A)\,(|x-a| < r).$$

Then ϕ is a formula in the language $\mathcal{L}_{\mathcal{R}}$ (see Section 4.1), with three free variables x, r, and A. Note that A is a set variable. Then the set $B = U_r A$ can be viewed as defined by the formula

$$x \in B \text{ iff } \phi(x, r, A) \text{ is true}.$$

We now develop a hyperreal analog of this situation. Consider the corresponding formula

$$^*\phi \in \mathcal{L}_{^*\mathcal{R}}$$

in the language of the relational structure $^*\mathcal{R}$ defined in Section 4.1. We can replace the variables x, r, and A in $^*\phi$ by their hyperreal analogs. Thus $^*\phi([x_n], [r_n], [A_n])$ would be the sentence

$$(\exists a \in [A_n])\,\big(\big|[x_n] - a\big| < [r_n]\big).$$

It can be shown that this sentence is true if and only if

$$\{n \in \mathbb{N} : (\exists a \in A_n)\,(|x_n - a| < r_n)\} \in F$$

(the usual translation in terms of the ultrafilter). Then we obtain the following general fact.

Corollary 9.1.2. *The formula*

$$^*\phi([x_n], [r_n], [A_n]) \text{ is true} \tag{9.1.1}$$

if and only if

$$\{n \in \mathbb{N} : \phi(x_n, r_n, A_n) \text{ is true}\} \in F. \tag{9.1.2}$$

This leads to a new way of defining internal sets: holding the hyperreal $[r_n]$ and the internal set $[A_n]$ fixed and allowing the value of b to range over $^*\mathbb{R}^m$, we define the set

$$X = \{b \in {}^*\mathbb{R}^m : {}^*\phi(b, [r_n], [A_n]) \text{ is true}\}. \tag{9.1.3}$$

Correspondingly, for each $n \in \mathbb{N}$, we set

$$B_n = \{b \in \mathbb{R}^m : \phi(b, r_n, A_n) \text{ is true}\}.$$

Then the equivalence of (9.1.1) and (9.1.2) amounts to saying that for all $[b_n]$,

$$[b_n] \in X \text{ iff } \{n \in \mathbb{N} : r_n \in B_n\} \in F.$$

But this shows that X is the internal set $[B_n]$ determined by the sequence of real subsets $\langle B_n : n \in \mathbb{N}\rangle$.

Corollary 9.1.3. *Formula* (9.1.3) *can be seen as a definition of the internal set* $[B_n]$.

9.2. Internal set definition principle

Expressing this phenomenon in the most general form available at this stage, we have the following statement.

Theorem 9.2.1 (Internal Set Definition Principle). *Let*

$$\phi(x_0, x_1, \ldots, x_n, A_1, \ldots, A_k)$$

be an $\mathcal{L}_{\mathcal{R}}$*-formula with free variables* $x_0, x_1 \ldots, x_n$ *as well as set symbols* $A_1, \ldots, A_k$. *Then for any hyperreals* $c_1, \ldots, c_n$ *and any internal sets* $X_1, \ldots, X_k$, *the collection*

$$\{b \in {}^*\mathbb{R} : {}^*\phi(b, c_1, \ldots, c_n, X_1, \ldots, X_k)\}$$

is an internal subset of $^*\mathbb{R}$.

It provides a ready means of demonstrating that various sets are internal, including the examples from Section 8.1.

EXAMPLE 9.2.2. Taking $\phi(x_0, x_1, \ldots, x_n)$ as the formula

$$(x_0 = x_1 \vee \cdots \vee x_0 = x_n)$$

shows that any finite set

$$\{c_1, \ldots, c_n\} = \{b \in {}^*\mathbb{R} : {}^*\phi(b, c_1, \ldots, c_n)\}$$

of hyperreals is internal.

EXAMPLE 9.2.3. Taking $\phi(x_0, x_1, x_2)$ as the formula $(x_1 < x_0 < x_2)$ yields that any open hyperreal interval is internal.

Recall that $\mathbb{N}$ is an external set (see Example 8.8.3). We can now use internal set definition to show the following, as promised in Section 8.3.

THEOREM 9.2.4. *Every infinite set of real numbers is external. In other words, if $A \subseteq \mathbb{R}$ is internal, then A must be finite.*

PROOF. We assume that A is internal and argue by contradiction. If such an A were infinite, then it would contain an infinite sequence, i.e., there would be an injective function $f : \mathbb{N} \to A$. Put

$$X = \{f(n) : n \in \mathbb{N}\}.$$

Then X is internal, since it is a subset of the internal set A, and by Theorem 8.3.1, any subset of an internal set of real numbers is internal.

Now, the set X is a bijective copy of $\mathbb{N}$ by a standard function, so we should be able to show that $\mathbb{N}$ is internal if X is, thereby getting a contradiction because we already know that $\mathbb{N}$ is external. We therefore apply the internal set definition principle, applied with $\phi(x, S)$ as the formula

$$x \in \mathbb{N} \wedge f(x) \in S.$$

Consider the corresponding formula ${}^*\phi \in \mathcal{L}_{{}^*\mathcal{R}}$. By the internal set definition principle, the set

$$B = \{n \in {}^*\mathbb{R} : {}^*\phi(n, X)\}$$

is internal. 0bserve that

$$B = \{n \in {}^*\mathbb{N} : {}^*f(n) \in X\} = {}^*f^{-1}(X).$$

However, as f is injective, ${}^*f : {}^*\mathbb{N} \to {}^*A$ is an injective extension of f (by transfer). It follows that B is just $\mathbb{N}$ itself. The resulting contradiction proves the theorem. $\square$

9.3. The Underflow Principle

The underflow principle is the order-theoretic dual of the overflow principle of Theorem 8.6.1. Its proof requires the additional reasoning power provided by the internal set definition principle of Section 9.2.

THEOREM 9.3.1. *Let X be an internal subset of ${}^*\mathbb{N}$, and let $K \in {}^*\mathbb{N}$ be infinite. If every infinite hypernatural $H \leq K$ belongs to X, then there is some $k \in \mathbb{N}$ such that every finite n with $n \geq k$ belongs to X.*

To comment on terminology, our internal set of infinite hypernaturals "spills over" into the finite natural numbers.

PROOF. For $M, N \in {}^*\mathbb{N}$ with $M \leq N$, let

$$\lfloor M, N \rfloor = \{z \in {}^*\mathbb{N} : M \leq z \leq N\}$$

be the interval in ${}^*\mathbb{N}$ between M and N. By hypothesis, the condition $\lfloor H, K \rfloor \subseteq X$ is satisfied for all infinite hypernatural $H \leq K$. We need to show that $\lfloor k, K \rfloor \subseteq X$ for some (finite) $k \in \mathbb{N}$.

Equivalently, we want to show that the set

$$Y = \{k \in {}^*\mathbb{N} : \lfloor k, K \rfloor \subseteq X\}$$

has a finite member.

Now, if Y is internal, then by the internal least number principle it has a least element k, and such a k must belong to $\mathbb{N}$, because if it were infinite, then $k-1$ would be infinite, so by our hypothesis $k-1$ would also be in Y but less than k.

It thus suffices to show that Y is internal. Now let $\phi(x, y, A)$ be the formula

$$x \in \mathbb{N} \wedge x \leq y \wedge \forall z \in \mathbb{N}\, (x \leq z \leq y \rightarrow z \in A)$$

expressing "$x \in \mathbb{N}$ and $\lfloor x, y \rfloor \subseteq A$".[1] Then by the internal set definition principle, the set

$$\{k \in {}^*\mathbb{R} : {}^*\phi(k, K, X)\} = \\ \{k \in {}^*\mathbb{N} : k \leq K \text{ and } \forall z \in {}^*\mathbb{N}(k \leq z \leq K \rightarrow z \in X)\}$$

is internal. This set is just Y. □

9.4. Internal sets and permanence

Overflow/permanence was already discussed in Section 8.6. The following is an additional result in this direction. The novelty compared to the earlier results is that the parameter b below could be nonstandard, e.g., infinite.

[1] The relation $\subseteq$ is not part of the language we are working with.

THEOREM 9.4.1. *If an internal set $X \subseteq {}^*\mathbb{R}$ contains all points that are infinitely close to $b \in {}^*\mathbb{R}$, then there is a positive real e such that X contains all points that are within e of b.*

PROOF. By hypothesis, $hal(b) \subseteq X$. Since $hal(b)$ is an external set, we cannot use it in a formula to which the internal definition principle would be applicable. Instead, we consider for each $k \in {}^*\mathbb{N}$, the (internal) hyperreal interval $(b - \frac{1}{k}, b + \frac{1}{k})$, or more precisely

$$\{z \in {}^*\mathbb{R} : |z - b| < \tfrac{1}{k}\}.$$

Whenever k is infinite, $\frac{1}{k}$ is infinitesimal, and so

$$(b - \tfrac{1}{k}, b + \tfrac{1}{k}) \subseteq hal(b) \subseteq X$$

by our hypothesis. Now consider the set

$$Y = \{k \in {}^*\mathbb{N} : (b - \tfrac{1}{k}, b + \tfrac{1}{k}) \subseteq X\}.$$

Note that Y contains all infinite members of ${}^*\mathbb{N}$. Hence by underflow we could conclude that $(b - \frac{1}{k}, b + \frac{1}{k}) \subseteq X$ for some (finite) $k \in \mathbb{N}$, and thereby complete the proof by setting $e = \frac{1}{k}$, if we knew that Y were internal.

To show that Y is internal, we apply internal set definition principle with $\phi(x, y, A)$ as the formula

$$x \in \mathbb{N} \wedge (\forall z \in \mathbb{R})(|z - y| < \tfrac{1}{x} \to z \in A). \tag{9.4.1}$$

This formula expresses "$x \in \mathbb{N}$ and $(y - \frac{1}{x}, y + \frac{1}{x}) \subseteq A$".[2] It follows that the set

$$\{k \in {}^*\mathbb{R} : {}^*\phi(k, b, X)\} = \{k \in {}^*\mathbb{N} : (\forall z \in {}^*\mathbb{R})\, (|z - b| < \tfrac{1}{k} \to z \in X)\}$$

is internal, and this set is just Y. □

9.5. Introduction to saturation

Internal sets form a special collection whose members are related to each other in remarkable ways. For instance, it is impossible to construct a nested sequence of internal sets whose intersection is empty. This fact, which we will prove below, is known as countable saturation. The use of the term *saturation* is explained at the beginning of the Section 13.3. We will first treat a special case in Lemma 9.5.1 to provide motivation, and then the general case in Theorem 13.1.1.

[2]The relation $\subseteq$ is not part of the language we are working with, and therefore cannot be used in equation (9.4.1).

Lemma 9.5.1. *Consider a nested sequence of nonempty sets in* $\mathbb{R}$*:*

$$X^1 \supseteq X^2 \supseteq \cdots \supseteq X^k \supseteq \cdots$$

Then the intersection of their natural extensions is always nonempty:

$$\bigcap\nolimits_{k\in\mathbb{N}} {}^*X^k \neq \emptyset.$$

Proof. The argument involves a kind of diagonalisation procedure. We have ${}^*X^k = [A_n^k]$, where $\langle A_n^k : n \in \mathbb{N}\rangle$ is the constant sequence $A_n^k = X^k$. For each $n \in \mathbb{N}$, we choose some $s_n \in A_n^n$ (a kind of diagonalisation). Since the sequence X^k is nested, we have

$$s_n \in A_n^1 \cap \cdots \cap A_n^n.$$

Then the set $\{n \in \mathbb{N} : s_n \in X^k\}$ is a cofinite set in $\mathbb{N}$. Therefore the hyperreal $[s_n]$ belongs to each of the ${}^*X^k$. $\square$

Example 9.5.2. A nested sequence of open intervals $X^k = (0, \frac{1}{k}) \subseteq \mathbb{R}$ has empty intersection:

$$\bigcap\nolimits_{k\in\mathbb{N}} X^k = \varnothing.$$

On the other hand, the natural extension ${}^*X^k$ contains the infinitesimal $[\frac{1}{n}]$ for each k. The existence of such an infinitesimal follows from Lemma 9.5.1.

Definition 9.5.3. A family of sets is said to have the *finite intersection property* if the intersection of any finite subfamily is nonempty.

Corollary 9.5.4. *Suppose a countable family* $\{X^k\}$*,* $X_k \subseteq \mathbb{R}$*, has the finite intersection property. Then* $\bigcap_{k\in\mathbb{N}} {}^*X^k \neq \emptyset$*.*

Proof. We apply Lemma 9.5.1 to the nested sequence of intersections $X^1 \cap \cdots \cap X^k$. $\square$

The general case (nested sequence of internal sets) will be treated in Section 13.1.

This version of saturation for a countable family $\{X^k\}$ can be reformulated as follows.

$$\forall\, n \in \mathbb{N}\, \forall\, k = (k_1, \ldots, k_n) \in \mathbb{N}^k\, \exists\, x \in X^{k_1} \cap \cdots \cap X^{k_n} \longrightarrow \exists\, x \in \bigcap_{k\in\mathbb{N}} {}^*X^k. \tag{9.5.1}$$

This should be spelled out in terms of additional quantifiers so as to avoid using the ellipsis.

The implication (9.5.1) can be reformulated in terms of LSEQ operator provided a starring adjustment is made; see Section 7.1. The result is $\Phi \longrightarrow \text{LSEQ}({}^*\Phi)$. Here Φ is the formula

$$\forall\, n \in \mathbb{N}\, \forall\, k = (k_1, \ldots, k_n) \in \mathbb{N}^k\, \exists\, x \in X^{k_1} \cap \cdots \cap X^{k_n}$$

(as before, this should be reformulated in terms of additional quantifiers so as to avoid using the ellipsis).

For a countable family of internal subsets $X_k \subseteq {}^*\mathbb{R}$, we obtain the following version of saturation: $\Phi \longrightarrow \mathrm{LSEQ}(\Phi)$ (without the stars).

9.6. Characterisation of compactness

As an application of saturation, we give a characterisation of compact sets via infinitesimal analysis.

THEOREM 9.6.1 (Characterisation of compactness). *Let X be a metric space. Then the following two conditions are equivalent:*

(1) *X is compact;*
(2) *every $y \in {}^*X$ is infinitely close to a suitable point $x \in X$.*

DEFINITION 9.6.2. A point $y \in {}^*X$ is said to be *nearstandard in X* if it is infinitely close to a point of X.

PROOF OF (1) ⇒ (2). Assume X is compact, and let $y \in {}^*X$. Let us show that y is nearstandard in X (this direction does not require saturation).

We give a proof by contradiction. Suppose that y is not nearstandard in X. Namely, it is not in the halo of any point $p \in X$. Then every $p \in X$ has a (standard) open neighborhood U_p such that $y \notin {}^*U_p$. Consider the collection $\{U_p\}_{p\in X}$. This collection is an open cover[3] of X. Since X is compact, the collection has a finite subcover $U_{p_1}, \ldots, U_{p_n}$, so that

$$X = U_{p_1} \cup \ldots \cup U_{p_n}. \tag{9.6.1}$$

Due to the finiteness of the union (9.6.1), by the algebra of natural extensions we have

$${}^*X = {}^*U_{p_1} \cup \cdots \cup {}^*U_{p_n}.$$

In particular, $y \in {}^*X$ must belong to one of the sets ${}^*U_{p_1}, \ldots, {}^*U_{p_n}$, contradicting our hypothesis. □

PROOF OF (2) ⇒ (1). We will give a proof in the case when X is separable (for example, subset of $\mathbb{R}^n$) and therefore the topology of X admits a countable basis.[4] This direction exploits saturation. Assume that every $y \in {}^*X$ is nearstandard. Given a countable open cover $\{U_a\}$ of X, we need to find a finite subcover.

[3]kisui patuach

[4]Metrisable compact spaces are necessarily separable, but non-metrizable may not be. For examples see `https://math.stackexchange.com/questions/74923` (a compact Hausdorff space that is not metrizable). Davis [**3**, p. 78] uses concurrency.

The proof is by contradiction. Suppose that the union of any finite collection of U_a is not all of X. Then the complements of U_a form a collection of (closed) sets $\{S_a\}$, where $S_a = X - U_a$, with the finite intersection property. It follows that the collection $\{{}^*S_a\}$ similarly has the finite intersection property.

We now use the condition that the family is countable. By saturation and Corollary 9.5.4, the intersection of all ${}^*S_a$ is non-empty. Let y be a point in this intersection:

$$y \in \bigcap\nolimits_a {}^*S_a.$$

By hypothesis of nearstandardness in X, there is a point $p \in X$ such that $y \in hal(p)$. Now $\{U_a\}$ is a cover of X so there is a U_b such that $p \in U_b$. But y is in ${}^*S_a$ for all a, in particular $y \in {}^*S_b$. Thus $y \in {}^*S_b \cap {}^*U_b = \varnothing$, a contradiction. □

9.7. Application of saturation: Cantor's intersection theorem

We can now use saturation to prove Cantor's theorem on infinite nested sequences of compact sets.

THEOREM 9.7.1 (Cantor's intersection theorem). *A nested decreasing sequence of nonempty compact sets has a common point.*

PROOF. Given a nested sequence of compact sets, $\langle S^k : k \in \mathbb{N}\rangle$, we consider the corresponding decreasing nested sequence of internal sets, $\langle {}^*S^k : k \in \mathbb{N}\rangle$. This sequence has a common point x by saturation. By Theorem 9.6.1, for a compact set S^k, every point of ${}^*S^k$ is nearstandard in S^k, i.e., infinitely close to a standard point $x_k \in S^k$. In particular, $sh(x) \in S^k$ for all k. In more detail, we have $x_k \approx x \approx x_\ell$ and therefore $x_k = x_\ell\,(\forall k, \ell)$ is the common point of all the compact sets S^k. □

Part 2

Effective infinitesimals

CHAPTER 10

Effective infinitesimals

10.1. Axiom of choice

For simplicity we will restrict the discussion to families of subsets of $\mathbb{R}$. Let $(A_i)_{i\in I}$ be a family of nonempty disjoint subsets of $\mathbb{R}$. Thus for each i, we have $A_i \subseteq \mathbb{R}$, $A_i \neq \varnothing$, and $A_i \cap A_j = \varnothing$ whenever $i, j \in I$ with $i \neq j$. Note that I is not necessarily countable.

Definition 10.1.1. A *choice function* for the family (A_i) is a function $f\colon I \to \mathbb{R}$ such that $f(i) \in A_i$ for each $i \in I$.

Then the axiom of choice asserts the existence of a choice function. A more general statement of the axiom of choice asserts the existence of a choice function for any family of disjoint nonempty sets (not necessarily subsets of $\mathbb{R}$).

10.2. Set theories

Let ZF be the Zermelo–Fraenkel set theory. Let ZFC be the Zermelo–Fraenkel set theory with the axiom of choice. Let ACC be the axiom of countable choice, and ADC the axiom of (countable) dependent choice. The theories ZF, ZF+ACC, and ZF+ADC have the advantage (over ZFC) of not entailing set-theoretic paradoxes such as Banach–Tarski. Similarly, ZF, ZF+ACC, and ZF+ADC do not prove the existence of nonprincipal ultrafilters.

The theories SPOT and SCOT developed in [**8**] provide frameworks for analysis with infinitesimals that are conservative respectively over ZF and ZF+ADC, and therefore share the same advantage (the axioms of SPOT and SCOT appear in Section 11). Mathematicians generally consider theorems provable in ZF as more effective than results that require the full ZFC for their proof, and many feel this way not only about ZF but about ZF+ADC, as well. In this sense, the theories SPOT and SCOT enable an effective development of analysis based on infinitesimals. Some applications were already presented in [**8**], such as (local) Peano's existence theorem for first-order differential equations [**8**, Example 3.5] and infinitesimal construction of Lebesgue measure via counting measures [**8**, Example 3.6].

We first consider the case of compactness. In Section 10.3, we present the traditional extension view. Following an outline of SPOT and SCOT in Sections 10.4 and 11, we deal with compactness in internal set theories in Section 11.3. After preliminaries on continuity in Section 12.1, we present an effective proof using infinitesimals of the compactness of a continuous image of a compact set in Section 12.2. After preliminaries on uniform continuity in Sections 12.3 and 12.4, we present an effective proof using infinitesimals of the Heine–Borel theorem in Section 12.5. In Section 12.6, we show that Nelson's *Radically Elementary Probability Theory* is a subtheory of SCOT.

10.3. Compactness in the extension view

In this section, we analyze compactness from the viewpoint of traditional extensions $\mathbb{R} \hookrightarrow {}^*\mathbb{R}$ to hyperreals. These cannot be constructed in ZF+ADC and cannot be described as effective in the sense of Section 10.2. In Section 11.3, we will present an effective treatment of compactness in axiomatic frameworks for analysis with infinitesimals.

For $\mathbb{N}$, $\mathbb{R}$, $\mathbb{P} = \mathcal{P}(\mathbb{R})$, or any set X, the corresponding nonstandard extensions ${}^*X$, etc. satisfying the transfer principle can be formed either via the compactness theorem of first-order logic, or via ultrapowers $X^{\mathbb{N}}/\mathcal{F}$, etc., in terms of a fixed nonprincipal ultrafilter $\mathcal{F}$.

Lemma 10.3.1. *For a finite union, the star of the union is the union of stars.*

Proof. Given sets $A, B \subseteq X$, we have

$$(\forall y \in X)\big[y \in A \cup B \longleftrightarrow (y \in A) \vee (y \in B)\big]. \tag{10.3.1}$$

Applying upward transfer to (10.3.1), we obtain

$$(\forall y \in {}^*X)\big[y \in {}^*(A \cup B) \longleftrightarrow (y \in {}^*A) \vee (y \in {}^*B)\big],$$

and the claim follows by induction. □

Theorem 10.3.2. *If $\langle A_n : n \in \mathbb{N}\rangle$ is a nested sequence of nonempty subsets of $\mathbb{R}$ then the sequence $\langle {}^*A_n : n \in \mathbb{N}\rangle$ (standard n) has a common point.*

Proof. We give an alternative argument to the one given in Section 9.5. Let $\mathbb{P} = \mathcal{P}(\mathbb{R})$ be the set of all subsets of $\mathbb{R}$. Consider a sequence $\langle A_n \in \mathbb{P} : n \in \mathbb{N}\rangle$ viewed as a function $f : \mathbb{N} \to \mathbb{P}$, $n \mapsto A_n$. By the extension principle we have a function ${}^*f : {}^*\mathbb{N} \to {}^*P$. Let $B_n =$

${}^*f(n)$. For each standard n, we have $B_n = {}^*\!A_n \in {}^*P$.[1] For a nonstandard value of the index $n = H$, the entity $B_H \in {}^*P$ is by definition internal but is in general not the natural extension of any subset of $\mathbb{R}$.

If $\langle A_n \rangle$ is a *nested* decreasing sequence in $\mathbb{P} \setminus \{\varnothing\}$ then by transfer $\langle B_n : n \in {}^*\mathbb{N} \rangle$ is nested in ${}^*P \setminus \{\varnothing\}$. Let H be a fixed nonstandard index. Since $n < H$ for each standard n, the set ${}^*\!A_n \subseteq {}^*\mathbb{R}$ includes B_H. Choose any element $c \in B_H$. Then c is contained in ${}^*\!A_n$ for each standard n:

$$c \in \bigcap_{n \in \mathbb{N}} {}^*\!A_n$$

as required.[2] □

Definition 10.3.3. Let I be a set. A collection $\mathcal{H} \subseteq \mathcal{P}(I)$ has the *finite intersection property* if the intersection of every nonempty finite subcollection of $\mathcal{H}$ is nonempty, i.e.,

$$B_1 \cap \cdots \cap B_n \neq \varnothing \text{ for all } n \in \mathbb{N} \text{ and all } B_1, \ldots, B_n \in \mathcal{H}.$$

Then Theorem 10.3.2 has the following equivalent formulation.

Corollary 10.3.4 (Countable Saturation). *If a family of subsets $\{A_n\}_{n \in \mathbb{N}}$ has the finite intersection property (see Definition* 10.3.3*) then the intersection $\bigcap_{n \in \mathbb{N}} {}^*\!A_n$ is nonempty.*

Recall that a topological space T is second-countable if its topology admits a countable base. Recall that a space is Lindelöf if every open cover includes a countable subcover. A second countable space is necessarily Lindelöf (over ZF+ACC). If T is a separable metric space then T is second countable and hence Lindelöf.

A point $y \in {}^*T$ is called *nearstandard in* T if y is infinitely close to a standard point $p \in T$, i.e., such that y is contained in the star ${}^*U$ of every open neighborhood U of p. The intersection of all such ${}^*U$ is called the *halo* of p. The relation $x \simeq y$ holds if and only if for all open sets O, $x \in {}^*O$ if and only if $y \in {}^*O$.

---

[1] The injective map $*\colon \mathbb{P} \to {}^*P$ sends A_n to ${}^*\!A_n$. For each standard natural n we have a symbol a_n in the appropriate language (including at least the names for all subsets of $\mathbb{R}$), whose standard interpretation is $A_n \in \mathbb{P}$. Meanwhile the nonstandard interpretation of a_n is the entity ${}^*\!A_n \in {}^*P$. The sequence $\langle A_n \colon n \in \mathbb{N} \rangle$ in $\mathbb{P}$ is the standard interpretation of the symbol $a = \langle a_n \rangle$. Meanwhile, the nonstandard interpretation of the symbol a is $\langle B_n \colon n \in {}^*\mathbb{N} \rangle$ in ${}^*P$. In particular, one has $B_n = {}^*\!A_n$ for standard n.

[2] The conclusion of non-empty intersection remains valid for any nested sequence of nonempty *internal* sets, i.e., members of ${}^*P$; see e.g., [**6**, Theorem 11.10.1, p. 138]. The proof is more involved and can be found in Section 13.1.

THEOREM 10.3.5. *Assume T is Lindelöf. Then the following two conditions are equivalent:*

1. *T is compact (i.e., every open cover admits a finite subcover);*
2. *every $y \in {}^*T$ is nearstandard in T.*

PROOF OF (1) $\Rightarrow$ (2). Assume T is compact, and let $y \in {}^*T$. Let us show that y is nearstandard in T (this direction does not require saturation).

Suppose on the contrary that y is not nearstandard in T, i.e., y is not in the halo of any (standard) point $p \in T$. Then we can form the open cover $\mathcal{U}$ of T containing all open sets U such that

$$y \notin {}^*U. \tag{10.3.2}$$

Since T is compact, $\mathcal{U}$ includes a finite subcover $U_1, \ldots, U_n$. Applying Lemma 10.3.1 to the finite union $T = U_1 \cup \cdots \cup U_n$, we obtain

$${}^*T = {}^*U_1 \cup \cdots \cup {}^*U_n.$$

Hence y is in one of the ${}^*U_i$, $i = 1, \ldots, n$, contradicting (10.3.2). The contradiction establishes that y is necessarily nearstandard in T. □

PROOF OF (2) $\Rightarrow$ (1). This direction exploits saturation. Assume each $y \in {}^*T$ is nearstandard in T. Given an open cover $\{U_a\}$ of T, we need to find a finite subcover. Since T is Lindelöf, we can assume that the cover is countable.

Suppose on the contrary that no finite subcollection of $\{U_a\}$ covers T. Then the complements S_a of U_a form a countable collection of (closed) sets $\{S_a\}$ with the finite intersection property. Applying countable saturation (Corollary 10.3.4) to this countable family, we conclude that the intersection of all ${}^*s_a$ is non-empty. Let $y \in \bigcap_a {}^*s_a$. By assumption, there is a point $p \in T$ such that

$$y \simeq p. \tag{10.3.3}$$

Since $\{U_a\}$ is a cover of T, it contains a set U_b such that $p \in U_b$, and hence $y \in {}^*U_b$ since U is open. But $y \in {}^*s_a$ for all a, in particular $y \in {}^*s_b$, so $y \notin {}^*U_b$ by Lemma 10.3.1, contradicting (10.3.3). The contradiction establishes the existence of a finite subcover. □

10.4. Internal set theories

In this section we explain in what sense analysis with infinitesimals does not require the axiom of choice any more than traditional non-infinitesimal analysis, following [**8**]. There are two popular approaches to Robinson's nonstandard mathematics (including analysis with infinitesimals):

(1) model-theoretic, and
(2) axiomatic/syntactic.

For a survey of the various approaches see [**4**].

The model-theoretic approach (including the construction of the ultrapower) typically relies on strong forms of the axiom of choice. The axiomatic/syntactic approach turns out to be more economical in the use of foundational material, and exploits a richer **st**-$\in$-language, as explained below.

The traditional set-theoretic foundation for mathematics is Zermelo–Fraenkel set theory (ZF). The theory ZF is a set theory formulated in the $\in$-language. Here "$\in$" is the two-place membership relation. In ZF, all mathematical objects are built up step-by-step starting from $\emptyset$ and exploiting the one and only relation $\in$.

For instance, the inequality $0 < 1$ is formalized as the membership relation $\emptyset \in \{\emptyset\}$, the inequality $1 < 2$ is formalized as the membership relation $\{\emptyset\} \in \{\emptyset, \{\emptyset\}\}$, etc. Eventually ZF enables the construction of the set of natural numbers $\mathbb{N}$, the ring of integers $\mathbb{Z}$, the field of real numbers $\mathbb{R}$, etc.

For the purposes of mathematical analysis, a set theory SPOT has been developed in the more versatile **st**-$\in$-language (its axioms are given in Section 11). Such a language exploits a predicate **st** in addition to the relation $\in$. Here "**st**" is the one-place predicate **standard** so that $\mathbf{st}(x)$ is read "x is standard".

Theorem 10.4.1 ([**8**]). *The theory* SPOT *is a conservative extension of* ZF.

This means that every statement in the $\in$-language provable in SPOT is provable already in ZF. In particular, the axiom of choice and the existence of non-principal ultrafilters are not provable is SPOT, because they are not provable in ZF. Thus SPOT does not require any additional foundational commitments beyond ZF.

Remark 10.4.2. The Separation Axiom of ZF asserts, roughly, that for any $\in$-formula ϕ and any set A, there exists a set S such that $x \in S$ if and only if $x \in A \wedge \phi(x)$ is true. This remains valid in SPOT which is a conservative extension of ZF. But Separation does not apply to formulas involving the new predicate **st**. Specifically, Separation does not apply to the predicate **st** itself.

Example 10.4.3. The collection of standard natural numbers is not a set that could be described as "$\{x \in \mathbb{N} : \mathbf{st}(x)\}$." Such external collections can be viewed informally as classes defined by the corresponding predicate. Thus, in [**8**] one uses the *dashed curly brace* notation $\{n \in \mathbb{N} : \mathbf{st}(n)\}$ for such a class, when convenient. Writing $k \in \{n \in \mathbb{N} : \mathbf{st}(n)\}$ is equivalent to writing "$\mathbf{st}(k)$ (is true)". In many cases the passage from a predicate to a set turns out to be unnecessary: as mentioned in the introduction, in SCOT (conservative over ZF+ADC) one can give an infinitesimal construction of the Lebesgue measure; in BST (a modification of Nelson's IST, possessing better meta-mathematical properties), the Loeb measure can be handled, as well; see [**9**].

Remark 10.4.4 (Sources in Leibniz). The predicate **st** formalizes the distinction already found in Leibniz between assignable and inassignable numbers. An inassignable (nonstandard) natural number μ is greater than every assignable (standard) natural number. One of the formulations of Leibniz's *Law of Continuity* posits that "the rules of the finite are found to succeed in the infinite and vice versa" (cf. Robinson [**14**, p. 266]), formalized by Robinson's transfer principle. See further in [**1**], [**2**], and [**11**].

If $\mu \in \mathbb{N}$ is a nonstandard integer, then its reciprocal $\varepsilon = \frac{1}{\mu} \in \mathbb{R}$ is a positive infinitesimal (smaller than every positive standard real). Such an ε is a nonstandard real number.

A real number smaller in absolute value than some standard real number is called *limited*, and otherwise *unlimited*. SPOT proves that every nonstandard natural number is unlimited [**8**, Lemma 2.1].

The theory SPOT enables one to take the standard part, or shadow, of every limited real number r, denoted $\mathbf{sh}(r)$. This means that the difference $r - \mathbf{sh}(r)$ is infinitesimal.

The derivative of the standard function $f(x)$ is then $\mathbf{sh}\left(\frac{f(x+\varepsilon)-f(x)}{\varepsilon}\right)$ for nonzero infinitesimal ε. In more detail, we have the following.

Definition 10.4.5. Let f be a standard function, and x a standard point. A standard number L is the *slope* of f at x if

$$(\forall^{in}\varepsilon)(\exists^{in}\lambda)\, f(x+\varepsilon) - f(x) = (L+\lambda)\varepsilon. \qquad (10.4.1)$$

where $\forall^{in}$ and $\exists^{in}$ denote quantification over infinitesimals.[3]

The Riemann integral of f over $[a, b]$ (with f, a, b standard), when it exists, is the shadow of the sum $\sum_{i=1}^{\mu} f(x_i)\varepsilon$ as i runs from 1 to μ, where the x_i are the partition points of an equal partition of $[a, b]$ into μ subintervals. For a fuller treatment see [**8**, Example 2.8].

The (external) relation of infinite proximity $x \simeq y$ for $x, y \in \mathbb{R}$ is defined by requiring $x - y$ to be infinitesimal.

[3] For further details, see note 1.

CHAPTER 11

The theories SPOT and SCOT

We will now present the axioms that enable this effective approach (conservative over ZF) to analysis with infinitesimals.

11.1. Axioms of the theory SPOT

SPOT is a subtheory of axiomatic (syntactic) theories developed in the 1970s independently by Hrbacek [**7**] and Nelson [**12**]. In addition to the axioms of ZF, SPOT has three axioms: Standard Part, Nontriviality, and Transfer (for the historical origins of the latter see Remark 10.4.4):

> T (Transfer) Let ϕ be an $\in$-formula with standard parameters. Then $\forall^{\mathbf{st}} x\, \phi(x) \to \forall x\, \phi(x)$.
>
> O (Nontriviality) $\exists \nu \in \mathbb{N}\, \forall^{\mathbf{st}} n \in \mathbb{N}\, (n \neq \nu)$.
>
> SP (Standard Part) Every limited real is infinitely close to a standard real.

An equivalent existential version of the Transfer axiom is
$\exists x\ \phi(x) \implies \exists^{\mathbf{st}} x\ \phi(x)$, for $\in$-formulas ϕ with standard parameters.

Nontriviality asserts simply that there exists a nonstandard integer.

An equivalent version of Standard Part is the following.

> SP$'$ (Standard Part)
>
> $$\forall A \subseteq \mathbb{N}\, \exists^{\mathbf{st}} B \subseteq \mathbb{N}\, \forall^{\mathbf{st}} n \in \mathbb{N}\, (n \in B \leftrightarrow n \in A).$$

Remark 11.1.1. The latter formulation can be motivated intuitively as follows. Given a real number $0 < r < 1$, consider its base-2 decimal expansion. Let A be the set of ranks where digit 1 appears. The set A is not standard if r is not standard. The corresponding standard set B (whose existence is postulated by SP$'$) can be thought of as the set of nonzero digits of the shadow $\mathbf{sh}(r)$ of r. The fact that r and $\mathbf{sh}(r)$ are infinitely close reflects the fact that A and B agree at all limited ranks. The detailed argument is a bit more technical because binary representation (like decimal representation) is not unique; see [**8**, Lemma 2.4].

In the model-theoretic frameworks one has three categories of sets: sets that are natural extensions of, say, subsets of $\mathbb{R}$, more general internal sets, as well as external sets. In the axiomatic frameworks, the standard and nonstandard sets correspond to the natural extensions and the internal sets, whereas there are no external sets.

11.2. Additional principles

The theory SPOT proves that standard integers are an initial segment of $\mathbb{N}$ [**8**, Lemma 2.1].

LEMMA 11.2.1 (Countable Idealisation). *Let ϕ be an $\in$-formula with arbitrary parameters. The theory* SPOT *proves the following:*

$$\forall^{\mathbf{st}} n \in \mathbb{N}\, \exists x\, \forall m \in \mathbb{N}\, (m \leq n \to \phi(m, x)) \longleftrightarrow \exists x\, \forall^{\mathbf{st}} n \in \mathbb{N}\, \phi(n, x).$$

This is proved in [**8**, Lemma 2.2]. One could elucidate Countable Idealization by means of an equivalent version with countable A in words as follows. If for every standard finite subset $a \subseteq A$ there is some x such that for all $z \in a$, one has $\phi(z, x)$, then there is a single x such that $\phi(z, x)$ holds for all standard $z \in A$ simultaneously (the converse is obvious given that all elements of a standard finite set are standard, which is a consequence of [**8**, Lemma 2.1]). This is analogous to saturation (see Corollary 10.3.4).

DEFINITION 11.2.2. SN is the standardisation principle for **st**-$\in$-formulas with no parameters. Namely, let $\phi(v)$ be an **st**-$\in$-formula with no parameters. Then

$$\forall^{\mathbf{st}} A\, \exists^{\mathbf{st}} S\, \forall^{\mathbf{st}} x\, (x \in S \longleftrightarrow x \in A \wedge \phi(x)). \tag{11.2.1}$$

It is proved in [**8**, Lemma 6.1] that SN is equivalent to standardisation for formulas with only standard parameters.

Although separation does not hold, SN is a kind of *approximation* to it in the following sense. The standard elements of S (but not all elements) are exactly those for which $\phi(x)$ holds. Note also that the assumption that all parameters are standard is necessary to maintain conservativity over ZF, because otherwise one could prove the existence of nonprincipal ultrafilters (see [**8**]).

Note that SPOT+SN is also conservative over ZF [**8**, Theorem B, p. 4]. The axiom SN enables one to give a simple infinitesimal definition of the derivative function conservatively over ZF.[1]

[1] To dot the i's, let $\phi_f(x, L)$ be the formula of (10.4.1) depending on the standard parameter f, a real-valued function. Let $A = \mathbb{R}^2$ in (11.2.1). Then passing from f to f' is enabled by the following consequence of (11.2.1) containing only standard

SCOT incorporates the following choice-type axiom CC (which is a strengthening of SP); see [**8**, Section 3, p. 10].

Definition 11.2.3. (CC) Let $\phi(u,v)$ be an **st**-$\in$-formula with arbitrary parameters. Then

$$\forall^{\mathbf{st}} n \in \mathbb{N}\, \exists x\, \phi(n,x) \;\longrightarrow\; \exists f\, (f \text{ is a function} \wedge \forall^{\mathbf{st}} n \in \mathbb{N}\, \phi(n, f(n))).$$

The following definition was given in [**8**, p. 10].

Definition 11.2.4. SCOT is the theory SPOT+ADC+SN+CC.

SCOT (in fact, its subtheory SPOT+CC) also proves the following statement SC [**8**, Lemma 3.1].

Definition 11.2.5. SC (Countable Standardisation) Let $\psi(v)$ be an **st**-$\in$-formula with arbitrary parameters. Then

$$\exists^{\mathbf{st}} S\, \forall^{\mathbf{st}} n\, (n \in S \longleftrightarrow n \in \mathbb{N} \wedge \psi(n)).$$

11.3. Compactness in internal set theories

In this section, we use infinitesimals to deal with compactness conservatively over ZF or ZF+ADC, as indicated below (the traditional extension view was already elaborated in Section 10.3).

Let T be a standard topological space. A point $x \in T$ is *nearstandard in T* if there is a standard $p \in T$ such that $p \in O$ implies $x \in O$ for every standard open set O (in other words, x is in the *halo* of p.)

Lemma 11.3.1. *Assume T is a standard Lindelöf space. If every x in T is nearstandard in T then T is compact.*

Proof. Suppose T is not compact. By downward transfer, there is a standard countable cover $\mathcal{U}$ of T by open sets such that for every (standard) finite k-tuple $O_1, \ldots, O_k \in \mathcal{U}$ there is a $p \in T \setminus \bigcup_{1\le i\le k} O_i$. By Countable Idealisation with the standard parameters T, $\mathcal{U}$, there is $x \in T$ such that $x \notin O$ for any standard $O \in \mathcal{U}$. Such an x is not nearstandard in T, because if x were in the halo of some standard $p \in T$, we would have a standard $O \in \mathcal{U}$ such that $p \in O$ ($\mathcal{U}$ is a cover) and hence $x \in O$, a contradiction.[2] $\square$

parameters: $\exists^{\mathbf{st}} f'\, \forall^{\mathbf{st}}(x,L)\, \big((x,L) \in f' \longleftrightarrow (x,L) \in \mathbb{R}^2 \wedge \phi_f(x,L)\big)$ where f' is thought of as its graph in the plane.

[2]An analogous proof goes through for arbitrary standard topological spaces if one has *full* idealisation (with standard parameters) such as in the theory BSPT$'$ [**8**], which is still conservative over ZF (unfortunately it is not known whether SN can be added to it conservatively over ZF). Note that BSPT$'$ proves the existence of a finite set containing all standard reals [**8**, p. 10]. Such sets are used in Benci's approach to measure theory.

Lemma 11.3.2. *Assume T is a standard second countable space. If T is compact then every $x \in T$ is nearstandard in T.*

Proof. Suppose $\mu \in T$ is not nearstandard in T. Then for every standard $p \in T$ there is a standard open set O such that $p \in O$ and $\mu \notin O$. Let $\mathcal{B}$ be a standard countable base for the topology of T, and let

$$\mathcal{U} = {}^{\mathbf{st}}\{O \in \mathcal{B} : \mu \notin O\}.$$

This set is obtained by SC (Countable Standardisation, see Definition 11.2.5) with a nonstandard parameter (namely, μ), available in SCOT [**8**, Lemma 3.1]. By the above and transfer, $\mathcal{U}$ is a standard open cover of T. If T were compact, $\mathcal{U}$ would have, by transfer, a standard finite open subcover $O_1, \ldots, O_k$. Then $\mu \in O_i$ for some $1 \leq i \leq k$, contradicting the definition of $\mathcal{U}$. $\square$

Since second countable implies Lindelöf, we have the equivalence of the two definitions of compactness, for second-countable spaces in SCOT.

CHAPTER 12

Continuity and uniform continuity

12.1. Continuity

Based on the results of Section 11.3, the following can be proved conservatively over ZF+ADC using infinitesimals. We will first discuss continuity over SPOT.

In this section, f is a standard map between standard topological spaces. f is said to be S-*continuous at* c if whenever $x \simeq c$, one has $f(x) \simeq f(c)$.

LEMMA 12.1.1. *If a standard map f from a first countable topological space into a topological space is S-continuous at a standard point c then f is continuous at c.*

PROOF. Let $\mathcal{B}_c$ be a standard countable base of open neighborhoods of c. Assume that f is not continuous at c. Then there is a standard open neighborhood U of $f(c)$ such that for every (standard) finite $O_1, \ldots, O_k \in \mathcal{B}_c$ there is $x \in \bigcap_{1 \le i \le k} O_i$ with $f(x) \notin U$. By Countable Idealization there is x such that $x \in O$ holds for all standard $O \in \mathcal{B}_c$ and $f(x) \notin U$. Then $x \simeq c$ and $f(x) \not\simeq f(c)$, a contradiction. $\square$

LEMMA 12.1.2. *If a standard function f is continuous at a standard point c then f is S-continuous at c.*

PROOF. Assume that f is not S-continuous at c. Then there is $x \simeq c$ for which $f(x) \not\simeq f(c)$, i.e., $f(x) \notin U$ holds for some standard neighborhood U of $f(c)$. By continuity of f there is a standard open neighborhood O of c such that $z \in O$ implies $f(z) \in U$. As $x \simeq c$, we have $x \in O$ and hence $f(x) \in U$, a contradiction. $\square$

12.2. Continuous image of compacts

We prove the following well-known result in SCOT.

THEOREM 12.2.1. *Let $f : T \to Y$ be a continuous map between second-countable topological spaces. Let $E \subseteq T$ be compact. Then $f(E)$ is compact.*

Proof. We prove the theorem under the assumption that f, T, Y are standard; its validity for arbitrary $f, T.Y$ follows by transfer. By Lemma 11.3.2, every point $x \in E$ is infinitely close to a standard point $p \in E$. By the nonstandard characterisation of continuity of f (Lemma 12.1.2), the point $f(x)$ is infinitely close to $f(p)$. Thus every point $f(x)$ in the image $f(E)$ is infinitely close to a standard point $f(p) \in f(E)$. By Lemma 11.3.1, applied to $f(E)$, the space $f(E)$ is compact. □

The proof compares favorably with the traditional proof using pullbacks of open covers, and is as effective (in the sense explained in Section 10.2) as the traditional proof.

12.3. Characterisation of uniform continuity

Let D, E be standard metric spaces; we will denote the distance functions by $|\cdot|$. A standard map $f\colon D \to E$ is *uniformly continuous* on D if

$$\begin{aligned}&(\forall \epsilon \in \mathbb{R}^+)(\exists \delta \in \mathbb{R}^+)\\ &\quad \underline{(\forall x \in D)(\forall x' \in D)\ \big[|x' - x| < \delta \to |f(x') - f(x)| < \epsilon\big].}\end{aligned} \tag{12.3.1}$$

f is *S-continuous on D* if

$$(\forall x \in D)(\forall x' \in D)\,[x \simeq x' \ \to\ f(x) \simeq f(x')]\,. \tag{12.3.2}$$

Lemma 12.3.1. *If f is uniformly continuous on D then it is S-continuous there.*

Proof. To show that condition (12.3.1) implies (12.3.2), fix a standard parameter ϵ. By downward transfer, there is a standard δ such that the underlined part of formula (12.3.1) holds:

$$(\forall x \in D)(\forall x' \in D)\ \big[|x' - x| < \delta \to |f(x') - f(x)| < \epsilon\big]. \tag{12.3.3}$$

If $x \simeq x'$ then the condition $|x - x'| < \delta$ is satisfied regardless of the value of the standard number $\delta > 0$. Therefore

$$|f(x') - f(x)| < \epsilon. \tag{12.3.4}$$

Since (12.3.4) is true for each standard $\epsilon > 0$, we conclude that $f(x') \simeq f(x)$, proving (12.3.2). □

Lemma 12.3.2. *If f is S-continuous on D then f is uniformly continuous there.*

Proof. We will show the contrapositive statement, namely that $\neg$(12.3.1) implies $\neg$(12.3.2). Assume the negation of (12.3.1). By downward transfer it follows that there exists a standard number $\epsilon > 0$ such that

$$(\forall \delta \in \mathbb{R}^+)(\exists x \in D)(\exists x' \in D)\ \big[|x' - x| < \delta\ \wedge\ |f(x') - f(x)| > \epsilon\big]. \tag{12.3.5}$$

The formula is true for all positive δ, so in particular it holds for an infinitesimal $\delta_0 > 0$. For this value, we obtain

$$(\exists x \in D)(\exists x' \in D)\ \big[|x' - x| < \delta_0\ \wedge\ |f(x') - f(x)| > \epsilon\big]. \tag{12.3.6}$$

Fix such x and x'. The condition $|x' - x| < \delta_0$ implies that $x \simeq x'$, while $|f(x') - f(x)| > \epsilon$. As the lower bound $\epsilon > 0$ is standard, it follows that $f(x') \not\simeq f(x)$. This violates condition (12.3.2) and establishes the required contrapositive implication $\neg(12.3.1) \implies \neg(12.3.2)$. $\square$

12.4. Continuity implies uniform continuity

As shown in Section 12.3, the theory SPOT proves that uniform continuity of a map between metric spaces amounts to S-continuity at all points (standard and nonstandard) of the domain.

Theorem 12.4.1. *A continuous map from a compact metric space to a metric space is uniformly continuous.*

Proof. Let $f : E \to Y$ where f, E, Y are standard. By the characterisation of the compactness of E (Lemma 11.3.2), if $x \in E$ then x is infinitely close to a standard point $p \in E$. For each $x' \simeq x$, one has $x' \simeq p \simeq x$. If f is continuous at p then $f(x') \simeq f(p) \simeq f(x)$, and therefore f is S-continuous at all points of E, establishing uniform continuity by Lemma 12.3.2. By transfer, the theorem holds for arbitrary f, E, Y. $\square$

This proof in SPOT compares favorably with the traditional proof: given $\epsilon > 0$, we need to find $\delta > 0$ such that if $d_E(x, y) < \delta$ then one has $d_Y(f(x), f(y)) < \epsilon$. By continuity, for each $x \in E$ there exists a $\delta_x > 0$ such that if $d(x, y) < \delta_x$ then $d(f(x), f(y)) < \frac{\epsilon}{2}$. Then

$$\big\{B(x, \tfrac{\delta_x}{2}) : x \in E\big\}$$

is an open cover of E. By compactness, there are points $x_1, \ldots, x_n \in E$ such that $\big\{B(x_1, \frac{\delta_1}{2}), \ldots, B(x_n, \frac{\delta_n}{2})\big\}$ is a finite subcover covering E. Let $\delta = \min(\frac{\delta_1}{2}, \ldots, \frac{\delta_n}{2})$. If $y, z \in E$ and $d(y, z) < \delta \le \frac{\delta_k}{2}$ for each $k = 1, \ldots, n$, then by the triangle inequality

$$d(x_k, z) \le d(x_k, y) + d(y, z) \le \tfrac{\delta_k}{2} + \tfrac{\delta_k}{2} < \delta_k.$$

Therefore

$$d(f(y), f(z)) \leq d(f(y), f(x_k)) + d(f(x_k), f(z)) < \tfrac{\epsilon}{2} + \tfrac{\epsilon}{2} = \epsilon,$$

establishing uniform continuity.

12.5. Heine–Borel theorem

Here we present an effective approach to the Heine–Borel theorem exploiting the characterisation of compactness of Section 11.3. A standard set C in a standard metric space is *closed* if every standard element near some element of C is actually in C. It is *bounded*[1] if it contains no unlimited elements. If standard x and y are infinitely close then $x = y$.

LEMMA 12.5.1. *If C is compact then it is closed and bounded.*

PROOF. Every element of C is nearstandard in C by compactness, so there are no unlimited elements, i.e., C is bounded. Let a standard point x be infinitely close to some $y \in C$. By compactness of C there is a standard $z \in C$ such that $z \simeq y$. Thus $x = z \in C$ and C is closed. □

LEMMA 12.5.2. *For standard n, if $C \subseteq \mathbb{R}^n$ is closed and bounded then it is compact.*

PROOF. For standard n, the condition of infinite proximity in $\mathbb{R}^n$ amounts to the condition of infinite proximity for each of the n coordinates. A similar remark applies to boundedness.

Let $x \in C$. Since C is bounded, x is limited and hence infinitely close to a standard $y \in \mathbb{R}^n$. By closure, $y \simeq x$ entails $y \in C$. Thus x is nearstandard in C. This proves that C is compact. □

12.6. Radically Elementary Probability Theory

Nelson's *Radically Elementary Probability Theory* is based on traditional mathematics plus axioms 1 through 5 stated in [**13**, Section 4, pp. 13–14]. The axioms 1 through 4 hold in SPOT. Axiom 5, which according to Nelson is rarely used, is the axiom CC (see Definition 11.2.3). It follows that *Radically Elementary Probability Theory* is conservative over ZF+ADC. Furthermore, it follows that all results from [**13**] automatically hold in SCOT. In particular, this includes Nelson's S-integral.

[1]Here C must be a subset of a metric space for the notion of boundedness to make sense. Note that Countable Idealisation (CI) is needed to prove its equivalence to the usual definition "There is a real r such that for all $x \in C$, $d(x, p) < r$ holds for some (equivalently, all) $p \in C$." CI is available in SPOT.

Further applications include proofs in SPOT of Peano and Osgood theorems for ordinary differential equations [**10**].

Our perspective fits with a relative view of the foundations of mathematics such as that provided by Hamkins' multiverse. For the relation between the Gitman–Hamkins "toy" model of the multiverse [**5**] and nonstandard analysis, see Fletcher et al. [**4**, Section 7.3].

Part 3

Topology, Universes, Superstructure

CHAPTER 13

Saturation, topology, hyperfinite sets

In Section 9.5 we treated countable saturation in the case of natural extensions of standard sets. Now we will deal with the general case of countable saturation for internal sets.

13.1. Saturation of internal sets: the general case

THEOREM 13.1.1. *The intersection of a decreasing sequence*

$$X^1 \supseteq X^2 \supseteq \cdots \supseteq X^k \supseteq \cdots$$

of nonempty internal sets is always nonempty:

$$\bigcap_{k\in\mathbb{N}} X^k \neq \emptyset.$$

PROOF. This is a delicate analysis of the ultrapower construction, involving a kind of diagonalisation argument.[1] It is a refinement of the proof of Lemma 9.5.1.

For each $k \in \mathbb{N}$, let $X^k = [A_n^k]$, so that X^k is the internal set defined by the sequence $\langle A_n^k : n \in \mathbb{N}\rangle$ of subsets of $\mathbb{R}$. By Section 8.2 (algebra of internal sets), both sets

$$\{n \in \mathbb{N} : A_n^k \neq \emptyset\} \text{ and } \{n \in \mathbb{N} : A_n^k \supseteq A_n^{k+1}\}$$

belong to F. Now let

$$J^k = \{n \in \mathbb{N} : A_n^1 \supseteq \cdots \supseteq A_n^k \neq \emptyset\}. \tag{13.1.1}$$

REMARK 13.1.2. Intuitively, the set J^k is the set of indices n for which the k real sets $A_n^1, \ldots A_n^k$ behave "as they should" in the sense of mimicking the nesting behavior of the internal sets $X^1, \ldots, X^k$ themselves.

Since the ultrafilter F is closed under finite intersections, for each $k \in \mathbb{N}$ we have $J^k \in F$. Note that $J^1 \supseteq J^2 \supseteq \cdots$.

[1]A possibly more conceptual argument was outlined by Andreas Blass at `https://math.stackexchange.com/a/442194/72694`.

We want to find a hyperreal $[s_n]$ that belongs to every X^k. This will require that for each k we have $s_n \in A_n^k$ for F-almost all n. We will arrange this to work for F-almost all $n \geq k$, in the sense that

$$\{n \in \mathbb{N} : k \leq n\} \cap J^k \subseteq \{n \in \mathbb{N} : s_n \in A_n^k\}. \tag{iii}$$

Note that

- the set $\{n \in \mathbb{N} : k \leq n\}$ is cofinite in $\mathbb{N}$, and so belongs to the nonprincipal ultrafilter F;
- as already mentioned, $J^k \in F$.

Hence formula (iii) yields $\{n \in \mathbb{N} : s_n \in A_n^k\} \in F$. Therefore $[s_n] \in X^k$ as required.

It thus remains to define s_n satisfying condition (iii). For $n \in J^1$ let

$$k_n = \max\{i : i \leq n \text{ and } n \in J^i\}. \tag{iv}$$

Then $n \in J^{k_n}$, so by the definition (13.1.1) of J^{k_n} we can choose some $s_n \in A_n^{k_n}$, and hence

$$s_n \in A_n^1 \cap \cdots \cap A_n^{k_n}. \tag{v}$$

For $n \notin J^1$, let s_n be arbitrary. Now, to prove (iii), observe that if $k \leq n$ and $n \in J^k$, then by (iv), $k \leq k_n$, and so by (v), $s_n \in A_n^k$. $\square$

13.2. Algebra of countable families of internal sets

Countable saturation has some important consequences for the nature of countable unions and intersections of internal sets:

COROLLARY 13.2.1. *If $\{X_n : n \in \mathbb{N}\}$ is a collection of internal sets and X is internal, then:*

(1) $\cap_{n\in\mathbb{N}} X_n \neq \emptyset$ *if $\{X_n : n \in \mathbb{N}\}$ has the finite intersection property.*
(2) *If $X \subseteq \cup_{n\in\mathbb{N}} X_n$, then already $X \subseteq \cup_{n\leq k} X_n$ for some $k \in \mathbb{N}$.*
(3) *If $\cap_{n\in\mathbb{N}} X_n \subseteq X$, then already $\cap_{n\leq k} X_n \subseteq X$ for some $k \in \mathbb{N}$.*
(4) *If $\cup_{n\in\mathbb{N}} X_n$ is internal, then it equals $\cup_{n\leq k} X_n$ for some $k \in \mathbb{N}$.*
(5) *If $\cap_{n\in\mathbb{N}} X_n$ is internal, then it equals $\cap_{n\leq k} X_n$ for some $k \in \mathbb{N}$.*

PROOF. (1) Let $Y^k = X_1 \cap \cdots \cap X_k$. Then $Y^1 \supseteq Y^2 \supseteq \cdots$, and each Y^k is internal by Section 8.2 (algebra of internal sets). The finite intersection property implies that $Y^k \neq \emptyset$, so by Theorem 13.1.1 there is some hyperreal that belongs to every Y^k, and hence to every X_k.

(2) We argue by contradiction. Suppose that for all $k \in \mathbb{N}$, we have $X \not\subseteq \cup_{n\leq k} X_n$ and hence the finite intersection satisfies

$$\cap_{n\leq k}(X - X_n) = X - (\cup_{n\leq k} X_n) \neq \emptyset.$$

Then $\{X - X_n : n \in \mathbb{N}\}$ is a collection of internal sets with the finite intersection property, so by (1) there is some x with

$$x \in \cap_{n \in \mathbb{N}}(X - X_n) = X - (\cup_{n \in \mathbb{N}} X_n).$$

Hence $X \nsubseteq \cup_{n \in \mathbb{N}} X_n$, contradicting our hypothesis.

(3) Exercise.

(4) Put $X = \cup_{n \in \mathbb{N}} X_n$ in (2).

(5) Similarly, from (3). □

Result (4) of Corollary 13.2.1 plays a crucial role in the nonstandard approach to measure theory discussed later.

13.3. Saturation creates nonstandard entities

The use of the term *saturation* is intended to convey that ${}^*\mathbb{R}$ is "full of elements". Countable saturation ensures the existence of those elements that can be characterised as belonging to the intersection of a decreasing sequence of internal sets.

EXAMPLE 13.3.1. Let X_n be the hyperreal interval $(0, \frac{1}{n}) \subseteq {}^*\mathbb{R}$. Then $\langle X_n : n \in \mathbb{N}\rangle$ is a decreasing sequence of nonempty internal sets. Its (nonempty) intersection $\cap_{n \in \mathbb{N}} X_n$ is precisely the set of positive infinitesimals.

Another consequence of saturation is the following property.

THEOREM 13.3.2. *Every sequence of infinitesimals has an infinitesimal upper bound.*

PROOF. Take $\langle e_n : n \in \mathbb{N}\rangle$ with $e_n \approx 0$ for all $n \in \mathbb{N}$.

Let X_n be the hyperreal interval $[e_n, \frac{1}{n})$. Then X_n is internal, and the collection $\{X_n : n \in \mathbb{N}\}$ has the finite intersection property. For in general, if e is the maximum element of $\{e_{n_1}, \dots, e_{n_k}\}$, then

$$e \in \left[e_{n_1}, \tfrac{1}{n_1}\right) \cap \dots \cap \left[e_{n_k}, \tfrac{1}{n_k}\right).$$

By saturation, the countable intersection is nonempty. But any member of $\bigcap_{n \in \mathbb{N}} X_n$ is an upper bound of the e_n's that is smaller than $\frac{1}{n}$ for all $n \in \mathbb{N}$, and hence is infinitesimal. □

Dually, we can use saturation to show the following.

THEOREM 13.3.3. *Every sequence $\langle s_n : n \in \mathbb{N}\rangle$ of infinite hypernatural numbers has an infinite hypernatural lower bound.*

PROOF. For each $n \in \mathbb{N}$, consider the internal interval $X_n = (n, s_n]$. Then any x belonging to $\cap_{n \in \mathbb{N}} X_n \neq \emptyset$ is a positive infinite lower bound of the terms s_n. By transfer, we can take a member of ${}^*\mathbb{N}$ between $x - 1$

and x to get an infinite hypernatural number that is less that s_n for all $n \in \mathbb{N}$. (Alternatively, put $X_n = (n, s_n] \cap {}^*\mathbb{N}$ in this argument.) $\square$

13.4. The cardinality of an internal set

Countable saturation implies that ${}^*\mathbb{R}$ has so many elements that an infinite internal set cannot be countably infinite.

THEOREM 13.4.1. *Every internal set is either finite or uncountable.*

In the case of subsets of $\mathbb{R}$, we already showed that any internal set of reals must be finite. In proving this in Section 9.2 we showed in effect that an internal subset of $\mathbb{R}$ cannot be put in one-to-one correspondence with $\mathbb{N}$. We can now can demonstrate this for any internal set whatsoever.

PROOF. Suppose $X = \{x_n : n \in \mathbb{N}\}$ (possibly with repetitions) is a countable internal set. We remove all the points from X one by one, by defining for each n the set $X_n = X - \{x_1, \ldots, x_n\}$, which is internal. Then the sets X_n form a decreasing sequence. If they were all nonempty, countable saturation would imply that their intersection would be nonempty, which is false. We must therefore conclude that there is an n for which $X_n = \emptyset$ and so $X = \{x_1, \ldots, x_n\}$,

This shows that any countable internal set must be finite. Hence an infinite internal set must be uncountable. $\square$

This observation has the following consequence for the structure of the set ${}^*\mathbb{N}$ of hypernatural numbers. If H is an infinite hypernatural, then the initial segment $\{1, 2, \ldots, H\}$ of ${}^*\mathbb{N}$ is internal, and is certainly infinite, since it includes all of $\mathbb{N}$, so is uncountable. It follows that there must be uncountably many infinite members of ${}^*\mathbb{N}$ that are less than H. The set of all infinite hypernaturals is partitioned into ${}^*\mathbb{N}$-galaxies, each of which looks like a copy of $\mathbb{Z}$. If H is infinite, then there are uncountably many of these ${}^*\mathbb{N}$-galaxies between $\mathbb{N}$ and H.

13.5. Closure of the shadow of an internal set

For any $X \subseteq {}^*\mathbb{R}$, let

$$sh(X) = \{sh(x) : x \in X \text{ and } x \text{ is finite}\}.$$

EXAMPLE 13.5.1. Let X be the interval $(a, b) \subseteq {}^*\mathbb{R}$. Let a, b be finite. Then $sh(X)$ is the closed interval $[sh(a), sh(b)]$ in $\mathbb{R}$. If a is finite but b infinite, then $sh(X) = [sh(a), +\infty) \subseteq \mathbb{R}$, again a topologically closed subset of $\mathbb{R}$.

THEOREM 13.5.2. *If X is internal, then $sh(X)$ is closed.*

PROOF. Let $r \in \mathbb{R}$ be a closure point of $sh(X)$. We need to show that $r \in sh(X)$, i.e., r is the shadow of some $y \in X$. For each $n \in \mathbb{N}$, the hyperreal open interval $(r - \frac{1}{n}, r + \frac{1}{n})$ meets $sh(X)$ in some real point s_n that must be the shadow of some $x_n \in X$. Hence $x_n \approx s_n \in (r - \frac{1}{n}, r + \frac{1}{n})$, so the internal set

$$X_n = X \cap (r - \tfrac{1}{n}, r + \tfrac{1}{n})$$

contains x_n and is therefore nonempty. The sets X_n form a decreasing sequence. By countable saturation, there is a point $y \in \bigcap_{n \in \mathbb{N}} X_n$ in their intersection. Then $y \in X$ and $|y - r| < \frac{1}{n}$ for all $n \in \mathbb{N}$, so $y \approx r$. Hence

$$r = sh(y) \in sh(X).$$

Hence $sh(X)$ contains all its closure points and so is closed. □

Topological closure of the shadow of an internal set plays an important role in the hyperreal "reconstruction" of the Lebesgue measure. This will be explained later.

13.6. Interval topology and hyper-open sets

We introduce the following three notions.

DEFINITION 13.6.1. A set A of hyperreals is *interval-open* if each of its points belongs to some hyperreal open interval (a, b) that is included in A. The family of interval-open sets is the *interval topology* on ${}^*\mathbb{R}$.

Thus the interval-open sets are precisely those that are unions of hyperreal open intervals. A "thinner" family of sets is the following.

DEFINITION 13.6.2. A *real-open* set is one that is a union of hyperreal open neighbourhoods $(r - c, r + c)$ having real radius $c > 0$.

Equivalently, a real-open set is a union of hyperreal open intervals of appreciable length. Each real-open set is interval-open, but not conversely: the real-open sets are not a topology on ${}^*\mathbb{R}$, since they are not closed under intersection.

EXAMPLE 13.6.3. Let $r = 2 - \varepsilon$ for a positive infinitesimal $\varepsilon > 0$. Then the intersection of hyperreal intervals $(-1, 1)$ and $(r - 1, r + 1)$ is of infinitesimal size in the sense that it is contained in $hal(1)$, and therefore not real-open.

An even "thinner" family of sets is the following.

DEFINITION 13.6.4. An S-*open* set is a union of S-*neighbourhoods* $((r-c, r+c))$ having real radius $c > 0$, where

$$((r-c, r+c)) = \{x \in {}^*\mathbb{R} : hal(x) \subseteq (r-c, r+c)\}.^2$$

The S-open sets form the S-*topology* on ${}^*\mathbb{R}$.

Every S-open set is real-open, but not conversely. Every S-open set is a union of halos, but not conversely.

EXAMPLE 13.6.5. The set

$$\mathbb{L} = \cup_{n\in\mathbb{N}}(-n, n)$$

of finite numbers is external. Thus, while a real-open set is always a union of internal sets (namely, open intervals), it may itself be external.

We now introduce a further class of subsets of ${}^*\mathbb{R}$.

DEFINITION 13.6.6. An internal set $[A_n]$ is *hyper-open* if

$$\{n \in \mathbb{N} : A_n \text{ is open in } \mathbb{R}\} \in F.$$

Each hyperreal interval (a, b) is hyper-open: if $a = [a_n]$ and $b = [b_n]$, then (a, b) is the internal set defined by the sequence $\langle A_n : n \in \mathbb{N}\rangle$, where A_n is the real interval (a_n, b_n), which is indeed open in $\mathbb{R}$.

LEMMA 13.6.7. *Every hyper-open set is a union of hyperreal open intervals.*

PROOF. Let $A = [A_n]$ be hyper-open. Take a point $r = [r_n] \in A$. Then we find that the set

$$J = \{n \in \mathbb{N} : r_n \in A_n \text{ and } A_n \text{ is open in } \mathbb{R}\}$$

belongs to the ultrafilter F. Our task is to show that r belongs to some hyperreal interval (a, b) that is included in A.

Now, if $n \in J$, then there is some real interval $(a_n, b_n) \subseteq \mathbb{R}$ with $r_n \in (a_n, b_n)$. Since $J \in F$, this is enough to specify a as the hyperreal number $[a_n]$ and b as $[b_n]$. Furthermore, working with the properties of F, we can show that

(1) $[a_n] < [r_n] < [b_n]$, and
(2) $[s_n] \in [A_n]$ whenever $[a_n] < [s_n] < [b_n]$.

Therefore

$$r \in (a, b) \subseteq A$$

as required. □

[2] Here we discard a pair of half-halos at the extremities of $(r-c, r+c)$.

This lemma implies that every hyper-open set is interval-open. But there are interval-open sets, like the set $\mathbb{L}$ of finite numbers, that are not hyperopen, simply because they are external, whereas hyper-open sets are always internal by definition. The example of $\mathbb{L}$ shows that the family of hyperopen sets is not a topology, because it is not closed under infinite unions. Instead, it is what is known as a *base* for the *interval topology*, because every interval-open set is a union of hyper-open sets (open intervals).

LEMMA 13.6.8. *The families of real-open sets and hyper-open sets are incomparable.*

PROOF. The set $\mathbb{L}$ is real-open (indeed S-open) but not hyper-open, while any infinitesimal length open interval is hyper-open but not real-open. $\square$

This latter example shows that even for internal sets the two classes remain distinguishable. There is a characterisation of S-openness of internal sets that corresponds to the nonstandard characterisation of openness of subsets of $\mathbb{R}$ and involves an interesting application of underflow.

THEOREM 13.6.9. *If B is an internal set, then B is S-open if and only if it contains the halo of each of its points.*

PROOF. The proof uses the underflow principle. Recall that an S-open set is a union of halos.

Conversely, assume that $hal(r) \subseteq B$ whenever $r \in B$. For such an r, consider the set

$$X = \left\{n \in {}^*\mathbb{N} : (\forall x \in {}^*\mathbb{R})\,(|r - x| < \tfrac{1}{n} \to x \in B)\right\}.$$

Since B is internal, it follows by the internal set definition principle that X is internal. Moreover, since $hal(r) \subseteq B$, it follows that X contains every infinite hypernatural n, because for such an n, the bound $|r - x| < \frac{1}{n}$ implies $x \in hal(r)$. Hence by underflow, X must contain some standard $n \in \mathbb{N}$. It follows that B includes the real-radius interval $(r - \frac{1}{n}, r + \frac{1}{n})$. But then, since $\frac{1}{n}$ is real,

$$r \in ((r - \tfrac{1}{n}, r + \tfrac{1}{n})) \subseteq (r - \tfrac{1}{n}, r + \tfrac{1}{n}) \subseteq B.$$

It follows that B is the union of S-neighbourhoods, and is thereby S-open. $\square$

13.7. Internal functions

Let $\langle f_n : n \in \mathbb{N}\rangle$ be a sequence of functions $f_n : A_n \to \mathbb{R}$, with domains $A_n \subseteq \mathbb{R}$.

DEFINITION 13.7.1. An *internal function* is an ${}^*\mathbb{R}$-valued function $[f_n]$ defined on the internal set $[A_n]$ by setting

$$[f_n]([r_n]) = [f_n(r_n)].$$

Observe that if $[r_n] \in [A_n]$, then the set $J = \{n \in \mathbb{N} : r_n \in A_n\}$ belongs to F, and for each $n \in J$, $f_n(r_n)$ is defined. This is enough to make $[f_n]([r_n])$ well-defined. We have

$$dom[f_n] = [dom f_n].$$

In the case that (f_n) is a constant sequence, with $f_n = f : A \to \mathbb{R}$ for all n, then $[f_n]$ is just the function ${}^*f : {}^*A \to {}^*\mathbb{R}$ extending f.

The following result shows that we only need to specify almost all of the real functions f_n in order to determine the internal function $[f_n]$.

THEOREM 13.7.2. *Let $\langle f_n : n \in \mathbb{N}\rangle$ and $\langle g_n : n \in \mathbb{N}\rangle$ be sequences of partial functions from $\mathbb{R}$ to $\mathbb{R}$. Then the internal functions $[f_n]$ and $[g_n]$ are equal if and only if*

$$\{n \in \mathbb{N} : f_n = g_n\} \in F.$$

PROOF. Let

$$J_{fg} = \{n \in \mathbb{N} : f_n = g_n\}. \tag{13.7.1}$$

Suppose $J_{fg} \in F$. Now in general, two functions are equal precisely when they have the same domain and assign the same values to all members of that domain. Thus

$$J_{fg} \subseteq \{n \in \mathbb{N} : dom f_n = dom\, g_n\},$$

leading to the conclusion that the internal sets $[dom f_n]$ and $[dom\, g_n]$ are equal, i.e., $dom[f_n] = dom[g_n]$. But for $[r_n] \in dom[f_n]$,

$$J_{fg} \cap \{n \in \mathbb{N} : r_n \in dom f_n\} \subseteq \{n \in \mathbb{N} : f_n(r_n) = g_n(r_n)\},$$

which, by (13.7.1), leads to $[f_n]([r_n]) = [g_n]([r_n])$. Hence $[f_n] = [g_n]$.

For the converse, suppose that $J_{fg} \notin F$. Now, J_{fg}^c is a subset of the union

$$\{n \in \mathbb{N} : dom f_n \neq dom\, g_n\} \cup \{n \in \mathbb{N} : dom f_n = dom\, g_n \text{ but } f_n \neq g_n\},$$

so either $\{n \in \mathbb{N} : dom f_n \neq dom\, g_n\} \in F$, whence $dom[f_n] \neq dom[g_n]$ and so $[f_n] \neq [g_n]$, or else

$$J = \{n \in \mathbb{N} : dom f_n = dom g_n \text{ but } f_n \neq g_n\} \in F.$$

But for $n \in J$ there exists some r_n with $f_n(r_n) \neq g_n(r_n)$. This leads to $[f_n]([r_n]) \neq [g_n]([r_n])$, and so $[f_n] \neq [g_n]$. $\square$

13.8. Hyperfinite sets

If A_n is finite for (almost) all $n \in \mathbb{N}$, then $[A_n]$ may nevertheless be infinite (and then in fact uncountable by Theorem 13.4.1) but will have many properties that are similar to those of finite sets.

DEFINITION 13.8.1. An internal set $A = [A_n]$ is called *hyperfinite* if almost all of the sets A_n are finite, i.e., if

$$\{n \in \mathbb{N} : A_n \text{ is finite}\} \in F.$$

In that case, we may as well assume that all the sets A_n are finite and have finite integer size $|A_n|$.

DEFINITION 13.8.2. The internal cardinality (or size) of A is the hyperinteger

$$|A| = [\langle |A_n| : n \in \mathbb{N}\rangle].$$

More succinctly, $|[A_n]| = [|A_n|]$.

EXAMPLE 13.8.3. Let $A_n = \{1, \ldots, n\} \subseteq \mathbb{N}$. The resulting hyperfinite set A includes $\mathbb{N}$. Being internal, it must therefore be an uncountable subset of ${}^*\mathbb{N}$. To see that $\mathbb{N} \subseteq A$, observe that if $m \in \mathbb{N}$, then the set $\{n \in \mathbb{N} : m \in A_n\}$ is cofinite, being equal to $\{m, m+1, \ldots\}$, so belongs to F. Hence ${}^*m \in A$.

EXAMPLE 13.8.4. In the previous example, we saw that

$$\mathbb{N} \subseteq [A_n] \subseteq {}^*\mathbb{N}.$$

We can refine the example by replacing $\mathbb{N}$ by B. If B is any countable subset of $\mathbb{R}$, then there exists a hyperfinite set A with

$$B \subseteq A \subseteq {}^*B.$$

For if $B = \{x_n : n \in \mathbb{N}\}$, let $A = [A_n]$ where $A_n = \{x_1, \ldots, x_n\}$. In this case the internal size of A is $\omega = [(1, 2, 3, \ldots)]$.

REMARK 13.8.5. gLater we will see that the restriction to countability here can be removed: any subset B of $\mathbb{R}$ has a "hyperfinite approximation" A satisfying $B \subseteq A \subseteq {}^*B$.[3]

EXAMPLE 13.8.6. Any finite set of hyperreals is hyperfinite: as observed in Section 8.1, if $X = \{[r_n^1], \ldots, [r_n^k]\} \subseteq {}^*\mathbb{R}$, then X is the hyperfinite set $[A_n]$, where

$$A_n = \{r_n^1, \ldots, r_n^k\}.$$

[3]This involves more advanced models than those we have constructed so far.

If $H = [H_n] \in {}^*\mathbb{N}$, then the set

$$\{k \in {}^*\mathbb{N} : k \leq H\} = \{1, 2, \ldots, H\}$$

discussed in Section 8.1 is hyperfinite and has internal cardinality H, since it is equal to $[A_n]$, where $A_n = \{1, 2, \ldots, H_n\}$ and $|A_n| = H_n$.

EXAMPLE 13.8.7. If $H = [H_n] \in {}^*\mathbb{N}$, then the set

$$\left\{\frac{k}{H} : k \in {}^*\mathbb{Z} \text{ and } 0 \leq k \leq H\right\} = \left\{0, \frac{1}{H_n} \ldots, \frac{H_n - 1}{H_n}, 1\right\}$$

is hyperfinite of internal cardinality $H + 1$, since it is equal to $[A_n]$, where

$$A_n = \left\{0, \frac{1}{H_n}, \cdots, \frac{H_n - 1}{H_n}, 1\right\}.$$

EXAMPLE 13.8.8. More generally, for any hyperreals a, b, and any $H \in {}^*\mathbb{N}$, the uniform partition

$$\left\{a + k\frac{(b-a)}{H} : k \in {}^*\mathbb{Z} \text{ and } 0 \leq k \leq H\right\}$$

is hyperfinite of internal cardinality $H + 1$.

CHAPTER 14

Universes

14.1. Counting a hyperfinite set

The results of Section 13.8 are indicative of ways in which hyperfinite sets behave like finite sets. More fundamentally, a finite set can be defined as one that has n elements for some $n \in \mathbb{N}$, and so is in bijective correspondence with the set $\{1, \ldots, n\}$. Correspondingly, for hyperfinite sets we have the following result.

THEOREM 14.1.1. *An internal set A is hyperfinite with internal cardinality H if and only if there is an internal bijection $f : \{1, \ldots, H\} \to A$.*

For the proof see the note.[1]

An important feature of this result is that it gives a characterisation of hyperfinite sets that makes no reference to the ultrafilter F, but requires only the hypernatural numbers $^*\mathbb{N}$ and the notion of an internal function. Superstructures provide a systematic way of adopting an approach which sidesteps ultrafilters.

---

[1]Let $A = [A_n]$. If A is hyperfinite with internal cardinality $H = [H_n]$, then we may suppose that for each $n \in \mathbb{N}$, A_n is a finite set of cardinality H_n. Thus there is a bijection $f_n : \{1, \ldots, H_n\} \to A_n$. Let $f = [f_n]$. Then f is an internal function with domain $\{1, \ldots, H\}$ that is injective (Goldblatt's 12.2(4)) and has range A (Goldblatt's 12.2(1)). Conversely, suppose that $f = [f_n]$ is an internal bijection from $\{1, \ldots, H\}$ onto A. Then $[dom f_n] = dom[f_n] = \{1, \ldots, H\} = [\{1, \ldots, H_n\}]$, so for F-almost all n,

$$dom f_n = \{1, \ldots, H_n\} \tag{i}$$

Also, as A is the image of $\{1, \ldots, H\}$ under $[f_n]$, Goldblatt's Exercise 12.2(1) implies that $A = [f_n(\{1, \ldots, H_n\})]$, so

$$f_n(\{1, \ldots, H_n\}) = A_n \tag{ii}$$

for F-almost all n. Finally, by Goldblatt's 12.2(4),

$$f_n \text{ is injective} \tag{iii}$$

for F-almost all n. Then the set J of those $n \in \mathbb{N}$ satisfying (i)-(iii) must belong to F. But for $n \in J$, the set A_n is finite of cardinality H_n. Hence A is hyperfinite of internal cardinality H.

14.2. Motivating superstructures

Remark 14.2.1. The goal of Sections 14.2 through 14.8 is to motivate the introduction of superstructures, defined in Section 14.9.

We formulate some questions concerning the extension of the power of our techniques so as to make it applicable in a variety of fields in mathematics.

Remark 14.2.2 (Proofs based on transfer; range of quantifiers). In earlier sections, we proved internal versions of

(1) induction,
(2) least number principle,
(3) order-completeness,

etc. The proofs of these results reverted to ultrafilter calculations. Could one, instead, obtain such results by a transfer principle, involving an extended version of the formal language of Chapter 4? The idea would be to use a more expressive language that would allow the quantifiers $\forall$ and $\exists$ to range over suitable collections of sets or functions rather than just over numbers.

Remark 14.2.3 (Generalizing internality; power set). Now that we see how to identify certain subsets and functions in ${}^*\mathbb{R}$ as being internal, can we do the same for other more complex entities? If a set A is hyperfinite, is its powerset

$$\mathcal{P}(A)$$

also hyperfinite, or is it the collection of internal subsets of A that should be hyperfinite?

14.3. What do we need in the mathematical world?

In developing a mathematical theory, or analysing a particular structure, access may be needed to a wide range of entities: sets, members of sets, sequences, relations, functions, etc.

Remark 14.3.1 ($\mathbb{U}$, $\mathcal{L}_{\mathbb{U}}$, ${}^*\mathbb{U}$). We will posit the existence of a "universe"

$$\mathbb{U}$$

that contains all such entities that might be required. This will have an associated formal language

$$\mathcal{L}_{\mathbb{U}}$$

whose sentences express properties of the members of $\mathbb{U}$. Then $\mathbb{U}$ will be enlarged to another universe

$$^*\mathbb{U}$$

that contains certain new (nonstandard) entities whose behaviour can be used to establish results about $\mathbb{U}$ by the use of transfer and other principles.

What kind of entities and closure properties should $\mathbb{U}$ have?

14.3.1. Individuals; set $\mathbb{X}$. Although a real number might be viewed as a set of Cauchy sequences, or a pair of sets of rationals, when studying real analysis we generally regard real numbers as individuals,[2] i.e., as "points" or entities that have no internal structure. The same applies to the basic elements of any other structure that might concern us, be they elements of an algebraic number field, complex numbers, vectors in some Hilbert space, and so on. The universe $\mathbb{U}$ will contain a set

$$\mathbb{X}$$

of entities that are viewed as individuals in this way. An element of $\mathbb{X}$ will be taken to have no members within $\mathbb{U}$. It will be assumed that $\mathbb{R} \subseteq \mathbb{X}$.

14.3.2. Functions; B^A. If two sets A and B belong to $\mathbb{U}$, then we may wish to have all functions $f : A \to B$ available in $\mathbb{U}$, along with

- the range of f,
- the f-image $f(C) \subseteq B$ of any $C \subseteq A$,
- the inverse image of any subset of B under f.

Moreover, the set

$$B^A$$

of all functions from A to B should itself be in $\mathbb{U}$. Also, we should be able to compose functions in $\mathbb{U}$.

14.3.3. Relations; Cartesian products. An m-ary relation is a certain set of m-tuples $\langle a_1, \ldots, a_m \rangle$. Such a relation is usually presented as a subset of some Cartesian product $A_1 \times \cdots \times A_m$, the latter being the set of all m-tuples that have $a_1 \in A_1, \ldots, a_m \in A_m$. Thus $\mathbb{U}$ should be closed under the formation of tuples, and of Cartesian products and their subsets. For binary relations ($m = 2$) the domain and range should be available, and the operations of composing and inverting relations should be possible within our universe.

[2] Atoms - atomim? pritim? urelements?

14.3.4. Set Operations. All the usual set operations of intersection $A \cap B$, union $A \cup B$, difference $A - B$, and power set $\mathcal{P}(A)$, when performed on sets in $\mathbb{U}$, should produce entities that belong to $\mathbb{U}$. In fact, some important constructions will require the union

$$\cup Y$$

and intersection $\cap Y$ of any (possibly infinite) collection $Y \in \mathbb{U}$ to be available. Also, if a set A belongs to $\mathbb{U}$, then all subsets of A should too.

14.4. Transitivity

If a set A is in $\mathbb{U}$ (i.e., $A \in \mathbb{U}$), we will want all members of A to be present in $\mathbb{U}$ as well, i.e.,

$$A \subseteq \mathbb{U}.$$

Definition 14.4.1. The condition $(A \in \mathbb{U}) \to (A \subseteq \mathbb{U})$ is called transitivity of $\mathbb{U}$, because it takes the form

$$a \in A \in \mathbb{U} \text{ implies } a \in \mathbb{U}. \tag{14.4.1}$$

Remark 14.4.2. This has an important bearing on the interpretation of a bounded quantifier $(\forall x \in A)$. We naturally read this as "for all x in A", but when used to express a property of an entity of $\mathbb{U}$, there is a potential issue as to whether this means "for all x in A that belong to $\mathbb{U}$", or whether the variable x is ranging over all members of A absolutely. When $\mathbb{U}$ is transitive, this is not an issue: the members of A that belong to $\mathbb{U}$ are simply all the members of A that there are.

Transitivity thus ensures that quantified variables always range over members of $\mathbb{U}$ when given their natural interpretation.

14.4.1. Subset and relation closure. Transitivity of $\mathbb{U}$ together with closure under the power set operation will guarantee that $\mathbb{U}$ has the property mentioned above of closure under subsets of its members. For then if $A \subseteq B \in \mathbb{U}$, we get

$$A \in \mathcal{P}(B) \in U, \tag{14.4.2}$$

and hence $A \in \mathbb{U}$ by transitivity (14.4.1) applied to (14.4.2).

Then closure of $\mathbb{U}$ under Cartesian products will lead to closure under relations between given sets in general.

Lemma 14.4.3. *Let A, B be sets in $\mathbb{U}$, and $R \subseteq A \times B$. If $A \times B \in \mathbb{U}$, then $R \in \mathbb{U}$.*

This follows by the argument just given for subset closure.

14.5. Specifying primitive concepts: pairs are enough

The more we assume about the entities that exist and constructions that can be performed within $\mathbb{U}$, the more powerful will be this universe as a tool for applications. On the other hand, for demonstrating properties of $\mathbb{U}$ itself or showing that it exists (and $^*\mathbb{U}$ does too), the following is desirable.

REMARK 14.5.1 (Primitive concepts). It is desirable to have very few primitive concepts, so that we can minimize the number of cases and the amount and complexity of work required in carrying out proofs.

Studies of the foundations of mathematics have shown that these opposing tendencies can be effectively balanced by basing our conceptual framework on set theory. To see this we will first show that apart from purely set-theoretic operations, the other notions described in Section 14.3 can be reduced to the construction of sets of ordered pairs.

LEMMA 14.5.2 (Functions). *A function $f : A \to B$ can be identified with the set of pairs*

$$\{\langle a, b\rangle : b = f(a)\},$$

which is a subset of the Cartesian product set $A \times B$.

DEFINITION 14.5.3. Set-theoretically, we define a function from A to B to be a set f of pairs satisfying

(i) if $\langle a, b\rangle \in f$ then $a \in A$ and $b \in B$;

(ii) if $\langle a, b\rangle, \langle a, c\rangle \in f$, then $b = c$ (functionality);

(iii) for each $a \in A$ there exists $b \in B$ with $\langle a, b\rangle \in f$ (the domain of f is A).

LEMMA 14.5.4 (m-tuples). *Given a construction for ordered pairs (2-tuples), the case $m > 2$ can be handled, as well.*

This is done by defining

$$\langle a_1 \ldots, a_m\rangle = \{\langle 1, a_1\rangle, \ldots, \langle m, a_m\rangle\}.$$

Thus an m-tuple becomes a set of ordered pairs (and actually is a function with domain $\{1, \ldots, m\}$).[3]

LEMMA 14.5.5 (Relations). *An m-ary relation is a certain set of m-tuples $\langle a_1, \ldots, a_m\rangle$, and hence becomes a set of sets of ordered pairs.*

[3]An alternative approach would be to inductively put $\langle a_1, \ldots, a_{m+1}\rangle = \langle\langle a_1, \ldots, a_m\rangle, a_{m+1}\rangle$, so that an m-tuple becomes a pair of pairs of $\cdots$ of pairs. This works just as well, but would be more complex set-theoretically than the definition given.

The Cartesian product $A_1 \times \cdots \times A_m$ is a particular case of this, being the set of all such m-tuples that have $a_1 \in A_1, \ldots, a_m \in A_m$.

14.6. Actually, sets are enough

In Section 14.5, we argued that ordered pairs are sufficient for the kinds of purposes we have in mind. But what is an ordered pair? Well, one of the most effective ways to explain a mathematical concept is to give an account of when two instances of the concept are equal, and for ordered pairs the condition is that

$$\langle a, b\rangle = \langle c, d\rangle \text{ iff } a = c \text{ and } b = d.$$

In fact, this condition is all that is ever needed in handling pairs, and it can be fulfilled by setting

$$\langle a, b\rangle = \{\{a\}, \{a, b\}\}.$$

In this way pairs are represented as certain sets, and therefore so too are m-tuples, relations, and functions. When it comes to the study of a particular structure whose elements belong to some given set $\mathbb{X}$, all the entities we need can be obtained by applying set theory to $\mathbb{X}$. This demonstrates the power of set theory, and explains the sense in which it provides a foundation for mathematics.

Lemma 14.6.1 (Product closure). *Closure of $\mathbb{U}$ under Cartesian products can be derived set-theoretically from*

(1) *transitivity and*
(2) *closure under unions and power sets.*

Proof. If $A, B \in \mathbb{U}$ and $\langle a, b\rangle \in A \times B$, then both $\{a\}$ and $\{a, b\}$ are subsets of $A \cup B$, i.e., members of $\mathcal{P}(A \cup B)$. Hence

$$\langle a, b\rangle = \{\{a\}, \{a, b\}\} \in \mathcal{PP}(A \cup B).$$

This shows that $A \times B \subseteq \mathcal{PP}(A \cup B)$, and so

$$A \times B \in \mathcal{PPP}(A \cup B).$$

Thus, closure under $\cup$ and $\mathcal{P}$ and transitivity of $\mathbb{U}$ give $A \times B \in \mathbb{U}$. $\square$

14.7. Strong transitivity

Before giving the axioms for a universe, there is a further important property to be explained, which we do with the following example.

Example 14.7.1 (Question). If a binary relation R belongs to $\mathbb{U}$, then its domain $dom\, R$ should be available in $\mathbb{U}$ as well. Now, if $a \in$

dom R, then there is some entity b with $\langle a, b\rangle \in R$. According to our new definition of pairs, we then have the "membership chain"

$$a \in \{a\} \in \langle a, b\rangle \in R \in \mathbb{U}.$$

Transitivity of $\mathbb{U}$ will ensure that it is closed downwards under such membership chains, giving $a \in \mathbb{U}$. But this leads only to the conclusion that $dom\, R \subseteq \mathbb{U}$, whereas we want $dom\, R \in \mathbb{U}$. Is $dom\, R$ perhaps too "big" to be an element of $\mathbb{U}$?[4]

Now, if R itself were transitive, we would get $a \in R$, showing $dom\, R \subseteq R \in \mathbb{U}$, from which our desired conclusion would result by subset closure. But of course R need not be transitive. On the other hand, it is reasonable to suppose that R can be extended to a transitive set B that belongs to $\mathbb{U}$ (i.e., $R \subseteq B \in \mathbb{U}$). Then we can reason that $dom\, R \subseteq B \in \mathbb{U}$, leading to $dom\, R \in \mathbb{U}$, as desired, by subset closure. The justification for this is that any set A has a transitive closure

$$Tr(A),$$

whose members are precisely the members of members of $\cdots$ of members of A.

Definition 14.7.2. $Tr(A)$ is the smallest transitive set that includes A, so that any transitive set including A will also include $Tr(A)$.

We are going to require that $\mathbb{U}$ be "big enough" to have room for the transitive closure of any set $A \in \mathbb{U}$. For this to hold it is enough that some transitive set including A belong to $\mathbb{U}$. Thus our requirement is strong transitivity.

Definition 14.7.3. *Strong Transitivity* is the property that for any set $A \in \mathbb{U}$ there exists a transitive set $B \in \mathbb{U}$ with $A \subseteq B \subseteq \mathbb{U}$.

Note that the stipulation that $B \subseteq \mathbb{U}$ is superfluous if $\mathbb{U}$ were assumed transitive, since the inclusion $B \subseteq \mathbb{U}$ would then follow from $B \in \mathbb{U}$. But the definition of strong transitivity itself implies that $\mathbb{U}$ is transitive. Indeed, we get $A \subseteq \mathbb{U}$ when $A \in \mathbb{U}$ because $A \subseteq B \subseteq \mathbb{U}$.[5] So this single statement captures all that is needed. In a strongly transitive $\mathbb{U}$ we can assume that any set we are dealing with is located within a large transitive set. This will be the "key to the universe", as will become apparent.

[4]That this is not an idle question is evident from the discussion in Section 15.1.

[5]In more detail, let $A \in \mathbb{U}$. Consider the transitive closure $B \in \mathbb{U}$ of A, so that $A \subseteq B$. By the last hypothesis of Definition 14.7.3, we have $B \subseteq \mathbb{U}$. Thus $A \subseteq \mathbb{U}$, i.e., every member of A belongs to $\mathbb{U}$. This proves the transitivity of $\mathbb{U}$.

14.8. Universes

Definition 14.8.1. A *universe* is any strongly transitive set $\mathbb{U}$ such that

- if $a, b \in \mathbb{U}$, then $\{a, b\} \in \mathbb{U}$;
- if A and B are sets in $\mathbb{U}$, then $A \cup B \in \mathbb{U}$;
- if A is a set in $\mathbb{U}$, then $\mathcal{P}(A) \in \mathbb{U}$.

We can further specify the individuals of $\mathbb{U}$.

Definition 14.8.2. Such a $\mathbb{U}$ is a *universe over* $\mathbb{X}$ if

(1) $\mathbb{X}$ is a set that belongs to $\mathbb{U}$ ($\mathbb{X} \in \mathbb{U}$), and
(2) the members of $\mathbb{X}$ are regarded as individuals that are not sets and have no members:

$$(\forall x \in \mathbb{X})[x \neq \emptyset \wedge (\forall y \in \mathbb{U})(y \notin x)].$$

It will always be assumed further that a universe contains at least one set, and also contains the positive integers $1, 2, \dots$ to ensure that m-tuple formation can be carried out. In practice we will be using universes that have $\mathbb{R} \in \mathbb{U}$, with each member of $\mathbb{R}$ being an individual, so these conditions will hold.

Here now is a list of the main closure properties of such universes, many of which have been indicated already. Uppercase letters A, B, A_i, etc. are reserved for members of $\mathbb{U}$ that are sets.

14.8.1. Set theory.

- If $a \in \mathbb{U}$, then $\{a\} \in \mathbb{U}$.
- $A_1, \dots, A_m \in \mathbb{U}$ implies $A_1 \cup \cdots \cup A_m \in \mathbb{U}$.
- $\mathbb{U}$ contains all its finite subsets: if $A \subseteq \mathbb{U}$ and A is finite, then $A \in \mathbb{U}$.
- $A \subseteq B \in \mathbb{U}$ implies $A \in \mathbb{U}$.
- $\emptyset \in \mathbb{U}$.
- If $\{A_i : i \in I\} \subseteq A \in \mathbb{U}$, then $\cup_{i \in I} A_i \in \mathbb{U}$. (Note: this uses strong transitivity.)
- $\mathbb{U}$ is closed under unions of sets of sets: if $B = \{A_i : i \in I\} \in \mathbb{U}$ and each A_i is a set, then $\cup B = \cup_{i \in I} A_i \in \mathbb{U}$.
- $\mathbb{U}$ is closed under arbitrary intersections: if $\{A_i : i \in I\} \subseteq \mathbb{U}$, then $\cap_{i \in I} A_i \in \mathbb{U}$, whether or not the set $\{A_i : i \in I\}$ itself belongs to $\mathbb{U}$.

14.8.2. Relations.

- If $a, b \in \mathbb{U}$, then also the pair $\langle a, b\rangle \in \mathbb{U}$
- If $A, B \in \mathbb{U}$ and $R \subseteq A \times B$, then $R \in \mathbb{U}$.
- If $a_1 \dots, a_m \in \mathbb{U}$ $(m > 2)$, then $\langle a_1, \dots, a_m\rangle \in \mathbb{U}$.

- $\mathbb{U}$ is closed under finitary relations: if $A_1, \ldots, A_m \in \mathbb{U}$ and $R \subseteq A_1 \times \cdots \times A_m$, then $R \in \mathbb{U}$.
- If $R \in \mathbb{U}$ is a binary relation, then $\mathbb{U}$ contains the domain $dom\, R$, the range $ran\, R$, the R-image $R^i(C)$ of any set $C \subseteq dom\, R$, and the inverse relation R^{-1}, where

$$dom\, R = \{a : \exists b\, (\langle a, b\rangle \in R)\},$$

$$ran\, R = \{b : \exists a\, (\langle a, b\rangle \in R)\},$$

$$R^i(C) = \{b : \exists a \in C\, (\langle a, b\rangle \in R)\},$$

$$R^{-1} = \{\langle b, a\rangle : \langle a, b\rangle \in R\}.$$

- If $R, S \in \mathbb{U}$ are binary relations, then $\mathbb{U}$ contains their composition

$$R \circ S = \{\langle a, c\rangle : \exists b\, (\langle a, b\rangle \in R \text{ and } \langle b, c\rangle \in S)\}.$$

14.8.3. Functions.

- If $f : A \to B$ is a function with $A, B \in \mathbb{U}$, then also $f \in \mathbb{U}$. Moreover, for any $C \subseteq A$ and $D \subseteq B$, the universe $\mathbb{U}$ contains the image

$$f^i(C) = \{f(a) : a \in C\}$$

and the inverse image

$$f^{-1}(D) = \{a \in A : f(a) \in D\}.$$

- If $A, B \in \mathbb{U}$, then the set B^A of all functions from A to B belongs to $\mathbb{U}$.
- If $\{A_i : i \in I\} \in \mathbb{U}$ and $I \in \mathbb{U}$, then $(\prod_{i \in I} A_i) \in \mathbb{U}$.

The discussion until now has aimed to motivate the introduction of superstructures in Section 14.9.

14.9. Superstructures

It is time to demonstrate that there are such things as universes. Let $\mathbb{X}$ be a set with $\mathbb{R} \subseteq \mathbb{X}$.

Definition 14.9.1. The nth cumulative power set $\mathbb{U}_n(\mathbb{X})$ of $\mathbb{X}$ is defined inductively by

$$\mathbb{U}_0(\mathbb{X}) = \mathbb{X},$$

$$\mathbb{U}_{n+1}(\mathbb{X}) = \mathbb{U}_n(\mathbb{X}) \cup \mathcal{P}(\mathbb{U}_n(\mathbb{X})),$$

so that

$$\mathbb{U}_0(\mathbb{X}) \subseteq \mathbb{U}_1(\mathbb{X}) \subseteq \cdots \subseteq \mathbb{U}_n(\mathbb{X}) \subseteq \cdots$$

Definition 14.9.2. The *superstructure over* $\mathbb{X}$ is the union of all these cumulative power sets:

$$\mathbb{U}(\mathbb{X}) = \bigcup_{n=0}^{\infty} \mathbb{U}_n(\mathbb{X}).$$

Definition 14.9.3. The *rank* of an entity a is the least n such that $a \in \mathbb{U}_n(\mathbb{X})$.

The rank 0 entities (members of $\mathbb{X}$) will be regarded as individuals:

$$(\forall x \in \mathbb{U}_0(\mathbb{X}))\,\big(x \neq \emptyset \wedge (\forall y \in \mathbb{U}(\mathbb{X}))\,(y \notin x)\big).$$

All other members of $\mathbb{U}(\mathbb{X})$ (those with positive rank) are sets, and so $\mathbb{U}(\mathbb{X})$ has just these two types of entity. We can show the following.

(1) $\mathbb{U}_{n+1}(\mathbb{X}) = \mathbb{X} \cup \mathcal{P}(\mathbb{U}_n(\mathbb{X}))$.
(2) $\mathbb{U}_n(\mathbb{X}) \in \mathbb{U}_{n+1}(\mathbb{X})$. Hence $\mathbb{U}_n(\mathbb{X}) \in \mathbb{U}(\mathbb{X})$, and in particular, $\mathbb{X} \in \mathbb{U}(\mathbb{X})$.
(3) $\mathbb{U}_{n+1}(\mathbb{X})$ is transitive. Indeed, $a \in B \in \mathbb{U}_{n+1}(\mathbb{X})$ implies $a \in \mathbb{U}_n(\mathbb{X})$.[6]
(4) If $a, b \in \mathbb{U}_n(\mathbb{X})$, then $\{a, b\} \in \mathbb{U}_{n+1}(\mathbb{X})$.
(5) If $A, B \in \mathbb{U}_n(\mathbb{X})$, then $A \cup B \in \mathbb{U}_{n+1}(\mathbb{X})$.
(6) $A \in \mathbb{U}_n(\mathbb{X})$ implies $\mathcal{P}(A) \in \mathbb{U}_{n+2}(\mathbb{X})$.

From item (3) it follows that $\mathbb{U}(\mathbb{X})$ is strongly transitive, since every element of $\mathbb{U}(\mathbb{X})$ belongs to some $U_{n+1}(\mathbb{X})$. Properties (4)–(6) then ensure that $\mathbb{U}(\mathbb{X})$ is a universe, and by (2) it is a universe over $\mathbb{X}$.

In fact, $\mathbb{U}(\mathbb{X})$ is the smallest universe containing $\mathbb{X}$, in the sense that if any universe $\mathbb{U}$ has $\mathbb{X} \in \mathbb{U}$, then $\mathbb{U}(\mathbb{X}) \subseteq \mathbb{U}$. Another description of this superstructure over $\mathbb{X}$ is the following.

Lemma 14.9.4. $\mathbb{U}(\mathbb{X})$ *is the smallest transitive set that contains* $\mathbb{X}$ *and is closed under binary unions* $A \cup B$ *and power sets* $\mathcal{P}(A)$.

[6] For example, if $a \in B \in \mathbb{U}_1(\mathbb{X})$, then by definition of $\mathbb{U}_1(\mathbb{X})$, either $B \in \mathbb{X}$ or $B \in \mathcal{P}(\mathbb{X})$, i.e, $B \subseteq \mathbb{X}$. In the former case B is an individual and hence contains no elements a. In the latter case, $a \in B \subseteq \mathbb{X}$ so that $a \in \mathbb{X} = \mathbb{U}_0(\mathbb{X})$.

CHAPTER 15

Superstructure, language, NS framework, measure

15.1. Boundedness

Let $\mathbb{X}$ be a set of individuals (atoms). In Section 14.9, we defined a particular type of universe called superstructure $\mathbb{U}(\mathbb{X})$ over $\mathbb{X}$ inductively as follows:

$$\begin{aligned} \mathbb{U}_0(\mathbb{X}) &= \mathbb{X} \\ \mathbb{U}_{n+1}(\mathbb{X}) &= \mathbb{U}_n(\mathbb{X}) \cup \mathcal{P}(\mathbb{U}_n(\mathbb{X})), \\ \mathbb{U}(\mathbb{X}) &= \bigcup\nolimits_{n=0}^{\infty} \mathbb{U}_n(\mathbb{X}). \end{aligned}$$

A universe is not closed under arbitrary subsets: if $A \subseteq \mathbb{U}$, it need not follow that $A \in \mathbb{U}$ (e.g., consider $A = \mathbb{U}$). The relevant requirement is that of boundedness.

LEMMA 15.1.1. *In a superstructure, A belongs to $\mathbb{U}(\mathbb{X})$ iff there is an upper bound $n \in \mathbb{N}$ on the ranks of the members of A, i.e., iff $A \subseteq \mathbb{U}_n(\mathbb{X})$ for some n.*

LEMMA 15.1.2. *If $A, B \in \mathbb{U}_n(\mathbb{X})$, then any subset of $A \times B$, and in particular any function from A to B, is in $\mathbb{U}_{n+2}(\mathbb{X})$.*

All the entities typically involved in studying the analysis of $\mathbb{X}$ can be obtained in $\mathbb{U}(\mathbb{X})$ using only rather low ranks. By Lemma 15.1.2, constructing a function between given sets increases the rank by at most 2. Using this, we see the following.

(1) A *topology* on $\mathbb{X}$ is a subset of $\mathcal{P}(\mathbb{X})$, hence a subset of $\mathbb{U}_1(\mathbb{X})$, so belongs to $\mathbb{U}_2(\mathbb{X})$. Thus the *set of all topologies* on $\mathbb{X}$ is itself a member of $\mathbb{U}_3(\mathbb{X})$.
(2) An $\mathbb{R}$-valued *measure* on $\mathbb{X}$ is a function $\mu : \mathcal{A} \to \mathbb{R}$ with $\mathcal{A}$ a collection of subsets of $\mathbb{X}$. Thus $\mathcal{A}$ is of rank 2 and μ of rank 4. Thus the *set of all measures* on $\mathbb{X}$ is also an element of $\mathbb{U}_5(\mathbb{X})$ (i.e., of rank 5).
(3) A *metric* on $\mathbb{X}$ is a function $d : \mathbb{X}\times\mathbb{X} \to \mathbb{R}$ of rank 5 (since $\mathbb{X}\times\mathbb{X}$ has rank 3). The set of all metrics on $\mathbb{X}$ has rank 6.

(4) The *Riemann integral* on a closed interval $[a, b]$ can be viewed as a function
$$\int_a^b : \mathcal{R}[a, b] \to \mathbb{R}.$$
where $R[a, b]$ is the set of integrable functions $f : [a, b] \to \mathbb{R}$. Such an f is of rank 3, since $[a, b]$ and $\mathbb{R}$ have rank 1, so $\mathcal{R}[a, b]$ has rank 4 and therefore the integral is an entity of rank 6.

15.2. The language of a universe

Given a denumerable list of variables, a language
$$\mathcal{L}_{\mathbb{U}}$$
associated with the universe $\mathbb{U}$ is generated much as the language $\mathcal{L}_{\mathcal{R}}$ of Chapter 4, by defining

- terms,
- atomic formulas,
- formulas, and
- sentences.

The first significant difference is that one is starting with a larger collection of constant terms (see Section 4.3), namely all the entities belonging to $\mathbb{U}$. We will point out additional significant differences as we go along; see e.g., Section 15.4.

15.3. Nonstandard framework; starring a formula

Let $* : \mathbb{U} \to \mathbb{U}'$ be a mapping between two universes taking each $a \in \mathbb{U}$ to an element ${}^*a \in \mathbb{U}'$.

(1) (Terms) Each $\mathcal{L}_{\mathbb{U}}$-term τ has an associated *-transform ${}^*\tau$, which is the $\mathcal{L}_{\mathbb{U}'}$-term obtained by replacing each constant symbol a by ${}^*a$.

(2) (Formulas) A constant a occurring in an $\mathcal{L}_{\mathbb{U}}$-formula ϕ will occur as part of a term τ that appears either in an atomic formula or within one of the quantifier forms $(\forall x \in \tau)$ and $(\exists x \in \tau)$. Applying the replacement $a \mapsto {}^*a$ to all such constants transforms ϕ into an $\mathcal{L}_{\mathbb{U}'}$-formula ${}^*\phi$. If ϕ is a sentence, then so too is ${}^*\phi$.

Definition 15.3.1. A *nonstandard framework* for a set $\mathbb{X}$ is a pair
$$\mathbb{U}, *$$
where $\mathbb{U}$ is a universe over $\mathbb{X}$ and $* : \mathbb{U} \to \mathbb{U}'$ is a map with the following three properties:

(1) ${}^*a = a$ for all $a \in \mathbb{X}$.
(2) ${}^*\emptyset = \emptyset$.
(3) Transfer: an $\mathcal{L}_{\mathbb{U}}$-sentence ϕ is true if and only if the $\mathcal{L}_{\mathbb{U}'}$-sentence ${}^*\phi$ is true.

Such a map will be called a *universe embedding* or *transfer map*. It preserves many set-theoretic operations:

- $a = b$ iff ${}^*a = {}^*b$. Hence $a \mapsto {}^*a$ is injective.
- $a \in B$ iff ${}^*a \in {}^*B$.
- $A \subseteq B$ iff ${}^*A \subseteq {}^*B$.
- If $A \subseteq \mathbb{X}$, then $A \subseteq {}^*A \subseteq {}^*\mathbb{X}$. In particular, $\mathbb{X} \subseteq {}^*\mathbb{X}$.
- ${}^*(A \cap B) = {}^*A \cap {}^*B$.
- ${}^*(A \cup B) = {}^*A \cup {}^*B$.
- ${}^*(A - B) = {}^*A - {}^*B$.
- ${}^*\{a_1, \dots, a_m\} = \{{}^*a_1, \dots, {}^*a_m\}$. Thus ${}^*A = \{{}^*a : a \in A\}$ if A is finite.

LEMMA 15.3.2. *We have* ${}^*\mathcal{P}(A) \subseteq \mathcal{P}({}^*A)$.

PROOF. We apply the transfer principle to

$$(\forall x \in \mathcal{P}(A))(\forall y \in x)\,(y \in A).$$

This shows that if $x \in {}^*\mathcal{P}(A)$, then $y \in x$ implies $y \in {}^*A$, and so $x \subseteq {}^*A$, whence $x \in \mathcal{P}({}^*A)$. The exact relationship between ${}^*\mathcal{P}(A)$ and $\mathcal{P}({}^*A)$ is mentioned in Section 15.5. □

15.4. Transforming functions

The hyperreal extension of a real-valued function f was denoted by f, as well. In the more general situation of a universe, a transformed function ${}^*f$ need not agree with f where their domains overlap. Therefore more caution is needed with notation for functions than for the language $\mathcal{L}_R$. In general, if $a \in dom f$, then ${}^*f({}^*a) = {}^*(f(a))$, but even when ${}^*a = a$ this will reduce to ${}^*f(a) = f(a)$ only when ${}^*(f(a))$ is equal to $f(a)$. This need not hold in general.

EXAMPLE 15.4.1. Let $f : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ be defined by

$$f(r) = \{x \in \mathbb{R} : x > r\}.$$

For a given $r \in \mathbb{R}$, transfer of the sentence

$$(\forall x \in \mathbb{R})(x \in f(r) \leftrightarrow x > r)$$

shows (since ${}^*r = r$) that

$${}^*f(r) = {}^*(f(r)) = \{x \in {}^*\mathbb{R} : x > r\}.$$

In particular, $f(0) = \mathbb{R}^+$, while ${}^*f(0) = {}^*\mathbb{R}^+$.

15.5. Punchlines

The punchlines[1] are as follows.

(1) The nonstandard universe ${}^*\mathbb{U}(\mathbb{X})$ is generated by an ultrapower of $\mathbb{U}(\mathbb{X})$.[2]
(2) An internal set is an element of the star of a standard set. Thus, a subset $I \subseteq {}^*\mathbb{R}$ is internal iff $I \in {}^*\mathcal{P}(\mathbb{R})$.
(3) Transfer is valid for formulas involving quantification over sets, so long as those sets are internal.

EXAMPLE 15.5.1. Internal order-completeness of the hyperreals was already discussed in Theorem 8.7.2. If $c \in \mathbb{R}$ is an upper bound for a set $A \subseteq \mathbb{R}$, we will abbreviate the formula $(\forall x \in A)(x < c)$ as

$$A \leq c.$$

The completeness property of $\mathbb{R}$ asserts that if A is bounded from above, then there is a least upper bound $d \in \mathbb{R}$ for A, or in formulas

$$(\forall A \subseteq \mathbb{R})\big[(\exists c \in \mathbb{R})(A \leq c) \to (\exists d \in \mathbb{R})(A \leq d) \wedge (\forall e \in \mathbb{R})(A \leq e \to d \leq e)\big]. \tag{15.5.1}$$

To reformulate the completeness property (15.5.1) in a way amenable to an application of the transfer principle, we write

$$(\forall A \in \mathcal{P}(\mathbb{R}))\big[(\exists c \in \mathbb{R})(A \leq c) \to (\exists d \in \mathbb{R})(A \leq d) \wedge (\forall e \in \mathbb{R})(A \leq e \to d \leq e)\big]. \tag{15.5.2}$$

To make this more readable, we introduce the collection of bounded sets $\mathcal{P}_{bd}(\mathbb{R})$. Then the completeness property becomes

$$(\forall A \in \mathcal{P}_{bd}(\mathbb{R}))\big[(\exists d \in \mathbb{R})\big((A \leq d) \wedge (\forall e \in \mathbb{R})(A \leq e \to d \leq e)\big)\big]. \tag{15.5.3}$$

Applying the transfer principle to (15.5.3), we obtain

$$(\forall A \in {}^*\mathcal{P}_{bd}(\mathbb{R}))\big[(\exists d \in {}^*\mathbb{R})\big((A \leq d) \wedge (\forall e \in {}^*\mathbb{R})(A \leq e \to d \leq e)\big)\big]. \tag{15.5.4}$$

Here formula (15.5.4) expresses the internal completeness of ${}^*\mathbb{R}$.

15.6. Rings and algebras

Developing the Loeb measure (starting in Section 16.1) will require some algebraic preliminaries.

---

[1] shurat machatz

[2] By using an ultrafilter over a large index set, one can obtain higher saturation properties for the nonstandard extension. We will not pursue this since countable saturation is sufficient for Loeb measures; see Chapter 16.

DEFINITION 15.6.1. Let K be a set. A *ring of sets* is a nonempty collection $\mathcal{A}$ of subsets of K that is closed under set differences and unions:

If $A, B \in \mathcal{A}$ then $A - B,\ A \cup B \in \mathcal{A}$.

It follows that $\emptyset \in \mathcal{A}$, since $A - A = \emptyset$, and that $\mathcal{A}$ is closed under symmetric differences $A \Delta B$ and intersections $A \cap B$, since

$$A \Delta B = (A - B) \cup (B - A),$$

and

$$A \cap B = A - (A - B).$$

DEFINITION 15.6.2 (Algebra of sets). An *algebra* is a ring $\mathcal{A}$ that has $K \in A$ and hence (indeed equivalently) is closed under complements $A^c = K - A$.

If $\mathcal{A}$ is a ring, then $\mathcal{A} \cup \{K - A : A \in \mathcal{A}\}$ is an algebra, the smallest one including $\mathcal{A}$.

DEFINITION 15.6.3. A σ-ring is a ring that is closed under countable unions:

If $A_n \in \mathcal{A}$ for all $n \in \mathbb{N}$, then $\cup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.

The equation

$$\cap_{n \in \mathbb{N}} A_n = A_1 - \cup_{n \in \mathbb{N}} (A_1 - A_n))$$

shows that a σ-ring is also closed under countable intersections.

DEFINITION 15.6.4. A σ-algebra is a σ-ring that is an algebra.

The intersection of any family of σ-algebras is a σ-algebra.

COROLLARY 15.6.5. *For any $\mathcal{A} \subseteq \mathcal{P}(K)$, there is a smallest σ-algebra*

$$S(\mathcal{A}) \subseteq \mathcal{P}(K)$$

that includes $\mathcal{A}$, called the σ-algebra generated by $\mathcal{A}$.

15.7. Examples of rings and algebras, $C_{\mathbb{R}}$

EXAMPLE 15.7.1. The powerset $\mathcal{P}(K)$ itself is a σ-algebra.

EXAMPLE 15.7.2. If K is infinite, then

- the collection of all finite subsets of K is a ring that is not an algebra;
- the collection of all finite or cofinite subsets of K is an algebra that is not a σ-algebra;

- the collection of all countable subsets of K is a σ-ring that is *not* an algebra when K is uncountable.

EXAMPLE 15.7.3. Let $C_{\mathbb{R}}$ be the collection of all subsets of $\mathbb{R}$ that are finite unions of left-open intervals

$$(a, b] = \{x \in \mathbb{R} : a < x \leq b\}$$

with $a, b \in \mathbb{R}$ and $a \leq b$. (Thus $\emptyset = (a, a] \in C_{\mathbb{R}}$.) Then $C_{\mathbb{R}}$ is a ring in which each member is in fact a disjoint union of left-open intervals $(a, b]$. Note that $C_{\mathbb{R}}$ is not an algebra, and is not closed under countable unions: $(0, 1)$ is not in $C_{\mathbb{R}}$, since each member of $C_{\mathbb{R}}$ will have a greatest element, but $(0, 1)$ is the union of the intervals $(0, 1 - \frac{1}{n}]$ for $n \in \mathbb{N}$.

The ring $C_{\mathbb{R}}$ does, however, contain certain significant countable unions. For instance $(0, 1]$ is the union of the pairwise disjoint intervals $(\frac{1}{n+1}, \frac{1}{n}]$. Any reasonable notion of measure should thus assign to $(0, 1]$ the infinite sum of the measures of the intervals $(\frac{1}{n+1}, \frac{1}{n}]$.

15.8. Borel algebra

DEFINITION 15.8.1 (Borel sets). Let $B_{\mathbb{R}}$ be the σ-algebra generated by $C_{\mathbb{R}}$.

Thus $B_{\mathbb{R}} = S(C_{\mathbb{R}})$. Each open interval (a, b) in $\mathbb{R}$ is in $B_{\mathbb{R}}$, being the union of the countably many left-open intervals $(a, b - \frac{1}{n}]$ for $n \in \mathbb{N}$. Hence every open subset of $\mathbb{R}$ is in $B_{\mathbb{R}}$, being the union of countably many open intervals (take ones with rational end points).

On the other hand, if a σ-algebra contains all open intervals, it must contain any left-open $(a, b]$ as the intersection of all $(a, b + \frac{1}{n})$ for $n \in \mathbb{N}$.

COROLLARY 15.8.2. *$B_{\mathbb{R}}$ is also the σ-algebra generated by the open intervals, as well as the σ-algebra generated by the open sets of $\mathbb{R}$.*

The members of $B_{\mathbb{R}}$ are called the Borel sets.

15.9. Algebras of hyperfinite sets

Let $K = \{1, \ldots, N\} \subseteq {}^*\mathbb{N}$ with N an infinite hypernatural. Then K is hyperfinite. Consider the collection $\mathcal{P}_I(K)$ of all internal subsets of K. Note that $\mathcal{P}_I(K) \subseteq {}^*\mathcal{P}(\mathbb{N})$. Then $\mathcal{P}_I(K)$ is an algebra (also hyperfinite) that by transfer of the finite case will be closed under hyperfinite unions, i.e., unions of internal sequences $\langle A_n : n \leq H \rangle$ for $H \in {}^*\mathbb{N}$.

The family $P_I(K)$ is not, however, a σ-algebra. Indeed, it contains each initial segment $\{1, \ldots, n\}$ with $n \in \mathbb{N}$, but does not contain their union because that is the external set $\mathbb{N}$.

This same analysis applies to the algebra of internal subsets of any nonstandard hyperfinite set $K = \{s_n : n \leq N\}$ indexed by the set K above.

Example 15.9.1. Let $\mathcal{A}$ be an algebra in some universe $\mathbb{U}$. In any enlargement of $\mathbb{U}$, the collection $^*\!\mathcal{A}$ will be an algebra, by transfer, but in a countably saturated enlargement, the algebra $^*\!\mathcal{A}$ will not in general be a σ-algebra, even if $\mathcal{A}$ is. To see this, let $\langle A_n : n \in \mathbb{N}\rangle$ be a sequence of members of $^*\!\mathcal{A}$ with union a set A. Each A_n is internal, and if A were in $^*\!\mathcal{A}$, it would also be internal and hence by countable saturation would already be equal to $\cup_{n \leq k} A_n$ for some $k \in \mathbb{N}$ (cf. Corollary 13.2.1).

Thus if A is a genuinely infinite union of the sets A_n, it cannot be in $^*\!\mathcal{A}$. This will happen, for example, if the sets A_n are strictly increasing ($An \subsetneq A_{n+1}$) or pairwise disjoint. For instance, in the case of the Borel algebra, the internal sets $^*(-n, n)$ belong to $^*B_{\mathbb{R}}$ for all $n \in \mathbb{N}$, but their union is not in $^*B_{\mathbb{R}}$ because it is the external set of all finite hyperreals.

The closure condition that we do get for $^*\!\mathcal{A}$ is that the sequence $\langle A_n : n \in \mathbb{N}\rangle$ extends to an internal sequence $\langle A_n : n \in {}^*\mathbb{N}\rangle$ whose union can be shown by transfer to be in $^*\!\mathcal{A}$. In this sense $^*\!\mathcal{A}$ is a "hyper-σ-algebra", but that is not the type of structure on which a standard measure is defined.

This reasoning in fact shows that for any internal algebra of sets (not just one of the form $^*\!\mathcal{A}$),

> The union of a countable sequence of sets can belong to the algebra only if it is equal to the union of finitely many of its terms.

It is this feature upon which Loeb measure is founded.[3]

15.10. Measures

Classical measure theory employs the extended real numbers

$$[-\infty, +\infty] = \{-\infty\} \cup \mathbb{R} \cup \{+\infty\},$$

with $-\infty < r < +\infty$ for $r \in \mathbb{R}$, with rules such as $r \pm \infty = \pm\infty$, etc. We will usually put ∞ for $+\infty$, and also make use of the set $[0, \infty] = \{r \in \mathbb{R} : r \geq 0\} \cup \{\infty\}$.

[3]It shows that a certain Caratheodory condition is vacuously satisfied.

DEFINITION 15.10.1 (Measure). Let $\mathcal{A}$ be a ring of subsets of a set K, and μ a function from $\mathcal{A}$ to $[0,\infty]$ that has $\mu(\emptyset)=0$. Then μ is called a *measure* if it satisfies the following condition:

(M1) If $\langle A_n : n \in \mathbb{N}\rangle$ is a sequence of pairwise disjoint elements of $\mathcal{A}$ whose union is in the ring $\mathcal{A}$, then

$$\mu\left(\cup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

This condition is called countable additivity.

REMARK 15.10.2. The condition is not required to hold for all (pairwise disjoint) sequences $\langle A_n : n \in \mathbb{N}\rangle$, but only those whose union happens to belong to $\mathcal{A}$ (which is not guaranteed when $\mathcal{A}$ is not a σ-algebra).

DEFINITION 15.10.3. The function μ is called *finitely additive* if in place of M1 it satisfies the following condition:

(M2) $\mu(A\cup B) = \mu(A)+\mu(B)$ whenever $A, B \in A$ with $A\cap B=\emptyset$.

Since a ring is closed under finite unions, condition M2 implies that

$$\mu\left(\cup_{i=1}^{n} A_n\right) = \sum_{i=1}^{n} \mu(A_i)$$

whenever $A_1,\dots,A_n$ is a finite sequence of pairwise disjoint members of $\mathcal{A}$. Condition M2 also implies that μ is monotonic:

- $A\subseteq B$ implies $\mu(A)\le\mu(B)$, for all $A,B\in\mathcal{A}$;
  as well as being subtractive:
- $A\subseteq B$ and $\mu(B)<\infty$ implies $\mu(B-A)=\mu(B)-\mu(A)$, for all $A,B\in\mathcal{A}$.

Countable additivity implies the following important fact.

PROPOSITION 15.10.4. *Suppose μ satisfies* M1. *If $\langle A_n : n\in\mathbb{N}\rangle$ is an increasing sequence of elements of $\mathcal{A}$ whose union is in $\mathcal{A}$, then*

$$\mu(\cup_{n\in\mathbb{N}} A_n) = \lim_{n\to\infty} \mu(A).$$

DEFINITION 15.10.5. An element $A\in\mathcal{A}$ is called μ-finite if $\mu(A)<\infty$, and μ-null if $\mu(A)=0$.

DEFINITION 15.10.6. The function μ is called σ-finite if the set K is the union of countably many μ-finite subsets.

We give two examples.

EXAMPLE 15.10.7 (Counting measure). If $A \subseteq K$, put

$$\mu_c(A) = \begin{cases} |A| \text{ if } A \text{ is finite,} \\ \infty \text{ if } A \text{ is infinite.} \end{cases}$$

Then μ_c is a measure on $\mathcal{P}(K)$, the counting measure, which is σ-finite iff K is countable. The restriction of μ_c to the ring of finite subsets of K, or to the algebra of finite or cofinite sets, is also a measure.

EXAMPLE 15.10.8. On the ring $C_{\mathbb{R}}$ of disjoint unions of left-open intervals $(a, b]$, put

$$\lambda((a, b]) = b - a$$

and extend λ additively to all members of $C_{\mathbb{R}}$.

LEMMA 15.10.9. *λ is a measure on $C_{\mathbb{R}}$, and λ is σ-finite.*

PROOF. Indeed, $\mathbb{R}$ is the union of intervals $(-n, n]$. $\square$

This will be used in Section 16.3 to define the Lebesgue measure.

Here the symbol λ may be thought of as denoting "length", but it also stands for "Lebesgue".

15.11. Counting measure on a hyperfinite set

Consider a countably saturated enlargement of a universe over a set $\mathbb{X}$ that has $[-\infty, +\infty] \subseteq \mathbb{X}$. Then the set

$${}^*[0, \infty] = \{x \in {}^*\mathbb{R} : x \geq 0\} \cup \{\infty\}$$

is internal. Now let K be a hyperfinite set and $\mathcal{P}_I(K)$ the algebra of all internal subsets of K. For each $A \in \mathcal{P}_I(K)$ put

$$\mu(A) = \frac{|A|}{|K|},$$

where $|A|$ is the internal cardinality of the hyperfinite set A.

Then μ is finitely additive (with values in ${}^*\mathbb{R}$), because $|A \cup B| = |A| + |B|$ when $A \cap B = \emptyset$ (but note that we are referring to $+$ in ${}^*\mathbb{R}$ rather than $\mathbb{R}$). Since $|A| \leq |K|$ whenever $A \subseteq K$, μ takes finite values between 0 and 1, i.e., we have

$$\mu : \mathcal{P}_I(K) \to {}^*[0, 1].$$

DEFINITION 15.11.1. Setting

$${}^o\mu_L(A) = sh(\mu(A))$$

defines ${}^{o}\mu_L : P_I(K) \to [0,1]$ as a real-valued finitely additive measure on $P_I(K)$, with ${}^{o}\mu_L(K) = 1$.[4]

REMARK 15.11.2. Here μ is an internal entity but ${}^{o}\mu_L$ is not.

PROPOSITION 15.11.3. *${}^{o}\mu_L$ is a (σ-additive) measure.*

PROOF. The reason, based on internality, was already explained in Example 15.9.1. If $\langle A_n : n \in \mathbb{N}\rangle$ is a sequence of pairwise disjoint elements of $\mathcal{P}_I(K)$ whose union A belongs to $\mathcal{P}_I(K)$, then A must already be equal to $\cup_{n\le k} A_n$ for some k. But then whenever $m > k$, we necessarily have $A_m = \emptyset$, since $\cup_{n\le k} A_n$ and the A_m are disjoint, so that ${}^{o}\mu_L(A_m) = 0$. Hence a countable union reduces to a finite union: $\bigcup_{n\in\mathbb{N}} A_n = A_1 \cup \cdots \cup A_k$. Therefore we obtain $\sum_{n\in\mathbb{N}} {}^{o}\mu_L(A_n) = {}^{o}\mu_L(A_1) + \cdots + {}^{o}\mu_L(A_k)$, from which it follows that μ satisfies M1. $\square$

REMARK 15.11.4. The σ-additivity of ${}^{o}\mu_L$ results from countable saturation applied to families of internal sets via the reduction of being σ-additive to being finitely-additive. This enables the application of Caratheodory's construction of outer measure, as in Section 16.1. Indeed, Caratheodory's starting point is a σ-additive measure on a ring.

15.12. Additional examples

EXAMPLE 15.12.1 (Generalisation of counting measure). Let $\mathcal{A}$ be an internal ring of subsets of some internal set K in a countably saturated enlargement. Let $\mu : \mathcal{A} \to {}^*[0,\infty]$ be a finitely additive internal function. Adapting the construction of Section 15.11, put

$$ {}^{o}\mu_L(A) = \begin{cases} sh(\mu(A)) \text{ if } \mu(A) \text{ is finite,} \\ \infty \text{ if } \mu(A) \text{ is infinite or } \infty. \end{cases} $$

Reasoning as in Section 15.11, we show that ${}^{o}\mu_L : \mathcal{A} \to [0,\infty]$ is σ-additive, and so is a measure on the ring $\mathcal{A}$.

This last construction has the example of Section 15.11 as a special case, and also covers other natural extensions that involve hyperfinite summation, such as the following.

EXAMPLE 15.12.2. Let $w : K \to {}^*\mathbb{R}$ be an internal "weighting" function on a hyperfinite set K. For each $A \in \mathcal{P}_I(K)$ put

$$ \mu^w(A) = \sum_{s\in A} w(s) \tag{15.12.1} $$

[4]Goldblatt uses the notation μ_L which may not be sufficiently suggestive of taking standard part. Robinson used the notation ${}^{o}r$ for the standard part of r.

(recall the definition of hyperfinite sums in Section 6.5). Then μ^w is a "weighted counting function" that is finitely additive and induces the measure ${}^o\mu^w_L$ on $\mathcal{P}_I(K)$. In fact, every internal finitely additive function $\mu : \mathcal{P}_I(K) \to {}^*[0,\infty]$ arises in this way: put $w(s) = \mu(\{s\})$. The example of Section 15.11 itself is the special case of a uniform weighting in which each point is assigned the same weight, namely $w(s) = \frac{1}{|K|}$.

Example 15.12.3. Weighted hyperfinite summation will be used to relate the Loeb measure to the Lebesgue measure in Section 16.5. It turns out that for every Lebesgue-measurable set $B \subseteq \mathbb{R}$ of finite Lebesgue measure, the latter can be obtained as the infinimum in $\mathbb{R}$ of $\mu^w(A)$ over all internal supersets

$$A \supseteq sh^{-1}(B) \cap K$$

in a suitable hyperfinite set K.[5] For details, see Section 16.6.

[5]This is due to μ-approximability; see [**6**, Section 16.6].

CHAPTER 16

Caratheodory, Lebesgue, Loeb, and beyond

16.1. Caratheodory's outer measure

The classical procedure of Caratheodory extends a measure μ on a ring of sets $\mathcal{A}$ to a measure on a σ-algebra including $\mathcal{A}$. Classically, Caratheodory's procedure is used to construct the Lebesgue measure and the family of Lebesgue-measurable sets. We will summarize the classical procedure, and then show how to use Caratheodory's procedure to construct the Loeb measure in Section 16.4.

If B is an arbitrary subset of the set K on which $\mathcal{A}$ is based, put

$$\mu^+(B) = \inf\left\{\sum_{n\in\mathbb{N}} \mu(A_n) : A_n \in \mathcal{A},\ B \subseteq \cup_{n\in\mathbb{N}} A_n\right\}. \tag{16.1.1}$$

Here the infimum is taken over all sequences $\langle A_n : n \in \mathbb{N}\rangle$ of members of $\mathcal{A}$ that cover B.

Definition 16.1.1. The function $\mu^+ : \mathcal{P}(K) \to [0,\infty]$ is called the *outer measure* defined by μ (although it may not actually be a measure).

The outer measure has the following properties:

- μ^+ agrees with μ on $\mathcal{A}$: if $B \in \mathcal{A}$, then $\mu^+(B) = \mu(B)$.
- $\mu^+(\emptyset) = 0$.
- Monotonicity: if $A \subseteq B$, then $\mu^+(A) \leq \mu^+(B)$.
- Countable subadditivity: for any sequence $\langle A_n\rangle$ of subsets of K, one has $\mu^+(\cup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} \mu^+(A_n)$.
- For any $B \subseteq K$ and any $\varepsilon \in \mathbb{R}^+$ there is an increasing sequence $A_1 \subseteq A_2 \subseteq \cdots$ of $\mathcal{A}$-elements that covers B and has $\mu^+(\cup_{n\in\mathbb{N}} A_n) \leq \mu^+(B) + \varepsilon$.

16.2. μ^+-measurable sets via additive splittings

We summarize the properties of μ^+ and the associative algebra of measurable sets. We start with the following data.

(1) $\mathcal{A}$ is a ring of subsets of a set K;
(2) μ is a measure on $\mathcal{A}$;

(3) μ^+ is the associated outer measure (16.1.1).

DEFINITION 16.2.1. A set $B \subseteq K$ is called μ^+-measurable[1] if it splits[2] every set $E \subseteq K$ in a μ^+-additive way, in the sense that

$$\mu^+(E) = \mu^+(E \cap B) + \mu^+(E - B).$$

For this to hold it is enough that

$$\mu^+(E) \geq \mu^+(E \cap B) + \mu^+(E - B).$$

whenever $\mu^+(E) < \infty$.

DEFINITION 16.2.2. We denote by $\mathcal{A}(\mu)$ the set of all μ^+-measurable sets.

It has the following properties.

- $\mathcal{A}(\mu)$ is a σ-algebra.
- $\mathcal{A} \subseteq \mathcal{A}(\mu)$, i.e., all members of $\mathcal{A}$ are μ^+-measurable.
- It follows that the algebra $\mathcal{A}(\mu)$ includes the σ-algebra $S(\mathcal{A})$ generated by $\mathcal{A}$.
- All μ^+-null sets belong to $\mathcal{A}(\mu)$.
- μ^+ is a measure on $\mathcal{A}(\mu)$, and hence is a measure on $S(\mathcal{A})$.
- If μ is σ-finite on $\mathcal{A}$, and $\mathcal{A}$ is an algebra, then μ^+ is the *only* extension of μ to a measure on $S(\mathcal{A})$ or on $\mathcal{A}(\mu)$.

LEMMA 16.2.3. *μ^+ is a complete measure on $\mathcal{A}(\mu)$, meaning that if $A \subseteq B \in \mathcal{A}(\mu)$ and $\mu^+(B) = 0$, then $A \in \mathcal{A}(\mu)$.*

PROOF. This follows from the fact that $\mathcal{A}(\mu)$ contains all μ^+-null sets. □

This entails the following corollary.

COROLLARY 16.2.4. *Let $A, B \in \mathcal{A}(\mu)$. Suppose we have $A \subseteq B$ and $\mu^+(A) = \mu^+(B)$. Then*

(1) *any subset of $B - A$ belongs to $\mathcal{A}(\mu)$ (and is μ^+-null); hence*
(2) *any set C with $A \subseteq C \subseteq B$ belongs to $\mathcal{A}(\mu)$ and moreover $\mu^+(C) = \mu^+(A) = \mu^+(B)$.*

16.3. Lebesgue measure

To construct the Lebesgue measure, we start with the measure λ on the ring $C_{\mathbb{R}}$ of Example 15.10.8. This measure satisfies $\lambda((a, b]) = b - a$.

DEFINITION 16.3.1. The *Lebesgue measure* is the outer measure λ^+ constructed from λ.

[1] medida
[2] pitzul

Letting $\mathcal{A} = C_{\mathbb{R}}$, we obtain the corresponding σ-algebra $\mathcal{A}(\lambda) = C_{\mathbb{R}}(\lambda)$ as in Definition 16.2.2.

DEFINITION 16.3.2. The members of the σ-algebra $C_{\mathbb{R}}(\lambda)$ are the *Lebesgue measurable sets* and include all members of the σ-algebra $B_{\mathbb{R}}$ of Borel sets generated by $C_{\mathbb{R}}$. We will write $\lambda(B)$ for $\lambda^+(B)$ whenever B is Lebesgue measurable.

The following material is classical.

THEOREM 16.3.3. *The Lebesgue measure has the following properties.*

(1) *(Uniqueness)* λ *is the only measure on* $B_{\mathbb{R}}$ *that has* $\lambda\big((a, b)\big) = b - a$.
(2) *Thus, any measure on an algebra including* $B_{\mathbb{R}}$ *that agrees with* λ *on open intervals must agree with* λ *on all Borel sets.*
(3) *(Approximation by Borel sets) For any Lebesgue measurable set* B *there exist Borel sets* C, D *with* $C \subseteq B \subseteq D$ *such that* $\lambda(D - C) = 0$, *hence* $\lambda(B) = \lambda(C) = \lambda(D)$.
(4) *(Approximation by open and closed sets) A set* $B \subseteq \mathbb{R}$ *is Lebesgue measurable iff for each* $\varepsilon \in \mathbb{R}^+$ *there is a closed set* $C_\varepsilon \subseteq B$ *and an open set* $D_\varepsilon \supseteq B$ *such that* $\lambda(D_\varepsilon - C_\varepsilon) < \varepsilon$.

By using the axiom of choice it can be shown that there is a subset of $\mathbb{R}$ that is not Lebesgue measurable.

16.4. Loeb measures

Loeb measures are defined by applying Caratheodory's outer measure construction to measures of the type ${}^o\mu_L$ introduced in Example 15.12.1. From now on we work in a nonstandard framework that is countably saturated. We start with the following data.

(1) $(K, \mathcal{A}, \mu)$ is any "measure space", meaning the following:
(2) $\mu : \mathcal{A} \to {}^*[0, \infty]$ is an internal finitely additive function (of "counting" type) on an internal ring $\mathcal{A}$ of subsets of an internal set K.
(3) ${}^o\mu_L : \mathcal{A} \to [0, \infty]$ is the (external) measure defined via standard part as in Example 15.12.1.
(4) ${}^o\mu_L^+$ is its associated outer measure on $\mathcal{P}(K)$.
(5) the family of ${}^o\mu_L^+$-measurable sets is defined via additive splitting as in Section 16.2.

DEFINITION 16.4.1. Members of the set $\mathcal{A}({}^o\mu_L)$ of ${}^o\mu_L^+$-measurable subsets of K will be called the *Loeb measurable sets* determined by μ.

Note that $\mathcal{A}({}^o\mu_L)$ may contain sets that are not internal (see Example 16.5.8).

REMARK 16.4.2. We write ${}^o\mu_L(B)$ for ${}^o\mu_L^+(B)$ whenever B is Loeb measurable.[3]

DEFINITION 16.4.3. ${}^o\mu_L$ is the *Loeb measure* and $(K, \mathcal{A}({}^o\mu_L), {}^o\mu_L)$ is the *Loeb measure space* determined by μ.

REMARK 16.4.4. This definition of Loeb measure via the outer measure construction is the way that the notion was first arrived at. By analysing its properties, one can show that $\mathcal{A}({}^o\mu_L)$ has a characterisation that would allow it and its measure ${}^o\mu_L$ to be defined in a more direct way.

See note.[4]

[3] As mentioned in note 5, for sets of finite measure one can obtain ${}^o\mu_L^+(B)$ is the $\mathbb{R}$-infinimum of $\mu(A)$ as the internal set A ranges over $\mathcal{P}_I(T)$ for a suitable grid T. Therefore the notation is coherent.

[4]

LEMMA 16.4.5. *If B is Loeb measurable with respect to μ, then* ${}^o\mu_L(B) = \inf\{{}^o\mu_L(A) : B \subseteq A \in \mathcal{A}\}$.

PROOF. By monotonicity, ${}^o\mu_L(B)$ is a lower bound of the values ${}^o\mu_L(A)$ for $B \subseteq A \in \mathcal{A}$. If ${}^o\mu_L(B) = \infty$, then the result follows. If ${}^o\mu_L(B) < \infty$, to show that it is the greatest lower bound it suffices to show that for any $c \in \mathbb{R}^+$ there is some set $A_c \in \mathcal{A}$ with $B \subseteq A_c$ and ${}^o\mu_L(A_c) \leq {}^o\mu_L(B) + c$. Now, for such an c, by properties of the outer measure ${}^o\mu_L^+$, there is an increasing sequence $A_1 \subseteq A_2 \subseteq \cdots$ of $\mathcal{A}$-elements whose union includes B and has ${}^o\mu_L^+(\cup_{n\in\mathbb{N}} A_n) < {}^o\mu_L(B) + c$. The sequence $\langle A_n : n \in \mathbb{N}\rangle$ extends by sequential comprehensiveness[5] to an internal sequence $\langle A_n : n \in {}^*\mathbb{N}\rangle$ of elements of A. Then for each $k \in \mathbb{N}$ we have

$$(\forall n \in {}^*\mathbb{N})\big(n \leq k \text{ implies } A_n \subseteq A_k \text{ and } \mu(A_n) < {}^o\mu_L(B) + c\big), \tag{i}$$

since $\mu(A_n) \approx {}^o\mu_L(A_n) \leq {}^o\mu_L^+(\cup_{n\in\mathbb{N}} A_n)$. But (i) is an internal assertion, since μ and the extended sequence are internal, while k, c, and ${}^o\mu_L(B)$ are fixed internal entities (real numbers). Therefore by overflow (i) must be true with some infinite $K \in {}^*\mathbb{N}$ in place of k. For such a K we have $A_K \in \mathcal{A}$ and $A_n \subseteq A_K$ for all $n \in \mathbb{N}$, so that $B \subseteq \cup_{n\in\mathbb{N}} A_n \subseteq A_K$, while $\mu(A_K) < {}^o\mu_L(B) + c$. Hence as $\mu(A_K) \approx {}^o\mu_L(A_K)$, we obtain ${}^o\mu_L(A_K) \leq {}^o\mu_L(B) + c$, establishing that A_K is the set A_c we are looking for. $\square$

LEMMA 16.4.6. *If B is Loeb measurable and also ${}^o\mu_L$-finite, then* ${}^o\mu_L(B) = \sup\{{}^o\mu_L(A) : A \subseteq B \text{ and } A \in \mathcal{A}\}$.

PROOF. Given any $c \in \mathbb{R}^+$, we will show that there is some set $A_c \in \mathcal{A}$ such that $A_c \subseteq B$ and ${}^o\mu_L(B) - c < \mu(A_c)$. Since ${}^o\mu_L(B) < \infty$, we know from the previous Lemma 16.4.5 that there is some $D \in A$ with $B \subseteq D$ and ${}^o\mu_L(D) < \infty$. The desired result is obtained by using complementation relative to D. Firstly, $D - B$ is Loeb measurable and ${}^o\mu_L$-finite, so by Lemma 16.4.5 there is a set C with $D-$

16.5. Hyperfinite time line

Let $H \in {}^*\mathbb{N}$ be a fixed infinite hypernatural number.

Definition 16.5.1. The *hyperfinite time line* is the set

$$\begin{aligned} T &= \{\tfrac{k}{H} : k \in {}^*\mathbb{Z} \text{ and } |k| \leq H^2\} \\ &= \{\tfrac{-k}{H} : 1 \leq k < H^2\} \cup \{0\} \cup \{\tfrac{k}{H} : 1 \leq k \leq H^2\}. \end{aligned}$$

The hyperfinite time line T is a hyperfinite set, of internal cardinality $2H^2 + 1$, forming a grid of points spread across the hyperreal line between $-H$ and H, with adjacent points being of infinitesimal distance $\frac{1}{H}$ apart.

Lemma 16.5.2. *Each real number r is approximated infinitely closely on either side by the grid points in T.*

Proof. Consider the statement

$$(\forall n \in \mathbb{N}) \big(|r| < n \to (\exists k \in \mathbb{Z}) \big[|k| < n^2 \text{ and } (\tfrac{k}{n} \leq r < \tfrac{k+1}{n})\big]\big)$$

and apply transfer. □

Now let $\mathcal{A} = \mathcal{P}_I(T)$ be the set of all internal subsets of T. Here $\mathcal{A}$ is an algebra, is itself internal and hyperfinite, and all its members are hyperfinite. The function $\mu : \mathcal{A} \to {}^*[0, \infty)$ given by

$$\mu(A) = \frac{|A|}{H} \tag{16.5.1}$$

is internal and finitely additive. This is similar to the example of the counting measure in Section 15.11. More precisely, μ is a weighted counting function in the sense of Example 15.12.2, determined by assigning the infinitesimal weight $\frac{1}{H}$ to each grid point. It induces the measure ${}^o\mu_L$ on $\mathcal{A}$ by setting

$${}^o\mu_L(A) = \begin{cases} sh\left(\frac{|A|}{H}\right) & \text{if } \frac{A}{H} \text{ is finite,} \\ \infty & \text{otherwise.} \end{cases}$$

Definition 16.5.3. Let $(T, \mathcal{A}({}^o\mu_L), {}^o\mu_L)$ be the associated Loeb measure space as defined in Section 16.4.

$B \subseteq C \in A$ and ${}^o\mu_L(C) < {}^o\mu_L(D - B) + c$. We may assume $C \subseteq D$ (since we could replace C by $C \cap D$ here). Let $A_c = D - C \in A$. Then $A_c \subseteq B$, and C is the disjoint union of $D - B$ and $B - A_c$, so ${}^o\mu_L(D-B) + {}^o\mu_L(B - A_c) = {}^o\mu_L(C) < {}^o\mu_L(D-B) + c$, implying that ${}^o\mu_L(B - A_e) < c$. Therefore ${}^o\mu_L(B) = {}^o\mu_L(A_c) + {}^o\mu_L(B - A_c) < {}^o\mu_L(A_c) + c$, so ${}^o\mu_L(B) - c < \mu(A_e)$ as required. □

First we show that the Lebesgue measure of any real interval can be obtained by using ${}^o\mu_L$ to count the weighted number of grid points between the endpoints of the interval.

THEOREM 16.5.4. *For any $a, b \in \mathbb{R}$ with $a < b$,*

$$ {}^o\mu_L(\{t \in T : a < t < b\}) = b - a. $$

PROOF. Let $A = \{t \in T : a < t < b\}$. Then $A = T \cap {}^*(a, b)$, so A is internal and belongs to the algebra $\mathcal{A}$, hence is Loeb measurable. Since A is hyperfinite, it has smallest and greatest elements, say a' and b'. Since a and b can be approximated infinitely closely by members of T, we must have $a \approx a'$ and $b \approx b'$. We have $a' = \frac{K+1}{H}$ and $b' = \frac{L}{H}$ for suitable $K, L \in {}^*\mathbb{Z}$. Thus

$$ A = \left\{\tfrac{K+1}{H}, \tfrac{K+2}{H}, \ldots, \tfrac{L}{H}\right\} = \left\{\tfrac{M}{H} : K < M \leq L\right\}. $$

The set A is hyperfinite of cardinality $L - K$, since the internal function $f(x) = \frac{K+x}{H}$ is a bijection from $\{1, \ldots, L - K\}$ onto A. It follows that

$$ \mu(A) = \frac{|A|}{H} = \frac{L-K}{H} = \frac{L}{H} - \frac{K}{H} \approx b - a, $$

and therefore ${}^o\mu_L(A) = b - a$ as required. □

Note that the proof of Theorem 16.5.4 shows readily that ${}^o\mu_L$ assigns measure $b - a$ as well to the sets

$$ T \cap {}^*(a, b],\ T \cap {}^*[a, b),\ T \cap {}^*[a, b]. $$

COROLLARY 16.5.5. *If B is any finite interval in $\mathbb{R}$, the Lebesgue measure λ of B is equal to the Loeb measure of the set $T \cap {}^*B$ of grid points that are (possibly nonstandard) members of B.*

One might wonder whether the equation

$$ \lambda(B) = {}^o\mu_L(T \cap {}^*B) $$

holds in general, but such a notion is quickly dispelled by considering the following case.

EXAMPLE 16.5.6. Let $B = \mathbb{Q}$. Since every grid point is a hyperrational number, we have $T \subseteq {}^*\mathbb{Q}$. Thus, ${}^o\mu_L(T \cap {}^*\mathbb{Q}) = {}^o\mu_L(T) = \infty$, while $\lambda(\mathbb{Q}) = 0$.

Rather than $T \cap {}^*B$, the appropriate set to represent B in T is the set of grid points that approximate members of B infinitely closely. This is the inverse shadow of B.

DEFINITION 16.5.7. The *inverse shadow* of B (relative to the hyperfinite set T) is the set

$$\begin{aligned} sh^{-1}(B) &= \{t \in T : t \text{ is infinitely close to some } r \in B\} \\ &= \{t \in T : t \text{ is finite and } sh(t) \in B\} \\ &= \{t \in T : (\exists b \in B)(t \approx b)\}. \end{aligned}$$

The definition of $sh^{-1}(B)$ uses a condition that is not internal, so the set itself cannot be guaranteed to be internal, and may not even be Loeb measurable, i.e., may not belong to the algebra $\mathcal{A}({}^o\mu_L)$ (see Definition 16.4.1).

EXAMPLE 16.5.8. One case in which the inverse shadow is not internal but nonetheless is Loeb measurable occurs when $B = \mathbb{R}$. We have

$$sh^{-1}(\mathbb{R}) = \{t \in T : t \text{ is finite}\} = \cup_{n\in\mathbb{N}} (T \cap {}^*(-n, n)).$$

Each set $T \cap {}^*(-n, n)$ is an internal subset of T and so belongs to $\mathcal{A}$. It follows that $sh^{-1}(\mathbb{R})$ belongs to $\mathcal{A}({}^o\mu_L)$ by closure under countable unions. But the set $sh^{-1}(\mathbb{R})$ cannot itself be internal, because it is bounded in ${}^*\mathbb{R}$ but has no least upper (or greatest lower) bound.

16.6. Lebesgue measure via Loeb measure

The precise relation between the Lebesgue and Loeb measures is as follows. We define $sh^{-1}(B) \subseteq T$ as in Section 16.5.

THEOREM 16.6.1. *A subset $B \subseteq \mathbb{R}$ is Lebesgue measurable if and only if $sh^{-1}(B)$ is Loeb measurable. When this holds, the Lebesgue measure of B is equal to the Loeb measure of the set of grid points infinitely close to points of B:*

$$\lambda(B) = {}^o\mu_L(sh^{-1}(B)).$$ [6]

PROOF. Consider the set of sets

$$M = \{B \subseteq \mathbb{R} : sh^{-1}(B) \in \mathcal{A}({}^o\mu_L)\}.$$

For $B \in M$, we define ν by setting

$$\nu(B) = {}^o\mu_L(sh^{-1}(B)).$$

We will show that M is precisely the class $C_{\mathbb{R}}(\lambda)$ of Lebesgue-measurable sets, and that ν is the Lebesgue measure λ.

We will use the fact that $\mathcal{A}$ is a σ-algebra to show that M is a σ-algebra.

[6] For the use of notation ${}^o\mu_L$ rather than ${}^o\mu_L^+$ see note 3.

By properties of inverse images of functions,

$$\begin{cases} sh^{-1}(\emptyset) = \emptyset, \\ sh^{-1}(A - B) = sh^{-1}(A) - sh^{-1}(B), \\ sh^{-1}(\cup_{n\in\mathbb{N}} A_n) = \cup_{n\in\mathbb{N}}\, sh^{-1}(A_n). \end{cases}$$

Since the algebra $\mathcal{A}({}^o\mu_L)$ contains $\emptyset$ and is closed under set differences and countable unions, these facts imply that M has the same closure properties. Since $sh^{-1}(\mathbb{R}) \in \mathcal{A}({}^o\mu_L)$, as was shown in Example 16.5.8, we also have $\mathbb{R} \in M$.

It follows that M is a σ-algebra, on which ν turns out to be a measure. To conclude the proof, we need the following lemma. □

LEMMA 16.6.2. *The σ-algebra M includes the Borel algebra $B_{\mathbb{R}}$, and ν agrees with Lebesgue measure on all Borel sets.*

PROOF. To show that each open interval $(a, b) \subseteq \mathbb{R}$ belongs to M, note that $sh^{-1}\big((a, b)\big)$ is the union of the sequence of internal sets $\langle A_n : n \in \mathbb{N}\rangle$, where

$$A_n = T \cap {}^*(a + \tfrac{1}{n}, b - \tfrac{1}{n}) \in \mathcal{A}.$$

But $B_{\mathbb{R}}$ is the smallest σ-algebra containing all open intervals (a, b), so this implies that $B_{\mathbb{R}} \subseteq M$. Also, by Theorem 16.5.4 on intervals, we have

$${}^o\mu_L(A_n) = (b - \tfrac{1}{n}) - (a + \tfrac{1}{n}) = b - a - \tfrac{2}{n}.$$

Since the sets A_n form an increasing sequence, it follows that

$$\nu\big((a, b)\big) = {}^o\mu_L\left(sh^{-1}\big((a, b)\big)\right) = \lim_{n\to\infty} {}^o\mu_L(A_n) = b - a.$$

Thus ν is a measure on $B_{\mathbb{R}}$ that agrees with λ on all open intervals. But any such measure must agree with λ on all Borel sets, as noted in Section 16.3. □

We now complete the proof of the fact that Lebesgue measureable sets are in M and that the function ν agrees with λ on all Lebesgue measurable sets.

If $B \subseteq \mathbb{R}$ is Lebesgue measurable, then by Theorem 16.3.3, there are Borel sets C, D with $C \subseteq B \subseteq D$ and $\lambda(C) = \lambda(B) = \lambda(D)$. Then

$$sh^{-1}(C) \subseteq sh^{-1}(B) \subseteq sh^{-1}(D).$$

Now by Lemma 16.6.2, we have $C, D \in M$, whence $sh^{-1}(C), sh^{-1}(D) \in \mathcal{A}({}^o\mu_L)$, and

$${}^o\mu_L(sh^{-1}(C)) = \nu(C) = \lambda(C) = \lambda(D) = \nu(D) = {}^o\mu_L(sh^{-1}(D)).$$

Since ${}^o\mu_L$ is a complete measure on $\mathcal{A}({}^o\mu_L)$ (by the general theory of outer measures), it follows that $sh^{-1}(B) \in \mathcal{A}({}^o\mu_L)$, and hence $B \in M$, with

$$\nu(B) = {}^o\mu_L(sh^{-1}(B)) = {}^o\mu_L(sh^{-1}(C)) = \lambda(C) = \lambda(B).$$

This establishes that every Lebesgue measurable set is in M (i.e., we have an inclusion of algebras $C_{\mathbb{R}}(\lambda) \subseteq M$) and that the function ν agrees with λ on all Lebesgue measurable sets.[7]

COROLLARY 16.6.3. *If B is a Lebesgue-measurable set of finite Lebesgue measure $\lambda(B)$, the latter can be retrieved as the real infimum of $\mu(A)$ where $A \subseteq T$ ranges over internal sets containing the inverse measure*

$$sh^{-1}(B) = \{t \in T : (\exists b \in B)(t \approx b)\},$$

and μ is the counting-type measure (16.5.1)*:*

$$\lambda(B) = \inf{}^{\mathbf{st}} \{\mu(A) : A \in \mathcal{P}_I(T) \text{ and } sh^{-1}(B) \subseteq A\}.$$

[7] The conclusion of the proof of Theorem 16.6.1 is on pages 218–219 in Goldblatt, where it is shown that M is no bigger than the set of Lebesgue-measurable sets.

Bibliography

[1] Bair, Jacques; Błaszczyk, Piotr; Ely, Robert; Katz, Mikhail; Kuhlemann, Karl. Procedures of Leibnizian infinitesimal calculus: An account in three modern frameworks. *British Journal for the History of Mathematics* **36** (2021), no. 3, 170–209. https://doi.org/10.1080/26375451.2020.1851120, https://arxiv.org/abs/2011.12628, https://mathscinet.ams.org/mathscinet-getitem?mr=4353153

[2] J. Bair, A. Borovik, V. Kanovei, M. Katz, S. Kutateladze, S. Sanders, D. Sherry, and M. Ugaglia, *Historical infinitesimalists and modern historiography of infinitesimals*, Antiquitates Mathematicae **16** (2022), 189–257. http://arxiv.org/abs/2210.14504, https://doi.org/10.14708/am.v16i1.7169

[3] Davis, Martin. *Applied nonstandard analysis.* Pure and Applied Mathematics. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1977. Reprinted by Dover, NY, 2005.

[4] P. Fletcher, K. Hrbacek, V. Kanovei, M. Katz, C. Lobry, and S. Sanders, *Approaches to analysis with infinitesimals following Robinson, Nelson, and others*, Real Analysis Exchange **42** (2017), no. 2, 193–252. https://arxiv.org/abs/1703.00425, http://doi.org/10.14321/realanalexch.42.2.0193

[5] V. Gitman and J. Hamkins, *A natural model of the multiverse axioms*, Notre Dame Journal of Formal Logic, **51** (2010), no. 4, 475–484.

[6] Goldblatt, Robert. *Lectures on the hyperreals. An introduction to Nonstandard Analysis.* Springer, 1998.

[7] Hrbacek, Karel. Axiomatic foundations for nonstandard analysis. *Fundamenta Mathematicae* **98** (1978), no. 1, 1–19.

[8] Hrbacek, Karel; Katz, Mikhail. Infinitesimal analysis without the Axiom of Choice. *Annals of Pure and Applied Logic* **172** (2021), no. 6, 102959. https://doi.org/10.1016/j.apal.2021.102959, https://arxiv.org/abs/2009.04980, https://mathscinet.ams.org/mathscinet-getitem?mr=4224071

[9] K. Hrbacek and M. Katz, *Constructing nonstandard hulls and Loeb measures in internal set theories*, Bulletin of Symbolic Logic **29** (2023), no. 1, 97–127. https://www.doi.org/10.1017/bsl.2022.43, https://arxiv.org/abs/2301.00367

[10] K. Hrbacek and M. Katz, *Peano and Osgood theorems via effective infinitesimals*, in preparation (2023).

[11] Katz, M.; Kuhlemann, K.; Sherry, D.; Ugaglia, M. Leibniz on bodies and infinities: rerum natura and mathematical fictions. *Review of Symbolic Logic* (2023). https://www.doi.org/10.1017/S1755020321000575, https://arxiv.org/abs/2112.08155

[12] Nelson, Edward. Internal set theory: a new approach to nonstandard analysis. *Bulletin of the American Mathematical Society* **83** (1977), no. 6, 1165–1198.
[13] E. Nelson, *Radically Elementary Probability Theory*, Annals of Mathematics Studies 117, Princeton University Press, Princeton, NJ, 1987, 98 pp.
[14] Robinson, Abraham. *Non-standard analysis.* North-Holland Publishing, Amsterdam, 1966.