J. M. Henle

Second-order Non-nonstandard Analysis

Abstract. Following [3], we build higher-order models of analysis resembling the frameworks of nonstandard analysis. The models are entirely canonical, constructed without Choice. Weak transfer principles are developed and the models are applied to topology, graph theory, and measure theory. A Loeb-like measure is constructed.

1. Introduction

In [3], we presented a number system with infinitesimals. We used the system to define the structures of calculus and prove its major theorems, all in the manner of nonstandard analysis. Originally developed by Detlef Laugwitz [7] and [8], the system is not nonstandard, does not require the Axiom of Choice, and uses ordinary sequences of real numbers.

The purpose of [3] was partly mathematical, partly pedagogical. The author was surprised to discover how much of the nonstandard development of calculus could be produced without its machinery and assumptions.

In this paper, we push the ideas further. The system of [3] could not handle second-order ideas to any great extent. We expand that system to allow sequences of sets and sequences of sequences. We produce, in effect, what one might obtain if one could iterate the process of taking nonstandard models into the transfinite. Like [3], the construction in entirely canonical.

We are interested in the question of how much of the second-order theory of nonstandard analysis can be reproduced without the Axiom of Choice, without superstructures, without enlargements, without saturation, and so on. The answer is that quite a lot can be done, not always nicely, but in a manner that retains much of the flavor of the original. There are some disappointments—at one time a non-nonstandard proof of the Baire Category Theorem seemed possible—but some interesting successes.

This paper was not imagined as a contribution to nonstandard analysis, nor as a challenge to it. We admire and respect the field and we especially enjoy its powerful and creative applications of logic. We see no mathematical or philosophical need to replace it, and we are intrigued with the recent

Presented by **Robert Goldblatt**; *Received* December 1, 2002

Studia Logica **68**: 1–28, 2001.

efforts by Di Nasso [1] and Hrbacek [6] to axiomatize it.

We also have no quarrel with the Axiom of Choice.

Similarly, this paper is not intended as a contribution to constructive analysis, which we also admire and respect. Serious efforts to constructivize nonstandard analysis have been made by in recent years by Palmgren [10], [11] and Schuster [14].

In truth, this paper is simply an exploration of ideas that have puzzled and entertained the author for several years.

* * * * *

The first two sections reproduce parts of [3], but in the expanded system. In section 3, we prove a few topological theorems that can be handled in a first-order way.

Sections 4 and 5 introduce sets and revisit the theorems of section 3 in a second-order fashion. "Open" and "closed" are defined for nonstandard sets using a relativized version of infinitely close.

Without the Axiom of Choice the full transfer principle is not possible, but in section 6 we prove a weak version somewhat similar to constructive transfer principles found by Palmgren [11]. This is applied in section 7 to yield some classical results in finite and infinite graph theory.

Lebesgue measure is constructed in section 8. Section 9 constructs a non-nonstandard version of Loeb measure. This is used in Section 10 to construct, for a second time, Lebesgue measure.

2. Background

Definition 2.1. By *sequence*, we mean any function with domain $\mathbb{N}$. A set $\mathcal{S}$ is *sequentially closed* if $\mathcal{S}$ contains all sequences with range contained in $\mathcal{S}$.

Any set is contained in a sequentially closed set. For any set A, and ordinal α, define inductively $A_\alpha = \{a | a : \mathbb{N} \to \bigcup_{\beta<\alpha} A_\beta\}$. Then

$$\bigcup_{\alpha<\omega_1} A_\alpha$$

is sequentially closed, indeed, it is the smallest such set containing A. Because we will be arguing inductively over this set, we will restrict our attention to a special class of sets A.

Definition 2.2. If A contains no sequences, The *sequential closure* of a set A is the smallest sequentially closed set containing A. We will denote the sequential closure of A by $\overset{+}{A}$. We will refer to members of A as *ordinary* and members of $\overset{+}{A} \setminus A$ as *extraordinary.*

A may be a set of numbers, such as $\mathbb{R}$, or something more complicated, such as $\mathbb{R}$ together with its power set. Our structures will be generous.

Note that the relation, "is a term of," is well-founded, hence there is a well–defined rank function on $\overset{+}{A}$, namely,

$$\text{rank}(a) = \min\{\alpha | a \in A_\alpha\}.$$

In a situation with finitely many elements, we choose

$$\text{rank}(\{a_0, \ldots, a_n\}) = \min\{\text{rank}(a_i)\}.$$

Definition 2.3. If a is extraordinary, then a_n is its nth element. If a is ordinary, then a_n is a. If $a^1, \ldots, a^k$ are sequences and f is a k-place function on A, then $f(a^1, \ldots, a^k)$ is the sequence $\left\{f(a_n^1, \ldots, a_n^k)\right\}_n$.

Definition 2.4. $\mathcal{L}_A$ is the first-order language consisting of constant symbols for all members of A, function symbols for all (k-place) functions from A^k to A, relation symbols for all relations on A^k, with variables x, y, z, x_1, x_2, …, put together in the usual way. $\overset{+}{\mathcal{L}}_A$ is $\mathcal{L}_A$ extended by including constant symbols for all members of $\overset{+}{A}$. Terms and well-formed formulas are put together in the usual way by using variables, constants, function symbols, connectives, and quantifiers.

Truth in $\overset{+}{A}$ for well-formed formulas of $\overset{+}{\mathcal{L}}_A$ is defined inductively.

Definition 2.5. We first define truth for simple predicates by induction on the rank of the terms:

If P is a k-place relation on A, and $t^1, \ldots, t^k$ are terms of $\overset{+}{\mathcal{L}}_A$, we say $P(t^1, \ldots, t^k)$ is true in $\overset{+}{A}$ if either

(1) $t^1, \ldots, t^k \in \mathcal{L}_A$ and $P(t^1, \ldots, t^k)$ is true in A, or

(2) $P(t_n^1, \ldots, t_n^k)$ is true in $\overset{+}{A}$ for sufficiently large n, otherwise.

We will sometimes write $\overset{+}{P}(a^1, \ldots, a^k)$ for $\overset{+}{A} \vDash P(a^1, \ldots, a^k)$, for example, $a \overset{+}{\leq} b$.

Notation: *We abbreviate "for sufficiently large n" by "$\forall_n$" (script, as opposed to $\forall n$). We write: $\overset{+}{A} \vDash \Phi$ for "Φ is true in $\overset{+}{A}$." If A is understood, we write simply Φ.*

We extend the definition of truth to more complex formulas by:

(1) $\overset{+}{A} \vDash \Phi \wedge \Psi$ iff $\overset{+}{A} \vDash \Phi$ and $\overset{+}{A} \vDash \Psi$

(2) $\overset{+}{A} \vDash \Phi \vee \Psi$ iff $\overset{+}{A} \vDash \Phi$ or $\overset{+}{A} \vDash \Psi$

(3) $\overset{+}{A} \vDash \Phi \Rightarrow \Psi$ iff $\overset{+}{A} \vDash \Phi \Rightarrow \overset{+}{A} \vDash \Psi$

(4) $\overset{+}{A} \vDash \neg\Phi(t^1, \ldots, t^k)$ iff $\overset{+}{A} \vDash \neg\Phi(t^1_n, \ldots, t^k_n)$ for sufficiently large n

(5) $\overset{+}{A} \vDash \forall x \Phi(x)$ iff for all $a \in \overset{+}{A}$, $\overset{+}{A} \vDash \Phi(a)$

(6) $\overset{+}{A} \vDash \exists x \Phi(x)$ iff for some $a \in \overset{+}{A}$, $\overset{+}{A} \vDash \Phi(a)$

We say $\overset{+}{A}$ *decides* Φ if either $\overset{+}{A} \vDash \Phi$ or $\overset{+}{A} \vDash \neg\Phi$.

Notice that there are statements which are neither true nor, in a sense, false. For $a = 0, 1, 0, 1, 0, 1, \ldots$, for example, $a = 0$ is not true, but neither are $\neg a = 0$ and $a \neq 0$. This example illustrates the fact that there is a difference between saying "$\neg\Phi$" and "Φ is not true."

As an example of the sort inductive reasoning we will frequently use, we prove:

PROPOSITION 2.6. *For all $a, b, c \in \overset{+}{A}$, if $a \overset{\pm}{=} b$ and $b \overset{\pm}{=} c$, then $a \overset{\pm}{=} c$.*

PROOF. This is immediate if a, b, c are members of A. Otherwise, $\overset{+}{A} \vDash a = b$ and $\overset{+}{A} \vDash b = c$ give us $\forall\!\!\!/\, n\, \overset{+}{A} \vDash a_n = b_n$ and $\forall\!\!\!/\, n\, \overset{+}{A} \vDash b_n = c_n$. Then for some n_0, $n > n_0$ implies $\overset{+}{A} \vDash a_n = b_n$ and for some n_1, $n > n_1$ implies $\overset{+}{A} \vDash b_n = c_n$. If n_2 is the larger of n_0, n_1, then $n > n_2$ implies both $\overset{+}{A} \vDash a_n = b_n$ and $\overset{+}{A} \vDash b_n = c_n$, hence, arguing inductively, $\overset{+}{A} \vDash a_n = c_n$. Thus, $\forall\!\!\!/\, n\, \overset{+}{A} \vDash a_n = c_n$ and so $\overset{+}{A} \vDash a = c$. ■

The plus notation, $\overset{+}{A}$, is in imitation of nonstandard theory, where the nonstandard extension of $/R$ is written: $^{*}\mathbb{R}$. In further imitation, we call $\overset{+}{\mathbb{R}}$ the *sortareal* numbers.

For the case of $A = \mathbb{R}$ we define a new relation:

DEFINITION 2.7. We say $a, b \in \mathbb{R}$ are *infinitely close* iff $|a - b| < \epsilon$ for all positive reals ϵ. We write this: $a \approx b$.

We proved in [3] that $\approx$ is transitive and that it respects arithmetic operations, e.g., $a \approx b$, $c \approx d$ imply $a + c \approx b + d$. The relation $\approx$ doesn't satisfy $a \approx b$ iff $\forall\!\!\!/\, n\, a_n \approx b_n$, however, since $\left\{\frac{1}{n}\right\}_n \approx 0$, but for no n is $\frac{1}{n} \approx 0$. We do have the converse.

Proposition 2.8. *$\rlap{/}V n\, a_n \approx b_n$ implies $a \approx b$.*

Proof. Given $\rlap{/}V n\, a_n \approx b_n$, then for any positive real ϵ, $\rlap{/}V n\, |a_n - b_n| < \epsilon$, hence $|a - b| < \epsilon$ and so $a \approx b$. ■

In nonstandard analysis, every finite hyperreal is infinitely close to a unique real. The non-nonstandard version of this is Prop. 3.13 in the next section.

3. Undersequences

Definition 3.9. For $a, b \in \overset{+}{A}$, we say b is an *undersequence* of a ($b \sqsubset a$) iff either $a = b \in A$, or if there is an increasing function f ($f(n+1) > f(n)$ for sufficiently large n) such that b_n is an undersequence of $a_{f(n)}$ for all n. The sequences, $\{b^1, \ldots, b^k\}$ are *undersequences jointly* of $\{a^1, \ldots, a^k\}$, written $\{b^1, \ldots, b^k\} \sqsubset \{a^1, \ldots, a^k\}$, iff either all are in A and $a^i = b^i$ for all i, or $\rlap{/}V n$, $\{b^1_n, \ldots, b^k_n\}$ are undersequences jointly of $\{a^1_{f(n)}, \ldots, a^k_{f(n)}\}$, for some increasing function f.

The following facts are easily verifiable:

- $\{a^1, \ldots, a^k\} \sqsubset \{b^1_n, \ldots, b^k_n\} \sqsubset \{c^1_n, \ldots, c^k_n\} \Rightarrow \{a^1, \ldots, a^k\} \sqsubset \{c^1_n, \ldots, c^k_n\}$
- $\{a^1, \ldots, a^k\} \sqsubset \{b^1_n, \ldots, b^k_n\} \not\Rightarrow \{c, a^1, \ldots, a^k\} \sqsubset \{c, b^1_n, \ldots, b^k_n\}$, but
- $\{a^1, \ldots, a^k\} \sqsubset \{b^1_n, \ldots, b^k_n\} \Rightarrow \forall c \exists d\ \{c, a^1, \ldots, a^k\} \sqsubset \{d, b^1_n, \ldots, b^k_n\}$

Proposition 3.10. *Suppose $\Phi(x^1, \ldots x^k)$ contains no implication symbols. Then Φ is preserved by undersequence, that is, if $\left\{b^i\right\}_{i \leq k}$ are undersequences jointly of $\left\{a^i\right\}_{i \leq k}$, then $\overset{+}{A} \models \Phi(a^1, \ldots, a^k)$ implies $\overset{+}{A} \models \Phi(b^1, \ldots, b^k)$.*

Proof. We first prove the case where Φ is a relation $P(a^1, \ldots, a^k)$. If $a^1, \ldots a^k$ are ordinary, then $b^i = a^i$ for each $i \leq k$, so the statement is immediate. Otherwise, proceeding inductively, suppose each b^i_n is an undersequence of $a^i_{f(n)}$, f increasing. Then since we have that $\rlap{/}V n\ \overset{+}{A} \models P(a^1_n, \ldots, a^k_n)$, we have $\rlap{/}V n\ \overset{+}{A} \models P(a^1_{f(n)}, \ldots, a^k_{f(n)})$. By induction then, $\rlap{/}V n\ \overset{+}{A} \models P(b^1_n, \ldots, b^k_n)$ and so $\overset{+}{A} \models P(b^1, \ldots, b^k)$.

The rest of the proof is an easy induction on the length of Φ. ■

A key proposition in nonstandard analysis relates the truth of $\Phi(a^1, \ldots, a^k)$ to the truth $\Phi(a^1_n, \ldots, a^k_n)$ for a "large" set of n. This fails utterly here, even for sentences as simple as $a = 0$. This problem can be negotiated, however, with undersequences.

Proposition 3.11. *Let P be a k-place relation on A with $a^1, \dots a^k \in \overset{+}{A}$. Then*

(1) *We can find $\{b^i\}_{i \le k}$ undersequences jointly of $\{a^i\}_{i \le k}$ such that $\overset{+}{A}$ decides $P(b^1, \dots, b^k)$.*

(2) *If $\overset{+}{A} \not\models P(a^1, \dots a^k)$, then we can require that $\overset{+}{A} \models \neg P(b^1, \dots, b^k)$.*

(3) *If $\overset{+}{A} \not\models \neg P(a^1, \dots a^k)$, then we can require that $\overset{+}{A} \models P(b^1, \dots, b^k)$.*

Proof. For (1), we proceed by induction over sequences. If $a^1, \dots, a^k$ are ordinary, we are done. Otherwise, for each n, choose $\{c_n^i\}_{i \le k}$ undersequences jointly of $\{a_n^i\}_{i \le k}$ such that $\overset{+}{A}$ decides $P(c_n^1, \dots, c_n^k)$. Then either $\overset{+}{A}$ decides $P(c_n^1, \dots, c_n^k)$ positively for infinitely many n or negatively for infinitely many n. In either case, we can form $\{b^i\}_{i \le k}$ by putting together an undersequence of $\{c_n^1, \dots, c_n^k\}_{i \le k}$.

The proofs of (2) and (3) are similar. Suppose, for example, that $\overset{+}{A} \not\models P(a^1, \dots a^k)$. Then there are infinitely many n such that $\overset{+}{A} \not\models P(a_n^1, \dots a_n^k)$. For each such n, we can find $\{b_n^i\}_{i \le k}$ undersequences jointly with $\overset{+}{A} \models \neg P(b_n^1, \dots, c_b^k)$, and these give us the joint undersequences of $a^1, \dots a^k$. ∎

Proposition 3.12. *For any countable collection $\mathcal{S}$ of k-place relations on A and $a^1, \dots a^k \in \overset{+}{A}$, we can find $\{b^i\}_{i \le k}$ undersequences jointly of $\{a^i\}_{i \le k}$ such that $\overset{+}{A}$ decides $P(b^1, \dots, b^k)$ for all $P \in \mathcal{S}$.*

Proof. Let $\mathcal{S}$ be enumerated by $\{P_n\}_{n<\infty}$. Using Prop. 3.11 (1), choose $\{c^{i,0}\}_{i \le k}$ undersequences jointly of $\{a^i\}_{i \le k}$, witnessed by f_0, such that $\overset{+}{A}$ decides $P_0(c^{1,0}, \dots, c^{k,0})$. Let l_0 be such that $n \ge l_0$ implies that $\left\{c_n^{i,0}\right\}_{i \le k}$ are undersequences jointly of $\left\{a_{f_0(n)}^i\right\}_{i \le k}$ and set $\{b_0^i\}_{i \le k}$ equal to $\left\{c_{f_0(l_0)}^{i,0}\right\}_{i \le k}$.

For $j = 1, 2, \dots$, choose $\{c^{i,j}\}_{i \le k}$ undersequences jointly of $\{c^{i,j-1}\}_{i \le k}$, witnessed by f_j, such that $\overset{+}{A}$ decides $P_j(c^{1,j}, \dots, c^{k,j})$. Choose $l_j > l_{j-1}$ such that $n \ge l_j$ implies that $\left\{c_n^{i,j}\right\}_{i \le k}$ are undersequences jointly of $\left\{c_{f_j(n)}^{i,j-1}\right\}_{i \le k}$ and set $\left\{b_j^i\right\}_{i \le k}$ equal to $\left\{c_{f_j(l_j)}^{i,j}\right\}_{i \le k}$. Note that $\left\{b_j^i\right\}_{i \le k}$ are undersequences jointly of $\left\{c_{f_{j-1}(f_j(l_j))}^{i,j-1}\right\}_{i \le k}$ and of $\left\{c_{f_{j-2}(f_{j-1}(f_j(l_j)))}^{i,j-2}\right\}_{i \le k}$, and so on.

When we are done, $\{b^i\}_{i\leq k}$ will be jointly undersequences of $\{c^{i,d}\}_{i\leq k}$ for all d, witnessed by the function:

$$f^d(q) = f_d(f_{d+1}(\ldots(f_q(l_q))\ldots)).$$

Then by Prop. 3.10, since $\overset{+}{A}$ decides $P_d(c^{1,d},\ldots,c^{k,d})$, it follows that $\overset{+}{A}$ decides $P_d(b^1,\ldots,b^k)$ for all d. ■

Turning now to $\mathbb{R}$:

PROPOSITION 3.13. *For any sequence a, if $|a| \leq d$ for some real d, then there is a real r and an undersequence $b \sqsubset a$ with $b \approx r$.*

PROOF. We apply Prop. 3.12 to a and the relations: $R_q(x) : x \geq q$ where q ranges over all ordinary rational numbers. Any undersequence b of a which decides all $R_q(x)$ defines a real number $r \approx b$. ■

There may be many such r. We don't have the "standard part" ${}^{\circ}a$, of nonstandard analysis, the unique real number infinitely close to a. For most purposes, Prop. 3.13 is all we need, but uniqueness is important when we try to recover Loeb theory later in this paper. For that, we have:

DEFINITION 3.14. If a sequence a is finite ($|a| < r$ for some real r), we define the *ordinary part* of a, ${}^{\square}a$, inductively as:

$$ {}^{\square}a = \begin{cases} a & \text{if } a \text{ is real} \\ \liminf\{{}^{\square}a_n\} & \text{otherwise} \end{cases}$$

Ordinary part is not linear. If a is the sequence: $0,1,0,1,\ldots$ and b is $1,0,1,0,\ldots$, ${}^{\square}(a+b) = 1 \neq 0 = {}^{\square}a + {}^{\square}b$. Still, we can prove:

PROPOSITION 3.15. *For sequences a and b, ${}^{\square}(a+b) \geq {}^{\square}a + {}^{\square}b$.*

PROOF. This is certainly true for reals. Arguing inductively,

$$\begin{aligned} {}^{\square}(a+b) &= \liminf\{{}^{\square}(a+b)_n\} \\ &\geq \liminf\{{}^{\square}a_n + {}^{\square}b_n\} \\ &\geq \liminf\{{}^{\square}a_n\} + \liminf\{{}^{\square}b_n\} \\ &= {}^{\square}a + {}^{\square}b. \end{aligned}$$

■

Finally, it is easy to prove:

PROPOSITION 3.16. *For any sequence a, if $|a| \leq d$ for some real d, then there is an undersequence $b \sqsubset a$ with $b \approx {}^{\square}a$.*

4. Sets

We consider now the consequences when A is a richer structure than $\mathbb{R}$, a set containing sets of elements. The relations on A, $\in$ and $\subseteq$ extend to relations on $\overset{+}{A}$ inductively as in Definition 2.5, for example,

$$x \overset{+}{\in} y \text{ iff } \forall\!\!\!/_n\, x_n \overset{+}{\in} y_n.$$

We will write "$x \overset{+}{\notin} y$ for $\overset{+}{A} \vDash \neg(x \overset{+}{\in} y)$ (*not* $\overset{+}{A} \nvDash x \overset{+}{\in} y$).

The operations on A, $\cap$, $\cup$, ^{c}, and $\setminus$ extend to operations on $\overset{+}{A}$ inductively as in Definition 2.3, for example, $x \overset{+}{\cap} y$ is the sequence whose nth term is $x_n \overset{+}{\cap} y_n$.

We start by noting that Comprehension fails, though it comes very close. We have:

PROPOSITION 4.17. *For all $a, b \in \overset{+}{A}$,*

(1) $\overset{+}{A} \vDash a \overset{+}{\subseteq} b$ *implies* $\overset{+}{A} \vDash \forall z(z \overset{+}{\in} a \Rightarrow z \overset{+}{\in} b)$

(2) $\overset{+}{A} \vDash a \overset{+}{=} b$ *iff* $\overset{+}{A} \vDash a \overset{+}{\subseteq} b$ *and* $\overset{+}{A} \vDash b \overset{+}{\subseteq} a$

(3) *If* $\overset{+}{A} \vDash a \overset{+}{\neq} \emptyset$*, then* $\overset{+}{A} \vDash \forall z(z \overset{+}{\in} a \Rightarrow z \overset{+}{\in} b)$ *implies* $\overset{+}{A} \vDash a \overset{+}{\subseteq} b$*.*

PROOF. (1) and (2) follow easily by induction. We use induction as well for (3). It is clearly true for ordinary a, b. Otherwise, suppose that $\overset{+}{A} \vDash a \overset{+}{\neq} \emptyset$, $\overset{+}{A} \vDash \forall z(z \overset{+}{\in} a \Rightarrow z \overset{+}{\in} b)$, and $\overset{+}{A} \nvDash (a \overset{+}{\subseteq} b)$. This last statement means that for infinitely many n, $\overset{+}{A} \nvDash (a_n \overset{+}{\subseteq} b_n)$. By induction, for infinitely many n, we may find z_n such that $\overset{+}{A} \vDash z_n \overset{+}{\in} a_n \wedge z_n \overset{+}{\notin} b_n$. Since $\overset{+}{A} \vDash a_n \overset{+}{\neq} \emptyset$, we can complete the definition of z by setting all but finitely many of the remaining z_n to be elements of the corresponding a_n. Then $\overset{+}{A} \vDash z \overset{+}{\in} a$, but $\overset{+}{A} \nvDash z \overset{+}{\in} b$, a contradiction. ■

The restriction in (3) that a be non-empty is necessary, as the example:

$$a_n = \begin{cases} \emptyset & \text{if } n \text{ is odd} \\ \{1\} & \text{if } n \text{ is even,} \end{cases} \qquad b = \emptyset$$

shows. We do have ${}^{+}\!A \models \forall z(z \overset{+}{\in} a \Rightarrow z \overset{+}{\in} b)$, but ${}^{+}\!A \not\models a \overset{+}{\subseteq} b$.

The interface between $\overset{+}{\in}$ and ${}^{+}$set operations has similar problems.

Proposition 4.18. *For $b, c \in {}^{+}\!A$,*

(1) ${}^{+}\!A \models z \overset{+}{\in} (b \overset{+}{\cap} c)$ *iff* ${}^{+}\!A \models z \overset{+}{\in} b$ *and* ${}^{+}\!A \models z \overset{+}{\in} c$,

(2) ${}^{+}\!A \models z \overset{+}{\in} b$ *or* ${}^{+}\!A \models z \overset{+}{\in} c$ *implies* ${}^{+}\!A \models z \overset{+}{\in} (b \overset{+}{\cup} c)$,

(3) ${}^{+}\!A \models z \overset{+}{\notin} (b \overset{+}{\cup} c)$ *iff* ${}^{+}\!A \models z \overset{+}{\notin} b$ *and* ${}^{+}\!A \models z \overset{+}{\notin} c$,

(4) ${}^{+}\!A \models z \overset{+}{\notin} b$ *or* ${}^{+}\!A \models z \overset{+}{\notin} c$ *implies* ${}^{+}\!A \models z \overset{+}{\notin} (b \overset{+}{\cap} c)$.

These are all easily established by induction. Note that the missing implications are false. Consider in $\mathbb{R}$, for example,

$$b_n = \begin{cases} \{0\} & \text{if } n \text{ is odd} \\ \{1\} & \text{if } n \text{ is even} \end{cases} \quad \text{and} \quad c_n = \begin{cases} \{1\} & \text{if } n \text{ is odd} \\ \{0\} & \text{if } n \text{ is even.} \end{cases}$$

We have ${}^{+}\mathbb{R} \models 0 \overset{+}{\in} (b \overset{+}{\cup} c)$, but ${}^{+}\mathbb{R} \not\models 0 \overset{+}{\in} b$ and ${}^{+}\mathbb{R} \not\models 0 \overset{+}{\in} c$.

The difficulties here suggest that we isolate a class of more well-behaved sets. We do this in the context of the real numbers.

Definition 4.19. For any $X \in {}^{+}\mathcal{P}(\mathbb{R})$, let $Re(X)$ be the collection of real (ordinary) numbers in X. $X \in {}^{+}\mathcal{P}(\mathbb{R})$ is *real-definite* if $Re(X) \cup Re(X^c) = \mathbb{R}$. Equivalently, X is real-definite if for every real r, either ${}^{+}\mathbb{R} \models r \overset{+}{\in} X$ or ${}^{+}\mathbb{R} \models r \overset{+}{\notin} X$.

While $Re(b \overset{+}{\cap} c) = Re(b) \cap Re(c)$ is true in general, $Re(b \overset{+}{\cup} c) = Re(b) \cup Re(c)$ is not always the case as the example above shows. On the other hand, we do have:

Proposition 4.20. *For all real-definite sets a and c, the sets $b \overset{+}{\cap} c$, $b \overset{+}{\cup} c$, $b \overset{+}{\setminus} c$, and b^c are real-definite and*

(1) $Re(b \overset{+}{\cap} c) = Re(b) \cap Re(c)$

(2) $Re(b \overset{+}{\cup} c) = Re(b) \cup Re(c)$

(3) $Re(b \overset{+}{\setminus} c) = Re(b) \setminus Re(c)$, *and*

(4) $Re(b)^c = Re(b^c)$

PROOF. We will just show (2). One direction, $\supseteq$, is trivial. For the other, we proceed by induction. For real b, c, it is immediate. For sequences, suppose that $r \in Re(b \cup c)$, but $r \notin Re(b)$ and $r \notin Re(c)$. Then by real-definiteness, there is a k such that for all $n > k$, ${}^{+}\mathbb{R} \models r \stackrel{+}{\notin} b_n$ and ${}^{+}\mathbb{R} \models r \stackrel{+}{\notin} c_n$, and consequently ${}^{+}\mathbb{R} \models r \stackrel{+}{\notin} b_n \stackrel{+}{\cup} c_n$. Then ${}^{+}\mathbb{R} \models r \stackrel{+}{\notin} (b \stackrel{+}{\cup} c)$, so ${}^{+}\mathbb{R} \models r \notin Re(b) \cup Re(c)$, a contradiction. ∎

5. First-order Real Topology

Some aspects of the topology of the real line can be discussed in a first-order way, since every subset of $\mathbb{R}$ can be treated as a relation on $\mathbb{R}$. The nonstandard topological definitions of topological properties in [4] translate easily to non–nonstandard definitions which in turn are equivalent to the standard definitions. For example:

DEFINITION 5.21. A set $X \subseteq \mathbb{R}$ is *open* if for all reals r, $a \approx r \in X$ implies $a \stackrel{+}{\in} X$. X is *closed* if for all reals r, $r \approx a \stackrel{+}{\in} X$ implies $r \stackrel{+}{\in} X$.

PROPOSITION 5.22. *The properties of open and closed are preserved under finite unions and intersections.*

PROOF. This is immediate with one exception, the union of closed sets. Suppose that X and Y are closed and that $r \approx a \stackrel{+}{\in} X \cup Y$, r real. We can't conclude that either $a \stackrel{+}{\in} X$ or $a \stackrel{+}{\in} Y$, for example, if a is the sequence, $0, 1, 0, 1, \ldots$ and $X = \{0\}$, $Y = \{1\}$. We can, however, prove by a simple induction on a that $a \stackrel{+}{\in} X \cup Y$ implies that for some $b \sqsubset a$, either $b \stackrel{+}{\in} X$ or $b \stackrel{+}{\in} Y$. With such a b, we still have $r \approx b$, since the relation, $|r - x| < \epsilon$, is covered by Prop. 3.10. This gives us $r \in X \cup Y$. ∎

PROPOSITION 5.23. *For any sequence of bounded, non-empty closed sets, $F_0 \supseteq F_1 \supseteq F_1 \supseteq \ldots$, the intersection, $\bigcap_{k<\infty} F_k$, is non-empty.*

PROOF. For each k, choose $a_k \in F_k$. Then $a \stackrel{+}{\in} F_k$ for each k. Choose $b \sqsubset a$, r real such that $b \approx r$. Then $b \stackrel{+}{\in} F_k$ for all k by Prop. 3.10 so $r \in F_k$ and $r \in \bigcap_{k<\infty} F_k$. ∎

The following definition is equivalent to the standard definition of compactness:

Definition 5.24. X is *compact* iff for all $a \stackrel{+}{\in} X$, there is an undersequence $b \sqsubset a$ and a real $r \stackrel{+}{\in} X$, $b \approx r$.

Theorem 5.25. **(The Bolzano-Weierstrass Theorem)** *A set is compact if and only if it is closed and bounded.*

Proof. ($\Rightarrow$) If $r \approx a \stackrel{+}{\in} X$, $r \in \mathbb{R}$, let $b \sqsubset a$, $b \approx s \in X$. Then since $b \approx r$, $r = s \in X$, hence X is closed. Now suppose that X is not bounded. For each n, choose a real $a_n \in (X \setminus [-n, n])$. Then there cannot be an undersequence of $\{a_n\}_n$ infinitely close to a real number.

($\Leftarrow$) Suppose $a \in X$. Since X is bounded, a is bounded and we can find $b \sqsubset a$ and real $r \approx b$. Since X is closed, $r \in X$. ∎

6. Second-order Real Topology

For this section, our underlying set will be $A = \mathbb{R} \cup \mathcal{P}(\mathbb{R})$, the reals together with its power set. In this setting, the definitions of "open" and "closed" extend inductively, that is, ordinary elements of ${}^{+}A$ are ${}^{+}$open if they are open and an extraordinary element X is ${}^{+}$open if X_n is ${}^{+}$open for sufficiently large n. Unfortunately, these sets fail to have the property described in Definition 5.21. Consider, for example, then ${}^{+}$open set $X = \left\{(-\frac{1}{n^2}, \frac{1}{n^2})\right\}_n$ and $a = \left\{\frac{1}{n}\right\}_n$. We have $a \approx 0 \stackrel{+}{\in} X$, yet $a \stackrel{+}{\notin} X$.

To define a more appropriate property, we need to look at relative simplicity.

Definition 6.26. For sequences, a and b, we say a is *at least as simple as* b ($a \preccurlyeq b$) if either a is ordinary or if b is extraordinary and $\forall n\, a_n \preccurlyeq b_n$.

If b is ordinary, and $a \preccurlyeq b$, then a must be ordinary. This means that whether b is ordinary or not, $a \preccurlyeq b$ implies $\forall n\, a_n \preccurlyeq b_n$.

Definition 6.26 allows us to define gradations of "infinitely close:"

Definition 6.27. For $a, b, c \in {}^{+}\mathbb{R}$, $a \approx_b c$ iff $|a - c| < \epsilon$, for all positive $\epsilon \preccurlyeq b$.

Note that $a \approx c$ is the same as $a \approx_b c$ for any real b.

Proposition 6.28. *For $a, b, c \in {}^{+}\mathbb{R}$, $\forall n\, a_n \approx_{b_n} c_n$ implies $a \approx_b c$. If b is extraordinary, then $a \approx_b c$ implies $\forall n\, a_n \approx_{b_n} c_n$.*

Proof. For the first statement, given $\epsilon > 0$, $\epsilon \preccurlyeq b$, we have $\rlap{/}\forall n\, \epsilon_n > 0$ and $\rlap{/}\forall n\, \epsilon_n \preccurlyeq b_n$. Then $\rlap{/}\forall n\, |a_n - c_n| < \epsilon_n$ so $|a - c| < \epsilon$.

For the second statement, assume the contrary. Then for infinitely many n, there exist $\epsilon_n > 0$, $\epsilon_n \preccurlyeq b_n$ with $\neg|a_n - c_n| < \epsilon_n$. For all other n, let $\epsilon_n = 1$. Then $\epsilon > 0$ and $\epsilon \preccurlyeq b$ by Definition 6.26 since b is real, so $|a - c| < \epsilon$. But then $\rlap{/}\forall n\, |a_n - c_n| < \epsilon_n$, a contradiction. ■

Now we can prove:

Proposition 6.29. *If X is open, then*

$$\forall b \preccurlyeq X, \text{ if } a \approx_X b \text{ and } b \overset{+}{\in} X, \text{ then } a \overset{+}{\in} X. \qquad (*)$$

If X is closed, then

$$\forall b \preccurlyeq X, \text{ if } a \approx_X b \text{ and } a \overset{+}{\in} X, \text{ then } b \overset{+}{\in} X. \qquad (**)$$

Furthermore, the converses hold for all $X \overset{+}{\neq} \emptyset$.

Proof. Suppose X is open. If X is ordinary, then $b \preccurlyeq X$ means b is real, $a \approx_X b$ means $a \approx b$, and so $a \overset{+}{\in} X$ implies $b \in X$. If X is not real, then $\rlap{/}\forall n\, b_n \preccurlyeq X_n$, $\rlap{/}\forall n\, a_n \approx_{X_n} b_n$, and $\rlap{/}\forall n\, a_n \overset{+}{\in} X_n$, so by induction, $\rlap{/}\forall n\, b_n \overset{+}{\in} X_n$ and so $b \overset{+}{\in} X$.

For the converse suppose that $X \overset{+}{\neq} \emptyset$ and X has property (*). We can prove that as X is non-empty, X must contain an element at least as simple as X (arguing inductively, this is certainly true if X is ordinary; otherwise we can choose $c_n \overset{+}{\in} X_n$, $c_n \preccurlyeq X_n\ \rlap{/}\forall n$, and so on). Let c be in X, with $c \preccurlyeq X$.

Now if X is ordinary, it follows as before that property (*) is equivalent to that in Definition 5.21 and so X is open. Suppose X is extraordinary and not open. Then for all n in some infinite set I, X_n is not open. By induction, for all n in I we can find $b_n \preccurlyeq X_n\ a_n \approx_{X_n} b_n$, $b_n \overset{+}{\in} X_n$, such that $a_n \overset{+}{\in} X_n$ is false. Form $d \preccurlyeq X$, $d \overset{+}{\in} X$, by setting $d_n = \begin{cases} b_n & \text{if } n \in I \\ c_n & \text{if } n \notin I \end{cases}$. Form $e \approx_X d$, by setting $e_n = \begin{cases} a_n & \text{if } n \in I \\ c_n & \text{if } n \notin I \end{cases}$. Then d, e, witness the failure of (*), since $e \overset{+}{\in} X$ fails as $e_n \overset{+}{\in} X_n$ is not true for all $n \in I$. Thus the assumption that X is not open is false.

The proofs for closed sets are nearly identical to the proofs for open. ■

The difficulty of generalization can be seen in Prop. 5.23. On the one hand, we can prove:

PROPOSITION 6.30. *If* $F_0 \overset{+}{\supseteq} F_1 \overset{+}{\supseteq} F_2 \overset{+}{\supseteq} \ldots$ *is any sequence of non-empty sets, closed or not, then* $\bigcap_{k<\infty} F_k$*, is non-empty, by which we mean that there is a* b *such that* $b \overset{+}{\in} F_k$ *for all* k*.*

Clearly, this is not a generalization. If we try to sharpen the statement by requiring that the real content of the sets be non-empty, it becomes false: Let $(F_n)_k = [\frac{1}{k}, \frac{1}{n} - \frac{1}{k}]$. Each F_n is closed, and contains $\frac{1}{2n}$. We have $F_n \overset{+}{\supseteq} F_{n+1}$ since $\forall k\, (F_n)_k \overset{+}{\supseteq} (F_{n+1})_k$, but $\bigcap_{n<\infty} Re(F_n) = \emptyset$.

To prove Prop. 6.30, choose $a_k \overset{+}{\in} F_k$ for each k. Construct an increasing sequence, $\{k_j\}_{j \in \mathbb{N}}$ such that $n \geq k_i$ implies that $(a_j)_n \bigcap_{m \leq j} (F_m)_n$. Then form b by setting $b_i = (a_q)_i$, where $k_q \leq i < k_{q+1}$. We have $b \overset{+}{\in} F_p$ for all p, since $b_i \overset{+}{\in} (F_p)_i$ for all $i \geq k_p$.

Generalizing compactness has similar problems. Boundedness, for example, is not preserved under taking sequences.

7. A Transfer Principle

The most fundamental property of nonstandard analysis is that a statement of $\mathcal{L}_A$ is true in the standard model A if and only if it is true in the non-standard model ${}^{*}A$. The analogous property is false for A and ${}^{+}A$. A simple example in $\mathbb{R}$ is $\forall x(x = 0 \vee x \neq 0)$. For the sequence $a : 0, 1, 0, 1, 0, 1, \ldots$, we have ${}^{+}\mathbb{R} \not\models a = 0$ and ${}^{+}\mathbb{R} \not\models \neg(a = 0)$, yet ${}^{+}\mathbb{R} \models \forall x(x = 0 \vee x \neq 0)$.

There is, however, what we will call the "Weak Transfer Principle," which applies to all formulas in a restricted class.

DEFINITION 7.31. Define the set of sentences $\mathcal{T}_A \subseteq \mathcal{L}_A$ recursively according to the rules :

(1) $\mathcal{T}_A$ includes all atomic sentences,
(2) $\mathcal{T}_A$ includes $\Phi \wedge \Psi$, if Φ and Ψ are in $\mathcal{T}_A$,
(3) $\mathcal{T}_A$ includes $\forall x \Phi(x)$ and $\exists x \Phi(x)$, if $\Phi(a)$ is in $\mathcal{T}_A$ for $a \in A$,
(4) $\mathcal{T}_A$ includes $\forall x(\Phi(x) \Rightarrow \Psi(x))$, if $\Phi(a)$ is in $\mathcal{T}_A$ for $a \in A$, and ${}^{+}A \models \exists x \Phi(x)$.

${}^{+}\mathcal{T}_A$ is formed from ${}^{+}\mathcal{L}_A$ in the same manner.

The Weak Transfer principle requires the Weak Łos Theorem:

THEOREM 7.32. **(The Weak Łos Theorem)** *For all sentences* Φ *in* $\overset{+}{\mathcal{T}}_A$,

$$\overset{+}{A} \models \Phi(a^1, a^2, \ldots, a^k) \text{ iff } \rlap{/}\forall n\, \overset{+}{A} \models \Phi(a^1_n, a^2_n, \ldots, a^k_n),$$

where $a^1, a^2, \ldots, a^k$ *includes all constants occuring in* Φ.

PROOF. We proceed by induction on the length of Φ. If Φ is a relation, this is simply true by Definition 2.5.

Suppose Φ is $\Psi \wedge \Theta$. Then $\overset{+}{A} \models \Psi(a^1, a^2, \ldots, a^k) \wedge \Theta(a^1, a^2, \ldots, a^k)$ iff $\overset{+}{A} \models \Psi(a^1, a^2, \ldots, a^k)$ and $\overset{+}{A} \models \Theta(a^1, a^2, \ldots, a^k)$ (by Definition 2.5) iff $\rlap{/}\forall n\, \overset{+}{A} \models \Psi(a^1_n, a^2_n, \ldots, a^k_n)$ and $\rlap{/}\forall n\, \overset{+}{A} \models \Theta(a^1_n, a^2_n, \ldots, a^k_n)$ (by induction) iff $\rlap{/}\forall n\, \overset{+}{A} \models \Psi(a^1_n, a^2_n, \ldots, a^k_n) \wedge \Theta(a^1_n, a^2_n, \ldots, a^k_n)$ (as in the proof of Prop. 2.6).

Of the two quantifiers, the case for $\exists$ is easier. We prove the case for $\forall$: Suppose Φ is $\forall x \Psi(x, a^1, a^2, \ldots, a^k)$ and $\overset{+}{A} \models \forall x \Psi(x, a^1, a^2, \ldots, a^k)$, but not $\rlap{/}\forall n\, \overset{+}{A} \models \forall x \Psi(x, a^1_n, a^2_n, \ldots, a^k_n)$. Then for infinitely many n there is a b_n such that $\overset{+}{A} \not\models \Psi(b_n, a^1_n, a^2_n, \ldots, a^k_n)$. For all other n, let b_n be arbitrary. Then since $\overset{+}{A} \models \Psi(b, a^1, a^2, \ldots, a^k)$, we have by induction that $\rlap{/}\forall n\, \overset{+}{A} \models \Psi(b_n, a^1_n, a^2_n, \ldots, a^k_n)$, a contradiction.

In the other direction, if $\rlap{/}\forall n\, \overset{+}{A} \models \forall x \Psi(x, a^1_n, a^2_n, \ldots, a^k_n)$ and b is given, then since $\rlap{/}\forall n\, \overset{+}{A} \models \Psi(b_n, a^1_n, a^2_n, \ldots, a^k_n)$, $\overset{+}{A} \models \Psi(b, a^1, a^2, \ldots, a^k)$ by induction. Thus, $\overset{+}{A} \models \forall x \Psi(x, a^1, a^2, \ldots, a^k)$.

Now suppose that Φ is $\forall x(\Psi(x, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(x, a^1, a^2, \ldots, a^k))$, $\overset{+}{A} \models \exists x \Psi(x, a^1, a^2, \ldots, a^k)$, and

$$\overset{+}{A} \models \forall x(\Psi(x, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(x, a^1, a^2, \ldots, a^k))$$

but it is false that $\rlap{/}\forall n\, \overset{+}{A} \models \forall x(\Psi(x, a^1_n, a^2_n, \ldots, a^k_n) \Rightarrow \Theta(x, a^1_n, a^2_n, \ldots, a^k_n))$. Then for all n in some infinite set I, we can find b_n such that

$$\overset{+}{A} \not\models (\Psi(b_n, a^1_n, a^2_n, \ldots, a^k_n) \Rightarrow \Theta(b_n, a^1_n, a^2_n, \ldots, a^k_n)),$$

that is, for all $n \in I$, $\overset{+}{A} \models \Psi(b_n, a^1_n, a^2_n, \ldots, a^k_n)$ and $\overset{+}{A} \not\models \Theta(b_n, a^1_n, a^2_n, \ldots, a^k_n)$. By induction, $\rlap{/}\forall n\, \overset{+}{A} \models \exists x \Psi(x, a^1_n, a^2_n, \ldots, a^k_n)$, so for $n \notin I$ sufficiently large, we may choose b_n so that $\overset{+}{A} \models \Psi(b_n, a^1_n, a^2_n, \ldots, a^k_n)$. Then

$$\rlap{/}\forall n\, \overset{+}{A} \models \Psi(b_n, a^1_n, a^2_n, \ldots, a^k_n)$$

so again by induction, $\overset{+}{A} \models \Psi(b, a^1, a^2, \ldots, a^k)$. Since

$$\overset{+}{A} \models (\Psi(b, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(b, a^1, a^2, \ldots, a^k)),$$

we have $\overset{+}{A} \models \Theta(b, a^1, a^2, \ldots, a^k)$ and once more by induction, $\forall\!\!\!/_n \overset{+}{A} \models \Theta(b_n, a_n^1, a_n^2, \ldots, a_n^k)$. This is a contradiction.

Finally, suppose $\forall\!\!\!/_n \overset{+}{A} \models \forall x(\Psi(x, a_n^1, a_n^2, \ldots, a_n^k) \Rightarrow \Theta(x, a_n^1, a_n^2, \ldots, a_n^k))$ and we are given $b \in \overset{+}{A}$. Then

$$\forall\!\!\!/_n \overset{+}{A} \models (\Psi(b_n, a_n^1, a_n^2, \ldots, a_n^k) \Rightarrow \Theta(b_n, a_n^1, a_n^2, \ldots, a_n^k)).$$

We claim this gives us $\overset{+}{A} \models (\Psi(b, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(b, a^1, a^2, \ldots, a^k))$ (if $\overset{+}{A} \models \Psi(b, a^1, a^2, \ldots, a^k)$, then by induction, $\forall\!\!\!/_n \overset{+}{A} \models \Psi(b_n, a_n^1, a_n^2, \ldots, a_n^k)$ so $\forall\!\!\!/_n \overset{+}{A} \models \Theta(b_n, a_n^1, a_n^2, \ldots, a_n^k))$, so $\overset{+}{A} \models \Theta(b, a^1, a^2, \ldots, a^k)))$. This proves $\overset{+}{A} \models \forall x(\Psi(x, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(x, a^1, a^2, \ldots, a^k))$. ■

The referee notes that the Theorem 7.32 can be used to improve a number of results in this paper. First, Prop. 3.11:

PROPOSITION 7.33. *Let $\Phi(x_1 \ldots, x_k)$ be a formula in $\mathcal{T}_A$ with $a^1, \ldots a^k \in \overset{+}{A}$. Then*

(1) *We can find $\{b^i\}_{i \leq k}$ undersequences jointly of $\{a^i\}_{i \leq k}$ such that $\overset{+}{A}$ decides $\Phi(b^1, \ldots, b^k)$.*

(2) *If $\overset{+}{A} \not\models \Phi(a^1, \ldots a^k)$, then we can require that $\overset{+}{A} \models \neg\Phi(b^1, \ldots, b^k)$.*

(3) *If $\overset{+}{A} \not\models \neg\Phi(a^1, \ldots a^k)$, then we can require that $\overset{+}{A} \models \Phi(b^1, \ldots, b^k)$.*

Second, Prop. 3.12:

PROPOSITION 7.34. *For any countable collection $\{\Phi^i(x_1 \ldots, x_k)\}_{i \in \mathbb{N}}$ of formulas in $\mathcal{T}_A$ and $a^1, \ldots a^k \in \overset{+}{A}$, we can find $\{b^i\}_{i \leq k}$ undersequences jointly of $\{a^i\}_{i \leq k}$ such that $\overset{+}{A}$ decides $\Phi^i(b^1, \ldots, b^k)$ for all $i \in \mathbb{N}$.*

We will find these useful later. In addition, the referee notes that Prop. 6.30 can be seen as a special case of what might be called a "sortasaturation" principle:

PROPOSITION 7.35. **(Sortasaturation Principle)** *Given a countable set $\{\Phi_i(x, a^{i,1}, \ldots, a^{i,k}) | i \in \mathbb{N}\}$ of formulas of $\mathcal{T}_A$, suppose that $\overset{+}{A} \models \exists x(\Phi_1 \wedge \ldots \wedge \Phi_n)$ for every n. Then there exists b such that*

$$\overset{+}{A} \models \Phi_i(b, a^{i,1}, \ldots, a^{i,k})$$

for all $i \in \mathbb{N}$.

We are ready now for the Weak Transfer Principle:

Theorem 7.36. **(The Weak Transfer Principle)** *For all Φ in $\mathcal{T}_A$, Φ is true in A iff Φ is true in ${}^{+}\!A$.*

Proof of Theorem 7.36:

Proof. Proceeding by induction on the length of Φ, note first that the statement is true for simple predicates by Definition 2.5. The case for conjunctions is trivial.

Suppose Φ is $\forall x \Psi(x, a^1, a^2, \ldots, a^k)$, $a^1, a^2, \ldots, a^k$ real. The difficult direction is showing that if $A \vDash \forall x \Psi(x, a^1, a^2, \ldots, a^k)$ then it follows that ${}^{+}\!A \vDash \forall x \Psi(x, a^1, a^2, \ldots, a^k)$. Suppose we are given $b \in {}^{+}\!A$. If b is real, ${}^{+}\!A \vDash \Psi(b, a^1, a^2, \ldots, a^k)$ is immediate by induction. Otherwise, Lemma 7.32 tells us that ${}^{+}\!A \vDash \Psi(b, a^1, a^2, \ldots, a^k)$ iff $\not\forall n\ {}^{+}\!A \vDash \Psi(b_n, a^1_n, a^2_n, \ldots, a^k_n)$. Since $a^1, a^2, \ldots, a^k$ are real, $\Psi(b_n, a^1_n, a^2_n, \ldots, a^k_n)$ is $\Psi(b_n, a^1, a^2, \ldots, a^k)$. It is easy, then, to show by induction on b that $A \vDash \forall x \Psi(x, a^1, a^2, \ldots, a^k)$ implies ${}^{+}\!A \vDash \Psi(b, a^1, a^2, \ldots, a^k)$, completing the argument.

The case for $\exists x \Psi(x, a^1, a^2, \ldots, a^k)$ is handled similarly.

Now suppose that Φ is $\forall x(\Psi(x, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(x, a^1, a^2, \ldots, a^k))$ with ${}^{+}\!A \vDash \exists x \Psi(x, a^1, a^2, \ldots, a^k)$. Suppose first that

$$ {}^{+}\!A \vDash \forall x(\Psi(x, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(x, a^1, a^2, \ldots, a^k)) $$

and we are given b real. Then ${}^{+}\!A \vDash \Psi(b, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(b, a^1, a^2, \ldots, a^k)$. We can't apply the induction hypothesis to

$$ \Psi(b, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(b, a^1, a^2, \ldots, a^k) $$

since this is not in ${}^{+}\mathcal{T}_A$, but applying it to $\Psi(b, a^1, a^2, \ldots, a^k)$ and also to $\Theta(b, a^1, a^2, \ldots, a^k)$, we have

$$ {}^{+}\!A \vDash \Psi(b, a^1, a^2, \ldots, a^k) \text{ iff } A \vDash \Psi(b, a^1, a^2, \ldots, a^k) $$

and similarly for Θ. Then Definition 2.5 gives us that

$$ A \vDash \Psi(b, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(b, a^1, a^2, \ldots, a^k). $$

Finally, suppose $A \vDash \forall x(\Psi(x, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(x, a^1, a^2, \ldots, a^k))$. We will prove that for all $b \in {}^{+}\!A$, ${}^{+}\!A \vDash \Psi(b, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(b, a^1, a^2, \ldots, a^k)$. We proceed by induction on the complexity of b. If b is real,

$$ A \vDash \Psi(b, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(b, a^1, a^2, \ldots, a^k) $$

and as before, the induction hypothesis applied to Φ and Definition 2.5 give us that ${}^{+}\!A \models \Psi(b, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(b, a^1, a^2, \ldots, a^k)$. For non-real b, suppose that ${}^{+}\!A \models \Psi(b, a^1, a^2, \ldots, a^k)$ and ${}^{+}\!A \not\models \Theta(b, a^1, a^2, \ldots, a^k)$, that is, it is false that ${}^{+}\!A \models \Theta(b, a^1, a^2, \ldots, a^k)$. Since $\Psi \in \mathcal{T}_A$, we can apply Lemma 7.32 to get $\not{V}_n\, {}^{+}\!A \models \Psi(b_n, a^1, a^2, \ldots, a^k)$. From ${}^{+}\!A \not\models \Theta(b, a^1, a^2, \ldots, a^k)$, we have that for infinitely many n, ${}^{+}\!A \not\models \Theta(b_n, a^1, a^2, \ldots, a^k)$ hence for infinitely many n, ${}^{+}\!A \not\models \Psi(b_n, a^1, a^2, \ldots, a^k) \Rightarrow \Theta(b_n, a^1, a^2, \ldots, a^k)$. This contradicts the induction hypothesis on the complexity of b, and we are done. ■

Prop. 2.6 is a special case of Theorem 7.36, since $\forall x \forall y \forall z((x = y \wedge y = z) \Rightarrow x = z)$ is in $\mathcal{T}_A$. A more interesting application is continuity.

Definition 7.37. A real function f is continuous at a real number r iff whenever $a \approx r$, $f(a) \approx f(r)$.

Proposition 7.38. *The non-nonstandard definition of continuity is equivalent to the standard definition.*

Proof. The proof is essentially the same as proof of the corresponding proposition of nonstandard analysis. Suppose that f is continuous by the standard definition and $a \approx r$. To show that $f(a) \approx f(r)$, let $\epsilon > 0$ be arbitrary. By hypothesis, there is a δ such that $\forall x(|x - r| < \delta \Rightarrow |f(x) - f(r)| < \epsilon)$. This statement is in $\mathcal{T}_{\mathbb{R}}$, so by Theorem 7.36, ${}^{+}\mathbb{R} \models \forall x(|x - r| < \delta \Rightarrow |f(x) - f(r)| < \epsilon)$, hence $|f(a) - f(r)| < \epsilon$.

Next, suppose that f is continuous by the non-nonstandard definition and let ϵ be an arbitrary positive real. Let h be any infinitesimal. By hypothesis, ${}^{+}\mathbb{R} \models \forall x(|x - r| < h \Rightarrow |f(x) - f(r)| < \epsilon)$, hence ${}^{+}\mathbb{R} \models \exists \delta(\delta > 0 \wedge \forall x(|x - r| < \delta \Rightarrow |f(x) - f(r)| < \epsilon))$. Again, this statement is in $\mathcal{T}_{\mathbb{R}}$, so by Theorem 7.36, this is true in $\mathbb{R}$. ■

The blemish in the definitions of $\mathcal{T}_A$ and ${}^{+}\mathcal{T}_A$, that when $\forall x(\Phi(x) \Rightarrow \Psi(x))$ is included, we must have ${}^{+}\!A \models \exists x \Phi(x)$, is unfortunately necessary. For Lemma 7.32, we have, for example, that ${}^{+}\mathbb{R} \models \forall x(x^2 = a \Rightarrow x = 17)$ for $a = 1, -1, 1, -1, \ldots$ (for no x does $x^2 = a$), but $\not{V}_n\, {}^{+}\mathbb{R} \models \forall x(x^2 = a_n \Rightarrow x = 17)$ fails since $a_n = 1$ infinitely often and, ${}^{+}\mathbb{R} \not\models \forall x(x^2 = 1 \Rightarrow x = 17)$. For Theorem 7.36, we have that $\forall y(\forall x(x^2 = y \Rightarrow x = 17) \Rightarrow y < 0)$ is true in $\mathbb{R}$ but not in ${}^{+}\mathbb{R}$ (if $y = 1, -1, 1, -1, \ldots$, then as before, ${}^{+}\mathbb{R} \models \forall x(x^2 = y \Rightarrow x = 17)$, but ${}^{+}\mathbb{R} \not\models y < 0$).

8. Graph Theory

We take a moment in this section to apply the weak transfer principle to finite and infinite graphs. For our purposes we will use the following terms:

DEFINITION 8.39. A *graph* G is a set of two-element sets. The *vertices* of G, V_G, is $\bigcup G$, the union of the two-elements sets in G. If $\{a, b\} \in G$, then we say a and b are *adjacent*.

A *subgraph* of a graph G is a subset of G.

A *q-clique* is a graph consisting of all possible pairs from a set with q elements. We will sometimes identify a q-clique C with V_C, the set of its vertices.

Our Transfer and Łos Theorems allow us to move back forth between finite and infinite graphs. Here is one direction:

PROPOSITION 8.40. *If every finite subgraph of a countable graph G can be k-colored, then G can be k-colored.*

PROOF. Let A include the vertices and edges of G, k colors, and all possible assignments of k colors to the vertices of G. This will enable us to say easily in $\mathcal{T}_A$ that a given assignment of colors to vertices is a k-coloring.

Let g_n be the subgraph of G consisting of the first n vertices of G and all the attendant edges. Let $K(y, c)$ denote: "y is a graph and c is a k coloring of the vertices of y." Let $Co(c, i, j)$ denote: "c assigns vertex i the color j."

Then the sentence: $\exists x K(g, x)$ is in $\overset{+}{\mathcal{T}}_A$ and $\exists x K(g_n, x)$ is true in A for every n, so $\exists x K(g, x)$ is true in $\overset{+}{A}$. Let $c \in \overset{+}{A}$ be such that $\overset{+}{A} \models K(g, c)$. Now use Prop. 3.12 to construct c', a undersequence of c, such that all the sentences: $Co(c', i, j)$ are decided by $\overset{+}{A}$. This defines a k-coloring of G. ■

In the other direction, we can prove the finite Ramsey theorem,

$$\forall q \forall k \forall l \exists m \left[m \longrightarrow (q)^k_l \right],$$

from the infinite Ramsey theorem:

$$\forall k \forall l \left[\omega \longrightarrow (\omega)^k_l \right].$$

To see how this goes, we show this for $k = 2 = l$, that is, we show that given q, there is an m such that every graph with vertices $\{0, 1, 2, \dots m-1\}$ contains or is disjoint from a q-clique.

PROOF. Let $Gr(g,x)$ denote: "g is a graph with vertex set x". Let $Cl^q(c,g)$ denote: "c is a clique with at least q members that is either disjoint from g or contained in g." Let $Adj^{i,j}(g)$ denote: "i and j are adjacent in g." Let A be large enough so that one can formalize in $\mathcal{T}_A$ the relations Gr, CL^q, and $Adj^{i,j}$ for all $q, i, j \in \mathbb{N}$. Let d be the sequence $d_n = \{0, 1, 2, 3, \ldots, n-1\}$.

We will show: $\exists m\ A \models \forall x(Gr(x, d_m) \Rightarrow \exists y Cl^q(y, x))$.

By the Weak Transfer Principle, it is enough to show that $\exists m\ \overset{+}{A} \models \forall x(Gr(x, d_m) \Rightarrow \exists y Cl^q(y, x))$. Suppose, however, this is false, that is, $\overset{+}{A} \not\models \forall x(Gr(x, d_m) \Rightarrow \exists y Cl^q(y, x))$ for all m. Then by Prop. 7.32, $\overset{+}{A} \not\models \forall x(Gr(x, d) \Rightarrow \exists y Cl^q(y, x))$. So for some $g \in \overset{+}{A}$, $\overset{+}{A} \models Gr(g, d)$ and $\overset{+}{A} \not\models \exists y Cl^q(y, g))$.

Using Prop. 7.33, choose g', d', undersequences jointly of g, d, with $\overset{+}{A} \models \neg \exists y Cl^q(y, g'))$.

Now, by Prop. 7.34, we can find g'', d'' be undersequences jointly of g', d' such that $\overset{+}{A}$ decides all statements $Adj^{i,j}(g'')$. This defines a graph on $\mathbb{N}$. By the infinite Ramsey's theorem, there is an infinite clique c which is either disjoint or contained in G. Let $c_n = c \cap d_n$. Then $\overset{+}{A} \models Cl^q(c''_n, g''_n)$ for sufficiently large n. Thus $\overset{+}{A} \models Cl^q(c'', g'')$, and so $\overset{+}{A} \models \exists y Cl^q(y, g'')$. But from $\overset{+}{A} \models \neg\exists y Cl^q(y, g)$ we have $\overset{+}{A} \models \neg\exists y Cl^q(y, g'')$ by Prop. 3.10, a contradiction. ■

The same ideas can be used to prove the Paris-Harrington extension of the finite Ramsey theorem from the infinite Ramsey theorem. The corresponding special case can be stated as: given q, there is an m such that every graph with vertices $\{0, 1, 2, \ldots m-1\}$ contains or is disjoint from a clique c whose size is at least as large as q and at least as large as the least element of c. The significance of this "infinite" proof is that Paris and Harrington show in [12] that this proposition is unprovable in the finiteness of Peano Arithmetic.

9. Lebesgue Measure

Lebesgue measure has a fairly simple non-nonstandard representation. The motivation comes from one of Littlewood's principles[1], specifically, that every measurable set is "almost" the union of intervals. We will use $A = \mathbb{R} \cup \mathcal{P}(\mathbb{R})$ as our underlying set.

[1] [9], as quoted in [13].

Definition 9.41. A set $S \in {}^{+}\mathcal{P}(\mathbb{R})$ is a *spatter* if either S is a finite disjoint union of real, bounded intervals or if S is real-definite and $\forall n\,(S_n$ is a spatter). The *length* of S, $\ell en(S)$ is the sum of the lengths of the disjoint intervals of S if $S \in \mathcal{P}(\mathbb{R})$, and $\{\ell en(S_n)\}_n$ otherwise. The *weight* of S, $wt(S)$, is $\left\{\ell en\left(S \overset{+}{\cap} [-n,n]\right)\right\}_n$

We want to measure sets of reals in such a way that a spatter such as $S = \{[n, n+1]\}_n$, which has no real members has measure 0. This is the reason we use weight ($wt(S) = 0$) instead of length ($\ell en(S) = 1$).

The set a, where $a_n = \begin{cases} (0,1) & \text{if } n \text{ is odd} \\ (1,2) & \text{if } n \text{ is even} \end{cases}$ also has no real members. a would have both length and weight equal to 2 if we allowed it to be a spatter. This is the reason we restrict spatters to real-definite sets (a is not real-definite since both $.5 \overset{+}{\in} a$ and $.5 \overset{+}{\notin} a$ are false).

Note the following easily verifiable facts:

- For a spatter S, $Re(S) = \bigcup_{k<\infty} \bigcap_{n \geq k} Re(S_n)$,
- For every spatter S, $Re(S)$ is a Borel set.

To measure a set, we sandwich it between two spatters. We use a weak version of containment.

Definition 9.42. $A \subseteq_{\mathbb{R}} B$ iff every real in A is in B.

Definition 9.43. X is *measurable* if there are spatters S, T, such that $S \subseteq_{\mathbb{R}} X \subseteq_{\mathbb{R}} T$, with $wt(S) \approx wt(T)$.

The following proposition is key:

Proposition 9.44. *For any spatter S, $wt(S) \approx \lambda(Re(S))$, where $\lambda(X)$ is the Lebesgue measure of X.*

Proof. From the definition of weight and the additivity of λ, it is enough to show that for any spatter contained inside a finite interval, $\ell en(S) \approx \lambda(Re(S))$. We show this by induction.

If S is real, this is immediate. Otherwise, we have that

$$\lambda(Re(S)) = \lambda\left(\bigcup_{k<\infty} \bigcap_{n \geq k} Re(S_n)\right) = \lim_{k\to\infty} \lambda\left(\bigcap_{n\geq k} Re(S_n)\right)$$

using the completeness of λ.

We claim that $\lambda(Re(S)) \overset{+}{\leq} \ell en(S) + \epsilon$ for all real $\epsilon > 0$. Suppose otherwise, that $\lambda(Re(S)) \leq \ell en(S) + \epsilon$ is false. Then for infinitely many n,

$$\lambda(Re(S)) \leq \ell en(S_n) + \epsilon$$

is false. Using the induction hypothesis ($\forall n\, \ell en(S_n) \approx \lambda(Re(S_n))$), it follows that for infinitely many n

$$\lambda(Re(S)) \leq \lambda(Re(S_n)) + \frac{\epsilon}{2}$$

is false. Since these are all real numbers we can then say that

$$\lambda(Re(S)) > \lambda(Re(S_n)) + \frac{\epsilon}{2}$$

is *true* for infinitely many n. Thus, for all k,

$$\lambda(Re(S)) > \lambda\left(\bigcap_{n\geq k} Re(S_n)\right) + \frac{\epsilon}{2}.$$

Consequently $\lambda(Re(S)) > \lim_{k\to\infty} \lambda(\bigcap_{n\geq k} Re(S_n))$, a contradiction.

Since $Re(S) = \bigcap_{k<\infty} \bigcup_{n>k} Re(S_n)$, we have $\lambda(Re(S)) = \lim_{k\to\infty} \lambda(\bigcup_{n>k} Re(S_n))$ by boundedness and the completeness of λ. In the same manner as above we can show $\lambda(Re(S)) \overset{+}{>} \ell en(S) - \epsilon$ and the result follows. ■

Definition 9.45. If a set of reals X is measurable, we define its measure, $M(X)$, to be $r \in \mathbb{R}$ if $r \approx wt(S) \approx wt(T)$, where S and T are spatters such that $S \subseteq_{\mathbb{R}} X \subseteq_{\mathbb{R}} T$. If no such r exists, we set $M(X) = \infty$.

To complete our work, we need the following lemma:

Lemma 9.46. *B is Borel if and only if $B = Re(S)$ for some spatter S.*

Proof. It is sufficient to show this for bounded Borel sets, since if there are spatters $\{S_n\}$ such that $Re(S_n) = B \cap [-n, n]$, then $S = \{S_n\}_n$ is a spatter (it's real-definite) and $Re(S) = B$.

The collection of Borel subsets of $[0, 1]$, say, is the closure of the set of finite subintervals of $[0, 1]$ under the operations of set subtraction and countable union. The operation of set subtraction is covered by Prop. 4.20. For countable unions, if we have spatters $\{S_n\}$ such that $Re(S_n) \subseteq Re(S_{n+1})$ for all n, then $S = \{S_n\}_n$ is a spatter and $Re(S) = \bigcup_{n<\infty} Re(S_n)$. ■

THEOREM 9.47. *A set of reals is measurable if and only if it is Lebesgue-measurable, and* $M = \lambda$.

PROOF. If S and T are spatters such that $S \subseteq_{\mathbb{R}} X \subseteq_{\mathbb{R}} T$, and $r \approx wt(S) \approx wt(T)$, then X differs from a Borel set, $Re(T)$, by a subset, $Re(T) \setminus X$ of a Lebesgue measure-zero set, $Re(T) \setminus Re(S)$, hence is Lebesgue-measurable. Since $M(X) \approx \lambda(Re(S)) \leq \lambda(X) \leq \lambda(Re(T)) \approx M(X)$, $M(X) = \lambda(X)$.

For the other direction if X is Lebesgue-measurable, then we can find an F_σ, A and a G_δ, B, with $A \subseteq X \subseteq B$ and $\lambda(A) = \lambda(B)$. An appeal to Lemma 9.46 completes the proof. ■

10. Loeb-like Measures

In a series of papers that began appearing almost thirty years ago, Peter Loeb developed a mechanism for defining real measures over real spaces using nonstandard analysis. In each instance, the measures were constructed from a real-valued measure, now known as a *Loeb* measure, on a set that included both standard and nonstandard entities. The "Loeb-like" measure we present here serves a similar role. For a good reference, see Hurd and Loeb, [5]. For another, see Goldblatt, [2]; our development here is patterned after his treatment of Loeb measure.

Our object will be to construct a measure on a given set $\mathcal{X}$. Our underlying set, therefore, will be $A = \mathcal{X} \cup \mathcal{P}(\mathcal{X})$.

DEFINITION 10.48. A member a of ${}^{+}\!A$ is *sortafinite* iff either $a \in A$ and a is finite, or else $\not{V} n\, a_n$ is sortafinite. For any sortafinite a, ${}^{+}\|a\|$ is the ${}^{+}$cardinality of a (the extension of the standard cardinality function).

DEFINITION 10.49. For $x \in {}^{+}\!A$, sortafinite, $x \stackrel{+}{\subseteq} \mathcal{X}$, let $\mathfrak{A}_x = \{y \in {}^{+}\!A : y \stackrel{+}{\subseteq} x\}$. For $d \stackrel{+}{\in} {}^{+}\mathbb{R}$, define $\mu^{x,d}$ on $\mathfrak{A}_x$ by: $\mu^{x,d}(a) = d \cdot \|a\|$.

When x and d are understood, we will write only $\mathfrak{A}$ and μ. The following example illustrates the definition. Let $x_n = \{1, 2, \ldots, n\}$ and $d_n = \frac{1}{n}$. We have $\mu^{x,d}(x) \stackrel{+}{=} 1$. For $y_n = \{1, 2, \ldots, n-1\}$, $\mu^{x,d}(y) \stackrel{+}{=} d\|y\| \stackrel{+}{=} \left\{\frac{n-1}{n}\right\}_n \approx 1$. For $y_n = \{1, 3, 5, \ldots, 2\lfloor\frac{n-1}{2}\rfloor + 1\}$, $\mu^{x,d}(y) \approx .5$.

Note that x, while sortafinite, can be infinite. In the example above x contains, as ${}^{+}$members, every natural number.

PROPOSITION 10.50. *The measure* μ *is a finitely additive, non-negative, monotone, sortareal-valued function on* $\mathfrak{A}$.

By "finitely additive," we mean that if $a \overset{+}{\cap} b = \emptyset$, then $\mu(a) + \mu(b) = \mu(a \overset{+}{\cup} b)$. The proposition follows easily by induction.

For particular choices of x and d, the value of μ is always finite. Thus we can usually generate a real-valued measure on $\mathfrak{A}$.

Definition 10.51. For $a \in \mathfrak{A}$, let ${}^{\square}\mu(a) = {}^{\square}(\mu(a))$.

Using ${}^{\square}\mu$, we now define a measure on a more general set. For the rest of this paper, let $\mathfrak{B}$ be the algebra of sets B satisfying the property: $y \in B \in \mathfrak{B} \Rightarrow y \overset{+}{\in} x$.

Our goal is still to measure subsets of $\mathcal{X}$. $\mathfrak{B}$ may contain only some members of $\mathcal{P}(\mathcal{X})$. Nonetheless, our next step is to measure the subsets of $\mathfrak{B}$. Our measure is constructed from ${}^{\square}\mu$ in a manner resembling the Carathéodory "outer measure." We define a hybrid subset relation:

Notation For $B \in \mathfrak{B}$ and $b \in \mathfrak{A}$, we write $B \subseteq_{\overset{+}{\in}} b$ if $y \in B$ implies $y \overset{+}{\in} b$. We will also write $B \subseteq_{\overset{+}{\in}} \underset{n\in\mathbb{N}}{\cup} b_n$ if $y \in B$ implies $y \overset{+}{\in} b_n$ for some n.

Definition 10.52. For all $B \in \mathfrak{B}$, we define:

$$\mu_{LL}^{x,d}(B) = \inf \left\{ \sum_{n\in\mathbb{N}} {}^{\square}\mu(b_n) | B \subseteq_{\overset{+}{\in}} \underset{n\in\mathbb{N}}{\cup} b_n \text{ where all } b_n \in \mathfrak{A} \right\}.$$

The "LL" is for "Loeb-like." As before, we will write only μ_{LL} when there is no confusion.

Proposition 10.53.

(1) *For all* $B, C \in \mathfrak{B}$, $B \subseteq C$ *implies* $\mu_{LL}(B) \leq \mu_{LL}(C)$,

(2) $\mu_{LL}(B \cup C) \leq \mu_{LL}(B) + \mu_{LL}(C)$, *and*

(3) *For all sequences* $\{B^k\}_{k\in\mathbb{N}} \subseteq \mathfrak{B}$, $\mu_{LL}\left(\underset{k\in\mathbb{N}}{\cup} B^k\right) \leq \sum_{k\in\mathbb{N}} \mu_{LL}(B^k)$.

Proof. (1) follows from the easily verified fact that $B \subseteq C \subseteq_{\overset{+}{\in}} \underset{n\in\mathbb{N}}{\cup} c_n$ implies $B \subseteq_{\overset{+}{\in}} \underset{n\in\mathbb{N}_n}{c}$.

For (2), let ϵ be a positive real and choose $b, c \in \mathfrak{A}$ with $B \subseteq_{\overset{+}{\in}} \underset{n\in\mathbb{N}}{\cup} b_n$, $C \subseteq_{\overset{+}{\in}} \underset{n\in\mathbb{N}}{\cup} c_n$ and $\sum_{n\in\mathbb{N}} {}^{\square}\mu(b_n) < \mu_{LL}(B) + \epsilon$, and $\sum_{n\in\mathbb{N}} {}^{\square}\mu(c_n) < \mu_{LL}(C) + \epsilon$.

Let $d \in \mathfrak{A}$ be the sequence: $b_0, c_0, b_1, c_1, \ldots$. It is clear that $B \cup C \subseteq_{\in}^{+} \bigcup_{n\in\mathbb{N}} d_n$ and so

$$\begin{aligned} \mu_{LL}(B \cup C) &\leq \sum_{n\in\mathbb{N}} {}^{\square}\mu(d_n) \\ &= \sum_{n\in\mathbb{N}} {}^{\square}\mu(b_n) + {}^{\square}\mu(c_n) \\ &= \sum_{n\in\mathbb{N}} {}^{\square}\mu(b_n) + \sum_{n\in\mathbb{N}} {}^{\square}\mu(c_n) \\ &< \mu_{LL}(B) + \mu_{LL}(C) + 2\epsilon. \end{aligned}$$

As this is true for any ϵ, the result follows.

The proof of (3) follows that of (2). Choosing $b^k \in \mathfrak{A}$ for each k such that $B^k \subseteq_{\in}^{+} \bigcup_{n\in\mathbb{N}} (b^k)_n$ and $\sum_{n\in\mathbb{N}} {}^{\square}\mu((b^k)_n) < \mu_{LL}(B^k) + \epsilon \cdot 2^{-k}$, define a sequence d such that $\{d_n\}_n$ consists of all $\{(b^k)_n\}_{k,n}$ so $\bigcup_{k\in\mathbb{N}} B^k \subseteq_{\in}^{+} \bigcup_{n\in\mathbb{N}} d_n$. Then

$$\mu_{LL}\left(\bigcup_{k\in\mathbb{N}} B^k\right) \leq \sum_{n\in\mathbb{N}} {}^{\square}\mu(d_n) = \sum_{k\in\mathbb{N}} \sum_{n\in\mathbb{N}} {}^{\square}\mu((b^k)_n) = \left(\sum_{k\in\mathbb{N}} \mu_{LL}(B^k)\right) + 2\epsilon.$$

■

Definition 10.54. A set $B \in \mathfrak{B}$ is *measurable* iff for all $E \in \mathfrak{B}$,

$$\mu_{LL}(E) = \mu_{LL}(E \cap B) + \mu_{LL}(E \setminus B).$$

Let $\mathfrak{M}^{x,d}$ be the collection of μ_{LL}-measurable sets.

Proposition 10.55. *$\mathfrak{M}^{x,d}$ is a σ-algebra. μ_{LL} is a σ-additive, extended real-valued measure. Furthermore, all null sets are in $\mathfrak{M}^{x,d}$.*

The proof is identical to the standard proof which uses only the facts about μ_{LL} established above (see, for example, Royden's *Real Analysis* [13], pp. 253–259).

Open Question: In [2], an approximability theory is developed for Loeb measures. It is open whether Loeb-like measures have a similar theory. Is it true, for example, that B has finite μ_{LL} measure iff for all $\epsilon \in \mathbb{R}^+$ there are $C, D \in \mathfrak{B}$ such that $C \subseteq_{\in}^{+} B \subseteq_{\in}^{+} D$ and $\mu_{LL}(D \setminus C) < \epsilon$?

11. Lebesgue Measure via a Loeb-Like Measure

Let the underlying set A include $\mathbb{R}$ and $\mathcal{P}(\mathbb{R})$. Let x be the sequence: $\left\{\{\frac{i}{n} : |i| < n^2\}\right\}_n$ and d: $\left\{\frac{1}{n}\right\}_n$. To measure subsets of $\mathbb{R}$, we lift $\mu_{LL}^{x,d}$ using the ordinary part function.

DEFINITION 11.56. For any $Y \subseteq \mathbb{R}$, let $Y^{\approx} = \{y \overset{+}{\in} x : \exists a \in Y, a \approx y\}$. Note that $Y^{\approx} \in \mathfrak{B}$. Let $m(Y) = \mu_{LL}(Y^{\approx})$. We define Y to be *m-measurable* iff $Y^{\approx}$ is measurable.

PROPOSITION 11.57. *The collection of m-measurable sets is a σ-algebra of subsets of $\mathbb{R}$. m is a σ-additive, extended real-valued measure on the collection of m-measurable sets. Every subset of a m-measurable set of measure 0 is again an m-measurable set.*

The proof is straight-forward, based on Prop. 10.55. Note that we easily have $(Y \cup Z)^{\approx} = Y^{\approx} \cup Z^{\approx}$ and $(Y \setminus Z)^{\approx} = Y^{\approx} \setminus Z^{\approx}$ (all real operations, not $^{+}$-operations).

PROPOSITION 11.58. *All Lebesgue-measurable sets are m-measurable. m and λ agree on the Lebesgue-measurable sets.*

PROOF. In view of Prop. 11.57, it is sufficient to show that $m([c,d]) = d-c$ for c, d real, and that $[c,d]$ is m-measurable.

LEMMA 11.59. *For c, d real,*

(1) $^{\square}\mu([c,d]) = {}^{\square}\mu([c,d] \overset{+}{\cap} x) = d - c$

(2) *If* $[c,d]^{\approx} \subseteq_{\overset{+}{\in}} a \in \mathfrak{A}$, *then* $[c,d] \overset{+}{\cap} x \subseteq_{\overset{+}{\in}} a$.

(3) *If* $[c,d]^{\approx} \subseteq_{\overset{+}{\in}} \bigcup_{n\in\mathbb{N}} a_n$, $a \overset{+}{\in} \mathfrak{A}$, *then* $[c,d]^{\approx} \subseteq_{\overset{+}{\in}} \bigcup_{m\leq n} a_m$ *for some* n.

PROOF. (1) If n is greater than $|c|$ and $|d|$ and $d-c$, then $\|[c,d] \overset{+}{\cap} x_n\|$ is either $\lfloor n(d-c) \rfloor$ or $\lfloor n(d-c) \rfloor + 1$, so $\mu([c,d] \overset{+}{\cap} x) = \left\{ \frac{1}{n} \|[c,d] \cap x_n\| \right\}_n \approx d-c$ hence, $^{\square}\mu([c,d] \overset{+}{\cap} x) = d - c$.

(2) If a is real, the hypothesis cannot be satisfied. A real a would have to be finite and $[c,d]^{\approx} \subseteq_{\overset{+}{\in}} a$ would be false.

Proceeding inductively, suppose that for infinitely many n, $[c,d] \overset{+}{\cap} x_n \subseteq_{\overset{+}{\in}} a_n$ is false. For each such n, choose $z_n \in [c,d] \cap x_n$ such that $z_n \overset{+}{\in} a_n$ is false. We can do this by Prop. 4.17, since $[c,d] \cap x_n$ is non-empty. Let $\{z_{n_k}\}_k$ be a undersequence of $\{z_n\}_n$ and $r \in [c,d]$ real such that $\{z_{n_k}\}_k \approx r$. For each k, let $w_{n_k} = z_{n_k}$ and choose values for all other w_i in $[c,d] \cap x_i$ so that $w \approx r$. Then we have $w \in [c,d]^{\approx}$, but $w \overset{+}{\in} a$ is false, a contradiction.

(3) Suppose this is not the case. Then for each n, we can choose $z_n \in [c,d]^{\approx}$ such that $\neg\left[z_n \overset{+}{\in} \bigcup_{m\leq n} a_m\right]$. For each n find a real $r_n \in [c,d]$ with $z_n \approx r_n$. Let $\{r_{n_k}\}_k$ be a undersequence infinitely close to $s \in [c,d]$ with the additional restriction that $|r_{n_k} - s| < 2^{-k}$. We will form $w \in [c,d]^{\approx}$ such that $\neg\left[w \overset{+}{\in} \bigcup_{n\in\mathbb{N}} a_n\right]$, contradicting $[c,d]^{\approx} \subseteq_{\overset{+}{\in}} \bigcup_{n\in\mathbb{N}} a_n$.

We begin by defining $w_{(i_k)}$ for an increasing sequence, $\{i_k\}_k$.

For each k, we have $\neg\left[(z_{n_k})_i \overset{+}{\in} \overset{+}{\cup} m \leq n_k (a_m)_i\right]$ for infinitely many i. Let each i_k be one of these with the restriction that $i_k > i_{k-1}$ and

$$\left|\left(z_{(n_k)}\right)_{(i_k)} - r_{(n_k)}\right| < 2^{-k}.$$

Set $w_{(i_k)} = \left(z_{(n_k)}\right)_{(i_k)}$, so we have $\left|w_{(i_k)} - s\right| < 2 \cdot 2^{-k}$.

When this is done, fill out the rest of the values of w in any way so that $w \approx s$ and $w \overset{+}{\in} x \overset{+}{\cap} [a,b]$. This gives us $w \in [c,d]^{\approx}$. We also have $\neg\left[w \overset{+}{\in} \bigcup_{n\in\mathbb{N}} a_n\right]$, since for any n, $n_k > n$ implies that

$$\neg\left[w_{(i_k)} \overset{+}{\in} \overset{+}{\cup} m \leq n_k (a_m)_{(i_k)}\right]$$

so

$$\neg\left[w_{(i_k)} \overset{+}{\in} (a_n)_{(i_k)}\right],$$

hence $\neg\left[w_i \overset{+}{\in} (a_n)_i\right]$ for infinitely many i, and hence $\neg\left[w \overset{+}{\in} a_n\right]$.

■

By (1) of Lemma 11.59, we now have $\mu_{LL}([c,d]^{\approx}) \leq d-c$, since $[c,d]^{\approx} \subseteq_{\overset{+}{\in}} [c-\epsilon, d+\epsilon]$ for any $\epsilon \in \mathbb{R}^+$, so $\mu_{LL}([c,d]^{\approx}) \leq {}^{\square}\mu([c-\epsilon, d+\epsilon]) = d-c+2\epsilon$. Further, we have $\mu_{LL}([c,d]^{\approx}) \geq d-c$, since if $[c,d]^{\approx} \subseteq_{\overset{+}{\in}} \bigcup_{n\in\mathbb{N}} a_n$, $a \overset{+}{\in} \mathfrak{A}$, then $[c,d]^{\approx} \subseteq_{\overset{+}{\in}} \bigcup_{m\leq n} a_m \overset{+}{\in} A$ for some n by (2) of Lemma 11.59, and then $[c,d] \overset{+}{\cap} x \subseteq_{\overset{+}{\in}} \bigcup_{m\leq n} a_m$ by (3) of Lemma 11.59, so ${}^{\square}\mu\left(\bigcup_{m\leq n} a_m\right) \geq {}^{\square}\mu([c,d] \overset{+}{\cap} x) = d-c$.

To show that $[c,d]$ is m-measurable, we must show that

$$\mu_{LL}(E) = \mu_{LL}(E \cap [c,d]^{\approx}) + \mu_{LL}(E \setminus [c,d]^{\approx})$$

for any $E \in \mathfrak{B}$. In view of Prop. 10.53, we need only show $\mu_{LL}(E) \geq \mu_{LL}(E \cap [c,d]^{\approx}) + \mu_{LL}(E \setminus [c,d]^{\approx})$.

Let $\epsilon > 0$ be a given real. If $\mu_{LL}(E)$ is infinite, we are done. Otherwise, choose $e \in \mathfrak{A}$, $E \subseteq_{\overset{+}{\in}} \bigcup_{n\in\mathbb{N}} e_n$, with $\mu_{LL}(E) + \epsilon \geq \sum_{n\in\mathbb{N}} {}^{\square}\mu(e_n)$. For each n, let

$$\begin{aligned} a_n &= e_n \overset{+}{\cap} [c - \epsilon 2^{-n}, d + \epsilon 2^{-n}], \\ b_n &= e_n \overset{+}{\setminus} (c + \epsilon 2^{-n}, d - \epsilon 2^{-n}), \text{ and} \\ d_n &= e_n \overset{+}{\cap} \left([c - \epsilon 2^{-n}, c + \epsilon 2^{-n}] \cup [d - \epsilon 2^{-n}, d + \epsilon 2^{-n}]\right). \end{aligned}$$

Since $e_n = a_n \overset{+}{\cup} (b_n \overset{+}{\setminus} d_n)$ and $a_n \overset{+}{\cap} (b_n \overset{+}{\setminus} d_n) = \emptyset$, the additivity of μ gives us $\mu(e_n) = \mu(a_n) + \mu(b_n \overset{+}{\setminus} d_n)$. We have similarly that $\mu(b_n \overset{+}{\setminus} d_n) = \mu(b_n) - \mu(d_n)$, so $\mu(e_n) + \mu(d_n) = \mu(a_n) + \mu(b_n)$. We have that $E \cap [c, d]^{\approx} \subseteq_{\overset{+}{\in}} \bigcup_{n\in\mathbb{N}} a_n$ and $E \setminus [c, d]^{\approx} \subseteq_{\overset{+}{\in}} \bigcup_{n\in\mathbb{N}} b_n$. Therefore,

$$\begin{aligned} & \mu_{LL}(E \cap [c, d]^{\approx}) + \mu_{LL}(E \setminus [c, d]^{\approx}) \\ \leq\ & \sum_{n\in\mathbb{N}} {}^{\square}\mu(a_n) + \sum_{n\in\mathbb{N}} {}^{\square}\mu(b_n) \\ \leq\ & \sum_{n\in\mathbb{N}} {}^{\square}(\mu(a_n) + \mu(b_n)) && \text{by Prop. 3.15} \\ =\ & \sum_{n\in\mathbb{N}} {}^{\square}(\mu(e_n) + \mu(d_n)) \\ \leq\ & \sum_{n\in\mathbb{N}} {}^{\square}(\mu(e_n) + 4\epsilon 2^{-n}) \\ =\ & \sum_{n\in\mathbb{N}} ({}^{\square}\mu(e_n) + 4\epsilon 2^{-n}) \\ =\ & \left(\sum_{n\in\mathbb{N}} {}^{\square}\mu(e_n) \right) + 8\epsilon \\ \leq\ & \mu_{LL}(E) + 9\epsilon. \end{aligned}$$

As ϵ is arbitrary, we have $\mu_{LL}(E \cap [c, d]^{\approx}) + \mu_{LL}(E \setminus [c, d]^{\approx} \leq \mu_{LL}(E)$ and the result is proved. ∎

A nice application of Loeb measure is the construction of a measure on the set $\mathcal{Y}$ of infinite sequences of 0s and 1s. This can be done with an appropriate Loeb-like measure.

Let x be defined by

$$x_n = \{\text{all infinite sequences } s \text{ such that } s_k = 0 \text{ for all } k \geq n\}.$$

Let d be defined by $d_n = 2^{-n}$. We have, as before, μ, ${}^{\square}\mu$, and μ_{LL}. For $a \in \mathcal{Y}$ and $b \overset{+}{\in} {}^{+}\mathcal{Y}$, define $a \approx b$ iff $a(k) \overset{+}{=} b(k)$ for all natural numbers k. With this we can define $Z^{\approx}$ for all $Z \subseteq \mathcal{Y}$ and use this as in section 9 to define a measure on $\mathcal{Y}$. It is not difficult to show that this gives the correct measure to sets defined by initial segments (for example, the measure of the set of all sequences beginning $0, 1, 1$ is $\frac{1}{8}$.)

References

[1] Di Nasso, Mauro, "An axiomatic presentation of the nonstandard methods in mathematics," *J. Symbolic Logic*, vol 67 (2002), 315-325.

[2] Goldblatt, Robert, *Lectures on the Hyperreals*, Springer-Verlag, New York, 1998.

[3] Henle, J. M., "Non–nonstandard Analysis: *Real* Infinitesimals", *The Mathematical Intelligencer*, 21, no. 1 (1999), Springer-Verlag.

[4] Henle, J. M. and E. M. Kleinberg, *Infinitesimal Calculus*, M.I.T. Press, Cambridge, MA, 1979.

[5] Hurd, A. E. and P. A. Loeb, *An Introduction to Nonstandard Real Analysis*, Academic Press, Orlando, Florida, 1985.

[6] Hrbacek, Karel, "Realism, Nonstandard Set Theory, and Large Cardinals," *Ann. Pure and Appl. Logic*, vol. 109 (2001) 15-48.

[7] Laugwitz, D. "Anwendungen unendlich kleiner Zahlen I.", *Journal für die reine und angewandte M.*, vol. 207 (1960) 53-60.

[8] Laugwitz D. "Anwendungen unendlich kleiner Zahlen II.", *Journal für die reine und angewandte M.*, vol. 208 (1961) 22-34.

[9] Lebesgue, Henri, *Lectures on the Theory of Functions*, Oxford Press, 1944.

[10] Palmgren, Erik, "A constructive approach to nonstandard analysis,", *Annals of Pure and Applied Logic*, 73 (1995) 297-325.

[11] Palmgren, Erik, "Developments in constructive nonstandard analysis," *Bull. Symb. Logic*, vol. 4 (1998), 223-271.

[12] Paris, Jeffrey, and Leo Harrington, "A Mathematical Incompleteness in Peano Arithmetic" in *Handbook of Mathematical Logic*, Jon Barwise, ed., North-Holland Publishing Co., 1977.

[13] Royden, H. L., *Real Analysis*, second edition, The Macmillan Company, 1968.

[14] Schuster, Peter M., *Constructive Non-nonstandard Reals Without Choice*, preprint.

J. M. Henle
Clark Science Center
Smith College Northampton, MA USA
jhenle@math.smith.edu