



Relative set theory: Some external issues

KAREL HRBACEK

Abstract: The paper establishes equivalence of two axiomatizations of the relative set theory **GRIST**, examines dependence of nonstandard concepts on the choice of the level, and proves that several nonstandard definitions of Lebesgue measure are equivalent. The last section extends **GRIST** to a theory of external sets.

2000 Mathematics Subject Classification [26E35](#) (primary); [03E70](#), [03H05](#), [28E05](#) (secondary)

Keywords: nonstandard analysis, nonstandard set theory, IST, RIST, level, relative infinitesimal, external set, S-continuity, Lebesgue measure

Introduction

Relative set theory is an axiomatic framework for nonstandard analysis distinguished by many “levels of standardness.” It was proposed by Péraire in [22] and several subsequent papers and given the acronym **RIST**; for early mathematical applications see for example [23]. The present author extended **RIST** further, to **FRIST** [14, 15] and then to **GRIST** [16], a theory that is, in a technical sense, complete over **ZFC**. In 2004, O’Donovan and the author realized that a simple fragment of relative set theory might serve as a vehicle for presentation of nonstandard methods at a very elementary level. In a joint paper with Lessmann [17] we propose such an elementary theory of “relative analysis” and show how to develop some basic calculus concepts in it. Calculus courses based on relative analysis have been implemented in two high schools in Geneva; for a report on the pedagogical aspects of the project see O’Donovan [21].

The purpose of this article is to bridge the gap between the technical papers [14, 16], concerned with metamathematics of **GRIST**, and the mathematical and pedagogical applications of **GRIST**. The common theme is the employment of external sets and classes. **GRIST**, like Nelson’s **IST**, is an *internal set theory* in the sense that objects of the theory are what nonstandard analysts think of as internal sets; we call them just sets and identify them with the “usual” sets of traditional mathematics. However, in relative set theory (as in **IST**) it is possible, and often useful, to describe and put to use

also collections that are not internal. In the attempts to present a fragment of **GRIST** as simply as possible it was found that an axiomatization in terms of certain external classes (“levels of standardness”) is more intuitive. Section 1 of this paper states the axioms of **GRIST** both in the original \in - \sqsubseteq -language and in the language of levels, and provides a translation between the two. Thus it formally establishes the consistency of the elementary system employed in [17].

External concepts, such as infinitesimals, S-continuity, S-integrability and many others, are the tool characteristic of every version of nonstandard analysis. In relative set theory with its many “levels of standardness,” such concepts are always relative to some level. In Section 2 we showcase the techniques available in **GRIST** for the study of dependence of external concepts on the choice of level. We focus on S-continuity and prove that, for a fixed function, it can change only finitely many times as the level varies.

Another feature of **GRIST** which is not present in the usual treatments of nonstandard analysis is the possibility to define concepts in ways that involve quantification over levels. One example is the notion of superfine partition, suggested by the problem of integrating an arbitrary derivative. In the Appendix to Section 3 we show that superfine partitions in the strong sense do not exist. But there is a weaker sense of relative standardness, proposed by Benninghofen and Richter [6] and Gordon [9, 10], which can be used to define superfine partitions and develop a theory of integration. As the focus of this paper is on the methods of relative set theory, we show in Section 3 only that the usual definition of Lebesgue measure is equivalent to two definitions using superfine partitions, and also to one based on the well-known Loeb measure approach; a more systematic development of integration in **GRIST** can be found in [18].

In order to enable construction of Loeb measures and many other important nonstandard concepts, **GRIST** has to be extended to a substantial formal theory of external sets. In Section 4 we propose such a theory, in four stages of increasing power, and prove its relative consistency with **ZFC**. The paper concludes with a discussion of why a more comprehensive theory would be desirable.

1 Levels

This section establishes mutual interpretability of two formulations of relative set theory. Our treatment of logic follows Enderton [8].

In Péraire [22] and the author’s [14, 15, 16], relative set theory is formulated as a first-order theory in the language with equality and two binary predicate symbols, \in

and \sqsubseteq (also denoted **st**). $x \sqsubset y$ is shorthand for $(x \sqsubseteq y \wedge \neg y \sqsubseteq x)$ and $x \boxplus y$ for $(x \sqsubseteq y \wedge y \sqsubseteq x)$.

If $\mathcal{P}(x_1, \dots, x_k)$ is a formula (with all its free variables among x_1, \dots, x_k) and u is a variable, the formula $\mathcal{P}^u(x_1, \dots, x_k)$ is obtained by replacing each occurrence of \sqsubseteq by \sqsubseteq_u , defined by

$$x \sqsubseteq_u y \text{ if and only if } (x \sqsubseteq y \vee x \sqsubseteq u).$$

Below, we state the axioms of **GRIST** as given in [16]. 0 is the empty set and $\mathcal{P}^{\text{fin}}A$ is the set of all finite subsets of A .

GRIST includes the axioms of **ZFC** and the following additional axioms.

(R) Relativization

The conjunction of:

- (i) $(\forall u)(0 \sqsubseteq u \wedge u \sqsubseteq u)$;
- (ii) $(\forall u, v, w)((v \sqsubseteq u \wedge w \sqsubseteq v) \rightarrow w \sqsubseteq u)$;
- (iii) $(\forall u, v)(u \sqsubseteq v \vee v \sqsubseteq u)$;
- (iv) $(\forall u)(\exists v)(u \sqsubset v)$;
- (v) $(\forall u, v)(u \sqsubset v \rightarrow (\exists w)(u \sqsubset w \sqsubset v))$.

(T) Transfer

For all $u \sqsubseteq v$ and all $x_1, \dots, x_k \sqsubseteq u$,

$$\mathcal{P}^u(x_1, \dots, x_k) \leftrightarrow \mathcal{P}^v(x_1, \dots, x_k).$$

(S) Standardization

For all $u \sqsupset 0$ and all A, x_1, \dots, x_k , there exist $v \sqsubset u$ and $B \sqsubseteq v$ such that, for every w with $v \sqsubseteq w \sqsubset u$,

$$(\forall y \sqsubseteq w)(y \in B \leftrightarrow y \in A \wedge \mathcal{P}^w(y, x_1, \dots, x_k)).$$

(I) Idealization

For all $A \sqsubset v$ and all x_1, \dots, x_k ,

$$\begin{aligned} &(\forall a \in \mathcal{P}^{\text{fin}}A)[a \sqsubset v \rightarrow (\exists y)(\forall x \in a) \mathcal{P}^v(x, y, x_1, \dots, x_k)] \\ &\leftrightarrow (\exists y)(\forall x \in A) [x \sqsubset v \rightarrow \mathcal{P}^v(x, y, x_1, \dots, x_k)]. \end{aligned}$$

(G) Granularity

For all x_1, \dots, x_k , if $(\exists u)\mathcal{P}^u(x_1, \dots, x_k)$, then

$$(\exists u)[\mathcal{P}^u(x_1, \dots, x_k) \wedge (\forall v)(v \sqsubset u \rightarrow \neg \mathcal{P}^v(x_1, \dots, x_k))].$$

As a vehicle for elementary presentations of nonstandard methods, and the first step towards enriching it by external objects, it is convenient to formulate relative set theory as a second-order theory, or, equivalently [8], a first-order theory in a two-sorted language. In this formulation (see [17]), there are individual variables, intended to range over sets, and 1-place predicate variables (class variables) denoted by $\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{U}, \mathbf{V}, \mathbf{W}, \dots$, intended to range over certain proper classes called **levels**. Equality, both between individual variables and between class variables, is denoted by $=$. We write $x \in \mathbf{V}$ [read: x **appears at level \mathbf{V}**] instead of $\mathbf{V}x$. $\mathbf{U} \subseteq \mathbf{V}$ [read: “ \mathbf{U} is **coarser** than \mathbf{V} ,” or “ \mathbf{V} is **finer** than \mathbf{U} ”] is shorthand for $(\forall x)(x \in \mathbf{U} \rightarrow x \in \mathbf{V})$, and $\mathbf{U} \subset \mathbf{V}$ for $(\mathbf{U} \subseteq \mathbf{V} \wedge \neg \mathbf{V} \subseteq \mathbf{U})$.

We say that a formula $\mathcal{P}(x_1, \dots, x_k, \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n)$ is a **\mathbf{V} -formula** [or: a **formula about \mathbf{V}**] if all quantifiers over levels are of the form $(\forall \mathbf{U} \supseteq \mathbf{V})$ or $(\exists \mathbf{U} \supseteq \mathbf{V})$, and we indicate it by a semicolon thus: $\mathcal{P}(x_1, \dots, x_k; \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n)$.

The axioms of **GRIST** in this language are given below; the superscript \heartsuit is used to distinguish them from their first-order counterparts. \mathcal{P} is an arbitrary \mathbf{V} -formula, with at most one free variable over levels.

GRIST ^{\heartsuit} postulates **ZFC** and

(R ^{\heartsuit}) Relativization

The conjunction of:

- (o) $(\forall \mathbf{U}, \mathbf{V})[(\forall x)(x \in \mathbf{U} \leftrightarrow x \in \mathbf{V}) \rightarrow \mathbf{U} = \mathbf{V}]$;
- (i) $(\forall x)(\exists \mathbf{V})[x \in \mathbf{V} \wedge (\forall \mathbf{U})(x \in \mathbf{U} \rightarrow \mathbf{V} \subseteq \mathbf{U})]$;
- (ii) $(\forall \mathbf{V})[0 \in \mathbf{V} \wedge (\exists x \in \mathbf{V})(\forall \mathbf{U})(x \in \mathbf{U} \rightarrow \mathbf{V} \subseteq \mathbf{U})]$;
- (iii) $(\forall \mathbf{U}, \mathbf{V})(\mathbf{U} \subseteq \mathbf{V} \vee \mathbf{V} \subseteq \mathbf{U})$;
- (iv) $(\forall \mathbf{U})(\exists \mathbf{V})(\mathbf{U} \subset \mathbf{V})$;
- (v) $(\forall \mathbf{U}, \mathbf{V})(\mathbf{U} \subset \mathbf{V} \rightarrow (\exists \mathbf{W})(\mathbf{U} \subset \mathbf{W} \subset \mathbf{V}))$.

(T ^{\heartsuit}) Transfer (or Stability)

For all $\mathbf{U} \subseteq \mathbf{V}$ and all $x_1, \dots, x_k \in \mathbf{U}$,

$$\mathcal{P}(x_1, \dots, x_k; \mathbf{U}) \leftrightarrow \mathcal{P}(x_1, \dots, x_k; \mathbf{V}).$$

(S[♡]) Standardization

For all \mathbf{U} and all A, x_1, \dots, x_k , either $(\forall \mathbf{V})(\mathbf{U} \subseteq \mathbf{V})$ or there exist $\mathbf{V} \subset \mathbf{U}$ and $B \in \mathbf{V}$ such that, for every \mathbf{W} with $\mathbf{V} \subseteq \mathbf{W} \subset \mathbf{U}$,

$$(\forall y \in \mathbf{W})(y \in B \leftrightarrow y \in A \wedge \mathcal{P}(y, x_1, \dots, x_k; \mathbf{W})).$$

(I[♡]) Idealization

For all $\mathbf{U}, \mathbf{V}, A$ such that $A \in \mathbf{U} \subset \mathbf{V}$, and all x_1, \dots, x_k ,

$$\begin{aligned} &(\forall a \in \mathcal{P}^{\text{fin}} A)(\forall \mathbf{W} \subset \mathbf{V}) [a \in \mathbf{W} \rightarrow (\exists y)(\forall x \in a)\mathcal{P}(x, y, x_1, \dots, x_k; \mathbf{V})] \leftrightarrow \\ &(\exists y)(\forall x \in A)(\forall \mathbf{W} \subset \mathbf{V})[x \in \mathbf{W} \rightarrow \mathcal{P}(x, y, x_1, \dots, x_k; \mathbf{V})]. \end{aligned}$$

(G[♡]) Granularity

For all x_1, \dots, x_k , if $(\exists \mathbf{U})\mathcal{P}(x_1, \dots, x_k; \mathbf{U})$, then

$$(\exists \mathbf{U})[\mathcal{P}(x_1, \dots, x_k; \mathbf{U}) \wedge (\forall \mathbf{V})(\mathbf{V} \subset \mathbf{U} \rightarrow \neg \mathcal{P}(x_1, \dots, x_k; \mathbf{V}))].$$

At the semantic level, every structure \mathfrak{M} for the first-order language can be made into a general pre-structure \mathfrak{N} for the second-order language [8, §4.4], by letting predicate variables \mathbf{V} range over subsets of $|\mathfrak{M}|$ of the form $\{x \in |\mathfrak{M}| : x \sqsubseteq^{\mathfrak{M}} a\}$, where $a \in |\mathfrak{M}|$ (and dropping $\sqsubseteq^{\mathfrak{M}}$). Conversely, every general pre-structure \mathfrak{N} for the second-order language can be made into a structure \mathfrak{M} for the first-order language via defining $\sqsubseteq^{\mathfrak{M}}$ by: $a \sqsubseteq^{\mathfrak{M}} b$ if and only if $\mathfrak{N} \models (\forall \mathbf{V})(b \in \mathbf{V} \rightarrow a \in \mathbf{V})$ (and dropping the levels). Moreover, $\mathfrak{M} \models \mathbf{GRIST}$ if and only if $\mathfrak{N} \models \mathbf{GRIST}^{\heartsuit}$. In the following we obtain stronger syntactic results that give an algorithm for translating statements from one language into the other in such a way that theorems get translated into theorems.

To facilitate the translations, we partition the variables of **GRIST** and the individual variables of **GRIST**[♡] into two infinite classes and use $x_1, x_2, \dots, x, y, z, a, A, \dots$ for variables of the first kind and $v_1, v_2, \dots, u, v, w, \dots$ for the second kind. We fix a one-one correspondence $v \mapsto \mathbf{V}$ between the **GRIST** variables of the second kind and the class variables of **GRIST**[♡]. A **GRIST** formula $\mathcal{P}(x_1, \dots, x_k, v_1, \dots, v_n)$ is **suitable** if the variables of the second kind do not occur in the scope of \in or $=$. Formally: $x_i = x_j, x_i \in x_j, x_i \sqsubseteq x_j, x_i \sqsubseteq v_j, v_i \sqsubseteq x_j, v_i \sqsubseteq v_j$ are suitable, and applications of logical connectives and quantifiers to suitable formulas yield suitable formulas. A **GRIST**[♡] formula $\mathcal{Q}(x_1, \dots, x_k, \mathbf{V}_1, \dots, \mathbf{V}_n)$ is **suitable** if no individual variables of the second kind occur in it.

We define a syntactic translation Γ of the suitable formulas of **GRIST** language into the language of **GRIST**[♡], and a syntactic translation Δ in reverse direction. We use \doteq for metamathematical identity, and \ulcorner, \urcorner as metaparentheses.

Definition 1.1

$$\begin{array}{ll}
\Gamma x = y \doteq x = y & \Delta \Gamma x = y \doteq x = y \\
\Gamma x \in y \doteq x \in y & \Delta \Gamma x \in y \doteq x \in y \\
\Gamma x \sqsubseteq v \doteq x \in \mathbf{V} & \Delta \Gamma x \in \mathbf{V} \doteq x \sqsubseteq v \\
\Gamma x \sqsubseteq y \doteq (\forall \mathbf{V})(y \in \mathbf{V} \rightarrow x \in \mathbf{V}) & \\
\Gamma u \sqsubseteq v \doteq \mathbf{U} \subseteq \mathbf{V} & \\
\Gamma v \sqsubseteq x \doteq (\forall \mathbf{U})(x \in \mathbf{U} \rightarrow \mathbf{V} \subseteq \mathbf{U}) & \\
\Gamma \neg \mathcal{P} \doteq \neg \Gamma \mathcal{P} & \Delta \Gamma \mathbf{U} = \mathbf{V} \doteq u \boxplus v \\
\Gamma \mathcal{P}_1 \rightarrow \mathcal{P}_2 \doteq (\Gamma \mathcal{P}_1 \rightarrow \Gamma \mathcal{P}_2) & \Delta \Gamma \neg \mathcal{Q} \doteq \neg \Delta \Gamma \mathcal{Q} \\
\Gamma (\forall x) \mathcal{P} \doteq (\forall x) \Gamma \mathcal{P} & \Delta \Gamma \mathcal{Q}_1 \rightarrow \mathcal{Q}_2 \doteq (\Delta \Gamma \mathcal{Q}_1 \rightarrow \Delta \Gamma \mathcal{Q}_2) \\
\Gamma (\forall v) \mathcal{P} \doteq (\forall \mathbf{V}) \Gamma \mathcal{P} & \Delta \Gamma (\forall x) \mathcal{Q} \doteq (\forall x) \Delta \Gamma \mathcal{Q} \\
& \Delta \Gamma (\forall \mathbf{V}) \mathcal{Q} \doteq (\forall v) \Delta \Gamma \mathcal{Q}
\end{array}$$

Let \mathbf{R}_b denote the conjunction of \mathbf{R} (i) and (ii), and \mathbf{R}_b^\heartsuit the conjunction of \mathbf{R}^\heartsuit (o), (i) and (ii).

Proposition 1.2 $\mathbf{R}_b \vdash (\Delta \Gamma \mathcal{P} \leftrightarrow \mathcal{P})$ and $\mathbf{R}_b^\heartsuit \vdash (\Gamma \Delta \mathcal{Q} \leftrightarrow \mathcal{Q})$, for all suitable \mathcal{P} and \mathcal{Q} (of the appropriate languages).

Proof By induction on the complexity of formulas. We verify the cases that are not entirely trivial.

$\Delta \Gamma x \sqsubseteq y \doteq \Delta (\forall \mathbf{V})(y \in \mathbf{V} \rightarrow x \in \mathbf{V}) \doteq (\forall v)(y \sqsubseteq v \rightarrow x \sqsubseteq v)$. The last formula implies $y \sqsubseteq y \rightarrow x \sqsubseteq y$, and then $x \sqsubseteq y$, using $\mathbf{R}(i)$; the converse direction follows from $\mathbf{R}(ii)$.

$\Delta \Gamma u \sqsubseteq v \doteq \Delta \Gamma \mathbf{U} \subseteq \mathbf{V} \doteq \Delta (\forall x)(x \in \mathbf{U} \rightarrow x \in \mathbf{V}) \doteq (\forall x)(x \sqsubseteq u \rightarrow x \sqsubseteq v)$. This is equivalent to $u \sqsubseteq v$. [Let $x = u$ and use $\mathbf{R}(i)$; for the other direction use $\mathbf{R}(ii)$.]

$\Delta \Gamma v \sqsubseteq x \doteq \Delta (\forall \mathbf{U})(x \in \mathbf{U} \rightarrow \mathbf{V} \subseteq \mathbf{U})$ is equivalent to $(\forall u)(x \sqsubseteq u \rightarrow v \sqsubseteq u)$, as in the preceding case. This is further equivalent to $v \sqsubseteq x$.

$\Gamma \Delta \mathbf{U} = \mathbf{V} \doteq \Gamma u \sqsubseteq v \wedge v \sqsubseteq u \doteq (\mathbf{U} \subseteq \mathbf{V} \wedge \mathbf{V} \subseteq \mathbf{U})$, which is equivalent to $\mathbf{U} = \mathbf{V}$ by $\mathbf{R}^\heartsuit(o)$. \square

Proposition 1.3 If $\mathbf{R}_b^\heartsuit \vdash \mathcal{Q}$, then $\mathbf{R}_b \vdash \Delta \Gamma \mathcal{Q}$, for all suitable \mathcal{Q} .

Proof Wlog we can assume that all formulas occurring in the proof of \mathcal{Q} are suitable (if an individual variable of the second kind appears in the proof, replace it by some variable of the first kind that does not appear in the proof). It is trivial to verify that Δ preserves logical axioms and modus ponens. The axioms for equality between classes are translated into easy consequences of $\mathbf{R}(i, ii)$.

Δ translates $\mathbf{R}^\heartsuit(o)$ into $(\forall u, v)[(\forall x)(x \sqsubseteq u \leftrightarrow x \sqsubseteq v) \rightarrow (u \sqsubseteq v \wedge v \sqsubseteq u)]$. Given u, v , let $x \doteq u$ and use $\mathbf{R}(i, ii)$ to deduce $u \sqsubseteq v$. Similarly, letting $x \doteq v$ gives $v \sqsubseteq u$.

Next, note that $\Delta^\ulcorner \forall \mathbf{V} \subseteq \mathbf{U}^\urcorner \doteq \Delta^\ulcorner (\forall x)(x \in \mathbf{V} \rightarrow x \in \mathbf{U})^\urcorner \doteq (\forall x)(x \sqsubseteq v \rightarrow x \sqsubseteq u)$ is equivalent to $v \sqsubseteq u$, again by $\mathbf{R}(i, ii)$.

Δ translates $\mathbf{R}^\heartsuit(i)$ into an equivalent of $(\forall x)(\exists v)[x \sqsubseteq v \wedge (\forall u)(x \sqsubseteq u \rightarrow v \sqsubseteq u)]$. Let $v \doteq x$.

As for $\mathbf{R}^\heartsuit(ii)$, $\Delta^\ulcorner (\forall \mathbf{V})(0 \in \mathbf{V})^\urcorner \doteq (\forall v)(0 \sqsubseteq v)$ is true by $\mathbf{R}(i)$.

$\Delta^\ulcorner (\forall \mathbf{V})(\exists x)[x \in \mathbf{V} \wedge (\forall \mathbf{U})(x \in \mathbf{U} \rightarrow \mathbf{V} \subseteq \mathbf{U})]^\urcorner \doteq (\forall v)(\exists x)[x \sqsubseteq v \wedge (\forall u)(x \sqsubseteq u \rightarrow v \sqsubseteq u)]$. Let $x \doteq v$ and use $\mathbf{R}(i, ii)$. \square

The analogous proposition for Γ is not so immediate.

Proposition 1.4 *If $\mathbf{R}_b \vdash \mathcal{P}$, then $\mathbf{R}_b^\heartsuit \vdash \Gamma^\ulcorner \mathcal{P}^\urcorner$, for all suitable \mathcal{P} .*

The primary difficulty is that, while \mathcal{P} is assumed to be a suitable formula, some other formulas that appear in the proof of \mathcal{P} may not be suitable, and the translation Γ may be undefined for them. Renaming variables does not help here; if v occurs in \mathcal{P} and say $v \in x$ appears somewhere in the proof, replacing v with \mathbf{V} does not make sense. To circumvent this difficulty, we define another, more “literal” translation Γ_0 that does not suffer from this problem; on the other hand, Γ_0 does not satisfy Proposition 1.2. In Proposition 1.6 we establish a relationship between the two translations from which Proposition 1.4 follows.

If \mathcal{P} is any formula of **GRIST**, $\Gamma_0^\ulcorner \mathcal{P}^\urcorner$ is the formula obtained from \mathcal{P} by replacing each occurrence of $\xi \sqsubseteq \eta$ by $(\forall \mathbf{U})(\eta \in \mathbf{U} \rightarrow \xi \in \mathbf{U})$ [ξ, η are variables of either kind].

Proposition 1.5 *If $\mathbf{R}_b \vdash \mathcal{P}$, then $\mathbf{R}_b^\heartsuit \vdash \Gamma_0^\ulcorner \mathcal{P}^\urcorner$, for all \mathcal{P} .*

Proof Trivially, Γ_0 preserves logical axioms and modus ponens. $\Gamma_0^\ulcorner u \sqsubseteq u^\urcorner$ and $\Gamma_0^\ulcorner (v \sqsubseteq u \wedge w \sqsubseteq v) \rightarrow w \sqsubseteq u^\urcorner$ are logically true. $\Gamma_0^\ulcorner 0 \sqsubseteq u^\urcorner \doteq (\forall \mathbf{U})(u \in \mathbf{U} \rightarrow 0 \in \mathbf{U})$ follows from $\mathbf{R}^\heartsuit(ii)$. \square

The translation Γ_0 , unlike Γ and Δ , does not take suitable formulas into suitable formulas, and Proposition 1.2 does not hold with Γ_0 in place of Γ : Let \mathcal{Q} be the formula $x \in \mathbf{V}$; then $\Delta^\Gamma \mathcal{Q}^\neg \doteq x \in v$ and $\Gamma_0^\Gamma \Delta^\Gamma \mathcal{Q}^{\neg\neg} \doteq (\forall \mathbf{U})(v \in \mathbf{U} \rightarrow x \in \mathbf{U})$, which does not even have the same free variables as \mathcal{Q} . In general, if \mathcal{P} is a suitable formula with free variables $x_1, \dots, x_k, v_1, \dots, v_n$, then $\Gamma_0^\Gamma \mathcal{P}^\neg$ has the same free variables, while $\Gamma^\Gamma \mathcal{P}^\neg$ has free variables $x_1, \dots, x_k, \mathbf{V}_1, \dots, \mathbf{V}_n$.

We next establish the relationship between Γ and Γ_0 .

Axioms (o) and (i) of \mathbf{R}^\heartsuit imply that for every x there is a unique coarsest level \mathbf{V} such that $x \in \mathbf{V}$. In \mathbf{R}_b^\heartsuit we can thus define a function $x \mapsto \mathbf{V}(x)$ with arguments of the individual sort and values of the class sort such that

$$(\forall x)[x \in \mathbf{V}(x) \wedge (\forall \mathbf{U})(x \in \mathbf{U} \rightarrow \mathbf{V}(x) \subseteq \mathbf{U})].$$

Proposition 1.6 *Let $\mathcal{P}(x_1, \dots, x_k, v_1, \dots, v_n)$ be a suitable formula of GRIST. Then $\mathbf{R}_b^\heartsuit \vdash \Gamma_0^\Gamma \mathcal{P}^\neg(x_1, \dots, x_k, v_1, \dots, v_n) \leftrightarrow \Gamma^\Gamma \mathcal{P}^\neg(x_1, \dots, x_k, \mathbf{V}(v_1), \dots, \mathbf{V}(v_n))$.*

In particular, $\mathbf{R}_b^\heartsuit \vdash \Gamma_0^\Gamma \mathcal{P}^\neg \leftrightarrow \Gamma^\Gamma \mathcal{P}^\neg$ if \mathcal{P} is a suitable sentence of GRIST.

Proof Proceed by induction on complexity of suitable formulas. Let \bar{x} be shorthand for x_1, \dots, x_k . The interesting cases are:

$$\Gamma_0^\Gamma x \sqsubseteq y^\neg \doteq (\forall \mathbf{U})(y \in \mathbf{U} \rightarrow x \in \mathbf{U}) \text{ is equivalent to } \Gamma^\Gamma x \sqsubseteq y^\neg;$$

$$\Gamma_0^\Gamma x \sqsubseteq v^\neg \doteq (\forall \mathbf{U})(v \in \mathbf{U} \rightarrow x \in \mathbf{U}) \text{ is equivalent to } x \in \mathbf{V}(v) \text{ and } \Gamma^\Gamma x \sqsubseteq v^\neg \doteq x \in \mathbf{V};$$

$$\Gamma_0^\Gamma u \sqsubseteq v^\neg \doteq (\forall \mathbf{U})(v \in \mathbf{U} \rightarrow u \in \mathbf{U}) \text{ is equivalent to } \mathbf{V}(u) \subseteq \mathbf{V}(v), \text{ and } \Gamma^\Gamma u \sqsubseteq v^\neg \doteq \mathbf{U} \subseteq \mathbf{V};$$

$$\Gamma_0^\Gamma v \sqsubseteq x^\neg \doteq (\forall \mathbf{U})(x \in \mathbf{U} \rightarrow v \in \mathbf{U}) \text{ is equivalent to } (\forall \mathbf{U})(x \in \mathbf{U} \rightarrow \mathbf{V}(v) \subseteq \mathbf{U}), \text{ and } \Gamma^\Gamma v \sqsubseteq x^\neg \doteq (\forall \mathbf{U})(x \in \mathbf{U} \rightarrow \mathbf{V} \subseteq \mathbf{U}).$$

If \mathcal{P} is $\mathcal{P}(\bar{x}, v, v_1, \dots, v_n)$, then $\Gamma_0^\Gamma (\forall v)\mathcal{P}^\neg \doteq (\forall v)\Gamma_0^\Gamma \mathcal{P}^\neg$, $\Gamma^\Gamma (\forall v)\mathcal{P}^\neg \doteq (\forall \mathbf{V})\Gamma^\Gamma \mathcal{P}^\neg$, and $\Gamma_0^\Gamma \mathcal{P}^\neg(\bar{x}, v, v_1, \dots, v_n) \leftrightarrow \Gamma^\Gamma \mathcal{P}^\neg(\bar{x}, \mathbf{V}(v), \mathbf{V}(v_1), \dots, \mathbf{V}(v_n))$ by the inductive assumption. But $(\forall v)\Gamma_0^\Gamma \mathcal{P}^\neg(\bar{x}, v, v_1, \dots, v_n) \leftrightarrow (\forall v)\Gamma^\Gamma \mathcal{P}^\neg(\bar{x}, \mathbf{V}(v), \mathbf{V}(v_1), \dots, \mathbf{V}(v_n)) \leftrightarrow (\forall \mathbf{V})\Gamma^\Gamma \mathcal{P}^\neg(\bar{x}, \mathbf{V}, \mathbf{V}(v_1), \dots, \mathbf{V}(v_n))$, because for every \mathbf{V} there is some v such that $\mathbf{V} = \mathbf{V}(v)$, by \mathbf{R}^\heartsuit (ii). \square

Proof of Proposition 1.4 If $\mathbf{R}_b \vdash \mathcal{P}$ where $\mathcal{P} \doteq \mathcal{P}(\bar{x}, v_1, \dots, v_n)$, then $\mathbf{R}_b^\heartsuit \vdash \Gamma_0^\Gamma \mathcal{P}^\neg(\bar{x}, v_1, \dots, v_n)$ and by Proposition 1.6, $\mathbf{R}_b^\heartsuit \vdash \Gamma^\Gamma \mathcal{P}^\neg(\bar{x}, \mathbf{V}(v_1), \dots, \mathbf{V}(v_n))$, ie, $\mathbf{R}_b^\heartsuit \vdash (\forall v_1, \dots, v_n)\Gamma^\Gamma \mathcal{P}^\neg(\bar{x}, \mathbf{V}(v_1), \dots, \mathbf{V}(v_n))$. But by \mathbf{R}^\heartsuit (ii), for every \mathbf{V} there is v such that $\mathbf{V} = \mathbf{V}(v)$, so $\mathbf{R}_b^\heartsuit \vdash (\forall \mathbf{V}_1, \dots, \mathbf{V}_n)\Gamma^\Gamma \mathcal{P}^\neg(\bar{x}, \mathbf{V}_1, \dots, \mathbf{V}_n)$, ie, $\mathbf{R}_b^\heartsuit \vdash \Gamma^\Gamma \mathcal{P}^\neg(\bar{x}, \mathbf{V}_1, \dots, \mathbf{V}_n)$. \square

Before extending Propositions 1.3 and 1.4 to all of **GRIST**, we need a technical result. We use \bar{x} as shorthand for x_1, \dots, x_k .

Lemma 1.7 (a) If $\mathcal{P}(\bar{x})$ is a formula where only variables of the first kind occur, then $\mathbf{R}^\heartsuit \vdash (\Gamma^\ulcorner \mathcal{P}^v(\bar{x})^\urcorner \leftrightarrow \mathcal{Q}(\bar{x}; \mathbf{V}))$, for some suitable \mathbf{V} -formula $\mathcal{Q}(\bar{x}; \mathbf{V})$.

(b) If $\mathcal{Q}(\bar{x}; \mathbf{V})$ is a suitable \mathbf{V} -formula, then $\mathbf{R} \vdash (\Delta^\ulcorner \mathcal{Q}(\bar{x}; \mathbf{V})^\urcorner \leftrightarrow \mathcal{P}^v(\bar{x}))$, for some formula $\mathcal{P}(\bar{x})$ where only variables of the first kind occur.

Proof (a) $\Gamma^\ulcorner (x \sqsubseteq y)^v \urcorner \doteq \Gamma^\ulcorner x \sqsubseteq y \vee x \sqsubseteq v \urcorner \doteq$

(*) $(\forall \mathbf{U})(y \in \mathbf{U} \rightarrow x \in \mathbf{U}) \vee x \in \mathbf{V}$.

We show that, in \mathbf{R}^\heartsuit , (*) is equivalent to the \mathbf{V} -formula

(**) $(\forall \mathbf{U} \supseteq \mathbf{V})(y \in \mathbf{U} \rightarrow x \in \mathbf{U}) \vee x \in \mathbf{V}$.

(*) \Rightarrow (**) is clear. The other direction is clear if $x \in \mathbf{V}$ holds, so assume $(\forall \mathbf{U} \supseteq \mathbf{V})(y \in \mathbf{U} \rightarrow x \in \mathbf{U}) \wedge x \notin \mathbf{V}$. Letting $\mathbf{U} \doteq \mathbf{V}$, we deduce $y \notin \mathbf{V}$. For any \mathbf{U} such that $y \in \mathbf{U}$ now $\mathbf{U} \supseteq \mathbf{V}$ [by \mathbf{R}^\heartsuit (iii)], so $(y \in \mathbf{U} \rightarrow x \in \mathbf{U})$. This proves (*).

The rest of (a) follows trivially by induction on the complexity of \mathcal{P} .

(b) We show by induction on the complexity of formulas that for every suitable \mathbf{V} -formula $\mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n)$ there is a suitable formula $\mathcal{P}(\bar{x}, v_1, \dots, v_n)$, in which all variables except v_1, \dots, v_n are of the first kind, such that

$$\mathbf{R} \vdash (\forall v_1, \dots, v_n \sqsupseteq v)[\Delta^\ulcorner \mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n)^\urcorner \leftrightarrow \mathcal{P}^v(\bar{x}, v_1, \dots, v_n)].$$

In particular, letting $n = 0$ gives $\mathbf{R} \vdash \Delta^\ulcorner \mathcal{Q}(\bar{x}; \mathbf{V})^\urcorner \leftrightarrow \mathcal{P}^v(\bar{x})$ and proves (b).

We focus on the nontrivial cases.

$\Delta^\ulcorner x \in \mathbf{V}^\urcorner \doteq x \sqsubseteq v$ and $x \sqsubseteq v \leftrightarrow x \sqsubseteq_v 0 \leftrightarrow (x \sqsubseteq 0)^v$ holds.

$\Delta^\ulcorner x \in \mathbf{V}_i^\urcorner \doteq x \sqsubseteq v_i$ and $(\forall v_i \sqsupseteq v)[x \sqsubseteq v_i \leftrightarrow x \sqsubseteq_v v_i \leftrightarrow (x \sqsubseteq v_i)^v]$ holds.

Wlog we consider $(\forall \mathbf{V}_1 \supseteq \mathbf{V})\mathcal{Q}$. We have, in \mathbf{R}^\heartsuit , $(\forall \mathbf{V}_1 \supseteq \mathbf{V})\mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n) \leftrightarrow [\mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}, \mathbf{V}_2, \dots, \mathbf{V}_n) \wedge (\forall \mathbf{V}_1 \supseteq \mathbf{V})\mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n)]$; hence, by Proposition 1.3, \mathbf{R} proves $\Delta^\ulcorner (\forall \mathbf{V}_1 \supseteq \mathbf{V})\mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n)^\urcorner \leftrightarrow [\Delta^\ulcorner \mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}, \mathbf{V}_2, \dots, \mathbf{V}_n)^\urcorner \wedge \Delta^\ulcorner (\forall \mathbf{V}_1 \supseteq \mathbf{V})\mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n)^\urcorner]$. By the inductive assumption, we can prove in \mathbf{R} that, for all $v_1, \dots, v_n \sqsupseteq v$, $\Delta^\ulcorner \mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n)^\urcorner \leftrightarrow \mathcal{P}^v(\bar{x}, v_1, \dots, v_n)$ and $\Delta^\ulcorner \mathcal{Q}(\bar{x}; \mathbf{V}, \mathbf{V}, \mathbf{V}_2, \dots, \mathbf{V}_n)^\urcorner \leftrightarrow \mathcal{R}^v(\bar{x}, v_2, \dots, v_n)$, where \mathcal{P}, \mathcal{R} are suitable formulas with all variables except v_1, \dots, v_n of the first kind. It follows that, for $v_2, \dots, v_n \sqsupseteq v$, $\Delta^\ulcorner (\forall \mathbf{V}_1 \supseteq \mathbf{V})\mathcal{Q} \urcorner$ is \mathbf{R} -equivalent to

$$(*) \quad \mathcal{R}^v(\bar{x}, v_2, \dots, v_n) \wedge (\forall v_1 \sqsupseteq v)\mathcal{P}^v(\bar{x}, v_1, \dots, v_n).$$

Observe that, using **R**(iii, i), $v \sqsubset v_1 \leftrightarrow \neg(v_1 \sqsubseteq v) \leftrightarrow \neg(v_1 \sqsubseteq_v 0) \leftrightarrow [\neg(v_1 \sqsubseteq_v 0) \wedge 0 \sqsubseteq_v v_1] \leftrightarrow (0 \sqsubset v_1)^v$. Hence, for $v_2, \dots, v_n \sqsupseteq v$,

$$(*) \leftrightarrow [\mathcal{R}(\bar{x}, v_2, \dots, v_n) \wedge (\forall v_1 \sqsupset 0) \mathcal{P}(\bar{x}, v_1, \dots, v_n)]^v.$$

Replace the bound variable v_1 in the second conjunct by some variable of the first kind that does not occur in \mathcal{P} , and the proof is complete. \square

Theorem 1.8 (a) **GRIST** $\vdash \mathcal{P}$ if and only if **GRIST**[♡] $\vdash \Gamma^\ulcorner \mathcal{P}^\urcorner$,
for all suitable \mathcal{P} .

(b) **GRIST**[♡] $\vdash \mathcal{Q}$ if and only if **GRIST** $\vdash \Delta^\ulcorner \mathcal{Q}^\urcorner$, for all suitable \mathcal{Q} .

Proof We can assume that all variables that occur in the axioms of **GRIST** (in particular, in the formula \mathcal{P}), except for u, v, w , are of the first kind. It is then easy to verify, with the help of Lemma 1.7(a), that each of the remaining axioms of **GRIST**: **R**(iii, iv, v), **T**, **S**, **I**, **G**, translates into a formula that is equivalent to an instance of the corresponding axiom of **GRIST**[♡]. This establishes the “only if” direction in (a).

For the converse, we can assume that all individual variables occurring in the axioms of **GRIST**[♡] are of the first kind. Then all of the remaining axioms of **GRIST**[♡] translate into equivalents of the corresponding axioms of **GRIST**, using Lemma 1.7(b). This proves the “only if” direction in (b).

The “if” directions then follow from Proposition 1.2. \square

Corollary 1.9 *Corollaries 12.1 – 12.10 in [16] are valid for **GRIST**[♡] [in place of **SST**[#]], with obvious adjustments. In particular, **GRIST**[♡] is a conservative extension of **ZFC**.*

From now on, we do not formally distinguish between **GRIST** and **GRIST**[♡].

A number of consequences of **GRIST** that are useful in relative analysis have been derived in [16]. Below we give a translation of these consequences into the language of levels. It follows easily from **GRIST** that for every x_1, \dots, x_k there is a coarsest level where x_1, \dots, x_k appear; see Axiom **I** below. We denote it $\mathbf{V}(x_1, \dots, x_k)$ and call it **the level of** x_1, \dots, x_k . In particular, $\mathbf{V}(\cdot)$ is the coarsest level [\cdot is the empty list].

Proposition 1.10 *The following are theorems of **GRIST**:*

(1) **FRIST Standardization** [16, 12.17]: *Given \mathbf{V}, \bar{x}, A :*

$$(\exists B \in \mathbf{V}) (\forall y \in \mathbf{V}) (y \in B \leftrightarrow y \in A \wedge \mathcal{P}(y, \bar{x}; \mathbf{V})).$$

- (2) **FRIST Idealization** [16, 12.17]: Given $\mathbf{U} \subset \mathbf{V}$, $A, B \in \mathbf{U}$ and \bar{x} :

$$(\forall a \in \mathcal{P}^{\text{fin}}A \cap \mathbf{U})(\exists y \in B)(\forall x \in a)\mathcal{P}(x, y, \bar{x}; \mathbf{V}) \leftrightarrow (\exists y \in B)(\forall x \in A \cap \mathbf{U})\mathcal{P}(x, y, \bar{x}; \mathbf{V}).$$

- (3) **External Induction** [16, 12.19]: If, for all ordinals $\xi \in \mathbf{V}$, $(\forall \eta \in \xi \cap \mathbf{V})\mathcal{P}(\eta, \bar{x}; \mathbf{V}) \rightarrow \mathcal{P}(\xi, \bar{x}; \mathbf{V})$ holds, then $(\forall \xi \in \mathbf{V})\mathcal{P}(\xi, \bar{x}; \mathbf{V})$ holds. In particular, if $\mathcal{P}(0, \bar{x}; \mathbf{V})$ holds and $(\forall n \in \omega \cap \mathbf{V})[\mathcal{P}(n, \bar{x}; \mathbf{V}) \rightarrow \mathcal{P}(n+1, \bar{x}; \mathbf{V})]$, then $(\forall n \in \omega \cap \mathbf{V})\mathcal{P}(n, \bar{x}; \mathbf{V})$ holds.
- (4) **Finite Choice** [16, 12.20]: If $a \in \mathbf{V}$ is finite and $(\forall x \in a)(\exists y)\mathcal{P}(x, y, \bar{x}; \mathbf{V})$, then there exists a function f with $\text{dom} f = a$ such that $(\forall x \in a)\mathcal{P}(x, f(x), \bar{x}; \mathbf{V})$.
- (5) **Support Principle** [16, 12.21]: Given a \mathbf{V} -formula $\mathcal{P}(\bar{x}; \mathbf{V})$ and sets x_1, \dots, x_k , there is a finite set $\{v_0, v_1, \dots, v_n\}$ such that $\mathbf{V}(\cdot) = \mathbf{V}(v_0) \subset \mathbf{V}(v_1) \subset \dots \subset \mathbf{V}(v_n)$ and for all $i \leq n$ and all \mathbf{V} with $\mathbf{V}(v_i) \subseteq \mathbf{V} \subset \mathbf{V}(v_{i+1})$ [$\mathbf{V}(v_i) \subseteq \mathbf{V}$ if $i = n$], $\mathcal{P}(\bar{x}; \mathbf{V}(v_i)) \leftrightarrow \mathcal{P}(\bar{x}; \mathbf{V}) \leftrightarrow \neg \mathcal{P}(\bar{x}; \mathbf{V}(v_{i+1}))$.
- (6) **Local Transfer** [16, 12.22]: For any sets x_{k+1}, \dots, x_n and any \mathbf{V}_0 there is $\mathbf{V}' \supset \mathbf{V}_0$ such that, for all $\mathbf{V}_0 \subseteq \mathbf{V} \subset \mathbf{V}'$ and all $x_1, \dots, x_k \in \mathbf{V}_0$, $\mathcal{P}(x_1, \dots, x_k, \dots, x_n; \mathbf{V}_0) \leftrightarrow \mathcal{P}(x_1, \dots, x_k, \dots, x_n; \mathbf{V})$.
- (7) **Standardization** [16, 12.23]: For any $\mathbf{V} \supset \mathbf{V}(\cdot)$ and any A, \bar{x} , there exists B such that $\mathbf{V}(B) \subset \mathbf{V}$ and $(\forall y)[\mathbf{V}(y) \subset \mathbf{V} \rightarrow (y \in B \leftrightarrow y \in A \wedge \mathcal{P}(y, \bar{x}; \mathbf{V}))]$.
- (8) **Polytransfer** [16, 12.26]:
Let $\mathbf{V} \subset \mathbf{V}_1 \subset \dots \subset \mathbf{V}_n$ and $\mathbf{V}' \subset \mathbf{V}'_1 \subset \dots \subset \mathbf{V}'_n$. Then $(\forall \bar{x} \in \mathbf{V} \cap \mathbf{V}') [\mathcal{P}(\bar{x}; \mathbf{V}, \mathbf{V}_1, \dots, \mathbf{V}_n) \leftrightarrow \mathcal{P}(\bar{x}; \mathbf{V}', \mathbf{V}'_1, \dots, \mathbf{V}'_n)]$.
- (9) **Partial Transfer** [16, 12.27]: If $\mathbf{V} \subset \mathbf{V}'$, then $(\forall \bar{x} \in \mathbf{V}) [(\exists y)(\mathcal{P}(y, \bar{x}; \mathbf{V}) \wedge \mathcal{Q}(y, \bar{x}; \mathbf{V})) \rightarrow (\exists y)(\mathcal{P}(y, \bar{x}; \mathbf{V}) \wedge \mathcal{Q}(y, \bar{x}; \mathbf{V}'))]$.
- (10) **Standard Size Choice** [16, 12.28]: For every $A \in \mathbf{V}$ such that $(\forall x \in A \cap \mathbf{V})(\exists y)\mathcal{P}(x, y, \bar{x}; \mathbf{V})$ there exists a function f with $\text{dom} f = A$ such that $(\forall x \in A \cap \mathbf{V})\mathcal{P}(x, f(x), \bar{x}; \mathbf{V})$.
- (11) **Map Standardization** [16, 12.29]: For every $A \in \mathbf{V}$ there exists $f \in \mathbf{V}$ such that $\text{dom} f = A$ and $(\forall x \in A \cap \mathbf{V})[(\exists y \in \mathbf{V})\mathcal{P}(x, y, \bar{x}; \mathbf{V}) \rightarrow \mathcal{P}(x, f(x), \bar{x}; \mathbf{V})]$.
- (12) [16, 12.31]: If $A \neq \emptyset$, then there is $x \in A$ such that $\mathbf{V}(x) = \mathbf{V}(A)$.
- (13) **Levels of Elements of Sets** [16, 12.32]:
(a) If A is infinite and $A \in \mathbf{V}$, then $(\exists x)(x \in A \wedge \mathbf{V}(x) = \mathbf{V})$.
(b) $(\forall x \in A)(x \in \mathbf{V})$ if and only if A is finite and $A \in \mathbf{V}$.

Another useful consequence is Saturation, a general form of compactness.

Proposition 1.11 (Saturation) *If $\mathcal{F} \in \mathbf{U} \subset \mathbf{V}$ and \mathcal{F} has the finite intersection property, then there exists $y \in \mathbf{V}$ such that $y \in X$, for all $X \in \mathcal{F} \cap \mathbf{U}$.*

Proof Let $A := \mathcal{F}$, $B := \bigcup \mathcal{F}$, and let $\mathcal{P}(X, y; \mathbf{V})$ be the formula $(y \in X) \wedge (y \in \mathbf{V})$; then apply **FRIST** Idealization. \square

As outlined in [17], an elementary exposition of calculus in relative analysis does not need the full strength (and complexity) of **GRIST**; a much weaker system suffices. Here and in the next section we derive the axioms of [17] from **GRIST**.

Axiom I *For every x_1, \dots, x_k there is a level \mathbf{V} such that $x_1, \dots, x_k \in \mathbf{V}$ and, for all levels \mathbf{U} , $x_1, \dots, x_k \in \mathbf{U}$ implies $\mathbf{V} \subseteq \mathbf{U}$.*

Proof Given x_1, \dots, x_k , there are levels $\mathbf{V}_1, \dots, \mathbf{V}_k$ such that $x_i \in \mathbf{V}_i \wedge (\forall \mathbf{U})(x_i \in \mathbf{U} \rightarrow \mathbf{V}_i \subseteq \mathbf{U})$ [**R**[∇](i)]. Using **R**[∇](iii) repeatedly, one finds \mathbf{V}_n such that $\mathbf{V}_i \subseteq \mathbf{V}_n$ for all $i = 1, \dots, k$; clearly \mathbf{V}_n has the required properties. \square

Remark A stronger result follows from Proposition 1.10(13): For every finite set $\{x_1, \dots, x_k\}$ there is a coarsest level \mathbf{V} such that $\{x_1, \dots, x_k\} \subseteq \mathbf{V}$.

The axioms **II** and **VI** (Stability) are **R**[∇](iii) and **T**[∇], respectively, and **V** is a consequence of **VI**. The statements and proofs of axioms **III**, **IV** (Neighbor Principle) and **VIII** (Density of Levels) are given in Section 2.

In [17] a formula $\mathcal{P}(x_1, \dots, x_k)$ is called **internal** if all quantifiers over levels are of the form $(\forall \mathbf{U})(x_1, \dots, x_k \in \mathbf{U} \rightarrow \dots)$ or $(\exists \mathbf{U})(x_1, \dots, x_k \in \mathbf{U} \wedge \dots)$.

Let $\mathcal{P}(\bar{x}; \mathbf{V})$ be the \mathbf{V} -formula obtained from an internal $\mathcal{P}(\bar{x})$ by replacing each $(\forall \mathbf{U})$ with $(\forall \mathbf{U} \supseteq \mathbf{V})$ and each $(\exists \mathbf{U})$ with $(\exists \mathbf{U} \supseteq \mathbf{V})$.

Lemma 1.12 $(\forall \mathbf{V})(\forall \bar{x})(\mathcal{P}(\bar{x}) \leftrightarrow \mathcal{P}(\bar{x}; \mathbf{V}))$, for all internal \mathcal{P} .

Proof We first note that $\mathcal{P}(\bar{x})$ is equivalent to the formula $\mathcal{P}^*(\bar{x})$ obtained from \mathcal{P} by replacing each $(\forall \mathbf{U})(x_1, \dots, x_k \in \mathbf{U} \rightarrow \dots)$ with $(\forall \mathbf{U} \supseteq \mathbf{V}(x_1, \dots, x_k)) \dots$, and similarly for \exists . Let $\mathcal{P}^*(\bar{x}; \mathbf{V})$ be the \mathbf{V} -formula obtained from \mathcal{P}^* by replacing each $\mathbf{V}(x_1, \dots, x_k)$ with \mathbf{V} ; so that $\mathcal{P}^*(\bar{x}) \leftrightarrow \mathcal{P}^*(\bar{x}; \mathbf{V}(x_1, \dots, x_k))$. If now \mathbf{V} is an arbitrary level, $\mathbf{V} = \mathbf{V}(z)$ for some z by **R**[∇](ii), and we have $\mathcal{P}(\bar{x}) \leftrightarrow \mathcal{P}^*(\bar{x}) \leftrightarrow \mathcal{P}^*(\bar{x}; \mathbf{V}(x_1, \dots, x_k)) \leftrightarrow \mathcal{P}^*(\bar{x}; \mathbf{V}(x_1, \dots, x_k, z))$ [by Transfer] $\leftrightarrow \mathcal{P}(\bar{x}; \mathbf{V})$, because $\mathbf{U} \supseteq \mathbf{V}(x_1, \dots, x_k, z) \leftrightarrow \mathbf{U} \supseteq \mathbf{V}(z) \wedge x_1, \dots, x_k \in \mathbf{U} \leftrightarrow \mathbf{U} \supseteq \mathbf{V} \wedge x_1, \dots, x_k \in \mathbf{U}$. \square

Axiom VII (Definition Principle) *If $\mathcal{P}(y, x_1, \dots, x_k)$ is internal and A is a set, then there is a set B such that $(\forall y)(y \in B \leftrightarrow y \in A \wedge \mathcal{P}(y, x_1, \dots, x_k))$.*

Proof Let $\mathcal{P}(y, \bar{x}; \mathbf{V})$ be the \mathbf{V} -formula obtained from the internal $\mathcal{P}(y, \bar{x})$ as above, and let $\mathbf{V}_0 := \mathbf{V}(A, \bar{x})$. By **FRIST** Standardization, there is $B \in \mathbf{V}_0$ such that

$$(*) \quad (\forall y \in \mathbf{V}_0)(y \in B \leftrightarrow y \in A \wedge \mathcal{P}(y, \bar{x}; \mathbf{V}_0)).$$

By Transfer (Stability) applied to $(*)$, for any $\mathbf{V}_1 \supseteq \mathbf{V}_0$

$$(**) \quad (\forall y \in \mathbf{V}_1)(y \in B \leftrightarrow y \in A \wedge \mathcal{P}(y, \bar{x}; \mathbf{V}_1)).$$

Given an arbitrary y , let $\mathbf{V}_1 := \mathbf{V}(y, A, \bar{x})$; Lemma 1.12 gives $y \in B \leftrightarrow y \in A \wedge \mathcal{P}(y, \bar{x})$. \square

Mathematical practice enriches the set-theoretic language by new defined concepts. We conclude this section by proving that the axioms of **GRIST** remain valid for formulas in the language enriched by internal predicates, where a *predicate* $R(x_1, \dots, x_k)$ defined by $R(x_1, \dots, x_k) \leftrightarrow \mathcal{R}(x_1, \dots, x_k)$ is **internal** if its defining formula \mathcal{R} is internal.

\mathbf{V} -formulas and internal formulas of the language of **GRIST**[♥] with an additional predicate symbol R are defined in the same way as for the original language.

Proposition 1.13 (a) *If $\mathcal{P}(\bar{x}; \mathbf{V})$ is a \mathbf{V} -formula in the language with an additional predicate R and $\mathcal{P}^\sharp(\bar{x}, \mathbf{V})$ is obtained from \mathcal{P} by replacing each occurrence of R by its internal defining formula \mathcal{R} , then \mathcal{P}^\sharp is equivalent to a \mathbf{V} -formula in the original language.*

(b) *If $\mathcal{P}(x_1, \dots, x_n)$ is an internal formula in the language with an additional predicate R and $\mathcal{P}^\sharp(\bar{x})$ is obtained from \mathcal{P} by replacing each occurrence of R by its internal defining formula \mathcal{R} , then \mathcal{P}^\sharp is equivalent to an internal formula in the original language.*

Proof (a) is obvious from Lemma 1.12 [replace $\mathcal{R}(\bar{x})$ by $\mathcal{R}(\bar{x}; \mathbf{V})$].

(b) Refer to the proof of Lemma 1.12 and the notation therein. We have $\mathcal{R}(\bar{y}) \leftrightarrow \mathcal{R}^*(\bar{y}) \leftrightarrow \mathcal{R}^*(\bar{y}; \mathbf{V}(\bar{y})) \leftrightarrow \mathcal{R}^*(\bar{y}; \mathbf{V}(\bar{y}, \bar{x}))$ [by Transfer]. The last formula is equivalent to the formula $\mathcal{R}(\bar{y}, \bar{x})$ obtained from \mathcal{R} by replacing each occurrence of the universal quantifier $(\forall \mathbf{U})(y_1, \dots, y_k \in \mathbf{U} \dots)$ by $(\forall \mathbf{U})(y_1, \dots, y_k, x_1, \dots, x_n \in \mathbf{U} \dots)$ [and similarly for \exists]. It is easy to see that replacing each occurrence of $R(\bar{y})$ in \mathcal{P} by $\mathcal{R}(\bar{y}, \bar{x})$ yields an internal formula equivalent to \mathcal{P}^\sharp . \square

2 S-continuity in GRIST.

If \mathbf{V} is a fixed level and $\text{st}(x)$ [x is **standard**] is defined as $x \in \mathbf{V}$, all axioms of **BST** become provable in **GRIST** [16, Corollary 12.18]. Therefore, nonstandard analysis in the style of Internal Set Theory of Nelson [20, 7] can be practiced in **GRIST**, with the additional advantage that the notion of standardness is not fixed: Every set a can be considered as standard relative to any level with $a \in \mathbf{V}$. The paper [17] outlines an elementary presentation of nonstandard methods in analysis in the framework of **GRIST**, and discusses its advantages.

The existence of many “levels of standardness” in **GRIST** raises the question of dependence of nonstandard concepts on the choice of the level. It also enables definitions of concepts that involve quantification over levels. In this and the next section we illustrate some of the techniques available in **GRIST** for handling of such issues.

Definition 2.1 Given a level \mathbf{V} :

- (1) A real number ϵ is **ultrasmall** relative to \mathbf{V} if $|\epsilon| < r$ for all $r > 0$, $r \in \mathbf{V}$.
- (2) A real number x is **ultralarge** relative to \mathbf{V} if $|x| > r$ for all $r > 0$, $r \in \mathbf{V}$;
 x is **limited** relative to \mathbf{V} if it is not ultralarge relative to \mathbf{V} .
- (3) Real numbers a and b are **ultraclose** relative to \mathbf{V} , written $a \simeq_{\mathbf{V}} b$, if $a - b$ is ultrasmall relative to \mathbf{V} .

Lemma 2.2 (a) For every level \mathbf{V} there is $k \in \mathbb{N}$ such that $\mathbf{V} = \mathbf{V}(k)$.
(b) If $k \in \mathbb{N}$ and $\mathbf{V} \subset \mathbf{V}(k)$, then k is ultralarge relative to \mathbf{V} .

Proof (a) For every $\mathbf{W} \subset \mathbf{V}$ and every finite $a \subseteq \mathbb{N}$, $a \in \mathbf{W}$, there is $k \in \mathbb{N} \cap \mathbf{V}$ such that $(\forall n \in a)(n < k)$. By Idealization there is $k \in \mathbb{N} \cap \mathbf{V}$ such that $(\forall \mathbf{W} \subset \mathbf{V})(\forall n \in \mathbb{N} \cap \mathbf{W})(n < k)$; evidently then $\mathbf{V}(k) = \mathbf{V}$.

(b) Let $\mathbf{V} \subset \mathbf{V}(k)$. By **FRIST** Standardization, there is $B \in \mathbf{V}$ such that $(\forall n \in \mathbf{V})(n \in B \leftrightarrow n \in \mathbb{N} \wedge n < k)$. Clearly $0 \in B$ and $(\forall n \in \mathbf{V})(n \in B \rightarrow n + 1 \in B)$ [otherwise, $k = 0$ or $k = n + 1$, so $k \in \mathbf{V}$]. By Transfer, $(\forall n)(n \in B \rightarrow n + 1 \in B)$, so $B = \mathbb{N}$ and we conclude $(\forall n \in \mathbf{V})(n < k)$. \square

Proposition 2.3 (Axiom III) For every level \mathbf{V} there exist real numbers $\epsilon \neq 0$ ultrasmall relative to \mathbf{V} .

Proof Let $\mathbf{V} \subset \mathbf{U}$ [$\mathbf{R}^{\heartsuit}(\text{iv})$] and $\mathbf{U} = \mathbf{V}(k)$ for $k \in \mathbb{N}$ [Lemma 2.2(a)]. Then k is ultralarge relative to \mathbf{V} [Lemma 2.2(b)] and $\epsilon := 1/k$ is ultrasmall relative to \mathbf{V} . \square

Proposition 2.4 (Axiom VIII, Density of Levels) *If a real number $\epsilon \neq 0$ is ultrasmall relative to \mathbf{V} , then there is \mathbf{V}^+ and a real number $\delta \in \mathbf{V}^+$, $\delta \neq 0$, such that δ is ultrasmall relative to \mathbf{V} and ϵ is ultrasmall relative to \mathbf{V}^+ .*

Proof By Local Transfer there is $\mathbf{V}^+ \supset \mathbf{V}$ such that ϵ is ultrasmall relative to \mathbf{V}^+ . Lemma 2.2(a) gives $k \in \mathbb{N}$ such that $\mathbf{V}^+ = \mathbf{V}(k)$. Let $\delta = 1/k$. \square

Proposition 2.5 (Axiom IV, Neighbor Principle) *For every real number x limited relative to \mathbf{V} there is a real number $r \in \mathbf{V}$ such that $x \simeq_{\mathbf{V}} r$. The number r is uniquely determined; we call it the **V-neighbor** of x and denote it $\mathbf{n}_{\mathbf{V}}(x)$.*

Proof By FRIST Standardization 1.10(1) there is $B \in \mathbf{V}$ such that $(\forall s \in \mathbf{V})(s \in B \leftrightarrow s \in \mathbb{R} \wedge s < x)$. If $x \leq \bar{r} \in \mathbf{V}$, then B is bounded above by \bar{r} [use Transfer], hence B has a least upper bound r , and $r \in \mathbf{V}$, again by Transfer. One easily verifies that $r \simeq_{\mathbf{V}} x$. \square

If $x \in \mathbb{R}$, $x \neq 0$, then x is not ultrasmall relative to $\mathbf{V}(x)$. Granularity implies that for every $x \neq 0$ there is a coarsest level \mathbf{V}_0 such that x is not ultrasmall relative to \mathbf{V}_0 . If $\mathbf{V}_0 = \mathbf{V}(\cdot)$, the coarsest level, then x is not ultrasmall relative to any level. Otherwise, x is ultrasmall relative to $\mathbf{V} \subset \mathbf{V}_0$ and is not ultrasmall relative to $\mathbf{V} \supseteq \mathbf{V}_0$.

More interesting behavior is exhibited by the various *S-concepts* that play a key role in nonstandard analysis: S-continuity, S-integrability, etc. Here we study the dependence of S-continuity on the choice of level.

Definition 2.6 Given a set $A \subseteq \mathbb{R}$ and a level \mathbf{V} , there is a unique set $B \in \mathbf{V}$ such that $(\forall r \in \mathbf{V})[r \in B \leftrightarrow r \in \mathbb{R} \wedge (\exists x \in A)(r \simeq_{\mathbf{V}} x)]$ [FRIST Standardization 1.10(1)]. We call this B the **V-shadow** of A and denote it $\mathbf{sh}_{\mathbf{V}}(A)$.¹

Proposition 2.7 (a) For $A \subseteq \mathbb{R}$, $A \in \mathbf{V}$: A is a closed set if and only if $\mathbf{sh}_{\mathbf{V}}(A) = A$.
 (b) For any $A \subseteq \mathbb{R}$: $\mathbf{sh}_{\mathbf{V}}(A)$ is a closed set.
 (c) If $\mathbf{V} \subseteq \mathbf{U}$ and $A \subseteq \mathbb{R}$, then $\mathbf{sh}_{\mathbf{U}}(\mathbf{sh}_{\mathbf{V}}(A)) = \mathbf{sh}_{\mathbf{V}}(A)$.
 (d) For $A \subseteq \mathbb{R}$, $A \in \mathbf{V} \subseteq \mathbf{U}$: $\mathbf{sh}_{\mathbf{V}}(A) = \mathbf{sh}_{\mathbf{U}}(A)$.

Proof (a) A well-known nonstandard characterization of closed subsets of \mathbb{R} is: $A \in \mathbf{V}$ is closed if and only if, for all $r \in \mathbb{R} \cap \mathbf{V}$, $(\exists x \in A)(r \simeq_{\mathbf{V}} x) \leftrightarrow r \in A$. This is precisely the statement that $A = \mathbf{sh}_{\mathbf{V}}(A)$.

¹In [16], $\mathbf{sh}_{\alpha}(x)$ is defined for all sets x , with a different meaning.

(b) Let $B := \mathbf{sh}_{\mathbf{V}}(A)$; by Local Transfer 1.10(6), there is a level $\mathbf{V}_1 \supset \mathbf{V}$ such that $B = \mathbf{sh}_{\mathbf{V}_1}(A)$. Let now $r \in \mathbb{R} \cap \mathbf{V}$ and $(\exists x \in B)(r \simeq_{\mathbf{V}} x)$. Then $(\exists x \in B \cap \mathbf{V}'_1)(r \simeq_{\mathbf{V}} x)$ holds for some $\mathbf{V}'_1 \supset \mathbf{V}$. By Polytransfer 1.10(8), also $(\exists x \in B \cap \mathbf{V}_1)(r \simeq_{\mathbf{V}} x)$ holds. Fix such x ; as $B = \mathbf{sh}_{\mathbf{V}_1}(A)$, there exists $\xi \in A$ such that $x \simeq_{\mathbf{V}_1} \xi$. So $r \simeq_{\mathbf{V}} \xi$ for $\xi \in A$, and $r \in \mathbf{sh}_{\mathbf{V}}(A) = B$. Hence B satisfies the characterization of closed sets from the proof of (a).

(c) follows from (a) and (b), and (d) follows from the definition of \mathbf{sh} by Transfer. \square

Remark The usual proof of (b) uses ϵ -neighborhoods of r and Idealization; the argument given here is more “nonstandard.”

Proposition 2.8 *If $\mathbf{U} \subset \mathbf{V}$ and $A \subseteq \mathbb{R}$, then $\mathbf{sh}_{\mathbf{U}}(\mathbf{sh}_{\mathbf{V}}(A)) = \mathbf{sh}_{\mathbf{U}}(A)$.*

Proof If $x \in \mathbf{sh}_{\mathbf{U}}(A) \cap \mathbf{U}$, then $x \simeq_{\mathbf{U}} \xi$ for some $\xi \in A$. Let $y := \mathbf{n}_{\mathbf{V}}(\xi)$ [it is defined because ξ is limited relative to \mathbf{U} , hence also relative to \mathbf{V}]. We have $\xi \simeq_{\mathbf{V}} y$, so $y \in \mathbf{sh}_{\mathbf{V}}(A)$, and $x \simeq_{\mathbf{U}} y$, so $x \in \mathbf{sh}_{\mathbf{U}}(\mathbf{sh}_{\mathbf{V}}(A))$.

For the converse, let $x \in \mathbf{sh}_{\mathbf{U}}(\mathbf{sh}_{\mathbf{V}}(A)) \cap \mathbf{U}$. Let $B := \mathbf{sh}_{\mathbf{V}}(A)$. We know that $x \simeq_{\mathbf{U}} y$ for some $y \in B$. It suffices to prove

Claim: $x \simeq_{\mathbf{U}} y$ for some $y \in B \cap \mathbf{V}$

because then $y \simeq_{\mathbf{V}} \xi$ for some $\xi \in A$ and so $x \simeq_{\mathbf{U}} \xi$ and $x \in \mathbf{sh}_{\mathbf{U}}(A)$.

Proof of Claim Let U vary over neighborhoods of x in \mathbf{U} ; then $(\exists y)(y \in B \cap U)$ is true [because $(\exists y \in B)(y \simeq_{\mathbf{U}} x)$]. By Transfer into \mathbf{V} , $(\exists y \in \mathbf{V})(y \in B \cap U)$. Hence $(\forall U \in \mathbf{U})(\exists y \in B)(y \in U \cap \mathbf{V})$. By FRIST Idealization 1.10(2), $(\exists y \in B)(\forall U \in \mathbf{U})(y \in U \cap \mathbf{V})$, ie, $(\exists y \in B)(x \simeq_{\mathbf{U}} y \wedge y \in \mathbf{V})$. \square \square

The notion of shadow makes sense for subsets of $\mathbb{R} \times \mathbb{R}$. We let $(x, y) \simeq_{\mathbf{V}} (x', y')$ if and only if $x \simeq_{\mathbf{V}} x' \wedge y \simeq_{\mathbf{V}} y'$. Definition 2.6 and Propositions 2.7, 2.8 then have obvious analogs for $A \subseteq \mathbb{R} \times \mathbb{R}$.

In the rest of this section we study real-valued functions. For simplicity, we consider only functions $f : [a, b] \rightarrow [c, d]$ where $a, b, c, d \in \mathbf{V}(\cdot)$.

Definition 2.9 f is (uniformly) \mathbf{V} -continuous if $x \simeq_{\mathbf{V}} x'$ implies $f(x) \simeq_{\mathbf{V}} f(x')$, for all $x, x' \in \text{dom} f$. If also $f \in \mathbf{V}$, then f is (uniformly) **continuous**.

Theorem 2.10 *For every function f there is a finite set $\{v_0, \dots, v_n\}$ such that $\mathbf{V}(\cdot) = \mathbf{V}(v_0) \subset \mathbf{V}(v_1) \subset \dots \subset \mathbf{V}(v_n)$ and, for all $\mathbf{V}(v_i) \subseteq \mathbf{V} \subset \mathbf{V}(v_{i+1})$ [all $\mathbf{V}(v_i) \subseteq \mathbf{V}$ if*

$i = n]$, the function f is $\mathbf{V}(v_i)$ -continuous if and only if f is \mathbf{V} -continuous if and only if f is not $\mathbf{V}(v_{i+1})$ -continuous.

Conversely, for every finite set $\{v_0, \dots, v_n\}$ as above there is a function f with the above properties.

The first part of Theorem 2.10 is an immediate consequence of the Support Principle 1.10(5). Here we prove a stronger result [Theorem 2.13] showing that every function has only finitely many distinct shadows as \mathbf{V} ranges over all levels. We recall that $\mathbf{sh}_{\mathbf{V}}(f)$ is a set in \mathbf{V} such that $(r, s) \in \mathbf{sh}_{\mathbf{V}}(f) \leftrightarrow (\exists x \in [a, b])(r \simeq_{\mathbf{V}} x \wedge s \simeq_{\mathbf{V}} f(x))$ holds for all $r, s \in \mathbf{V}$.

The S-versions of the following two facts are well-known (eg see [7]).

Proposition 2.11 f is \mathbf{V} -continuous if and only if $\mathbf{sh}_{\mathbf{V}}(f)$ is a function.

Proof Let $F := \mathbf{sh}_{\mathbf{V}}(f)$.

\Rightarrow Assume that f is \mathbf{V} -continuous. Suppose $(r, s_1), (r, s_2) \in F \cap \mathbf{V}$; then $r \simeq_{\mathbf{V}} x$, $s_1 \simeq_{\mathbf{V}} f(x)$, $r \simeq_{\mathbf{V}} x'$, $s_2 \simeq_{\mathbf{V}} f(x')$, for some x, x' . It follows that $x \simeq_{\mathbf{V}} x'$, so by \mathbf{V} -continuity, $f(x) \simeq_{\mathbf{V}} f(x')$, and finally $s_1 \simeq_{\mathbf{V}} s_2$. As $s_1, s_2 \in \mathbf{V}$, we get $s_1 = s_2$. Thus F is a function.

\Leftarrow Assume f is not \mathbf{V} -continuous. Then there are $x, x' \in [a, b]$ such that $x \simeq_{\mathbf{V}} x'$ and $f(x) \not\simeq_{\mathbf{V}} f(x')$. Let $r := \mathbf{nv}(x) = \mathbf{nv}(x')$, $s_1 := \mathbf{nv}(f(x))$, $s_2 := \mathbf{nv}(f(x'))$. Then $(r, s_1), (r, s_2) \in F$ and $s_1 \neq s_2$, so F is not a function. \square

Proposition 2.12 If f is \mathbf{V} -continuous, then $F := \mathbf{sh}_{\mathbf{V}}(f)$ is continuous.

Proof By Local Transfer 1.10(6), there is $\mathbf{V}' \supset \mathbf{V}$ such that $F = \mathbf{sh}_{\mathbf{V}'}(f)$. For $r, r' \in [a, b] = \text{dom } F$, $r \simeq_{\mathbf{V}} r'$ and $r, r' \in \mathbf{V}'$, this implies that there exist $x \simeq_{\mathbf{V}'} r$, $x' \simeq_{\mathbf{V}'} r'$ such that $F(r) \simeq_{\mathbf{V}'} f(x)$, $F(r') \simeq_{\mathbf{V}'} f(x')$. Then $x \simeq_{\mathbf{V}} x'$, hence $f(x) \simeq_{\mathbf{V}} f(x')$ by \mathbf{V} -continuity of f , and so $F(r) \simeq_{\mathbf{V}} F(r')$. The statement just proved for \mathbf{V} and a particular $\mathbf{V}' \supset \mathbf{V}$:

$$(\forall r, r' \in [a, b] \cap \mathbf{V}')(r \simeq_{\mathbf{V}} r' \rightarrow F(r) \simeq_{\mathbf{V}} F(r'))$$

is true for all $\mathbf{V}' \supset \mathbf{V}$ by Polytransfer 1.10(8). Hence $r \simeq_{\mathbf{V}} r' \rightarrow F(r) \simeq_{\mathbf{V}} F(r')$ holds for all $r, r' \in \text{dom } F$, and F is (uniformly) continuous. \square

Theorem 2.13 For every function f there is a finite set $\{F_0, \dots, F_n\}$ such that $\mathbf{V}(\cdot) = \mathbf{V}(F_0) \subset \mathbf{V}(F_1) \subset \dots \subset \mathbf{V}(F_n)$ and $F_i = \mathbf{sh}_{\mathbf{V}(F_i)}(f)$ for all $\mathbf{V}(F_i) \subseteq \mathbf{V} \subset \mathbf{V}(F_{i+1})$ [all $\mathbf{V}(F_i) \subseteq \mathbf{V}$ if $i = n$].

Proof We consider a statement about a variable level \mathbf{U} :

“There is a finite set $\{F_0, \dots, F_n\}$ such that $F_0 \in \mathbf{U} \subset \mathbf{V}(F_1) \subset \dots \subset \mathbf{V}(F_n)$, $F_0 = \mathbf{sh}_{\mathbf{V}}(f)$ for all $\mathbf{U} \subseteq \mathbf{V} \subset \mathbf{V}(F_1)$ [all $\mathbf{U} \subseteq \mathbf{V}$ if $n = 0$], and $F_i = \mathbf{sh}_{\mathbf{V}}(f)$ for all $\mathbf{V}(F_i) \subseteq \mathbf{V} \subset \mathbf{V}(F_{i+1})$ [all $\mathbf{V}(F_i) \subseteq \mathbf{V}$ if $i = n$], $1 \leq i \leq n$.”

The statement is true if $\mathbf{U} = \mathbf{V}(f)$: let $n := 0$ and $F_0 := \mathbf{sh}_{\mathbf{U}}(f)$, and notice that, for all $\mathbf{V} \supseteq \mathbf{U}$, $F_0 = \mathbf{sh}_{\mathbf{U}}(f) = \mathbf{sh}_{\mathbf{V}}(f)$, by Proposition 2.7(d).

By Granularity, there is a coarsest \mathbf{U} about which the statement is true. If $\mathbf{U} = \mathbf{V}(\cdot)$, we are done; so we assume $\mathbf{V}(\cdot) \subset \mathbf{U}$ and deduce a contradiction.

We have $F_0 \in \mathbf{U}$ and $F_0 = \mathbf{sh}_{\mathbf{V}}(f)$ for all $\mathbf{U} \subseteq \mathbf{V} \subset \mathbf{V}(F_1)$. If $\mathbf{V}(F_0) \subset \mathbf{U}$, let $\bar{\mathbf{U}} := \mathbf{V}(F_0)$; then also for all $\bar{\mathbf{U}} \subseteq \mathbf{V} \subset \mathbf{U}$ we have $\mathbf{sh}_{\mathbf{V}}(f) = \mathbf{sh}_{\mathbf{V}}(\mathbf{sh}_{\bar{\mathbf{U}}}(f))$ [Proposition 2.8] = $\mathbf{sh}_{\mathbf{V}}(F_0) = F_0$ [Proposition 2.7(b, a)]. So $\{F_0, \dots, F_n\}$ witnesses the validity of the statement about $\bar{\mathbf{U}} \subset \mathbf{U}$, a contradiction. We conclude that $\mathbf{V}(F_0) = \mathbf{U}$. By Standardization, there is $\bar{\mathbf{U}} \subset \mathbf{V}(F_0)$ and $G \in \bar{\mathbf{U}}$ such that, for all $\bar{\mathbf{U}} \subseteq \mathbf{V}' \subset \mathbf{V}(F_0)$,

$$(\forall r, s \in \mathbf{V}')[(r, s) \in G \leftrightarrow r, s \in \mathbb{R} \wedge (\exists x)(r \simeq_{\mathbf{V}'} x \wedge s \simeq_{\mathbf{V}'} f(x))];$$

in other words, $G = \mathbf{sh}_{\mathbf{V}'}(f)$ for all $\bar{\mathbf{U}} \subseteq \mathbf{V}' \subset \mathbf{V}(F_0)$. The finite set $\{G, F_0, \dots, F_n\}$ then witnesses the validity of the statement about $\bar{\mathbf{U}} \subset \mathbf{U}$, again a contradiction. \square

Before proving the second part of Theorem 2.10, we need some preliminary results.

Lemma 2.14 *Let $L = \{v_0, v_1, \dots, v_n\}$ be a finite set with $\mathbf{V}(v_0) \subset \mathbf{V}(v_1) \subset \dots \subset \mathbf{V}(v_n)$. For every level \mathbf{V} there is a set $L_{\mathbf{V}} := \{v_i \in L : v_i \in \mathbf{V}\}$.*

Remark This is not trivial, because $v_i \in \mathbf{V}$ is not an internal formula. For example, if $\ell \in \mathbb{N}$ and $\ell \notin \mathbf{V}$, then $\{i \in \mathbb{N} : i \in \mathbf{V}\}$ is not a set.

Proof The statement is true for every \mathbf{V} such that $\mathbf{V}(v_n) \subseteq \mathbf{V}$ with $L_{\mathbf{V}} = L$, hence by Granularity there is a coarsest level $\bar{\mathbf{V}}$ such that the statement is true for all $\mathbf{V} \supseteq \bar{\mathbf{V}}$. If $L_{\bar{\mathbf{V}}} = \{v_0, \dots, v_m\}$ with $m > 0$, then $L_{\mathbf{V}}$ exists also for all $\bar{\mathbf{V}} \supset \mathbf{V} \supseteq \mathbf{V}(v_{m-1})$ [it is either $\{v_0, \dots, v_{m-1}\}$ or $\{v_0, \dots, v_m\}$] and we have a contradiction with the choice of $\bar{\mathbf{V}}$. If $L_{\bar{\mathbf{V}}} = \{v_0\}$ or $L_{\bar{\mathbf{V}}} = \emptyset$, then $L_{\mathbf{V}}$ exists also for all $\mathbf{V} \subset \bar{\mathbf{V}}$ [it is either \emptyset or $\{v_0\}$] and again we have a contradiction, unless $\bar{\mathbf{V}} = \mathbf{V}(\cdot)$ and the statement holds for all \mathbf{V} . \square

Proposition 2.15 *For every finite set $\{v_0, \dots, v_n\}$ with $\mathbf{V}(v_0) \subset \dots \subset \mathbf{V}(v_n)$ there is a set $\{k_0, \dots, k_n\} \subseteq \mathbb{N}$ such that $\mathbf{V}(k_j) = \mathbf{V}(v_j)$ for all $j \leq n$.*

Proof We fix $\{v_0, \dots, v_n\}$ and consider the following statement about \mathbf{V} :

If i is the least such that $\mathbf{V} \subset \mathbf{V}(v_i)$ [or $i = n + 1$ if $\mathbf{V}(v_n) \subseteq \mathbf{V}$], then there exists $S = \{k_i, \dots, k_n\}$ [or $S = \emptyset$, resp] such that $\mathbf{V}(k_j) = \mathbf{V}(v_j)$ for all $i \leq j \leq n$.

The statement is trivially true for $\mathbf{V} \supseteq \mathbf{V}(v_n)$. So by Granularity there is a coarsest level $\bar{\mathbf{V}}$ for which it is true. If $\bar{\mathbf{V}} \subset \mathbf{V}(v_0)$ we are done. If $\mathbf{V}(v_0) \subseteq \bar{\mathbf{V}} \subset \mathbf{V}(v_1)$, we have $\{k_1, \dots, k_n\}$ such that $\mathbf{V}(k_j) = \mathbf{V}(v_j)$ for $1 \leq j \leq n$. We fix some k_0 such that $\mathbf{V}(k_0) = \mathbf{V}(v_0)$ [Lemma 2.2(a)], let $S = \{k_0, k_1, \dots, k_n\}$, and we are done. Finally, let $\mathbf{V}(v_i) \subseteq \bar{\mathbf{V}} \subset \mathbf{V}(v_{i+1})$ for $i > 0$ [by Lemma 2.14, these are the only possibilities]. Then there is a set $\{k_{i+1}, \dots, k_n\}$ such that $\mathbf{V}(k_j) = \mathbf{V}(v_j)$ for $i + 1 \leq j \leq n$. We fix k_i such that $\mathbf{V}(k_i) = \mathbf{V}(v_i)$ and let $S = \{k_i, k_{i+1}, \dots, k_n\}$. The statement holds for $\mathbf{V} := \mathbf{V}(v_{i-1}) \subset \bar{\mathbf{V}}$ with this S , a contradiction. \square

Corollary 2.16 For every finite set $\{v_0, \dots, v_n\}$ with $\mathbf{V}(v_0) \subset \dots \subset \mathbf{V}(v_n)$ there is a set $\{h_0, \dots, h_n\} \subseteq \mathbb{R}$ such that, for all $j \leq n$, $\mathbf{V}(h_j) = \mathbf{V}(v_j)$ and h_j is ultrasmall relative to all $\mathbf{V} \subset \mathbf{V}(v_j)$.

Proof Let $\{k_0, \dots, k_n\} \subseteq \mathbb{N}$ be as in Proposition 2.15. Let $h_j := 1/k_j$. Clearly $\mathbf{V}(h_j) = \mathbf{V}(k_j) = \mathbf{V}(v_j)$. If $\mathbf{V} \subset \mathbf{V}(v_j)$ and $k \in \mathbb{N} \cap \mathbf{V}$, then $k < k_j$ and $0 < h_j < 1/k$. This shows h_j is \mathbf{V} -ultrasmall. \square

We can now prove the second part of Theorem 2.10.

Proof Let $\{h_0, \dots, h_n\}$ be as in Corollary 2.16; as $\mathbf{V}(v_0) = \mathbf{V}(\cdot)$, we can take $h_0 = 1$.

For $1 \leq i \leq n$ let $f_i : [0, 1] \rightarrow [0, 1]$ be defined by

$$f_i(x) = \begin{cases} \frac{2h_{i-1}}{h_i} \cdot x & \text{if } 0 \leq x \leq h_i/2; \\ -\frac{2h_{i-1}}{h_i} \cdot x + 2h_{i-1} & \text{if } h_i/2 < x \leq h_i; \\ 0 & \text{if } h_i < x \leq 1. \end{cases}$$

If n is odd, let $f(x) := \sum_{i=1}^n f_i(x - \frac{1}{i+1})$. It is easy to check that f is not \mathbf{V} -continuous for $\mathbf{V} \subset \mathbf{V}(v_1)$ and for $\mathbf{V}(v_i) \subseteq \mathbf{V} \subset \mathbf{V}(v_{i+1})$ when i is even, and is \mathbf{V} -continuous for $\mathbf{V}(v_i) \subseteq \mathbf{V} \subset \mathbf{V}(v_{i+1})$ when i is odd.

If n is even, take $\sum_{i=2}^n$ instead; the pattern of continuity versus not-continuity is reversed. Both cases are easily modified to produce the opposite pattern. \square

The notion of \mathbf{V} -continuity has a natural generalization.

Definition 2.17 A function f is (uniformly) $(\mathbf{V}_1, \mathbf{V}_2)$ -**continuous** if $x \simeq_{\mathbf{V}_1} x'$ implies $f(x) \simeq_{\mathbf{V}_2} f(x')$, for all $x, x' \in \text{dom} f$.

We forgo the detailed study of $(\mathbf{V}_1, \mathbf{V}_2)$ -continuity and prove only the basic result.

Proposition 2.18 *If $f \in \mathbf{V}_2$ and f is $(\mathbf{V}_1, \mathbf{V}_2)$ -continuous, then f is continuous.*

Proof If $\mathbf{V}_1 \subseteq \mathbf{V}_2$ and f is $(\mathbf{V}_1, \mathbf{V}_2)$ -continuous, then f is \mathbf{V} -continuous for all $\mathbf{V}_1 \subseteq \mathbf{V} \subseteq \mathbf{V}_2$, in particular, it is \mathbf{V}_2 -continuous. If also $f \in \mathbf{V}_2$, then f is continuous.

The case $\mathbf{V}_2 \subset \mathbf{V}_1$ is less trivial. Let $x \simeq_{\mathbf{V}_1} x'$, $x, x' \in \text{dom} f$. By Density of Levels [Proposition 2.4], there is \mathbf{V}_1^+ such that $\mathbf{V}_1 \subset \mathbf{V}_1^+$ and $x \simeq_{\mathbf{V}_1^+} x'$, and by Polytransfer 1.10(8), f is also $(\mathbf{V}_1^+, \mathbf{V}_1)$ -continuous; hence $f(x) \simeq_{\mathbf{V}_1} f(x')$. \square

3 Lebesgue measure in GRIST.

This section is concerned with some aspects of Lebesgue measure and integral. Our goal is to showcase the tools available in **GRIST** for dealing with such matters, not to give a systematic development of the theory of integration. For this reason, we limit ourselves to the representative simplest case, that of Lebesgue measure on $[0, 1]$. A more complete treatment of the theory of measure and integration in (a weak subsystem of) **GRIST** can be found in [18].

We give three nonstandard definitions of Lebesgue measure on $[0, 1]$ and prove their equivalence to the usual one. Letters A, B denote subsets of $[0, 1]$, and I, J are finite non-degenerate intervals; $\ell(I)$ is the length of I .

Definition 3.1 (Lebesgue)

$$m_0(A) := \inf \left\{ \sum_{i=0}^{\infty} \ell(I_i) : A \subseteq \bigcup_{i=0}^{\infty} I_i, \text{ where } \{I_i\}_{i=0}^{\infty} \text{ is a system of open intervals} \right\}.$$

It is well-known that m_0 is an outer measure on $\mathcal{P}([0, 1])$, $m_0(I) = \ell(I)$ for all $I \subseteq [0, 1]$, and a set A is Lebesgue measurable if and only if $m_0(A) + m_0([0, 1] \setminus A) = 1$.

The most useful nonstandard approach to integration is due to Loeb (see eg [1]). Loeb showed that every finitely additive measure μ on an algebra of sets \mathcal{A} gives rise to an *external* (countably additive) measure $\mathbf{L}(\mu)$ on an *external* σ -algebra $\mathbf{L}(\mathcal{A})$ generated by \mathcal{A} . This construction uses external sets in an essential way (see Section 4). Loeb

also showed that Lebesgue measure can be obtained as an image of a suitable Loeb measure by the “standard part” mapping (“neighbor” \mathbf{n} in our notation). There are several choices of \mathcal{A} and μ natural for this purpose.

One elegant approach is to “approximate” $[0, 1]$ by a “hyperfinite time line.” We fix a level \mathbf{V} and $N \in \mathbb{N}$ ultralarge relative to \mathbf{V} , let $t_i := i/N$ for $i = 0, \dots, N$, $T := \{t_0, t_1, \dots, t_N\}$, $\mathcal{A} := \mathcal{P}(T)$, and let μ be the counting measure on \mathcal{A} , ie, $\mu(X) := |X|/|T|$ for all $X \in \mathcal{A}$. It can then be shown (see [1]) that a set $A \in \mathbf{V}$ is Lebesgue measurable if and only if $\mathbf{n}_{\mathbf{V}}^{-1}[A] \cap T$ is Loeb measurable, and the Lebesgue measure of A is the Loeb measure of $\mathbf{n}_{\mathbf{V}}^{-1}[A] \cap T$.

Loeb measure is obtained from the externally finitely additive measure $\mathbf{n}_{\mathbf{V}} \circ \mu$ on \mathcal{A} by Carathéodory extension theorem. Explicitly, the Loeb measure of an external set $\mathbf{E} \subseteq [0, 1]$ is the (external) infimum of $\mathbf{n}_{\mathbf{V}}(\mu(X))$ taken over all (internal) sets $X \supseteq \mathbf{E}$, $X \in \mathcal{A}$. Hence the Lebesgue measure $m(A)$ of $A \in \mathbf{V}$ according to Loeb is given by

$$(*) \quad m(A) := \mathbf{inf}_{\mathbf{V}} \left\{ \mathbf{n}_{\mathbf{V}}(\mu(X)) : \mathbf{n}_{\mathbf{V}}^{-1}[A] \cap T \subseteq X, X \in \mathcal{A} \right\}.$$

The collection $\mathbf{n}_{\mathbf{V}}^{-1}[A]$ is usually *external*—it is not a set. Similarly, the collection of numbers whose infimum needs to be taken to obtain $m(A)$ is external; we indicate this by the use of boldface braces. The notation $\mathbf{inf}_{\mathbf{V}}$ stands for the greatest lower bound in \mathbf{V} , that is, in the external set $\mathbb{R} \cap \mathbf{V}$, ordered by $<$. Extensions of **GRIST** that formally allow external sets are discussed in Section 4. However, it turns out that the use of external sets in (*) and elsewhere in this section is only for notational convenience, and could be eliminated (see Remark 3 below). Nevertheless the question arises whether such external collections do have a supremum and infimum. The positive answer is provided by the following proposition.

Proposition 3.2 *Let $\mathcal{P}(r, \bar{x}; \mathbf{V})$ be a \mathbf{V} -formula. If $\{r \in \mathbb{R} \cap \mathbf{V} : \mathcal{P}(r, \bar{x}; \mathbf{V})\}$ is nonempty and has a lower bound [upper bound, resp.] in \mathbf{V} , then it has a greatest lower bound [least upper bound, resp.] in \mathbf{V} .*

Proof By **FRIST** Standardization, there is a set $S \in \mathbf{V}$ such that $(\forall r \in \mathbf{V})(r \in S \leftrightarrow r \in \mathbb{R} \wedge \mathcal{P}(r, \bar{x}; \mathbf{V}))$. Using Transfer, it is easy to verify that S is nonempty and bounded below, and that $\inf S$ has the desired properties. \square

Our objective in this section is to compare Loeb’s construction with other definitions of Lebesgue measure, which are stated in terms of systems of intervals. It is trivial to see that (*) is equivalent to

$$(**) \quad m(A) := \mathbf{inf}_{\mathbf{V}} \left\{ \mathbf{n}_{\mathbf{V}} \left(\sum_{i=0}^n \ell(I_i) \right) : \mathbf{n}_{\mathbf{V}}^{-1}[A] \cap T \subseteq \bigcup_{i=0}^n I_i \right\},$$

where $\{I_i\}_{i=0}^n$ is any finite system of intervals of the form $[t_k, t_{k+1})$ (or $[t_{N-1}, t_N]$ if $k = N - 1$).

For the purposes of this section it is preferable to use a variant of $(**)$ where $\mathbf{n}_{\mathbf{V}}^{-1}[A] \cap T$ is replaced by $\mathbf{n}_{\mathbf{V}}^{-1}[A] \cap [0, 1]$ and arbitrary systems of intervals are allowed. This leads to the following definition.

Definition 3.3 (Loeb)

$$m_1(A) := \mathbf{inf}_{\mathbf{V}(A)} \left\{ \mathbf{n}_{\mathbf{V}(A)} \left(\sum_{i=0}^n \ell(I_i) \right) : \mathbf{n}_{\mathbf{V}(A)}^{-1}[A] \cap [0, 1] \subseteq \bigcup_{i=0}^n I_i \right\}.$$

Remarks

(1) In this definition, $\{I_i\}_{i=0}^n$ ranges over all finite systems of intervals. It is easy to see that $m_1(A)$ remains unchanged if we restrict it to finite systems of open intervals, or to finite systems of non-overlapping closed subintervals of $[0, 1]$.

(2) The above remarks about the existence of $\mathbf{inf}_{\mathbf{V}}$ apply.

(3) Because of the appearance of external sets in Definition 3.3, it is not immediately obvious that m_1 is a function (ie, a set). However, this follows from the Definition Principle, because the statement $m_1(A) = r$, when spelled out in full, is equivalent to an internal formula. Let $\mathcal{Q}(A, r; \mathbf{V})$ be the \mathbf{V} -formula expressing “for every finite system of intervals $\{I_i\}_{i=0}^n$, if $(\forall x, \xi \in [0, 1])(x \in A \cap \mathbf{V} \wedge \xi \simeq_{\mathbf{V}} x \rightarrow \xi \in \bigcup_{i=0}^n I_i)$, then $r \leq \mathbf{n}_{\mathbf{V}} \left(\sum_{i=0}^n \ell(I_i) \right)$.” Then $m_1(A) = r$ is equivalent to $(\forall \mathbf{V})(A, r \in \mathbf{V} \rightarrow [\mathcal{Q}(A, r; \mathbf{V}) \wedge (\forall s \in \mathbb{R} \cap \mathbf{V})(\mathcal{Q}(A, s; \mathbf{V}) \rightarrow s \leq r)])$, which is an internal formula.

(4) Henson [11] proved that Loeb’s definition and Lebesgue’s definition of outer measure agree, ie, $m_0 = m$.

(5) We prove that $m(A) = m_1(A)$.

Let $\mathbf{V} := \mathbf{V}(A)$. First, note that $\mathbf{n}_{\mathbf{V}}^{-1}[A] \cap T$ can be replaced by $\mathbf{n}_{\mathbf{V}}^{-1}[A] \cap [0, 1]$ in $(**)$. Indeed, if $\mathbf{n}_{\mathbf{V}}^{-1}[A] \cap T \subseteq \bigcup_{i=0}^n I_i$ and $\xi \in \mathbf{n}_{\mathbf{V}}^{-1}[A] \cap [0, 1]$, then $\xi \simeq_{\mathbf{V}} a$ for some $a \in A \cap \mathbf{V}$, and $t_j \leq \xi < t_{j+1}$ for some j . Then also $t_j \simeq_{\mathbf{V}} a$, so $t_j \in \mathbf{n}_{\mathbf{V}}^{-1}[A] \cap T$, the interval $[t_j, t_{j+1}) = I_i$ for some i , and $\xi \in \bigcup_{i=0}^n I_i$; hence $\mathbf{n}_{\mathbf{V}}^{-1}[A] \cap [0, 1] \subseteq \bigcup_{i=0}^n I_i$.

It is now clear that $m_1(A) \leq m(A)$. We assume $m_1(A) < m(A) =: s$ and obtain a contradiction. From the definition of $m_1(A)$ and Remark (1) it follows that there is a

finite system of non-overlapping intervals $\{J_j\}_{j=0}^m$ such that $\mathbf{n}_V^{-1}[A] \cap [0, 1] \subseteq \bigcup_{j=0}^m J_j$ and $\mathbf{n}_V(\sum_{j=0}^m \ell(J_j)) < s$. By Relativization (v), there is a level $\mathbf{V}' \supset \mathbf{V}$ such that $N \notin \mathbf{V}'$. By Polytransfer 1.10(8), we can assume that the system $\{J_j\}_{j=0}^m$ with the above properties is in \mathbf{V}' . Let $\{I_i\}_{i=0}^n$ be the collection of all intervals of the form $[t_k, t_{k+1})$ that have a nonempty intersection with some J_j . Then $\mathbf{n}_V^{-1}[A] \cap [0, 1] \subseteq \bigcup_{j=0}^m J_j \subseteq \bigcup_{i=0}^n I_i$ and $\sum_{j=0}^m \ell(J_j) \leq \sum_{i=0}^n \ell(I_i) \leq \sum_{j=0}^m \ell(J_j) + 2(m+1)/(N+1)$, so $m(A) \leq \mathbf{n}_V(\sum_{i=0}^n \ell(I_i)) < s$ [note $m \in \mathbf{V}$, $N \notin \mathbf{V}$], a contradiction.

In order to motivate the next two definitions of Lebesgue measure (Definitions 3.9 and 3.10), we briefly summarize the nonstandard approach to the Riemann integral and the integrals of McShane and Henstock-Kurzweil. For the standard theory of these integrals see for example Bartle [4] and Pfeffer [24].

Definition 3.4 A **tagged interval** is a pair (I, t) where I is a closed interval and $t \in \mathbb{R}$. A **tagged covering** is a finite system $\mathcal{J} := \{(I_i, t_i)\}_{i=1}^n$ of tagged intervals. A **tagged partition** is a tagged covering where $\{I_i\}_{i=1}^n$ are non-overlapping. It is a **partition of** $[0, 1]$ if $\bigcup_{i=1}^n I_i = [0, 1]$; similarly for coverings. A tagged covering is **anchored in** A if $\{t_i\}_{i=1}^n \subseteq A$. A tagged partition is a **Riemann partition** if $t_i \in I_i$ for all $i = 1, \dots, n$. A Riemann partition \mathcal{J} is **fine** relative to \mathbf{V} if all $\ell(I_i)$ are ultrasmall relative to \mathbf{V} .

Definition 3.5 Let $f : [0, 1] \rightarrow \mathbb{R}$ be a function.

For all tagged coverings \mathcal{J} , the **Riemann sum** $\sum(f; \mathcal{J})$ is defined as $\sum_{i=1}^n f(t_i) \cdot \ell(I_i)$. The function f on $[0, 1]$ is **Riemann integrable** if there is $R \in \mathbb{R} \cap \mathbf{V}(f)$ such that $\sum(f; \mathcal{J}) \simeq_{\mathbf{V}(f)} R$ for all Riemann partitions \mathcal{J} of $[0, 1]$ that are fine relative to $\mathbf{V}(f)$. If this is the case, we let $\int_0^1 f(x) \cdot dx := R = \mathbf{n}_V(\sum(f; \mathcal{J}))$.

One of the inadequacies of Riemann integration is that it is not a true inverse to the operation of differentiation: If f' is continuous, then f' is Riemann integrable and the (indefinite) integral gives back f (up to a constant), but f' need not be Riemann integrable in general. Relative analysis sheds some light on the reasons for this phenomenon. Let us consider a differentiable function f and a fine Riemann partition \mathcal{J} of $[0, 1]$, where $I_i = [x_{i-1}, x_i]$, $0 = x_0 < x_1 < \dots < x_n = 1$, and $x_{i-1} \leq t_i \leq x_i$, for all $i = 1, \dots, n$. If f is uniformly differentiable on $[0, 1]$, or equivalently, f' is continuous, one proves easily that

$$(*) \quad f(x_i) - f(x_{i-1}) = f'(t_i) \cdot \ell(I_i) + \epsilon_i \cdot \ell(I_i) \text{ where } \epsilon_i \simeq_{\mathbf{V}(f)} 0.$$

Adding these equations then gives

$$f(1) - f(0) = \sum_{i=1}^n f'(t_i) \cdot \ell(I_i) + \sum_{i=1}^n \epsilon_i \cdot \ell(I_i)$$

where $\sum \epsilon_i \cdot \ell(I_i) \simeq_{\mathbf{V}(f)} 0$, and hence f' is integrable and $\int_0^1 f'(x) \cdot dx = f(1) - f(0)$.

However, if f is merely “pointwise” differentiable, (*) is valid only under the assumption that $\ell(I_i)$ is ultrasmall relative to $\mathbf{V}(f, t_i)$; the assumption that $\ell(I_i)$ is ultrasmall relative to $\mathbf{V}(f)$ is not sufficient. These considerations suggest that every derivative would become integrable if we used, in place of fine Riemann partitions, *superfine partitions*, defined as those tagged partitions where each $\ell(I_i)$ is ultrasmall relative to the level $\mathbf{V}(f, t_i)$, dependent on t_i . It turns out that superfine partitions of $[0, 1]$ in this strong sense do not exist—see Theorem 3.12 in the Appendix to this section. The idea does work if one employs instead a weaker notion of relative ultrasmallness due to Benninghofen and Richter [6] and Gordon [9, 10].

Definition 3.6 Given $a \in \mathbb{R}$, we say that a real number ϵ is a -**ultrasmall** relative to \mathbf{V} if $|\epsilon| < \varphi(a)$ for all positive functions $\varphi \in \mathbf{V}$ defined on \mathbb{R} .

Definition 3.7 A tagged covering \mathfrak{J} is **superfine** relative to \mathbf{V} if $|x - t_i|$ is t_i -ultrasmall relative to \mathbf{V} for all $x \in I_i$ and all $i = 1, \dots, n$. For Riemann partitions, this is equivalent to $\ell(I_i)$ being t_i -ultrasmall, for all $i = 1, \dots, n$.

Let φ be a positive function on $[0, 1]$. We say that a tagged covering $\{(I_i, t_i)\}_{i=1}^n$ is **subordinate to** φ if $(\forall x \in I_i)(|x - t_i| < \varphi(t_i))$, for all $i = 1, \dots, n$. A well-known classical result (Cousin’s Lemma) states that for each positive φ there exist Riemann partitions of $[0, 1]$ subordinate to φ . The existence of Riemann partitions of $[0, 1]$ superfine relative to \mathbf{V} follows from this by Idealization.

If the word “fine” is replaced by “superfine” in Definition 3.5, one obtains a notion of integral that is equivalent to the one introduced by Henstock and Kurzweil. The standard definition is as follows.

Definition 3.8 A function f on $[0, 1]$ is **Henstock-Kurzweil integrable** if there is a real number R such that for every $\epsilon > 0$ there is a positive function φ such that $|\sum(f; \mathfrak{J}) - R| < \epsilon$ holds for all Riemann partitions \mathfrak{J} of $[0, 1]$ subordinate to φ .

Henstock-Kurzweil integral agrees with Lebesgue integral on nonnegative functions (more generally, on absolutely integrable functions), but there exist functions that are Henstock-Kurzweil integrable but not Lebesgue integrable; in particular, all derivatives are Henstock-Kurzweil integrable. The nonstandard theory of Henstock-Kurzweil integral is worked out in some detail in Benninghofen [5] and in [18].

A weaker notion of integral is obtained from the (standard or nonstandard) definition of Henstock-Kurzweil integral by replacing the words “Riemann partitions” by “partitions” [so the tags t_i need not belong to I_i]. This yields an integral due to McShane, which is known to be equivalent to Lebesgue integral.

The next two definitions of the outer measure of a set A are motivated by the nonstandard definitions of the integral of the characteristic function of A , in the sense of McShane and Henstock-Kurzweil, respectively.

Definition 3.9 (McShane)

$$m_2(A) := \sup_{\mathbf{V}(A)} \left\{ \mathbf{n}_{\mathbf{V}(A)} \left(\sum_{i=1}^n \ell(I_i) \right) : \{(I_i, t_i)\}_{i=1}^n \text{ is a tagged partition anchored in } A \text{ and superfine relative to } \mathbf{V}(A) \right\}.$$

Definition 3.10 (Henstock-Kurzweil) $m_3(A)$ is defined as above, with “partition” replaced by “Riemann partition.”

We illustrate the capabilities of **GRIST** by giving a direct proof of the fact that the above four definitions of Lebesgue measure are indeed equivalent.

Theorem 3.11 $m_0(A) = m_1(A) = m_2(A) = m_3(A)$, for all $A \subseteq [0, 1]$.

Proof $m_2(A) \leq m_0(A)$:

It suffices to show: If $A \subseteq \bigcup_{i=0}^{\infty} I_i$ where $I_i = (a_i, b_i)$ are open intervals and $\{I_i\}_{i=0}^{\infty} \in \mathbf{V}(A)$, and if $\{(J_j, t_j)\}_{j=1}^k$ is a tagged partition anchored in A and superfine relative to $\mathbf{V}(A)$, then $\sum_{j=1}^k \ell(J_j) \leq \sum_{i=0}^{\infty} \ell(I_i)$.

We define $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ as follows:

$$\varphi(x) := \begin{cases} \min\{x - a_i, b_i - x\} & \text{where } i \text{ is the least such that } x \in I_i; \\ 1 & \text{otherwise.} \end{cases}$$

Thus $\varphi \in \mathbf{V}(A)$ is a positive function. For every $1 \leq j \leq k$, $(\forall x \in J_j)(|x - t_j| \text{ is } t_j\text{-ultrasmall relative to } \mathbf{V}(A))$, hence $(\forall x \in J_j)(|x - t_j| < \varphi(t_j))$. As $t_j \in A \subseteq \bigcup I_i$, this implies that $J_j \subseteq I_i$, for the least i such that $t_j \in I_i$. The intervals J_j are non-overlapping, and thence $\sum_{j=1}^k \ell(J_j) \leq \sum_{i=0}^{\infty} \ell(I_i)$.

$m_1(A) \leq m_2(A)$:

Let us assume that $A \subseteq (0, 1)$ and $m_2(A) < m < m_1(A)$ holds for some $m \in \mathbf{V}(A)$, and deduce a contradiction. If for every positive $\varphi \in \mathbf{V}(A)$ there is a tagged partition $\{(I_i, t_i)\}_{i=1}^n$ subordinate to φ , anchored in A , and such that $\sum \ell(I_i) > m$, then by Idealization there is a tagged partition $\{(I_i, t_i)\}_{i=1}^n$ superfine relative to $\mathbf{V}(A)$, anchored in A , and such that $\sum \ell(I_i) > m$, a contradiction. Hence there is a positive $\varphi \in \mathbf{V}(A)$ such that $\sum \ell(I_i) \leq m$ for all tagged partitions subordinate to φ and anchored in A .

For each $a \in A$ let $J_a := [a - \frac{1}{2}\varphi(a), a + \frac{1}{2}\varphi(a)]$. Let F be a finite set such that $A \cap \mathbf{V}(A) \subseteq F \subseteq A$ [Idealization]. If $a \in A \cap \mathbf{V}(A)$ and $x \simeq_{\mathbf{V}(A)} a$, then $x \in J_a$, so $\mathbf{n}_{\mathbf{V}(A)}^{-1}[A] \subseteq \bigcup_{a \in F} J_a$. It is now easy to construct a tagged partition $\{(I_i, t_i)\}_{i=1}^n$ with $\{t_i\}_{i=1}^n \subseteq F$ such that $\bigcup_{i=1}^n I_i = \bigcup_{a \in F} J_a$ and $I_i \subseteq J_{t_i}$ for all $i = 1, \dots, n$. In particular, $(\forall x \in I_i)(|x - t_i| < \varphi(t_i))$, so $\{(I_i, t_i)\}_{i=1}^n$ is subordinate to φ and anchored in A , and hence $\sum \ell(I_i) \leq m$. As also $\mathbf{n}_{\mathbf{V}(A)}^{-1}[A] \subseteq \bigcup_{i=1}^n I_i$, we have a contradiction with $m < m_1(A)$.

$m_0(A) \leq m_1(A)$:

First let A be compact. Then $A \subseteq \mathbf{n}_{\mathbf{V}(A)}^{-1}[A]$, hence $\mathbf{n}_{\mathbf{V}(A)}^{-1}[A] \subseteq \bigcup_{i=0}^n I_i$ implies $A \subseteq \bigcup_{i=0}^n I_i$ and (using properties of Lebesgue measure) $m_0(A) \leq \sum_{i=0}^n m_0(I_i) = \sum_{i=0}^n \ell(I_i)$. Hence $m_0(A) \leq m_1(A)$.

Now let A be open. From properties of Lebesgue measure it follows that $m_0(A) = \sup\{m_0(F) : F \subseteq A, F \text{ compact}\} \leq \sup\{m_1(F) : F \subseteq A, F \text{ compact}\} \leq m_1(A)$, using monotonicity of m_1 , evident from Definition 3.3 and Transfer.

Finally, let A be arbitrary. We define the *inner measure* m_1^- by

$$m_1^-(A) := \sup_{\mathbf{V}(A)} \left\{ \mathbf{n}_{\mathbf{V}(A)} \left(\sum_{i=0}^n \ell(I_i) \right) : \bigcup_{i=0}^n I_i \subseteq \mathbf{n}_{\mathbf{V}(A)}^{-1}[A] \text{ where } I_i \text{ are non-overlapping closed subintervals of } [0, 1] \right\}.$$

It is immediate from the definitions that if $A \cup C = [0, 1]$ and $A \cap C = \emptyset$, then $m_1(A) + m_1^-(C) = 1$.

Let $\{I_i\}_{i=0}^n$ be a system of non-overlapping closed subintervals of $[0, 1]$ such that $\bigcup_{i=0}^n I_i \subseteq \mathbf{n}_{\mathbf{V}(A)}^{-1}[A]$. Let $F := \mathbf{sh}_{\mathbf{V}(A)}(\bigcup_{i=0}^n I_i) \in \mathbf{V}(A)$. F is compact [Proposition 2.7(b)], $\bigcup_{i=0}^n I_i \subseteq \mathbf{n}_{\mathbf{V}(A)}^{-1}[F]$, and $F \subseteq A$. We conclude that, for any A , $m_1^-(A) = \sup\{m_1^-(F) : F \subseteq A, F \text{ compact}\}$, hence, by taking complements, $m_1(A) = \inf\{m_1(U) : A \subseteq U, U \text{ open}\} \geq \inf\{m_0(U) : A \subseteq U, U \text{ open}\} = m_0(A)$.

$m_3(A) \leq m_2(A)$: Trivial.

$m_2(A) \leq m_3(A)$:

For every tagged partition $\mathfrak{J} = \{(I_i, t_i)\}_{i=1}^n$ anchored in A we construct a Riemann partition $\mathfrak{J} = \{(J_j, s_j)\}_{j=1}^m$ anchored in A with $\bigcup_{i=1}^n I_i \subseteq \bigcup_{j=1}^m J_j$, hence $\sum_{i=1}^n \ell(I_i) \leq \sum_{j=1}^m \ell(J_j)$, in such a way that if \mathfrak{J} is superfine relative to \mathbf{V} , then \mathfrak{J} is also superfine relative to \mathbf{V} . [More strongly, if \mathfrak{J} is subordinate to φ , then \mathfrak{J} is subordinate to φ , for any $\varphi > 0$.] From this result, $m_2(A) \leq m_3(A)$ follows immediately.

Let $I_i = [a_i, b_i]$ and let $\langle r_1, \dots, r_p \rangle$ be a one-one enumeration of $\{t_1, \dots, t_n\}$. If $r_\ell = t_{i_1} = \dots = t_{i_k}$ and $r_\ell \neq t_i$ for $i \neq i_1, \dots, i_k$, let

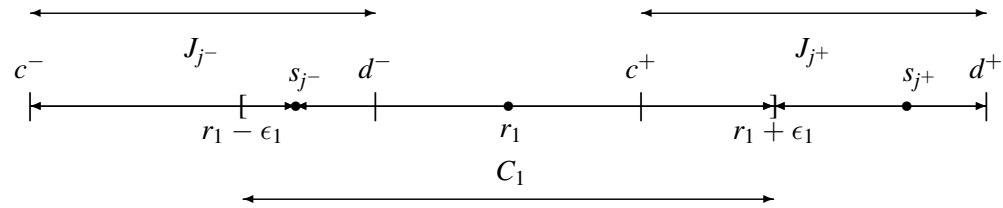
$$\epsilon_\ell := \max\{r_\ell - a_{i_1}, b_{i_1} - r_\ell, \dots, r_\ell - a_{i_k}, b_{i_k} - r_\ell\}.$$

Note that $\epsilon_\ell > 0$, and if \mathfrak{J} is superfine relative to \mathbf{V} , then ϵ_ℓ is r_ℓ -ultrasmall relative to \mathbf{V} [if \mathfrak{J} is subordinate to φ , then $\epsilon_\ell < \varphi(r_\ell)$]. The collection $\mathfrak{C} := \{(C_\ell, r_\ell)\}_{\ell=1}^p$, where $C_\ell := [r_\ell - \epsilon_\ell, r_\ell + \epsilon_\ell]$, is a covering anchored in A and $\bigcup_{i=1}^n I_i \subseteq \bigcup_{\ell=1}^p C_\ell$. Moreover, \mathfrak{C} is superfine relative to \mathbf{V} [subordinate to φ] whenever \mathfrak{J} is.

Claim: There is a Riemann partition $\mathfrak{J} = \{(J_j, s_j)\}_{j=1}^m$ such that $\bigcup_{\ell=1}^p C_\ell = \bigcup_{j=1}^m J_j$ and for each j there is ℓ such that $s_j = r_\ell$ and $J_j \subseteq C_\ell$.

Clearly, if \mathfrak{C} is superfine relative to \mathbf{V} [subordinate to φ], the same holds for \mathfrak{J} , and this concludes the proof.

Proof of Claim. The construction is by recursion. Wlog we assume that $\epsilon_1 = \max\{\epsilon_1, \dots, \epsilon_p\}$. By inductive assumption, for the covering $\{(C_\ell, r_\ell)\}_{\ell=2}^p$ there is $\mathfrak{J} = \{(J_j, s_j)\}_{j=1}^m$ satisfying the Claim. First, we omit from \mathfrak{J} those (J_j, s_j) where $J_j \subseteq C_1$. Let j^- and j^+ be such that $r_1 - \epsilon_1 \in J_{j^-}$ and $r_1 + \epsilon_1 \in J_{j^+}$, if they exist.



We note that $j^- \neq j^+$, because of maximality of ϵ_1 . Let $J_{j^-} = [c^-, d^-]$, $J_{j^+} = [c^+, d^+]$. We note that $s_{j^-} < r_1$ because $s_{j^-} - c^- \leq \epsilon_{j^-} \leq \epsilon_1$, by maximality of ϵ_1 ; similarly $r_1 < s_{j^+}$. Next, we replace in \mathfrak{J} the tagged interval (J_{j^-}, s_{j^-}) by $([c^-, \max\{s_{j^-}, r_1 - \epsilon_1\}], s_{j^-})$, (J_{j^+}, s_{j^+}) by $([\min\{s_{j^+}, r_1 + \epsilon_1\}, d^+], s_{j^+})$, and add to \mathfrak{J} the tagged interval $([\max\{s_{j^-}, r_1 - \epsilon_1\}, \min\{s_{j^+}, r_1 + \epsilon_1\}], r_1)$. If j^+ does not exist, we replace only (J_{j^-}, s_{j^-}) , and add $([\max\{s_{j^-}, r_1 - \epsilon_1\}, r_1 + \epsilon_1], r_1)$; similarly in case j^- does not exist. If neither exists, we only add $C_1 = [r_1 - \epsilon_1, r_1 + \epsilon_1]$. It is easy to see that the resulting Riemann partition has the required properties. \square \square

Appendix

Theorem 3.12 *Superfine coverings of $[0, 1]$ in the strong sense do not exist.*

Lemma 3.13 *If $\langle h_a : a \in A \rangle$ is a finite sequence, $A \in \mathbf{V}$, and h_a is ultrasmall relative to \mathbf{V} for every $a \in A$, then $\sum_{a \in A} h_a$ is ultrasmall relative to \mathbf{V} .*

Proof Let $h := \max\{|h_a| : a \in A\}$ and $n := |A|$. Note that $n \in \mathbf{V}$ and h is ultrasmall relative to \mathbf{V} ; hence $|\sum_{a \in A} h_a| \leq h \cdot n$ is ultrasmall relative to \mathbf{V} . \square

Proof of Theorem 3.12

Let $\{(I_i, t_i)\}_{i=1}^n$ be a tagged covering of $[0, 1]$ such that $(\forall x \in I_i)(|x - t_i|$ is ultrasmall relative to $\mathbf{V}(t_i)$). Wlog we assume that $t_i \in I_i$ and $t_i \neq t_j$, for all $1 \leq i \neq j \leq n$. Indeed, if $I_i = [a_i, b_i]$ and $t_{i_1} = \dots = t_{i_k} = t$, we can replace the tagged intervals $(I_{i_1}, t_{i_1}), \dots, (I_{i_k}, t_{i_k})$ with a single tagged interval (I, t) , where $I := [\min\{a_{i_1}, \dots, a_{i_k}, t\}, \max\{b_{i_1}, \dots, b_{i_k}, t\}]$. We let $T := \{t_i\}_{i=1}^n$ and write I_t in place of I_i for $t = t_i$.

Claim For every \mathbf{V} there is $A \in \mathbf{V}$, $A \subseteq T$, such that $\sum_{t \in T \setminus A} \ell(I_t) \simeq_{\mathbf{V}} 0$.

Proof of Claim We consider the following assertion about \mathbf{V} :

$$(*) \quad (\forall \mathbf{V}' \supseteq \mathbf{V})(\exists A \in \mathbf{V}')(A \subseteq T \wedge \sum_{t \in T \setminus A} \ell(I_t) \simeq_{\mathbf{V}'} 0).$$

Clearly $(*)$ is true if $T \in \mathbf{V}$ [let $A := T$]. By Granularity, there is a coarsest level \mathbf{V} for which $(*)$ is true. We assume $\mathbf{V}(\cdot) \subset \mathbf{V}$ and deduce a contradiction.

Let $A \in \mathbf{V}$, $A \subseteq T$ be such that $\sum_{t \in T \setminus A} \ell(I_t) \simeq_{\mathbf{V}} 0$. By **FRIST** Standardization, there is $B \in \mathbf{V}$ such that

$$(\forall t \in \mathbf{V})(t \in B \leftrightarrow t \in A \wedge \ell(I_t) \text{ is not ultrasmall relative to } \mathbf{V}).$$

We can assume $B \subseteq A$ [replace B with $B \cap A$ if necessary]. As $A \subseteq T$ is finite, $t \in A \rightarrow t \in \mathbf{V}$ [1.10(13)(b)] and we have $(\forall t)(t \in B \leftrightarrow t \in A \wedge \ell(I_t) \text{ is not ultrasmall relative to } \mathbf{V})$. Moreover, there is $\bar{\mathbf{V}} \subset \mathbf{V}$ such that $B \in \bar{\mathbf{V}}$ [if not, then \mathbf{V} would be the coarsest level where B appears, hence the coarsest level where some $t \in B$ appears [1.10(12)], but then $\ell(I_t)$ would be ultrasmall relative to \mathbf{V} , a contradiction].

Now $\sum_{t \in T \setminus B} \ell(I_t) = \sum_{t \in T \setminus A} \ell(I_t) + \sum_{t \in A \setminus B} \ell(I_t) \simeq_{\mathbf{V}} 0$; indeed, $\sum_{t \in T \setminus A} \ell(I_t) \simeq_{\mathbf{V}} 0$ by $(*)$ and $\sum_{t \in A \setminus B} \ell(I_t) \simeq_{\mathbf{V}} 0$ because $A \setminus B \in \mathbf{V}$ and each $\ell(I_t)$, $t \in A \setminus B$, is ultrasmall relative to \mathbf{V} (Lemma 3.13). Hence $\sum_{t \in T \setminus B} \ell(I_t) \simeq_{\mathbf{V}'} 0$ for all $\bar{\mathbf{V}} \subseteq \mathbf{V}' \subset \mathbf{V}$ and $(*)$ holds with $\bar{\mathbf{V}}$ in place of \mathbf{V} , a contradiction. \square

Let now $A \in \mathbf{V}(\cdot)$, $A \subseteq T$ be such that $\sum_{t \in T \setminus A} \ell(I_t) \simeq_{\mathbf{V}(\cdot)} 0$. We also have $\sum_{t \in A} \ell(I_t) \simeq_{\mathbf{V}(\cdot)} 0$ because $A \in \mathbf{V}(\cdot)$ and each $\ell(I_t)$, for $t \in A$, is ultrasmall relative to $\mathbf{V}(\cdot)$ (Lemma 3.13 again). Hence $\sum_{t \in T} \ell(I_t) \simeq_{\mathbf{V}(\cdot)} 0$, contradicting $\sum_{t \in T} \ell(I_t) \geq 1$. \square

4 GRIST and external sets.

Relative set theory enriches the \in -language of set theory by additional means: the binary relative standardness predicate \sqsubseteq , or equivalently, variables over levels. In this extended language it is possible to describe subcollections of sets that are not themselves sets. We use the term *external sets* in the inclusive sense, to refer to such collections as well as to the “usual” sets, sometimes called *internal sets* for emphasis.

There are two important reasons for expanding relative set theory to a theory of external sets. The first is foundational: the tendency to abstraction, so prominent in mathematics since at least the time of Cantor, makes us employ such collections almost automatically; “they are there.” The second reason is pragmatic: while the work of Nelson and the **IST** school shows that much can be accomplished by purely internal means, and while the techniques based on **GRIST** further facilitate and extend the internal approach, most of the practitioners of nonstandard analysis use some version of the Robinsonian model-theoretic framework, grounded in superstructures and characterized by heavy use of non-internal sets. If relative set theory is to serve as a universal vehicle for nonstandard analysis, it has to accommodate non-internal sets and the model-theoretic framework.

Here we consider four increasingly powerful extensions of **GRIST** to a theory of external sets, motivated mostly by pragmatic considerations. We believe that all arguments of current nonstandard practice can be formalized in (the strongest of) these systems. The issues at play in this section are similar to those that arise from attempts to extend **BST** to a theory of external sets. We rely heavily on the monograph [19] of Kanovei and Reeken, which contains a systematic comparative study of such extensions.

The most elementary use of external sets in relative set theory is as extensions of formulas of the language of **GRIST**. For example, given $r \in \mathbb{R}$, we can define

V-monad of r : $\mathbf{m}_{\mathbf{V}}(r) := \{x \in \mathbb{R} : |x - r| \text{ is ultrasmall relative to } \mathbf{V}\};$

V-galaxy of r : $\mathbf{g}_{\mathbf{V}}(r) := \{x \in \mathbb{R} : |x - r| \text{ is limited relative to } \mathbf{V}\};$

the **monad** of $r \in \mathbb{R}$ is then $\mathbf{m}_{\mathbf{V}(r)}(r)$, and the **galaxy** of $r \in \mathbb{R}$ is $\mathbf{g}_{\mathbf{V}(r)}(r)$;

V-proximity relation $\mathbf{E}_V \subseteq \mathbb{R} \times \mathbb{R}$: $(x, y) \in \mathbf{E}_V \leftrightarrow x \simeq_V y$;

V-neighbor function $\mathbf{n}_V \subseteq \mathbb{R} \times \mathbb{R}$: $y = \mathbf{n}_V(x)$ if and only if x is limited relative to V , $y \in V$, and $x \simeq_V y$.

External sets from these examples are typically used in nonstandard analysis for the sake of convenience; their use can be eliminated by replacing them with their defining formulas. We begin with an axiomatic system which is suitable for a formalization of such book-keeping use of external sets (and a bit more).

\mathbf{E}_1 –GRIST

We use boldface letters for variables of the theory; they are intended to range over **external sets**. There are two primitive binary predicates, \in and \sqsubseteq . We say that \mathbf{x} is **internal** [written $\mathbf{x} \in \mathbb{I}$] if $\mathbf{x} \sqsubseteq \mathbf{x}$, and use lightface letters as variables over internal sets, as was our practice hitherto. We continue to refer to internal sets as **sets**. The axioms of \mathbf{E}_1 –GRIST are:

- **GRIST ^{\mathbb{I}}** [the axioms of GRIST where all quantifiers are restricted to internal sets]. We can identify GRIST ^{\mathbb{I}} with GRIST.
- **Transitivity of \mathbb{I}** : $(\forall \mathbf{x} \in \mathbb{I})(\forall \mathbf{y} \in \mathbf{x})(\mathbf{y} \in \mathbb{I})$.
- **External Extensionality**: $(\forall \mathbf{x}, \mathbf{y})[(\forall \mathbf{z})(\mathbf{z} \in \mathbf{x} \leftrightarrow \mathbf{z} \in \mathbf{y}) \rightarrow \mathbf{x} = \mathbf{y}]$.
- **External Separation**: $(\forall \mathbf{A})(\exists \mathbf{B})(\forall \mathbf{x})(\mathbf{x} \in \mathbf{B} \leftrightarrow \mathbf{x} \in \mathbf{A} \wedge \mathcal{P}(\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_k))$, where \mathcal{P} is any formula of the \in - \sqsubseteq -language.

We say that \mathbf{x} is **standard** [written $\mathbf{x} \in \mathbb{S}$] if $\mathbf{x} \in \mathbb{I} \wedge \mathbf{x} \sqsubseteq 0$ [ie, $\mathbb{S} = \mathbf{V}(\cdot)$]. An external set \mathbf{X} is **\cap -closed** if $(\forall \mathbf{x}, \mathbf{y})(\mathbf{x}, \mathbf{y} \in \mathbf{X} \rightarrow \mathbf{x} \cap \mathbf{y} \in \mathbf{X})$. \mathbf{X} is **of standard size** if there is a function \mathbf{F} and $A \in \mathbb{S}$ such that $(\forall \mathbf{x})[\mathbf{x} \in \mathbf{X} \leftrightarrow (\exists a \in A \cap \mathbb{S})(\mathbf{x} = \mathbf{F}(a))]$.

- **External Restricted Standardization**:
 $(\forall \mathbf{X})(\forall A \in \mathbb{S})(\exists B \in \mathbb{S})(B \cap \mathbb{S} = A \cap \mathbf{X} \cap \mathbb{S})$.
- **External Saturation**: If $\mathbf{X} \subseteq \mathbb{I}$ is a \cap -closed external set of standard size and if $(\forall x \in \mathbf{X})(x \neq \emptyset)$, then $\bigcap \mathbf{X} \neq \emptyset$.
- **Standard Size Choice**: If \mathbf{X} is of standard size, then there is an external function \mathbf{F} such that $(\forall \mathbf{x} \in \mathbf{X})(\mathbf{x} \neq \emptyset \rightarrow \mathbf{F}(\mathbf{x}) \in \mathbf{x})$.

It is easy to check that the examples of external sets given at the beginning of this section can be formalized and proved to exist in \mathbf{E}_1 –GRIST, even without recourse to the last three axioms. It is also easy to obtain a model of \mathbf{E}_1 –GRIST minus the last

three axioms from any model $\mathfrak{M} = \langle |\mathfrak{M}|, \in^{\mathfrak{M}}, \subseteq^{\mathfrak{M}} \rangle$ of **GRIST**: add to $|\mathfrak{M}|$ all sets of the form $\{x \in |\mathfrak{M}| : \mathfrak{M} \models \underline{x} \in \underline{a} \wedge \mathcal{P}(x, \underline{x}_1, \dots, \underline{x}_k)\}$ where $a, x_1, \dots, x_k \in |\mathfrak{M}|$ and there is no $b \in |\mathfrak{M}|$ such that $\mathfrak{M} \models (\forall x)[x \in \underline{b} \leftrightarrow x \in \underline{a} \wedge \mathcal{P}(x, \underline{x}_1, \dots, \underline{x}_k)]$, and extend $\in^{\mathfrak{M}}$ and $\subseteq^{\mathfrak{M}}$ to this enlarged universe in the obvious way [let $\subseteq^{\mathfrak{M}}$ be false unless both arguments are in $|\mathfrak{M}|$]. Construction of an interpretation that satisfies also the last three axioms requires a bit more care and is described in the proof of Theorem 4.1.

The extra axioms allow for a more serious use of external sets than just book-keeping. In particular, one can define the set of external natural numbers \mathbf{N} and use External Restricted Standardization to show that it equals $\mathbb{N} \cap \mathbb{S}$. Similarly for the set of external integers and external rational numbers. External real numbers can be defined, and Standardization provides a natural correspondence between them and the standard real numbers [elements of $\mathbb{R} \cap \mathbb{S}$]. A detailed discussion of these issues can be found in the author's [13]. The actual *external set \mathbf{R} of external real numbers* cannot be proved to exist in \mathbf{E}_1 –**GRIST**, which does not guarantee existence of sets beyond the first level of the external cumulative hierarchy over \mathbb{I} . This defect is remedied by our second theory, an extension of \mathbf{E}_1 –**GRIST** that postulates that the universe of external sets satisfies **ZC**, Zermelo set theory with Choice.

\mathbf{E}_ω –**GRIST**

We add the following axioms to \mathbf{E}_1 –**GRIST**:

- **External Pair:** $(\forall \mathbf{x}, \mathbf{y})(\exists \mathbf{z})(\forall \mathbf{u})(\mathbf{u} \in \mathbf{z} \leftrightarrow \mathbf{u} = \mathbf{x} \vee \mathbf{u} = \mathbf{y})$.
- **External Union:** $(\forall \mathbf{x})(\exists \mathbf{y})(\forall \mathbf{u})[\mathbf{u} \in \mathbf{y} \leftrightarrow (\exists \mathbf{z})(\mathbf{u} \in \mathbf{z} \wedge \mathbf{z} \in \mathbf{x})]$.
- **External Power Set:** $(\forall \mathbf{x})(\exists \mathbf{y})(\forall \mathbf{u})(\mathbf{u} \in \mathbf{y} \leftrightarrow \mathbf{u} \subseteq \mathbf{x})$.
- **External Choice:** $(\forall \mathbf{X})(\exists \mathbf{F})(\forall \mathbf{x} \in \mathbf{X})(\mathbf{x} \neq \emptyset \rightarrow \mathbf{F}(\mathbf{x}) \in \mathbf{x})$.
- **Regularity over \mathbb{I} :** $(\forall \mathbf{x} \neq \emptyset)(\exists \mathbf{y} \in \mathbf{x})(\mathbf{y} \cap \mathbf{x} \subseteq \mathbb{I})$.

We note that there are infinite internal sets, for example \mathbb{N} , so “External Infinity” is not needed. There are external sets for which Regularity fails, for example $\{n \in \mathbb{N} : 0 \sqsubset n\} = \{n \in \mathbb{N} : n \notin \mathbb{S}\}$; Regularity over \mathbb{I} is a substitute. The classical proof of Zermelo establishes the Well-Ordering Principle: Every external set can be well-ordered.

In \mathbf{E}_ω –**GRIST** there is actually an external *set \mathbf{R} of all external real numbers*. The first presentation of Loeb measures in a framework of an axiomatic nonstandard set theory (**NST**) is given in great detail in the author's [13]; it carries over to \mathbf{E}_ω –**GRIST** without any essential changes. External Saturation and (Standard Size) Choice are instrumental

in proving countable additivity of Loeb measures. Kanovei and Reeken [19, Section 9.5] observe that the essence of the construction of Loeb measures can be carried out already in much weaker systems, such as $\mathbf{E}_1\text{-GRIST}$ [this is why we included External Saturation and Standard Size Choice in it], except that the Loeb algebra and the measure itself are not sets. $\mathbf{E}_\omega\text{-GRIST}$ adds the External Power Set axiom to remedy this last difficulty.

The universe of $\mathbf{E}_\omega\text{-GRIST}$ resembles a superstructure over \mathbb{I} ; with the addition of two more axioms we can obtain something like the familiar model-theoretic framework.

$\mathbf{E}_\Omega\text{-GRIST}$

Following [19, Definition 8.1.3], a set $\mathbf{x} \in \mathbb{S}$ is **condensable** if there is an external transitive set $\mathbf{T} \ni \mathbf{x}$ and a map $\mathbf{y} \mapsto \hat{\mathbf{y}}$ defined on $\mathbf{T} \cap \mathbb{S}$ and such that $\hat{\mathbf{y}} = \{\hat{\mathbf{z}} : \mathbf{z} \in \mathbf{y} \cap \mathbb{S}\}$ holds for all $\mathbf{y} \in \mathbf{T} \cap \mathbb{S}$.

$\mathbf{E}_\Omega\text{-GRIST}$ is $\mathbf{E}_\omega\text{-GRIST}$ plus the axioms:

- **Transitive Hulls:** $(\forall \mathbf{x})(\exists \mathbf{T})(\mathbf{x} \in \mathbf{T} \wedge \mathbf{T} \text{ is transitive}).$
- **Standard Condensation:** $(\forall \mathbf{x} \in \mathbb{S})(\mathbf{x} \text{ is condensable}).$

Let \mathbb{H} be the class of all external sets and \mathbb{WF} the subclass of all externally well-founded external sets. We let $\mathbb{W} := \{\hat{\mathbf{x}} : \mathbf{x} \in \mathbb{S}\}$ be the class of all **feasible** well-founded sets [$\mathbb{W} = \mathbb{WF}^{\text{feas}}$ in the notation of [19]]. It can be shown that either $\mathbb{W} = \mathbb{WF}$ or $\mathbb{W} = \mathbf{V}_\Omega$ for some external ordinal Ω , where $\langle \mathbf{V}_\xi, \xi \in \mathbf{Ord} \rangle$ is the external von Neumann cumulative hierarchy. It then follows [19, Exercise 8.2.6] that there is a uniquely determined \in -isomorphism $*$: $\mathbb{W} \rightarrow \mathbb{S} \subseteq \mathbb{I}$. Hence \mathbb{W} is an interpretation of **ZFC** (ie, $\mathcal{P}^{\mathbb{W}}$ holds for all axioms \mathcal{P} of **ZFC**) and $*$: $\mathbb{W} \rightarrow \mathbb{I}$ is an \in -elementary embedding [ie, $(\forall \bar{\mathbf{x}} \in \mathbb{W})(\mathcal{P}^{\mathbb{W}}(\bar{\mathbf{x}}) \leftrightarrow \mathcal{P}^{\mathbb{I}}(\bar{\mathbf{x}}))$ holds for all \in -formulas \mathcal{P}]. This is the scheme “ $\mathbb{WF}^{\text{feas}} \xrightarrow{*} \mathbb{I}$ [in \mathbb{H}]” in the terminology of [19]. The analogy with the model-theoretic framework is very close. In the model-theoretic terminology, elements of \mathbb{W} are the *standard sets*, elements of \mathbb{S} are the *standard copies*, \mathbb{I} is the universe of *internal sets*, and the entire universe \mathbb{H} of $\mathbf{E}_\Omega\text{-GRIST}$ is the “superstructure” of *external sets*. Most arguments of model-theoretic nonstandard analysis readily transfer into this setting; see Kanovei and Reeken [19, Chapter 2] for an exposition of nonstandard analysis in the closely related framework of “ $\mathbb{WF} \xrightarrow{*} \mathbb{I}$ [in \mathbb{H}].”

$\mathbf{E}\text{-GRIST}$

We extend the theory once more, again for both practical and fundamental reasons. The external universe of $\mathbf{E}_\Omega\text{-GRIST}$ satisfies all of **ZC**, but not necessarily Replacement.

From the point of view of applications there is also something missing. For example, the Loeb measure $\mathbf{L}(\mu)$ is an external set; but we might like to have a measure with the properties of $\mathbf{L}(\mu)$ in the standard universe \mathbb{S} , or in its well-founded counterpart \mathbb{W} . These shortcomings are overcome in the theory obtained from \mathbf{E}_Ω –**GRIST** by adding yet two more principles.

- **External Collection:** $(\forall \mathbf{A})(\exists \mathbf{B})(\forall \mathbf{x} \in \mathbf{A})[(\exists \mathbf{y})\mathcal{P}(\mathbf{x}, \mathbf{y}, \bar{\mathbf{x}}) \rightarrow (\exists \mathbf{y} \in \mathbf{B})\mathcal{P}(\mathbf{x}, \mathbf{y}, \bar{\mathbf{x}})]$, where \mathcal{P} is any formula.
- **External Transfer:** $(\forall \mathbf{x}_1, \dots, \mathbf{x}_k \in \mathbb{W})(\mathcal{P}^{\mathbb{W}}(\mathbf{x}_1, \dots, \mathbf{x}_k) \leftrightarrow \mathcal{P}^{\mathbb{WF}}(\mathbf{x}_1, \dots, \mathbf{x}_k))$, where \mathcal{P} is any \in -formula.

In the absence of Regularity, External Collection is stronger than External Replacement. \mathbf{E} –**GRIST** implies that every external set is in one-one correspondence with some element of \mathbb{WF} (in fact, with some external ordinal). Hence $\mathbf{L}(\mu)$ is isomorphic to some measure $\mathbf{m}_0 \in \mathbb{WF}$, in any appropriate sense of “isomorphic.” Next, by External Transfer, there is a measure $\mathbf{m} \in \mathbb{W}$ that has “the same properties” as \mathbf{m}_0 , at least as far as these properties can be expressed by \in -formulas in \mathbb{WF} . Finally, $*\mathbf{m} \in \mathbb{S}$ is a standard measure isomorphic to \mathbf{m} , via the isomorphism $*$.

One consequence of \mathbf{E} –**GRIST** is perhaps unexpected and worth pointing out. By [16, Proposition 12.33], in **GRIST** one can define a mapping of the (internal) set \mathbb{N} of natural numbers onto the class of all standard ordinals. The composition of this mapping with the internal von Neumann cumulative hierarchy $\xi \rightarrow V_\xi$ maps \mathbb{N} onto $\{V_\xi : \xi \in \mathbb{S}\}$, and External Collection implies that the latter is an external set. Hence [External Union] $\bigcup_{\xi \in \mathbb{S}} V_\xi = \mathbb{I}$ is an external set. We conclude that $\mathbb{I}, \mathbb{S} \subseteq \mathbb{I}$ and \mathbb{W} [the inverse image of \mathbb{S} by $*$] are external sets! [This conclusion does *not* hold in weaker theories, such as \mathbf{E}_Ω –**GRIST**; see the proof of Theorem 4.1.]

This observation suggests that, in \mathbf{E} –**GRIST**, the universe \mathbb{H} of external sets, or perhaps its well-founded part \mathbb{WF} , should be regarded as the “usual” universe of sets, while \mathbb{I}, \mathbb{S} , and \mathbb{W} are just “models” of set theory, with $* : \mathbb{W} \rightarrow \mathbb{S} \preccurlyeq \mathbb{I}$ being close to the Robinsonian model-theoretic framework.

Theorem 4.1 \mathbf{E} –**GRIST** is a conservative extension of **ZFC**.

Proof and further discussion of Theorem 4.1

We rely heavily on the material in Kanovei-Reeken [19, Sections 8.1, 8.2]. The preceding remarks indicate that \mathbf{E} –**GRIST** is similar to Kawai’s theory **KST**. The principal differences are:

- In **KST**, the internal universe $(\mathbb{I}, \in, \mathbb{S})$ is an interpretation of **IST**, while in **E-GRIST**, $(\mathbb{I}, \in, \sqsubseteq)$ is an interpretation of **GRIST**.
- **KST** postulates Strong Saturation [for \mathbb{S} -size external sets $\mathbf{X} \subseteq \mathbb{I}$], while **E-GRIST** postulates (External) Saturation only for standard-size external sets.
- The axiom of External Transfer is a new idea that is not part of **KST**. However, the interpretation of **KST** used by Kanovei–Reeken [19] to prove their Theorem 8.1.5 satisfies External Transfer; hence **KST**+ External Transfer is a conservative extension of **ZFC**.

To prove Theorem 4.1 we follow closely the proof of Theorem 8.1.5 in [19]. We argue in **ZFC** ϑ and let $V := \mathbf{V}_\vartheta$. In [19], the initial \mathbf{P}_0 is taken as *V , where ${}^*\mathbf{v} = \langle {}^*V, {}^*\in, {}^*\mathbf{st} \rangle$ is a ϑ^+ -saturated interpretation of **IST** with an \in -elementary standard core embedding $*$: $V \rightarrow {}^*V$. We start instead with an interpretation of **GRIST**, $\mathfrak{S}_\Lambda(V) = \langle V^*, =^*, \in^*, \sqsubseteq^* \rangle$, constructed inside $\mathbb{V} := V$ as in [16], and form the quotient structure ${}^*\mathbf{v} = \langle {}^*V, {}^*\in, {}^*\sqsubseteq \rangle$ modulo the congruence relation $=^*$. This is a ϑ -saturated interpretation of **GRIST**. Using this *V as \mathbf{P}_0 , we define the sets \mathbf{P}_ξ for all ξ as in [19], and let $\mathbf{P} := \bigcup_{\xi \in \mathbf{Ord}} \mathbf{P}_\xi$. [[19] has $\xi < \gamma$ in place of $\xi \in \mathbf{Ord}$; this appears to be a misprint.]

The claim that the structure $\langle \mathbf{P}, {}^*\in, {}^*\sqsubseteq \rangle$ interprets **E-GRIST** can be proved by arguments similar to those in [19]; in particular, ϑ -saturation of ${}^*\mathbf{v}$ suffices to prove that External Saturation holds. External Transfer follows from the axiom schema (*) of **ZFC** ϑ (Transfer from \mathbf{V}_ϑ to \mathbf{V}).

To obtain interpretations of **E₁-GRIST**, **E _{ω} -GRIST** and **E _{Ω} -GRIST**, it suffices to take $\mathbf{P}_1, \mathbf{P}_\omega$ and \mathbf{P}_ϑ , resp, in the place of \mathbf{P} . In all these interpretations, \mathbb{I} is an external set. While this is necessarily so in the case of **E-GRIST**, the other theories have also interpretations in which \mathbb{I} remains a proper class and, in fact, full Standardization holds in the form

- **External Standardization:** $(\forall \mathbf{X})(\exists B \in \mathbb{S})(B \cap \mathbb{S} = \mathbf{X} \cap \mathbb{S})$.

Addition of this axiom makes these theories resemble **NST** [12, 13, 19, Section 8.2]. To obtain interpretations that satisfy External Standardization, proceed as in [19, Exercise 8.2.15], again starting with $\mathfrak{S}_\Lambda(\mathbf{V}_\vartheta)$, and taking $\gamma = 1, \omega$ and ϑ , resp. \square

The theory **E-GRIST** accomodates arguments of the traditional model-theoretic non-standard analysis to the full extent. On the other hand, this theory is rather ad-hoc, and far from providing a foundationally satisfactory, philosophically coherent system.

Before discussing this matter further, we consider the simpler nonstandard set theory **IST**.

IST is a conservative extension of **ZFC**, and can thus be viewed as merely a formal tool for proving theorems of **ZFC**. However, most users of **IST** wish to identify the “usual” sets with some objects provided by **IST**. For this purpose there are two choices. The “official” philosophy of Nelson, enshrined by the **IST** terminology, is to regard the *internal sets* as the “usual” sets. This is the view we follow in **GRIST** as well. It has significant pedagogical advantages, as discussed in detail in [17]. However, one can equally well take the view that *standard sets* of **IST** are the “usual” sets. This is the philosophy of the author’s [13]; many working mathematicians seem to find it more palatable. The point we wish to stress is that there is no *mathematical* reason for preferring either alternative. The standard universe \mathbb{S} and the internal universe \mathbb{I} have equal claims to being the “usual” universe of set theory.

On the other hand, if we look at not just the universes themselves, but also at the way they are embedded in the wider “cosmos” of **IST**, we notice some essential differences; most important for mathematical applications, infinitesimals (ultrasmall numbers) exist relative to the standard universe (level) \mathbb{S} , but not relative to the internal universe \mathbb{I} . It is precisely this asymmetry that is remedied by relative set theory. In fact, the guiding principle behind the development of **GRIST** was the desire to make all levels have the same view of the surrounding “cosmos” (technically, to make Transfer hold for all \in - \sqsubseteq -formulas, a feature we refer to elsewhere as “Full Relativization”).

Coming back to **E-GRIST**: We have already noted that there are several universes that with some justification can be regarded as the “usual” universe of sets: \mathbb{I} , \mathbb{S} , \mathbb{W} , \mathbb{WF} , \mathbb{H} (and possibly any level $\mathbb{V} \subseteq \mathbb{I}$). If we follow the ideas that led from **IST** to **GRIST**, a picture of an extension of **E-GRIST** emerges wherein every universe can be regarded as the “usual” one, and all universes have the same view of the “cosmos”; in particular, every universe has a stratification into levels that satisfies **GRIST**, has its own external universe, and in this external universe it is isomorphic to a transitive universe, via its own $*$.

The “relativistic” perspective on axiomatic nonstandard set theory was advocated and developed by Ballard in [2]. In a later unpublished paper [3] Ballard realized that the foundational issues raised by many universes of axiomatic nonstandard analysis are similar to those raised by many universes obtainable by forcing in traditional set theory, and proposed a coherent, uncompromisingly relativistic theory of the mathematical “cosmos” that accommodates both nonstandard and forcing extensions. A truly universal and philosophically satisfying theory of the nonstandard appears to require

such relativism; however, the construction of an extension of **GRIST** to such a theory presents a number of technical, mathematical and philosophical challenges. We hope to address these issues in greater detail elsewhere.

References

- [1] S Albeverio, J E Fenstad, R Hoegh-Krohn and T Lindstrom, *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*, Academic Press, Inc., New York, 1986, xi + 514 pages.
- [2] D Ballard, *Foundational Aspects of “Non” standard Mathematics*, Contemporary Mathematics, vol. 176, American Mathematical Society, Providence, RI, 1994.
- [3] D Ballard, *A set theoretic cosmology for mathematics*, unpublished paper, January 2002, 22 pages.
- [4] R G Bartle, *A Modern Theory of Integration*, American Mathematical Society, Providence, RI, 2001, xiv + 458 pages.
- [5] B Benninghofen, *Superinfinitesimals and the calculus of the generalized Riemann integral*, in: *Models and Sets*, ed. by G H Müller and M M Richter, Lecture Notes in Math. 1103, Springer, Berlin, 1984, 9 - 52.
- [6] B Benninghofen and M M Richter, *A general theory of superinfinitesimals*, Fund. Math. 123 (1987), 199–215.
- [7] F Diener and M. Diener (eds.), *Nonstandard Analysis in Practice*, Springer-Verlag, Berlin 1995.
- [8] H B Enderton, *A Mathematical Introduction to Logic*, Academic Press, Inc., New York and London, 1972, xiii + 295 pages.
- [9] E I Gordon, *Relatively nonstandard elements in the theory of internal sets of E. Nelson*, Siberian Math. J. 30 (1989), 89–95 (in Russian).
- [10] E I Gordon, *Nonstandard Methods in Commutative Harmonic Analysis*, American Mathematical Society, Providence, Rhode Island, 1997.
- [11] C W Henson, *Analytic sets, Baire sets and the standard part map*, Canadian J. Math. 31 (1979), no. 3, 663–672.
- [12] K Hrbacek, *Axiomatic foundations for nonstandard analysis*, Fund. Math. 98 (1978), 1–19; *abstract* in J. Symbolic Logic 41 (1976), 285.
- [13] K Hrbacek, *Nonstandard set theory*, Amer. Math. Monthly 86 (1979), 1–19; doi:[10.2307/2321294](https://doi.org/10.2307/2321294).
- [14] K Hrbacek, *Internally iterated ultrapowers*, in *Nonstandard Models of Arithmetic and Set Theory*, ed. by A Enayat and R Kossak, Contemporary Math. 361, American Mathematical Society, Providence, RI, 2004, 87–120.

- [15] K Hrbacek, *Stratified analysis?*, in *The Strength of Nonstandard Analysis*, ed. by I van den Berg and V Neves, Springer Wien New York, 2007, 47–63.
- [16] K Hrbacek, *Relative set theory: Internal view*, *Journal of Logic and Analysis* 1 (2009), 1–108; doi:[10.4115/jla.2009.1.8](https://doi.org/10.4115/jla.2009.1.8).
- [17] K Hrbacek, O Lessmann and R O’Donovan, *Analysis with ultrasmall numbers*, *Amer. Math. Monthly* 117 (2010), 16 pages, to appear.
- [18] K Hrbacek, O Lessmann and R O’Donovan, *Analysis using Relative Infinitesimals*, 267 pages, March 19, 2008.
- [19] V Kanovei and M Reeken, *Nonstandard Analysis: Axiomatically*, Springer-Verlag Berlin Heidelberg New York, 2004, xvi + 408 pages.
- [20] E Nelson, *Internal set theory: a new approach to nonstandard analysis*, *Bull. Amer. Math. Soc.* 83 (1977), 1165–1198; doi:[10.1090/S0002-9904-1977-14398-X](https://doi.org/10.1090/S0002-9904-1977-14398-X).
- [21] R O’Donovan, *Teaching analysis with ultrasmall numbers*, *Mathematics Teaching-Research Journal Online*, 3, 3 (2009), 1 - 22.
- [22] Y Péraire, *Théorie relative des ensembles internes*, *Osaka J. Math.* 29 (1992), 267–297.
- [23] Y Péraire, *Infinitesimal approach of almost-automorphic functions*, *Ann. Pure Appl. Logic* 63 (1993), 283–297; doi:[10.1016/0168-0072\(93\)90152-4](https://doi.org/10.1016/0168-0072(93)90152-4).
- [24] W F Pfeffer, *The Riemann Approach to Integration: Local Geometric Theory*, Cambridge University Press, 1993, xvi + 302 pages.

The City College of New York

khrbacek@sci.cuny.cuny.edu

Received: 29 December 2009 Revised: 7 September 2010