MATHEMATICS IN RELATIVE SET THEORY

Karel Hrbacek

Department of Mathematics
The City College of CUNY
New York, NY 10031

July 31, 2009
Non-archimedean fields present a serious obstacle for attempts to teach calculus at an elementary level using Robinsonian framework for nonstandard analysis.

Axiomatic set theories, such as Nelson’s IST, can get around this obstacle. In the usual interpretation, IST does not add any new objects to those of traditional mathematics; it is only the mathematical language that gets extended, by a new undefined unary predicate “\(x\) is standard.” The standardness predicate allows us to define “infinitesimals” as real numbers smaller in absolute value than every standard positive real number. “Infinitesimals” are just ordinary real numbers that are “very small” compared to the standard ones. They have nothing to do with infinity. To avoid misunderstandings, we talk about “ultrasmall” and “ultralarge” numbers, rather than “infinitesimal” and “infinitely large” ones.
Another obstacle to the presentation of nonstandard methods at an elementary level remains:

The ultrasmall numbers can be used to define calculus concepts, such as $f'(a)$, but only for *standard* functions $f$ at *standard* points $a$. If say $a$ is nonstandard, no “infinitesimals relative to $a$” are available, and one has to fall back on the $\varepsilon-\delta$ method.

These issues are discussed in


and

KH, *Stratified Analysis?*, ibid, 47 - 63.
A solution to this problem is to replace the unary predicate “\(x\) is standard” with a binary “\(x\) is standard relative to \(y\)”. Such theory is

**RIST (Relative Internal Set Theory)**

developed by Péraire in


**RIST has been further extended in**


A simple subtheory of FRIST suitable for elementary expositions is formulated in


The framework described in this talk is based on the above papers.
ELEMEN TARY RELATIVE ANALYSIS

We add to the mathematical language a new sort of variables, denoted by $\mathbf{V}$ with various decorations, and assumed to range over levels.

Levels are not sets, and can appear only on the right side of $\in$; to stress this, we read “$x \in \mathbf{V}$” as “$x$ appears at the level $\mathbf{V}$,” and we read $\mathbf{V}_1 \subseteq \mathbf{V}_2$ as “$\mathbf{V}_1$ is coarser than $\mathbf{V}_2$” or “$\mathbf{V}_2$ is finer than $\mathbf{V}_1$.”

Axiom I: For every $x_1, \ldots, x_k$ there is a level $\mathbf{V}(x_1, \ldots, x_k)$ such that $x_1, \ldots, x_k \in \mathbf{V}(x_1, \ldots, x_k)$ and, for all levels $\mathbf{V}$, $x_1, \ldots, x_k \in \mathbf{V}$ implies $\mathbf{V}(x_1, \ldots, x_k) \subseteq \mathbf{V}$.

$\mathbf{V}(x_1, \ldots, x_k)$ is the coarsest level where $x_1, \ldots, x_k$ appear; we call it the level of $x_1, \ldots, x_k$.

Axiom II: For every $\mathbf{V}_1$ and $\mathbf{V}_2$ either $\mathbf{V}_1 \subseteq \mathbf{V}_2$ or $\mathbf{V}_2 \subseteq \mathbf{V}_1$. 
Definition 1.

(1) A real number $\varepsilon$ is \textit{ultrasmall} relative to $\mathbf{V}$ if $|\varepsilon| < r$ for all $r > 0$, $r \in \mathbf{V}$.

(2) A real number $x$ is \textit{ultralarge} relative to $\mathbf{V}$ if $|x| > r$ for all $r > 0$, $r \in \mathbf{V}$.

It is \textit{limited} relative to $\mathbf{V}$ if it is not ultralarge relative to $\mathbf{V}$.

(3) Real numbers $a$ and $b$ are \textit{ultraclose} relative to $\mathbf{V}$, written $a \simeq_\mathbf{V} b$, if $a - b$ is ultrasmall relative to $\mathbf{V}$.

Note that $0$ is the only number ultrasmall relative to $\mathbf{V}$ that appears at $\mathbf{V}$, and no numbers ultralarge relative to $\mathbf{V}$ appear at $\mathbf{V}$. A number $x$ is limited relative to $\mathbf{V}$ if and only if there are $r, s \in \mathbf{V}$ such that $r < x < s$. 

Axiom III:
*For every level $\mathbb{V}$ there exist nonzero real numbers ultrasmall relative to $\mathbb{V}$.*

Axiom IV (Neighbor Principle):
*For every real number $x$ limited relative to $\mathbb{V}$ there is a real number $r \in \mathbb{V}$ such that $x \simeq_{\mathbb{V}} r$."

Axiom V (Closure Principle):
*A number, function or set that is uniquely defined (in traditional mathematics, that is, without reference to levels) from parameters $x_1, \ldots, x_k$ appears at the level $\mathbb{V}(x_1, \ldots, x_k)$.*

All of these axioms are immediate consequences of RIST.
Proposition 2. *Relative to a given level V:*

1. *If* $x, y$ *are limited, then* $x \pm y$ *and* $x \cdot y$ *are limited.*
2. *If* $\delta, \varepsilon$ *are ultrasmall, then* $\delta \pm \varepsilon$ *are ultrasmall.*
3. *If* $\varepsilon$ *is ultrasmall and* $x$ *is limited, then* $\varepsilon \cdot x$ *is ultrasmall.*
4. $x \neq 0$ *is ultralarge if and only if* $\frac{1}{x}$ *is ultrasmall.*
5. *If* $a \simeq a$; *if* $a \simeq b$, *then* $b \simeq a$; *and* $a \simeq b$ *and* $b \simeq c$, *then* $a \simeq c$.
6. *If* $x \simeq a$ and $y \simeq b$, *then* $x \pm y \simeq a \pm b$.
7. *If* $x \simeq a$, $y \simeq b$, *and* $a$ *and* $b$ *are limited, then* $x \cdot y \simeq a \cdot b$.
8. *For* $a, b \in V$, $a \simeq b$ *if and only if* $a = b$.

**Convention.** *References to V can be omitted when the level is given, or understood from the context.*

By Proposition 2 (8), the number $r$ in Neighbor Principle is uniquely determined; we call it the *V-neighbor* of $x$ and denote it $n_v(x)$.

The cases (6) and (7) of Proposition 2, for $a, b$ limited, can now be written as:

$n(a \pm b) = n(a) \pm n(b)$ and

$n(a \cdot b) = n(a) \cdot n(b)$. 
As an example, we prove (3).

*Proof.* Let $|x| \leq r_0$, where $r_0 > 0$, $r_0 \in \mathbf{V}$.

For every $r > 0$, $r \in \mathbf{V}$, also $\frac{r}{r_0} > 0$ and $\frac{r}{r_0} \in \mathbf{V}$, by Closure Principle.

Hence $|\varepsilon| < \frac{r}{r_0}$, and $|\varepsilon \cdot x| < \frac{r}{r_0} \cdot r_0 = r$.

This shows $\varepsilon \cdot x$ is ultrasmall relative to $\mathbf{V}$. \qed

Convention. *In definitions where no level is specified, the context level is always the level of the parameters of the property or operation being defined.*

Definition 3. Let \( f \) be a function whose domain is an open interval about \( a \). The *derivative* of \( f \) at \( a \) is a real number \( L \) at the context level such that

\[
\frac{f(a + dx) - f(a)}{dx} \sim L
\]

for all ultrasmall \( dx \neq 0 \).

Note: The context level is \( V(f, a) \).

The derivative is "the best approximation, at the observation level" to the average rate of change of \( f \) in an ultrasmall interval about \( a \).

By Closure, the domain of \( f \) appears at the context level, and hence \( f \) is defined for all \( dx \) ultrasmall.

If the number \( L \) exists, it is uniquely determined; we denote it \( f'(a) \).
The definition of the derivative can be re-written in a number of ways; one of the most useful is the following.

**Theorem 4. (Increment Equation)**

*Let $f$ be a function defined on an open interval containing $x$. Then $f'(x)$ is a number $L$ at the context level $V(f, x)$ such that, for all ultrasmall $dx$,

$$f(x + dx) = f(x) + L \cdot dx + \varepsilon \cdot dx,$$

where $\varepsilon$ is ultrasmall.*
Example 5. Let \( f(x) = x^2 \) and \( a \in \mathbb{R} \). Compute \( f'(a) \).

We work relative to \( V(f, a) \). (Actually, \( V(f, a) = V(a) \) because the function \( f \) is defined without parameters, so it appears at every level.)

For \( dx \neq 0 \) ultrasmall,

\[
\frac{f(a + dx) - f(a)}{dx} = \frac{(a + dx)^2 - a^2}{dx} = 2a + dx \simeq 2a.
\]

We observe that \( 2a \in V; \) thus \( f'(a) = 2a \).
Example 5 and other similar examples show that the result of calculating $f'(a)$ does not depend on the choice of the level $V$, as long as $f$ and $a$ appear at $V$ (as long as $V$ is an “observation level” for the objects we are studying). We make this into a fundamental principle.

Axiom VI (Stability Principle):

Let $P(x_1, \ldots, x_k; V)$ be a statement about $V$. If $x_1, \ldots, x_k \in V_1$ and $x_1, \ldots, x_k \in V_2$, then

\[ P(x_1, \ldots, x_k; V_1) \iff P(x_1, \ldots, x_k; V_2). \]

A statement about $V$ is any statement (well-formed formula of the extended language) with free variables among $x_1, \ldots, x_k$ (also referred to as parameters) and $V$. Quantifiers over levels are allowed in the form $(\forall V' \supseteq V)$ and $(\exists V' \supseteq V)$, although in applications of Stability in analysis quantifiers over levels usually do not appear explicitly.

We call the level $V(x_1, \ldots, x_k)$ of the parameters of a statement $P(x_1, \ldots, x_k; V)$, the context level for that statement.

The Stability Principle asserts that if the statement is true about its context level, then it remains true about every finer level.
Hence in the definition of derivative, any level where \( f \) and \( a \) appear can be used in place of \( V(f, a) \).

Note: The Closure Principle is an easy consequence of the Stability Principle.

Recall the notation

\[
\Delta f(a) = f(a + dx) - f(a)
\]

and the increment equation

\[
\frac{\Delta f(a)}{dx} \approx f'(a) \quad \text{for } dx \neq 0 \text{ ultrasmall.}
\]

Theorem 6 (Derivative of a Product). Let \( f \) and \( g \) be functions differentiable at \( a \). Then \( (f \cdot g) \) is differentiable at \( a \) and

\[
(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a).
\]
Proof. The context level is $V(f, g, a)$.

Let $dx \neq 0$ be ultrasmall.

\[
\frac{\Delta(f \cdot g)(x)}{dx} = \frac{f(a + dx) \cdot g(a + dx) - f(a) \cdot g(a)}{dx}
\]

\[
= \frac{(f(a) + \Delta f(a)) \cdot (g(a) + \Delta g(a)) - f(a) \cdot g(a)}{dx}
\]

\[
= f(a) \cdot \frac{\Delta g(a)}{dx} + \frac{\Delta f(a)}{dx} \cdot g(a) + \frac{\Delta f(a)}{dx} \cdot \Delta g(a)
\]

\[
\simeq f(a) \cdot g'(a) + f'(a) \cdot g(a)
\]

because $\frac{\Delta g(a)}{dx} \simeq g'(a)$, $\frac{\Delta f(a)}{dx} \simeq f'(a)$,

and by the Increment Equation,

$\Delta g(a) \simeq 0$, so $\frac{\Delta f(a)}{dx} \cdot \Delta g(a) \simeq f'(a) \cdot 0 = 0$.

But $f'(a) \cdot g(a) + f(a) \cdot g'(a)$ appears at the context level by Closure, hence

\[
(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a).
\]
Definition 7. Let $f$ be a function whose domain is an interval $I$ and $a \in I$. We say that $f$ is continuous at $a$ if $f(x) \approx f(a)$ for all $x \approx a$, $x \in I$.

Theorem 8 (Maximum Value). Let $f$ be a function continuous at every $x \in [a, b]$. Then $f$ attains its maximum on $[a, b]$. 
Proof. Work relative to $V(f, a, b)$.
Let $N$ be an ultralarge positive integer, $dx = \frac{b-a}{N}$, and $x_i = a + i \cdot dx$, for $i = 0, \ldots, N$. The set \{f(x_0), \ldots, f(x_N)\} is finite, so it has a greatest element, say $f(x_j)$. Let $c = n(x_j)$ (it exists because $x_j$ is limited); clearly $c \in [a, b]$. By continuity of $f$ at $c$, we have $f(x_j) \simeq f(c)$, and by Closure, $f(c)$ appears at the context level.

Let $x \in [a, b]$ appear at the context level. There is $i$ such that $x_i \leq x < x_{i+1}$. Hence $x_i \simeq x$ and $f(x_i) \simeq f(x)$, because $f$ is continuous at $x$ and $x$ is at the context level. By definition of $x_i$ and $c$ we have

$$f(x) \simeq f(x_i) \leq f(x_j) \simeq f(c).$$

As $f(x)$ and $f(c)$ are at the context level, this implies that $f(x) \leq f(c)$.
We proved that $f(c)$ is the maximum value of $f(x)$ for all $x$ at the context level.

By Stability, the same is true about every finer level; hence $f(c)$ is the maximum value of $f$. $\square$
The Closure Principle implies that, for any level $\mathcal{V}$:

$1 \in \mathcal{V}$ and $(n \in \mathcal{V} \Rightarrow n + 1 \in \mathcal{V})$ for all $n \in \mathbb{N}$.

Hence not all statements in our extended language define sets or are subject to the Principle of Mathematical Induction;

“$n \in \mathbb{N} \land n \in \mathcal{V}$” is a simple counterexample. Fortunately, statements that behave in the conventional way are very easy to recognize by inspection.

A statement $Q(x_1, \ldots, x_k)$ is internal if either it does not mention levels, or all levels mentioned in it are at least as fine as its context level $\mathcal{V}(x_1, \ldots, x_k)$.

Technically, $Q(x_1, \ldots, x_k)$ is internal if all quantifiers over levels occurring in it (if any) are of the form

$(\exists \mathcal{V} \text{ such that } x_1, \ldots, x_k \in \mathcal{V})$ or
$(\forall \mathcal{V} \text{ such that } x_1, \ldots, x_k \in \mathcal{V})$. 

Axiom VII (Definition Principle):
If $Q(x, x_1, \ldots, x_k)$ is internal and $A$ is a set, then there is a set $B$ such that
\[(\forall x)(x \in B \iff x \in A \land Q(x, x_1, \ldots, x_k)).\]

By Closure Principle, if $A, x_1, \ldots, x_k \in V$, then also $B \in V$.

Corollary. The Principle of Mathematical Induction holds for internal statements.

Indeed, if $Q(x)$ is internal, then
\[\{n \in \mathbb{N} : Q(n)\}\] is a set.
Proofs involving “double limits” can often be simplified with the help of one more principle about levels.

Axiom VIII (Density of levels): If a real number $\varepsilon \neq 0$ is ultrasmall relative to $V$ and $\varepsilon \in V'$, then there is $V^+$ and a real number $\delta \in V^+$ such that $\delta$ is ultrasmall relative to $V$ and $\varepsilon$ is ultrasmall relative to $V^+$. Note that $V \subset V^+ \subset V'$.

Theorem 9 (L’Hospital’s rule for $\infty/\infty$). Let $f$ and $g$ be differentiable in a deleted neighborhood of $a$. Suppose that $\lim_{x \to a} |f(x)| = \lim_{x \to a} |g(x)| = \infty$ and $\lim_{x \to a} \frac{f'(x)}{g'(x)}$ exists. Then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$
Proof. We work relative to \( V := \mathcal{V}(f, g, a) \). Let \( L = \lim_{x \to a} \frac{f'(x)}{g'(x)} \) and let \( x \simeq a, x \neq a \). Assume that \( x > a \) (case \( x < a \) is similar).

It is easy to see that, by density of levels, we can choose \( y > a, y \simeq a \) relative to \( V \) such that \( x \simeq a \) relative to \( V^+ \supset V \) and \( y \in V^+ \). We use \( \overset{+}{\simeq} \) when we work relative to the finer level \( V^+ \). We have \( x \overset{+}{\simeq} a \) and necessarily \( a < x < y \).

By Cauchy Mean Value Theorem

\[
\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(c)}{g'(c)} \quad \text{for some } c \in (x, y).
\]

But \( c \simeq a \), so \( \frac{f'(c)}{g'(c)} \simeq L \).

Since \( \lim_{x \to a} |f(x)| = \lim_{x \to a} |g(x)| = \infty \), and \( x \overset{+}{\simeq} a \), we have \( f(x), g(x) \overset{+}{\simeq} \pm \infty \). Hence \( f(y)/f(x) \overset{+}{\simeq} 0 \) and \( g(y)/g(x) \overset{+}{\simeq} 0 \). The proof is completed by observing that

\[
L \simeq \frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f(x)}{g(x)} \cdot \frac{1 - \frac{f(y)}{f(x)}}{1 - \frac{g(y)}{g(x)}} \overset{+}{\simeq} \frac{f(x)}{g(x)}. \quad \square
\]
Advantages of RELATIVE ANALYSIS:

Nonstandard proofs can be presented with the same degree of (in)formality that is customary in traditional textbooks on analysis.

The book-keeping associated with the $\varepsilon$–$\delta$ method is eliminated; as a result, the proofs can focus on the “combinatorial” essence of the arguments.

Derivatives and definite integrals can be developed before limits, and independently of each other.

Proofs become more elementary; for example, the proof of the Maximum Value Theorem does not use the notion of supremum or compactness.

The relative framework allows arguments employing two or more levels. This can simplify proofs of results involving double limits.
Riemann theory of integration for bounded functions on $[a, b]$ works as follows:

Definition 10. A tagged partition is a pair $(\mathcal{P}, \mathcal{T})$, where $\mathcal{P} = \{x_0, x_1, \ldots, x_n\}$ with

$a = x_0 < x_1 < \ldots < x_n = b,$
and $\mathcal{T} = \{t_0, \ldots, t_{n-1}\}$ with $x_i \leq t_i \leq x_{i+1},$
for $i = 0, \ldots, n - 1.$

We let $dx_i = x_{i+1} - x_i.$

The Riemann sum $\sum(f; \mathcal{P}, \mathcal{T})$ is defined as

$$\sum(f; \mathcal{P}, \mathcal{T}) = \sum_{i=0}^{n-1} f(t_i) \cdot dx_i.$$ 

Let $f$ be a bounded function on $[a, b]$.

Relative to $V(f, a, b)$, a tagged partition $(\mathcal{P}, \mathcal{T})$ is fine if all $dx_i$ are ultrasmall.

Riemann integral of $f$ on $[a, b]$ is a number $R$ at the context level such that $\sum(f; \mathcal{P}, \mathcal{T}) \simeq R$ for all fine tagged partitions $(\mathcal{P}, \mathcal{T})$ of $[a; b]$.

We let

$$\int_a^b f(x) \cdot dx = n \left( \sum_{i=0}^{N-1} f(x_i) \cdot dx \right),$$

if it exists.
If \( f(x) = F'(x) \) for all \( x \in [a; b] \) and \( f \) is continuous, we have

**Theorem 11 (Uniform Increment Equation).**

*For all \( x \in [a, b] \) and all ultrasmall \( dx \) such that \( x + dx \in [a, b] \),

\[
F(x + dx) = F(x) + F'(x) \cdot dx + \varepsilon \cdot dx,
\]

where \( \varepsilon \) is ultrasmall.

Note: The context level here is \( V(F) \); it is independent of \( x \!\!\!\!.\)

From this one gets easily that

\[
F(x_{i+1}) - F(x_i) = F'(t_i) \cdot dx + \varepsilon_i \cdot dx_i = f(t_i) \cdot dx_i + \varepsilon_i \cdot dx_i
\]

where \( \varepsilon_i \simeq 0 \), for \( i = 0, \ldots, N - 1 \).

By adding up these equations we obtain

\[
F(b) - F(a) = \sum_{i=0}^{N-1} f(x_i) \cdot dx_i + \sum_{i=0}^{N-1} \varepsilon_i \cdot dx_i.
\]

An important observation is that

\[
\sum_{i=0}^{N-1} \varepsilon_i \cdot dx_i \simeq 0.
\]

Indeed, let \( \varepsilon = \max\{|\varepsilon_0|, \ldots, |\varepsilon_{N-1}|\}; \) then

\[
\left| \sum_{i=0}^{N-1} \varepsilon_i \cdot dx_i \right| \leq \sum_{i=0}^{N-1} |\varepsilon_i| \cdot dx_i \leq N \cdot \varepsilon \cdot dx_i = \varepsilon \cdot (b-a) \simeq 0.
\]
This implies that
\[ F(b) - F(a) \simeq \sum_{i=0}^{N-1} f(x_i) \cdot dx_i. \]

Since \( F(b) - F(a) \) appears at the context level, we have
\[ F(b) - F(a) = \int_a^b f(x) \cdot dx. \]

This is the First Fundamental Theorem of Calculus.

The Definition Principle justifies existence of the function \( F \) defined on \([a, b]\) by
\[ x \mapsto \int_a^x f(t) \cdot dt. \]

It is not hard to prove, from these definitions, that \( F'(x) = f(x) \) for \( x \in [a, b] \) where \( f \) is continuous (The Second Fundamental Theorem of Calculus).
The previous considerations explain why continuity of \( f = F' \) is essential for the recovery of \( F \) via Riemann integrals: it allows us to use the \textit{uniform} version of the Increment Equation, which in turn leads to the definition of the Riemann integral in terms of fine partitions, that is, partitions where each \( dx_i \) is ultrasmall relative to the context level, independent of the tag \( t_i \).

This suggests that to integrate an arbitrary derivative, one should replace fine partitions in the definition of Riemann integral by “superfine” partitions, where a tagged partition \((\mathcal{P}, \mathcal{T})\) is \textit{superfine} if each \( dx_i \) is ultrasmall relative to a level that also contains \( t_i \).
It turns out that superfine partitions in this sense do not exist.
One needs to use, for this purpose, a weaker notion of relative ultrasmall, due to

and

**Definition 12.** A real number $\varepsilon$ is an *ultrasmall* if $|\varepsilon| \leq \varphi(a)$ for all positive $\varphi : \mathbb{R} \to \mathbb{R}$ in the context level.
A tagged partition $(\mathcal{P}, \mathcal{T})$ is *superfine* if each $dx_i$ is $t_i$-ultrasmall.

If in the definition of the Riemann integral one replaces the word “fine” by “superfine,” one gets the generalized Riemann integral, also known as the Henstock-Kurzweil integral, which extends the Lebesgue integral. Every derivative is generalized Riemann integrable and its integral is the original function, up to a constant.
ULTIMATE INTERNAL THEORY: GRIST

Relativization:
The conjunction of

For every \( x \) there is \( V \) such that \( x \in V \).
For every \( V \) there is \( x \in V \) such that
\((\forall V')(x \in V' \to V \subseteq V')\).
For every \( V_1 \) and \( V_2 \) either \( V_1 \subseteq V_2 \) or \( V_2 \subseteq V_1 \).
For every \( V \) there is \( V' \) such that \( V \subseteq V' \).
For every \( V \subset V' \) there is \( V'' \) such that \( V \subset V'' \subset V' \).

Transfer (Stability):
If \( x_1, \ldots, x_k \in V \) and \( x_1, \ldots, x_k \in V' \), then
\[ P(x_1, \ldots, x_k; V) \leftrightarrow P(x_1, \ldots, x_k; V'). \]

Granularity:
For any \( x_1, \ldots, x_k \), if \( (\exists V)P(x_1, \ldots, x_k; V) \),
then
\[ (\exists V)[P(x_1, \ldots, x_k; V) \land (\forall V')(V' \subset V \to \neg P(x_1, \ldots, x_k; V'))]. \]
Standardization:

*Given V and arbitrary A, x₁, ..., xₖ:*

either (∀V')(V ⊆ V') or there exists V' ⊆ V and B ∈ V' such that, for every V'' with V' ⊆ V'' ⊆ V,

(∀y ∈ V'')(y ∈ B ↔ y ∈ A ∧ P(y, A, x₁, ..., xₖ; V'')).

Idealization:

*Given V, V', A such that A ∈ V ⊆ V', and arbitrary x₁, ..., xₖ:*

(∀V'' ⊆ V')(∀a ∈ P^{fin} A) [a ∈ V'' →

(∃y)(∀x ∈ a)P(x, y, A, x₁, ..., xₖ; V')] →

(∃y)(∀V'' ⊆ V')(∀x ∈ A)[x ∈ V'' → P(x, y, A, x₁, ..., xₖ; V')].
BST (Kanovei’s modification of IST) is the ultimate internal nonstandard set theory with two levels, in the following sense:

**Theorem 13** (Categoricity of BST over ZFC).
*Every countable model $\mathbb{M}$ of ZFC has a unique extension to a countable model of BST, up to an isomorphism which is the identity on $\mathbb{M}$.***

**Theorem 14** (Universality of BST over ZFC).
*If $\mathbb{N}_1$ satisfies ST, $\mathbb{N}_2$ satisfies BST, $\mathbb{N}_1$ and $\mathbb{N}_2$ have the same standard universe $\mathbb{M}$, and $|\mathbb{N}_1| = |\mathbb{N}_2| = \aleph_0$, then there exists an $\in$-elementary embedding of $\mathbb{N}_1$ into $\mathbb{N}_2$ which is the identity on $\mathbb{M}$.***

**Theorem 15** (Completeness of BST over ZFC).
*If $T \supseteq ZFC$ is a complete consistent theory (in the $\in$-language), then $T + BST$ is a complete consistent theory (in the $\in$-st-language).***

Analogs of the above results hold for GRIST. Hence GRIST is the ultimate internal nonstandard set theory with many levels, in a similar sense.
Let $f : [0; 1] \to [0; 1]$ (for simplicity).

Definition 16. $f$ is (uniformly) $\mathbf{V}$-continuous if $x \simeq_\mathbf{V} y$ implies $f(x) \simeq_\mathbf{V} f(y)$, for all $x, y \in [0, 1]$.

Of course, if $f$ is $\mathbf{V}$-continuous and $f \in \mathbf{V}$, then $f$ is continuous.

Theorem 17. For every $f : [0; 1] \to [0; 1]$ there is a finite set $\{v_0 = 0, v_1, \ldots, v_n\}$ such that, for $V_i := V(v_i)$, $V_0 \subset V_1 \subset \ldots \subset V_n$ and for all $V_i \subset V \subset V_{i+1} \ [V_n \subset V$ if $i = n]$, $f$ is $V_i$-continuous iff $f$ is $\mathbf{V}$-continuous iff $f$ is not $V_{i+1}$-continuous.

Definition 18. The $\mathbf{V}$-shadow of $f$ is the set $\text{sh}_\mathbf{V}(f) \in \mathbf{V}$ such that, for $(a, b) \in \mathbf{V}$,

$$(a, b) \in \text{sh}_\mathbf{V}(f) \iff (\exists x)(a \simeq_\mathbf{V} x \land b \simeq_\mathbf{V} f(x)).$$

Lemma 19. $f$ is $\mathbf{V}$-continuous if and only if $\text{sh}_\mathbf{V}(f)$ is a continuous function.

Lemma 20. For every $f : [0; 1] \to [0; 1]$ there is a finite set $\{v_0 = 0, v_1, \ldots, v_n\}$ such that, for $V_i := V(v_i)$, $V_0 \subset V_1 \subset \ldots \subset V_n$ and for all $V_i \subset V \subset V_{i+1} \ [V_n \subset V$ if $i = n ]$, $\text{sh}_{V_i}(f) = \text{sh}_\mathbf{V}(f)$. 
EXTERNAL VIEW OF RST

Why external sets?

• They are there.

Examples:

\[ I := \{ n \in \mathbb{N} : (\forall V)(n \in V) \} = \mathbb{N} \cap V(0) \]

\[ M_V(0) := \{ x \in \mathbb{R} : x \simeq_V 0 \}. \]

• External sets are essential for many important nonstandard constructions.

Examples: Loeb measures, nonstandard hulls.

• Most of the extant literature uses Robinsonian framework of standard, internal and external sets.
Extension of GRIST to Robinsonian framework:

Notation:
I is the collection of all sets of GRIST (from now on, internal sets).

S is a fixed level (standard copies).

Axiom E₁: There is a collection E of external sets such that I ⊆ E and E satisfies ZFC – Regularity.

Note: Every mathematical concept has an internal and an external version.

R(α) are the ranks of the cumulative hierarchy constructed in E;
R := ∪ external ordinals α R(α).

Axiom E₂: There is an ∈-isomorphism * : W → S, where W = R(ξ) for some external ordinal ξ.

Axiom E₃: Transfer between W and R and a slightly strengthened Idealization.
This setting closely resembles the superstructure framework.

One can construct Loeb measures and nonstandard hulls in it.

It is similar to Kawai’s nonstandard set theory. In particular, $\mathcal{V}$ and $\mathcal{I}$ are external sets.
Until now, we worked in the “internal picture,” that is, we identified the “usual” sets with internal sets. This picture is essential at elementary levels, in order to avoid non-archimedean fields.

In a theory that treats external sets seriously, as \( \text{sets} \), it seems more appropriate to regard \( \mathbb{N} \cap V(0) \) as the “usual” set of natural numbers, and more generally, to regard \( W \) (or \( R, E, \) or even \( S \)) as the “usual” set theoretic universe.

There appears to be no \textit{mathematical} reason to prefer one of these “pictures” to another. The choice of the “usual” universe should be fully relativized—any universe should be able to serve in this role and see “the same multiverse.”

This foundational philosophy was first proposed by Ballard.

Motivated by many universes obtainable by forcing, some traditional set theorists are now proposing “relativistic” theories with similar features.