A Brief Introduction to Non-standard Analysis

Joni Puljujärvi

DOMAST Student Seminar

March 31, 2023

[...] to find a tangent means to draw a line that connects two points of the curve at an infinitely small distance, or the continued side of a polygon with an infinite number of angles, which for us takes the place of the curve. This infinitely small distance can always be expressed by a known differential like $\mathrm{d}v$*, or by a relation to it, that is, by some known tangent.*

(G. W. Leibniz, 1684)

[. . .] the idea of infinitely small or infinitesimal quantities seems to appeal naturally to our intuition. At any rate, the use of infinitesimals was widespread during the formative stages of the Differential and Integral Calculus.

However, neither [Leibniz] nor his disciples and successors were able to give a rational development leading up to a system of this sort. As a result, the theory of infinitesimals gradually fell into disrepute and was replaced eventually by the classical theory of limits.

[. . .] Leibniz' ideas can be fully vindicated and [. . .] they lead to a novel and fruitful approach to classical Analysis and to many other branches of mathematics. The key to our method is provided by the detailed analysis of the relation between mathematical languages and mathematical structures which lies at the bottom of contemporary model theory.

(A. Robinson, 1966)

[Non-standard analysis] allows one to rigorously manipulate things such as "the set of all small numbers", or to rigorously say things like "η_1 is smaller than anything that involves η_0", while greatly reducing epsilon management issues by automatically concealing many of the quantifiers in one's argument. One has to take care as to which objects are standard, non-standard, sets of non-standard objects, etc., especially when transferring results between the standard and non-standard worlds, but as long as one is clearly aware of the underlying mechanism used to construct the non-standard universe and transfer back and forth [...], one can avoid difficulty.

(T. Tao, 2007)

Vocabularies and structures

- By a vocabulary, we mean any set of *constant symbols*, *function symbols* and *relation symbols*. Function and relation symbols S come with an arity, denoted by $\sharp S$. E.g. usually $\sharp + = 2$ and $\sharp < = 2$.
- If τ is the vocabulary $\{c_0, \ldots, c_{m-1}, f_0, \ldots, f_{n-1}, R_0, \ldots, R_{k-1}\}$, then a τ-structure is a tuple

$$\mathcal{M} = (M, c_0^{\mathcal{M}}, \ldots, c_{m-1}^{\mathcal{M}}, f_0^{\mathcal{M}}, \ldots, f_{n-1}^{\mathcal{M}}, R_0^{\mathcal{M}}, \ldots, R_{k-1}^{\mathcal{M}}),$$

  where
  - M is a set,
  - each $c_i^{\mathcal{M}} \in M$,
  - each $f_i^{\mathcal{M}}$ is a function $M^{\sharp f_i} \to \mathcal{M}$, and
  - each $R_i^{\mathcal{M}}$ is a subset of $M^{\sharp R_i}$.

Terms

Let τ be a vocabulary. The set of τ-terms is defined as follows.

- We fix an infinite set of variables v_i, $i \in \mathbb{N}$. Each v_i is a τ-term.
- If $c \in \tau$ is a constant symbol, then c is a τ-term.
- If $f \in \tau$ is a function symbol with arity n and $t_0, \ldots, t_{n-1}$ are τ-terms, then $f(t_0, \ldots, t_{n-1})$ is a τ-term.

Example

Let $\tau = \{+, 1\}$, where $+$ is a binary function symbol and 1 is a constant symbol. We denote $+(t, t')$ by $t + t'$ as is customary. Now the following are τ-terms:

- 1,
- $v_0 + v_6$,
- $(v_3 + 1) + v_2$.

Formulas

First-order logic over a vocabulary τ is a set of formulas defined as follows:

- If t and t' are variables, then $t = t'$ is a(n atomic) τ-formula.
- If $R \in \tau$ is a relation symbol of arity n and $t_0, \ldots, t_{n-1}$ are τ-terms, then $R(t_0, \ldots, t_{n-1})$ is a(n atomic) τ-formula.
- If φ is a τ-formula, then $\neg\varphi$ ("not φ") is a τ-formula.
- If φ and ψ are τ-formulas, then $\varphi \wedge \psi$ ("φ and ψ"), $\varphi \vee \psi$ ("φ or ψ"), $\varphi \rightarrow \psi$ ("if φ, then ψ") and $\varphi \leftrightarrow \psi$ ("φ if and only if ψ") are τ-formulas.
- If φ is a τ-formula and x is a variable, then $\exists x\varphi$ and $\forall x\varphi$ are τ-formulas.

Truth definition

If $\mathcal{M}$ is a τ-structure and $\varphi(x_0, \ldots, x_{n-1})$ a τ-formula, then we say that $\mathcal{M}$ and a tuple $(a_0, \ldots, a_{n-1}) \in M^n$ satisfy φ if the obvious thing holds. For example:

- $\mathcal{M}$ and (a, b) satisfy $x = y$ if $a = b$.
- $\mathcal{M}$ and $\bar{a}$ satisfy $\varphi \wedge \psi$ if they satisfy both φ and ψ.
- $\mathcal{M}$ and $(a_0, \ldots, a_{n-1})$ satisfy $\exists y \varphi(y, x_0, \ldots, x_{n-1})$ if there exists $b \in M$ such that $\mathcal{M}$ and $(b, a_0, \ldots, a_{n-1})$ satisfy φ.

If $\mathcal{M}$ and $(a_0, \ldots, a_{n-1})$ satisfy $\varphi(x_0, \ldots, x_{n-1})$, we write

$$\mathcal{M} \models \varphi(a_0, \ldots, a_{n-1}).$$

If Σ is a set of τ-sentences, we write $\mathcal{M} \models \Sigma$ if $\mathcal{M} \models \varphi$ for every $\varphi \in \Sigma$.

Real closed fields

- Let $\tau_{\mathbb{R}}$ be the vocabulary $\{0, 1, +, \cdot, <\}$ and denote by $\mathbb{R}$ the $\tau_{\mathbb{R}}$-structure
$$(\mathbb{R}, 0^{\mathbb{R}}, 1^{\mathbb{R}}, +^{\mathbb{R}}, \cdot^{\mathbb{R}}, <^{\mathbb{R}}),$$
where $\mathbb{R}$ is the set of real numbers $0^{\mathbb{R}}$ and $1^{\mathbb{R}}$ are the 0 and 1 of the reals, $+^{\mathbb{R}}$ and $\cdot^{\mathbb{R}}$ are the addition and multiplication of the reals, and $<^{\mathbb{R}}$ is the ordering of the reals. We will drop the superscript for convenience.
- Denote by T the first-order theory of $\mathbb{R}$, i.e. the set containing every first-order $\tau_{\mathbb{R}}$-sentence true in $\mathbb{R}$.
- What does T say about $\mathbb{R}$?

First-order vs. second-order logic

T says, for instance, the following things:

- $(\mathbb{R}, +, \cdot)$ is a field with 0 as the additive neutral element and 1 as the multiplicative neutral element.
- $(\mathbb{R}, <)$ is a dense linear order, i.e. for every $a, b \in \mathbb{R}$, if $a < b$, there is $c \in \mathbb{R}$ with $a < c < b$.
- $(\mathbb{R}, +, \cdot, <)$ is an ordered field, i.e. $(\mathbb{R}, <)$ is a linear order and for all $a, b, c \in \mathbb{R}$ we have

$$a < b \implies a + c < b + c \quad \text{and} \quad a, b > 0 \implies ab > 0.$$

- $(\mathbb{R}, +, \cdot)$ is *formally real*, i.e. if any sum of squares of elements of $\mathbb{R}$ equals zero, then each of those elements must be zero.
- Every polynomial of odd degree has at least one root, and for every element $a \in \mathbb{R}$ there is $b \in \mathbb{R}$ such that $a = b^2$ or $a = -b^2$.

T does not say any of the following things (which are true):

- $(\mathbb{R}, +, \cdot, <)$ is a complete ordered field, i.e. every bounded subset $A \subseteq \mathbb{R}$ has a least upper bound $a \in \mathbb{R}$.
- $(\mathbb{R}, +, \cdot, <)$ has the *Archimedean property*, i.e. for any $a \in \mathbb{R}$ there is $n \in \mathbb{N}$ such that $n > a$.
- There are no positive infinitesimally small elements in $\mathbb{R}$, i.e. for all $a \in \mathbb{R}$ if $a \geq 0$ and $a < 1/n$ for every $n \in \mathbb{N}$, then $a = 0$.

In addition T does not recognize whether an element is a natural number or not. However, it does recognize whether an element is *any particular natural number*.

Existence of non-standard models

- For any $n \in \mathbb{N}$, we denote by $\underline{n}$ the $\tau_{\mathbb{R}}$-term $\underbrace{1+1+\cdots+1}_{n \text{ times}}$. Then obviously $\underline{n}^{\mathbb{R}} = n$.
- Let c be a new constant symbol. Denote by T^* the theory

$$T \cup \{0 < c\} \cup \{\underline{n}c < 1 \mid n \in \mathbb{N}\},$$

which, in addition to the theory of $\mathbb{R}$, states that the interpretation of c will be an infinitesimal.

Theorem (Compactness)

Let Σ be any set of first-order τ-sentences such that for any finite $\Sigma' \subseteq \Sigma$, there exists a τ-structure $\mathcal{M}$ such that $\mathcal{M} \models \Sigma'$. Then there exists a τ-structure $\mathcal{M}$ such that $\mathcal{M} \models \Sigma$.

Corollary

T^* has a model.

Hyperreal numbers

The *hyperreal numbers* are a $\tau_{\mathbb{R}}$-structure ${}^*\mathbb{R} = ({}^*\mathbb{R}, 0, 1, +, \cdot, <)$ that is an *elementary extension* of $\mathbb{R}$ but contains infinitesimals. It satisfies the following:

1. ${}^*\mathbb{R} \models T$.
2. $\mathbb{R} \subseteq {}^*\mathbb{R}$.
3. For any $\tau_{\mathbb{R}}$-formula $\varphi(x_0, \ldots, x_{n-1})$ and $a_0, \ldots, a_{n-1} \in \mathbb{R}$,
$$ {}^*\mathbb{R} \models \varphi(a_0, \ldots, a_{n-1}) \iff \mathbb{R} \models \varphi(a_0, \ldots, a_{n-1}). $$
4. For any relation $R \subseteq \mathbb{R}^n$, there is a "natural" extension ${}^*R \subseteq {}^*\mathbb{R}^n$ (${}^*R$ looks like R and ${}^*R \cap \mathbb{R}^n = R$).
5. For any function $f\colon \mathbb{R}^n \to \mathbb{R}$, there is a "natural" extension ${}^*f\colon {}^*\mathbb{R}^n \to {}^*\mathbb{R}$ $\quad$ (${}^*f$ looks like f and ${}^*f \restriction \mathbb{R} = f$).

Theorem (Transfer principle, technical)

For a relation $R \subseteq \mathbb{R}^n$ *and a function* $f\colon \mathbb{R}^m \to \mathbb{R}$, *introduce a new relation symbol* $\underline{R}$ *and a new function symbol* $\underline{f}$. *Let*

$$\tau = \tau_{\mathbb{R}} \cup \bigcup_{n\in\mathbb{N}} (\{\underline{R} \mid R \subseteq \mathbb{R}^n\} \cup \{\underline{f} \mid f\colon \mathbb{R}^n \to \mathbb{R}\}).$$

Then for any first-order τ*-formula* $\varphi(x_0, \dots, x_{n-1})$ *and* $a_0, \dots, a_{n-1} \in \mathbb{R}$, *we have*

$$\mathbb{R} \models \varphi(a_0, \dots, a_{n-1}) \iff {}^*\mathbb{R} \models \varphi(a_0, \dots, a_{n-1}),$$

where $\underline{R}^{\mathbb{R}} = R$ *and* $\underline{f}^{\mathbb{R}} = f$, *and* $\underline{R}^{{}^*\mathbb{R}} = {}^*R$ *and* $\underline{f}^{{}^*\mathbb{R}} = {}^*f$.

Theorem (Tranfer principle, simplified)

If R_i *are relations on* $\mathbb{R}$, f_j *functions on* $\mathbb{R}$ *and* a_k *elements of* $\mathbb{R}$, *then for any reasonably simple statement* Φ:

Φ *holds for* R_i, f_j *and* a_k *in* $\mathbb{R}$ *iff* Φ *holds for* ${}^*R_i$, ${}^*f_j$ *and* a_k *in* ${}^*\mathbb{R}$.

Examples of transfer

Example

As $|\sin(x)| \leq 1$ for all $x \in \mathbb{R}$, we have $^*|^*\sin(x)| \, ^*\!\leq 1$ for all $x \in {}^*\mathbb{R}$.

Note: We will henceforth drop the star from $^*|\cdot|$ and $^*\!\leq$ because it's annoying and obfuscates notation.

Example

As $\mathbb{R}$ is Archimedean, for every $a \in \mathbb{R}$ there is $n \in \mathbb{N}$ with $a < n$, i.e. we have

$$\mathbb{R} \models \forall x \exists y (\underline{\mathbb{N}}(y) \wedge x < y).$$

Here $\underline{\mathbb{N}}$ is the name for the set of natural numbers. By the transfer principle,

$$^*\mathbb{R} \models \forall x \exists y (\underline{\mathbb{N}}(y) \wedge x < y).$$

Thus, if $a \in {}^*\mathbb{R}$ is unbounded, there is no *standard* natural number n such that $|a| < n$ (and hence $^*\mathbb{R}$ is not Archimedean), but there is a *non-standard* $\omega \in {}^*\mathbb{N}$ such that $|a| < \omega$. This ω must be infinite.

Infinitesimals and infinite numbers

Definition

- We call elements of $\mathbb{R}$ *standard* and elements of ${}^*\mathbb{R} \setminus \mathbb{R}$ *non-standard*.
- We say that $\varepsilon \in {}^*\mathbb{R}$ is an *infinitesimal* if for every $n \in \mathbb{N}$ (i.e. every standard natural number) $|\varepsilon| < 1/n$.
- We say that $a, b \in {}^*\mathbb{R}$ are infinitesimally close to each other if $a - b$ is an infinitesimal. We then write $a \simeq b$.
- We say that $\omega \in {}^*\mathbb{R}$ is *infinite* if for every $n \in \mathbb{N}$ (i.e. every standard natural number), $n < \omega$.
- We say that $\omega \in {}^*\mathbb{R}$ is *unbounded* if $|\omega|$ is infinite.

Easy facts:

- $\varepsilon \in {}^*\mathbb{R}$ is an infinitesimal if and only if $|\varepsilon| < r$ for every $r \in \mathbb{R}$, $r > 0$.
- $\varepsilon \in {}^*\mathbb{R}$ is an infinitesimal if and only if $1/\varepsilon$ is unbounded.
- If ε is an infinitesimal and $|\delta| \leq |\varepsilon|$, then also δ is an infinitesimal.
- Each $\simeq$-equivalence class of a bounded number contains a unique standard number.
- Let $F \subseteq {}^*\mathbb{R}$ be the set of bounded numbers. Then $F/\simeq$ is isomorphic to $\mathbb{R}$ as an ordered field.
- If $a \simeq c$ and $b \simeq d$, then $a + b \simeq c + d$, $ab \simeq cd$ and $a/b \simeq c/d$ (provided that $b \not\simeq 0$).
- If ε and δ are two distinct infinitesimals, then $1/\varepsilon \not\simeq 1/\delta$.
- If $a \simeq c$ and $b \simeq d$, then $a < b$ if and only if $c < d$.
- For every $r \in \mathbb{R}$ there is $q \in {}^*\mathbb{Q}$ such that $r \simeq q$.

Definition

- The $\simeq$-equivalence class of a bounded number $a \in {}^*\mathbb{R}$ is called the *monad* of a.
- The unique standard number residing in the monad of a is called the *standard part* of a and denoted by $\operatorname{st} a$.

Sequences

Since a sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ is a function $\mathbb{N} \to \mathbb{R}$, it has a natural extension $({}^*x_n)_{n\in{}^*\mathbb{N}}$ to ${}^*\mathbb{R}$, where ${}^*x_n = x_n$ for all $n \in \mathbb{N}$.

Theorem

- *A number $x \in \mathbb{R}$ is the limit of a sequence $(x_n) \in \mathbb{R}^{\mathbb{N}}$ in the sense of $\mathbb{R}$ if and only if ${}^*x_n \simeq x$ for all infinite $n \in {}^*\mathbb{N}$.*
- *A sequence $(x_n) \in \mathbb{R}^{\mathbb{N}}$ is Cauchy in the sense of $\mathbb{R}$ if and only if for all infinite $n, m \in {}^*\mathbb{N}$, ${}^*x_n \simeq {}^*x_m$.*

Functions and limits

Theorem

Let $f : \mathbb{R}^n \to \mathbb{R}$.

- *Given $x_0, L \in \mathbb{R}$, $f(x) \to L$ when $x \to x_0$, in the sense of $\mathbb{R}$, if and only if the following holds.*

  $$\text{For all } x \in {}^*\mathbb{R}, \text{ if } x \simeq x_0, \text{ then } {}^*f(x) \simeq L.$$

- *f is continuous at a point $x_0 \in \mathbb{R}$, in the sense of $\mathbb{R}$, if and only if the following holds.*

  $$\text{For all } x \in {}^*\mathbb{R}, \text{ if } x \simeq x_0, \text{ then } {}^*f(x) \simeq f(x_0).$$

- *f is uniformly continuous in the sense of $\mathbb{R}$ if and only if the following holds.*

  $$\text{For all } x, y \in {}^*\mathbb{R}, \text{ if } x \simeq y, \text{ then } {}^*f(x) \simeq {}^*f(y). \quad (*)$$

Note that ${}^*f$ is "continuous" at every point of ${}^*\mathbb{R}$ iff f is uniformly continuous in $\mathbb{R}$.

Proof.

We show the equivalence for uniform continuity. Suppose f is uniformly continuous. Then for any $n \in \mathbb{N} \setminus \{0\}$ there is $\delta_n > 0$ such that for all $x, y \in \mathbb{R}$ with $|x - y| < \delta_n$, we have $|f(x) - f(y)| < 1/n$. This means that

$$\mathbb{R} \models \forall x \forall y(|x - y| < \delta_n \to |\underline{f}(x) - \underline{f}(y)| < 1/n).$$

By the transfer principle,

$$^*\mathbb{R} \models \forall x \forall y(|x - y| < \delta_n \to |\underline{f}(x) - \underline{f}(y)| < 1/n).$$

Hence, for all $x, y \in {}^*\mathbb{R}$ with $|x - y| < \delta_n$, we have $|{}^*f(x) - {}^*f(y)| < 1/n$. Now if $x \simeq y$, then $|x - y| < \delta_n$ for all n and thus $|{}^*f(x) - {}^*f(y)| < 1/n$ for all n. But this means that $f(x) \simeq f(y)$. Hence $(*)$ holds.

Proof cont.

Suppose f is not continuous. Then there is $\varepsilon > 0$ such that for any $n \in \mathbb{N} \setminus \{0\}$ we can find $x_n, y_n \in \mathbb{R}$ with $|x_n - y_n| < 1/n$ but $|f(x_n) - f(y_n)| \geq \varepsilon$. This means that

$$\mathbb{R} \models \forall n \in \underline{\mathbb{N}} \setminus \{0\}(\left|\underline{x_n} - \underline{y_n}\right| < 1/n \land \left|\underline{f}(\underline{x_n}) - \underline{f}(\underline{y_n})\right| \geq \varepsilon).$$

By the transfer principle,

$$^*\mathbb{R} \models \forall n \in \underline{\mathbb{N}} \setminus \{0\}(\left|\underline{x_n} - \underline{y_n}\right| < 1/n \land \left|\underline{f}(\underline{x_n}) - \underline{f}(\underline{y_n})\right| \geq \varepsilon).$$

Hence, for all $n \in {}^*\mathbb{N} \setminus \{0\}$, $|{}^*x_n - {}^*y_n| < 1/n$ and $|{}^*f({}^*x_n) - {}^*f({}^*y)| \geq \varepsilon$. Now $x_n - y_n \to 0$ when $n \to \infty$, so by the previous theorem ${}^*x_\omega \simeq {}^*y_\omega$ for any infinite $\omega \in {}^*\mathbb{N}$. However, $|{}^*f({}^*x_\omega) - {}^*f({}^*y_\omega)| \geq \varepsilon$, so ${}^*f({}^*x_\omega) \not\simeq {}^*f({}^*y_\omega)$. Thus $(*)$ does not hold. □

Example

The function $f\colon (0,1) \to \mathbb{R}$, $f(x) = 1/x$, is continuous since for any $x \in (0,1)$ and $y \in {}^*\mathbb{R}$, $y \simeq x$, we have $f(x) = 1/x \simeq 1/y = {}^*f(y)$.

But f is not uniformly continuous, since for any infinite $\omega \in {}^*\mathbb{N}$, we have $1/\omega \simeq 1/2\omega \simeq 0$ but ${}^*f(1/\omega) = \omega \not\simeq 2\omega = {}^*f(1/2\omega)$.

Differentiation

Theorem

For a function $f : \mathbb{R} \to \mathbb{R}$ and a point $x_0 \in \mathbb{R}$, the following are equivalent.

- *f is differentiable at the point x_0, in the sense of $\mathbb{R}$.*
- *There is a unique $a \in \mathbb{R}$ such that whenever $x_0 \neq x \simeq x_0$, then*

$$\frac{{}^*f(x) - f(x_0)}{x - x_0} \simeq a.$$

- *There is a unique $a \in \mathbb{R}$ such that whenever h is infinitesimal, then*

$$\frac{{}^*f(x_0 + h) - f(x_0)}{h} \simeq a.$$

Note: The unique a above must be $f'(x_0)$.

Example

We show that $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$, is differentiable at any point $x_0 \in \mathbb{R}$ and $f'(x_0) = 2x_0$.

Fix $x_0 \in \mathbb{R}$ and let $x_0 \neq x \simeq x_0$. Now

$$\frac{{}^*f(x) - f(x_0)}{x - x_0} = \frac{x^2 - x_0^2}{x - x_0} = \frac{(x + x_0)(x - x_0)}{x - x_0} = x + x_0 \simeq 2x_0.$$

Hence f is differentiable at x_0 and $f'(x_0) = 2x_0$.

Example

We show that $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^x$, is differentiable at every point $x_0 \in \mathbb{R}$ and $f'(x_0) = e^{x_0}$.

Fix $x_0 \in \mathbb{R}$ and let $h \neq 0$ be an infinitesimal. Now

$$\frac{{}^*f(x_0 + h) - f(x_0)}{h} = \frac{e^{x_0+h} - e^{x_0}}{h} = \frac{e^{x_0}e^h - e^{x_0}}{h} = e^{x_0} \cdot \frac{e^h - 1}{h}.$$

Now let $\eta = (e^h - 1)/h$. If we can show that $\eta \simeq 1$, then

$$\frac{{}^*f(x_0 + h) - f(x_0)}{h} \simeq e^{x_0}\eta \simeq e^{x_0}$$

and we are done. We proceed to show that $\eta \simeq 1$.

As e^x is continuous and $e^0 = 1$, for any infinitesimal $\varepsilon \neq 0$, $e^\varepsilon \simeq 1$. Thus, as h is a non-zero infinitesimal, $\eta h = e^h - 1 \simeq 0$. Since ηh is a non-zero infinitesimal, $1/(\eta h)$ is an unbounded number.

Example (cont.)

Suppose that $h > 0$. Recall that $e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$. Thus for any infinite $\omega \in {}^*\mathbb{N}$, $(1+\frac{1}{\omega})^\omega \simeq e$. Now $1/(\eta h)$ is infinite and

$$e \simeq \left(1 + \frac{1}{1/(\eta h)}\right)^{\frac{1}{\eta h}} = (1+\eta h)^{\frac{1}{\eta h}} = (e^h)^{\frac{1}{\eta h}} = e^{\frac{1}{\eta}}.$$

Since $\log(x)$ is continuous and $e \simeq e^{1/\eta}$, we have $1 = \log(e) \simeq \log(e^{1/\eta}) = 1/\eta$. But then $\eta = 1/(1/\eta) \simeq 1/1 = 1$. We can conclude that $e^{x_0}\eta \simeq e^{x_0}$.

Suppose then that $h < 0$. Recall that $1/e = \lim_{n\to\infty}\left(1-\frac{1}{n}\right)^n$. Now $-1/(\eta h)$ is infinite and hence

$$1/e \simeq \left(1 - \frac{1}{-1/(\eta h)}\right)^{-\frac{1}{\eta h}} = (1+\eta h)^{-\frac{1}{\eta h}} = (e^h)^{-\frac{1}{\eta h}} = 1/e^{\frac{1}{\eta}}.$$

Then $e = 1/(1/e) \simeq 1/(1/e^{1/\eta}) = e^{1/\eta}$ and similarly as before we conclude $e^{x_0}\eta \simeq e^{x_0}$.

Definition

Whenever $x \in \mathbb{R}$ is given, denote by Δx a non-zero infinitesimal and if $y = f(x)$, then denote by $\Delta y = {}^*f(x + \Delta x) - f(x)$.

Note: f is differentiable at x iff there is a unique $a \in \mathbb{R}$ such that for any infinitesimal Δx, $\Delta y / \Delta x \simeq a$. It follows that

$$f'(x_0) = \operatorname{st}\left(\frac{\Delta y}{\Delta x}\right).$$

Theorem (Increment theorem)

Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable at $x \in \mathbb{R}$, and let $y = f(x)$. Then if Δx is infinitesimal, so is Δy, and moreover,

$$\Delta y = f'(x)\Delta x + \varepsilon \Delta x$$

for some infinitesimal ε.

Contrast this with the classical result: $f(x + h) - f(x) = f'(x)h + \varepsilon(h)\,|h|$ for some function $\varepsilon \colon \mathbb{R} \to \mathbb{R}$ such that $\varepsilon(h) \to 0$ when $h \to 0$.

Definition

Given $x \in \mathbb{R}$, denote $\mathrm{d}x = \Delta x$. When $y = f(x)$ and f is differentiable at x, denote by $\mathrm{d}y$ the number $f'(x)\mathrm{d}x$.

Rewriting the above, we get

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f'(x),$$

while (when Δx is infinitesimal)

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

Now the increment theorem gets the form

$$\Delta y = \mathrm{d}y + \varepsilon \mathrm{d}x.$$

Chain rule

Theorem

If f is differentiable at $x \in \mathbb{R}$ and g is differentiable at $f(x)$, then $g \circ f$ is differentiable at x and $(g \circ f)'(x) = g'(f(x))f'(x)$.

Proof.

Denote $y = f(x)$ and $z = g(y)$. Now by definition, $\mathrm{d}y = f'(x)\mathrm{d}x$ and $\mathrm{d}z = g'(y)\mathrm{d}y = g'(f(x))\mathrm{d}y$. If we cheat a little:

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\mathrm{d}z}{\mathrm{d}y} \cdot \frac{\mathrm{d}y}{\mathrm{d}x} = \frac{g'(f(x))\mathrm{d}y}{\mathrm{d}y} \cdot \frac{f'(x)\mathrm{d}x}{\mathrm{d}x} = g'(f(x))f'(x).$$

But here we did not show that $g \circ f$ actually is differentiable.

Proof cont.

Instead of $\mathrm{d}z$, $\mathrm{d}y$ and $\mathrm{d}x$, look at Δz, Δy and Δx. By the increment theorem, as Δx is infinitesimal, so is Δy. As Δy is infinitesimal, by the increment theorem again, so is Δz and $\Delta z = g'(y)\Delta y + \varepsilon \Delta y$ for some infinitesimal ε. Now

$$\frac{\Delta z}{\Delta x} = \frac{g'(y)\Delta y + \varepsilon \Delta y}{\Delta x} = g'(y)\frac{\Delta y}{\Delta x} + \varepsilon \frac{\Delta y}{\Delta x}.$$

Remember that $\Delta y / \Delta x \simeq f'(x)$. Then

$$g'(y)\frac{\Delta y}{\Delta x} \simeq g'(y)f'(x) \quad \text{and} \quad \varepsilon \frac{\Delta y}{\Delta x} \simeq 0 \cdot f'(x) \simeq 0,$$

so

$$\frac{\Delta z}{\Delta x} \simeq g'(y)f'(x) + 0 = g'(f(x))f'(x).$$

Thus $g \circ f$ is differentiable at x and $(g \circ f)'(x) = g'(f(x))f'(x)$. □

Integrals

- Given a function $f : \mathbb{R} \to \mathbb{R}$, we define the function $S_f : \mathbb{R}^3 \to \mathbb{R}$ by setting

$$S_f(a, b, \Delta x) = \sum_{i=0}^{n-1} f(x_i)\Delta x + f(x_n)(b - x_n)$$

for $a, b, \Delta x \in \mathbb{R}$, where n is the largest natural number such that $a + n\Delta x \leq b$, $x_0 = a$ and $x_{i+1} = x_i + \Delta x$ for $i < n$.

- In other words, $S_f(a, b, \Delta x)$ is the Riemann sum of f on the interval $[a, b]$ with respect to a partition into subintervals of length Δx.
- Now S has a natural extension ${}^*S_f : {}^*\mathbb{R}^3 \to {}^*\mathbb{R}$.
- If $\mathrm{d}x$ is infinitesimal, then $S_f(a, b, \mathrm{d}x)$ is a "sum"

$$\sum_{i=0}^{\omega} f(x_i)\mathrm{d}x + f(x_\omega)(b - x_\omega),$$

where $\omega \in {}^*\mathbb{N}$ is infinite.

Lemma

If $f : \mathbb{R} \to \mathbb{R}$ *is continuous on* $[a, b]$ *and* $\Delta x > 0$ *is infinitesimal, then* ${}^*S(a, b, \Delta x)$ *is a bounded number.*

Theorem

If $f : \mathbb{R} \to \mathbb{R}$ *is continuous on* $[a, b]$ *and* $\mathrm{d}x$ *is a positive infinitesimal, then*

$$\int_a^b f(x)\mathrm{d}x = \mathrm{st}\, {}^*S(a, b, \mathrm{d}x).$$

Note that it follows that for positive infinitesimals $\mathrm{d}x \neq \mathrm{d}y$, ${}^*S(a, b, \mathrm{d}x) \simeq {}^*S(a, b, \mathrm{d}y)$. Also:

$$\int_a^b f(x)\mathrm{d}x \simeq \sum_{i=0}^{1/\mathrm{d}x} f(x_i)\mathrm{d}x \text{ ``$=$'' } \sum_{x \in [a,b]} f(x)\mathrm{d}x$$

for any infinitesimal $\mathrm{d}x > 0$ such that $1/\mathrm{d}x \in {}^*\mathbb{N}$, where $x_0 = a$, $x_{1/\mathrm{d}x} = b$ and $x_{i+1} - x_i = \mathrm{d}x$.

Construction of "the" hyperreals

Definition

A family $U \subseteq \mathcal{P}(\mathbb{N})$ is a *non-principal ultrafilter* if

1. $\emptyset \notin U$,
2. $A \cap B \in U$ for all $A, B \in U$,
3. $A \supseteq B \in U$ implies $A \in U$,
4. for any $A \subseteq \mathbb{N}$, either $A \in U$ or $A^c \in U$, and
5. for any $m \in \mathbb{N}$, $\{n \in \mathbb{N} \mid n > m\} \in U$.

Definition

For $(x_n), (y_n) \in \mathbb{R}^{\mathbb{N}}$, denote $(x_n) \equiv_U (y_n)$ if

$$\{n \in \mathbb{N} \mid x_n = y_n\} \in U.$$

Lemma

$\equiv_U$ is an equivalence relation on $\mathbb{R}^{\mathbb{N}}$.

Definition (Ultrapower)

Denote by $\mathfrak{R}$ the set of $\equiv_U$-equivalence classes. For $(x_n) \in \mathbb{R}^\mathbb{N}$, denote by $[x_n]$ the equivalence class of (x_n). We make $\mathbb{R}^\mathbb{N}/U$ a $\tau_\mathbb{R}$-structure as follows.

- $0^\mathfrak{R} = [0]$ and $1^\mathfrak{R} = [1]$,
- $[x_n] +^\mathfrak{R} [y_n] = [x_n + y_n]$,
- $[x_n] \cdot^\mathfrak{R} [y_n] = [x_n \cdot y_n]$, and
- $[x_n] <^\mathfrak{R} [y_n]$ iff $\{n \in \mathbb{N} \mid x_n < y_n\} \in U$.

Lemma

$\mathfrak{R}$ is well-defined, i.e. the interpretations of the symbols of $\tau_\mathbb{R}$ do not depend on the choice of representatives for the equivalence classes.

We extend $\tau_{\mathbb{R}}$ to a new vocabulary τ by adding a new symbol $\underline{R}$ for *every* relation $R \subseteq \mathbb{R}^m$ and $\underline{f}$ for *every* function $f : \mathbb{R}^m \to \mathbb{R}$.[1] We then interpret $\underline{R}^{\mathbb{R}} = \mathbb{R}$ and $\underline{f}^{\mathbb{R}} = f$ like before. We make $\mathfrak{R}$ a τ-structure by setting

$$([x_n^0], \dots, [x_n^{m-1}]) \in \underline{R}^{\mathfrak{R}} \iff \{n \in \mathbb{N} \mid (x_n^0, \dots, x_n^{m-1}) \in R\} \in U$$

and

$$\underline{f}^{\mathfrak{R}}([x_n^0], \dots, [x_n^{m-1}]) = [f(x_n^0, \dots, x_n^{m-1})].$$

Theorem (Łoś' Theorem)

For any τ-formula $\varphi(v_0, \dots, v_{m-1})$ and $(a_n^0), \dots, (a_n^{m-1}) \in \mathbb{R}^{\mathbb{N}}$,

$$\mathfrak{R} \models \varphi([a_n^0], \dots, [a_n^{m-1}]) \iff \{n \in \mathbb{N} \mid \mathbb{R} \models \varphi(a_n^0, \dots, a_n^{m-1})\} \in U.$$

[1] This τ has a huge size, in fact $|\tau| = 2^{2^{|\mathbb{N}|}}$. This, however, doesn't cause an issue.

Lemma

The map $\iota\colon \mathbb{R} \to \mathfrak{R}$ *defined by*

$$\iota(x) = [x] = \{(x_n) \mid \{n \in \mathbb{N} \mid x_n = x\} \in U\},$$

is an elementary embedding, i.e. for any τ*-formula* $\varphi(v_0, \dots, v_{m-1})$ *and* $a_0, \dots, a_{m-1} \in \mathbb{R}$*,*

$$\mathbb{R} \models \varphi(a_0, \dots, a_{m-1}) \iff \mathfrak{R} \models \varphi(\iota(a_0), \dots, \iota(a_{m-1})).$$

Lemma

$\mathfrak{R}$ *contains infinite numbers and infinitesimals.*

Proof.

The equivalence class x of the sequence $(n)_{n\in\mathbb{N}}$ is an infinite number: as U is a non-principal ultrafilter, for every $m \in \mathbb{N}$, $\{n \mid m < n\} \in U$, and this set is the set of indices where (n) is larger than m. Due to Łoś' Theorem, $\mathfrak{R} \models x > \underline{m}$. But this holds for all m, so $x > m$ for all m. Then $1/x$ is an infinitesimal. □

V. Kanovei and S. Shelah.
A definable nonstandard model of the reals.
J. Symbolic Logic, 69(1):159–164, 2004.

H. J. Keisler.
Elementary Calculus: An Infinitesimal Approach.
Dover Books on Mathematics Series. Dover Publications, 2012.

P. A. Loeb.
An introduction to nonstandard analysis.
In *Nonstandard analysis for the working mathematician*, volume 510 of *Math. Appl.*, pages 1–95. Kluwer Acad. Publ., Dordrecht, 2000.
With an appendix by Horst Osswald.

A. Robinson.
Non-standard analysis.
Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1996.
Reprint of the second (1974) edition, With a foreword by Wilhelmus A. J. Luxemburg.