$\mathbb{R}, dx$ and ε

Derivatives and Infinitesimal Numbers

We use derivatives all day

- Looking for extrema:

$$f'(x) = 0$$

- Expressing conntection between quantities:

$$y' = f(y, x)$$

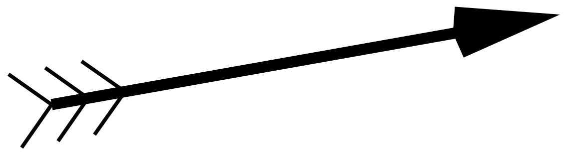
- Calculating norms or constaints:

$$\|f\| = \int |\nabla f|$$

Zenons Paradox

Zenon aus Elea [490 v. u. Z - 425 v. u. Z]:

A flying arrow isn't moving. Because at every point of time, the arrow takes over a place, which is similar to its shape. It can neither take over a bigger place, nor be at two places at the same time.
Since there is nothing between a moment and the next one and because the arrow can not move at a point of time, it isn't moving at all.



Different Lights on derivatives

- Dynamically as a limit:

$$f'(x) = \lim_{x_0 \to x} \frac{f(x_0) - f(x)}{x_0 - x}$$

- Static as a linear approximation:
  A linear mapping u is called derivative of f, if for every x_0 holds:

$$\forall \varepsilon > 0 \exists \delta > 0 : \|h\| < \delta \implies \|f(x_0 + h) - f(x_0) - u(h)\| < \varepsilon \|h\|$$

- Like a physicist as a quotient of infinitely small quantities:

$$f' = \frac{df}{dx}$$

Or (especially for Ronny):

- As an inverse problem:

$$f' = A^{-1} f \text{ where } A \text{ is the compact operator } A : f \mapsto \int f$$

The Discovery of the Derivative

- NEWTON [1643 – 1727]
- LEIBNITZ [1646 – 1716]
- CAUCHY [1789 – 1857], WEIERSTRASS [1815 – 1897]

NEWTON's Fluxions

NEWTON called a variable quantity a *fluent*. He called their velocity (i. e. their derivatives) *fluxions*.
The third important term in the fluxion calculus is the *moment* of a fluxion. NEWTON defines it as an just about noticeable increment to a quantity and denotes it with o.
So o is the moment of time, xo is the moment of the fluent x and $\dot{x}o$ is the moment of the fluxion. Today this is known as the differential dx. The moment of the fluxion $\dot{x}o$ is the velocity multiplied by an infinitesimal interval of time.

Example calculation in NEWTON's style:
A point moves along the curve described by $x^2 - y = 0$.
Substitute $x + \dot{x}o$ in place of x and $y + \dot{y}o$ in place of y and delete the term $x^2 - y$ which is zero:

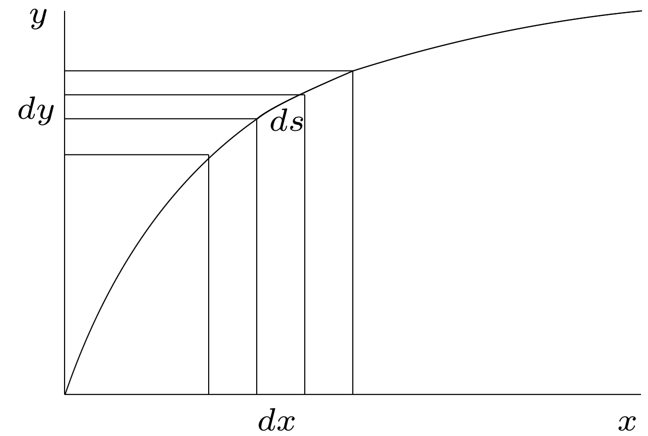
$$2x\dot{x}o + (\dot{x}o)^2 - \dot{y}o = 0 \text{ division by } o \text{ gives: } 2x\dot{x} + (\dot{x})^2 o - \dot{y} = 0.$$

Since o is an infinitesimal interval of time we have

$$2x\dot{x} - \dot{y} = 0 \text{ which is in todays notation } 2x = \frac{\dot{y}}{\dot{x}} = \frac{\dot{y}o}{\dot{x}o} = \frac{dy}{dx}.$$

LEIBNITZ and infinitesimal numbers

LEIBNITZ imagines a subdivision of the real x-axis into infinitely many infinitesimal intervals with extremes $x_1, x_2, x_3, \ldots$.
He defines $dx = x_{n+1} - x_n$.
On the curve and on the y-axis one has the corresponding succesions $s_1, s_2, s_3, \ldots$ and $y_1, y_2, y_3, \ldots$.
The area under the curve is the sum of infinitely many stripes ydx and denoted by $\int ydx$.

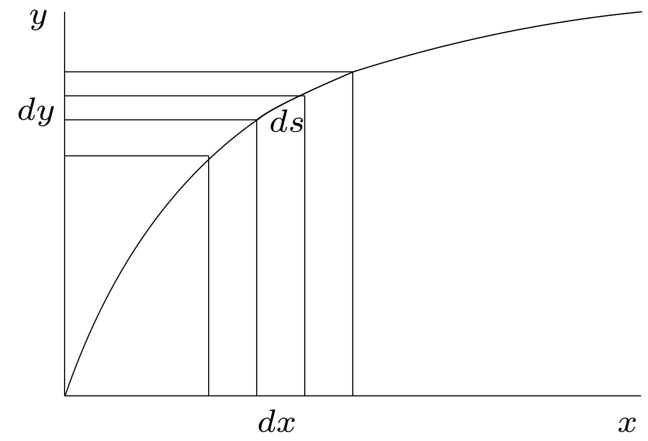


The symbols d and $\int$ applied to finite quantities generate infinitely small and infinitely great quantities. So if x is a finite angle, dx is a infinitely small angle.
Thus the symbols d and $\int$ change the order of infinity but preserve the geometrical dimensions. (NEWTON's dot-symbol do something else: If x is a flowing line, $\dot{x}$ is its velocity.)
The symbols d and $\int$ can be iterated to obtain higher-order infinitesimals or higher-order infinites. So ddx is infinitely small compared to dx and $\int\int x$ is infinitely great compared to $\int x$.

Example calculation in Leibnitz's style

The find the area under a curve we have to calculate $\int y dx$. This can be done by finding a z such that $dz = y dx$. Thus at once

$$\int y dx = \int dz = d \int z = z.$$



The Derivative in the 19th Century

The first definitions in infinitesimal calculus which are like today date back to CAUCHY [1789 – 1857]:

Definition 1 (Derivative, 1823). *When a function* $y = f(x)$ *remains continuous between two given limits of the variable* x*, and when one assigns to such a variable a value enclosed between the two limits at issue, then an infinitely small increment assigned to the variable produces an infinitely small increment in the function itself. Consequently, if one puts* $\Delta x = i$*, the two terms of the ratio of differences*

$$\frac{\Delta y}{\Delta x} = \frac{f(x+i) - f(x)}{i}$$

will be infinitely small quantities. But though these two terms will approach the limit zero indefinitely and simultaneously, the ration itself can converge towards another limit, be it positive or be it negative. This limit, when it exists, has a definite value for each perticular value of x*; but it varies with* $x \ldots$*; The form of the new function which serves as the limit of the ratio* $(f(x+i) - f(x))/i$ *will depend on the form of the proposed function* $y = f(x)$*. In order to indicate this dependance, one gives the new function the name* derived function, *and designates it with the aid of an accent by the notation* y' or $f'(x)$.

Today’s formalism: Epsilontik – Handcuffs for Analysis

CAUCHY and WEIERSTRASS developed the pecise notion with ε and δ.

Example Calculation: Differentiation of $f(x) = x^2$:

$$f'(x) = \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h\to 0} \frac{2xh - h^2}{h}$$
$$= \lim_{h\to 0} 2x + h.$$

For every $\varepsilon > 0$ we chose $h = \varepsilon/2$ and so we have $h = \varepsilon/2 < \varepsilon$ which shows $\lim_{h\to 0} h = 0$. Thus $f'(x) = 2x$.

Is a possible to do calculation with infinitesimal numbers in a mathematically justified way?

Non Standard Analysis

Construct a field ${}^*\mathbb{R}$ which contains $\mathbb{R}$ and in addition infinitesimal numbers $x \in {}^*\mathbb{R}$, $x \neq 0$, i. e.:

$$\forall n \in \mathbb{N} : 0 < x \leq \frac{1}{n}$$

What we know: From $\mathbb{Q}$ to $\mathbb{R}$:

1. Consider $M = \{(a_n) \in \mathbb{Q}^{\mathbb{N}} \mid (a_n)$ is a CAUCHY sequenz$\}$
2. Define an equivalence relation: $(a_n) \sim (b_n) \Leftrightarrow \lim_{n\to\infty}(a_n - b_n) = 0$.
3. Let $\mathbb{R} = M/\sim$
4. Define the operations $+, -, *, \div$ via the members of the sequences and verify the rules of calculation.

From $\mathbb{R}$ to ${}^*\mathbb{R}$

Essential tool for the contruction: **filters**.

Definition 2. *Let I be a set. $\mathcal{F} \subset \mathfrak{P}(I)$ is called filter, if*

1. $\mathcal{F} \neq \emptyset$ *and* $\emptyset \notin \mathcal{F}$
2. $A, B \in \mathcal{F} \implies A \cap B \in \mathcal{F}$
3. $A \in \mathcal{F}, A \subset B \implies B \in \mathcal{F}$

A filter which has no filter above (i. e. $\mathcal{F} \subset \mathcal{G} \implies \mathcal{F} = \mathcal{G}$) is called ultra filter.

Example 1. Let $a \in I$. The filter

$$\mathcal{U}_a = \{A \subset I \mid a \in A\}.$$

is called neighborhood filter for a (actually an ultra filter).

Example 2. Let I be an infinite set (e. g. $I = \mathbb{N}$). Then

$$\mathcal{F} = \{A \subset I \mid I \setminus A \text{ is finite}\}$$

is the so called filter of co-finite sets.

From $\mathbb{R}$ to ${}^*\mathbb{R}$

Theorem 1. *For every filter $\mathcal{F}$ the is an ultra filter which contains the filter $\mathcal{F}$.*

Beweis: Zorns lemma.

According to this theorem there is an ultra filter over $\mathbb{N}$ which contains the co-finite sets in $\mathbb{N}$. (Such a filter couldn't been given explicit but there are a lot of them: As much as $\mathfrak{P}(\mathfrak{P}(\mathbb{N}))$ has elements.)

In the following $\mathcal{F}$ is such an ultra filter.

Construction of the hyperreals:

1. Consider $\alpha \in \mathbb{R}^{\mathbb{N}}$.
2. Define $\alpha \sim \beta \Leftrightarrow \{i \in \mathbb{N} \mid \alpha(i) = \beta(i)\} \in \mathcal{F}$
3. Let ${}^*\mathbb{R} = \mathbb{R}^{\mathbb{N}} / \sim$.

Use representatives: For $\alpha \in \mathbb{R}^{\mathbb{N}}$ let

$$\overline{\alpha} = \{\beta \in \mathbb{R}^{\mathbb{N}} \mid \beta \sim \alpha\}$$

Arithmetic in $^*\mathbb{R}$

Define the operations $+, -, *, \div$ for representatives. E. g. the new addition $^*+$:

$$\overline{\alpha} \;^*{+}\; \overline{\beta} := \overline{\alpha + \beta}$$

$^*-$, $^*\cdot$, $^*\geq$ analogous.

$(^*\mathbb{R}, \;^*+ \;, \;^*\cdot \;, \;^*\geq\;)$ is a complete, ordered field.

Theorem 2. *1. There are non zero infinitesimal elements in* $^*\mathbb{R}$*, i. e. there is an* $\overline{\alpha}$*, so that* $|\overline{\alpha}| \leq 1/n$ *for every* $n \in \mathbb{N}$*.*

2. There are infinite elements in $^*\mathbb{R}$*, i. e. there is* $\overline{\alpha}$*, so that* $|\overline{\alpha}| \geq n$ *for every* $n \in \mathbb{N}$*.*

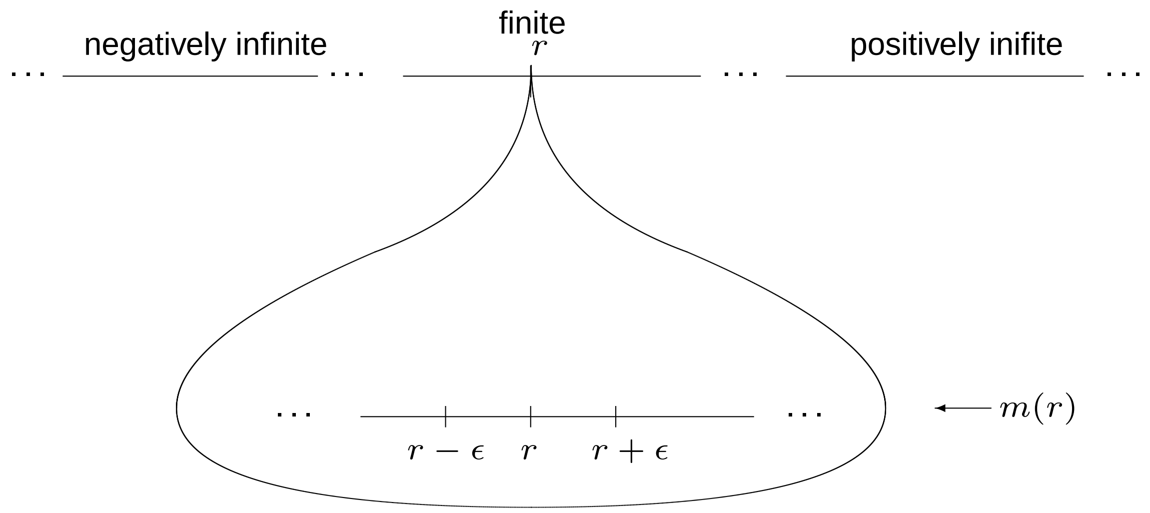
Namely:
Let $\alpha(i) = 1/i$. Then $\overline{\alpha} \neq 0$ and α is infinitesimal. $\overline{\beta}$ is infinite for $\beta(i) = i$.

Definition 3. *If* $\overline{\alpha} - \overline{\beta}$ *is infinitely small, we say that* $\overline{\alpha}$ *is an infinitesimal neighbor of* $\overline{\beta}$ *and we write* $\overline{\alpha} \approx \overline{\beta}$*.*

${}^*\mathbb{R}$

- For every finite $\overline{\alpha} \in {}^*\mathbb{R}$ there is one and only one $r \in \mathbb{R}$, which is an infinitesimal neighbor of $\overline{\alpha}$.
- For every $r \in \mathbb{R}$ there is the *monad* for r:

$$m(r) := \{r \pm \varepsilon \mid \varepsilon \approx 0\}$$



Notation:
$\mathrm{fin}({}^*\mathbb{R})$ is the set of finite hyperreal numbers (i. e. not infinitely large).
If $\overline{\alpha} \in \mathrm{fin}({}^*\mathbb{R})$, we write $\mathrm{st}(\overline{\alpha})$ for the real number r with $r \approx \overline{\alpha}$ and call it the standard part of $\overline{\alpha}$.

Remarks on the Construction of ${}^*\mathbb{R}$

Why do we need an ultra filter?

$$(\mathbb{R}^{\mathbb{N}}/\mathcal{F},\ {}^*\!+\ ,\ {}^*\!\cdot\ ) \text{ is a field} \implies \mathcal{F} \text{ is and ultra-filter.}$$

Why this special filter?

$\mathbb{R}^{\mathbb{N}}/\mathcal{F}$ has an infinite element $\Leftrightarrow$ $\mathcal{F}$ contains the co-finite sets

$\implies$ $\mathbb{R}^{\mathbb{N}}/\mathcal{F}$ has an infinitesimal element different from zero.

Functions on ${}^*\mathbb{R}$

For every function $f : \mathbb{R} \longrightarrow \mathbb{R}$ we define ${}^*f : {}^*\mathbb{R} \longrightarrow {}^*\mathbb{R}$ via

$$ {}^*f(\overline{\alpha}) = \overline{\beta} \ \text{ where } \ \beta(i) := f(\alpha(i)) \ \text{ for } \ i \in \mathbb{N}. $$

Definition 4. $f : \mathbb{R} \longrightarrow \mathbb{R}$ *is continuous in* x_0 *if*

$$ \big(x \in {}^*\mathbb{R} \ \textit{and} \ x \approx x_0\big) \implies {}^*f(x) \approx f(x_0). $$

Definition 5. $f : \mathbb{R} \longrightarrow \mathbb{R}$ *is differentiable in* x_0 *with derivative* $f'(x_0) = c$ *if*

$$ \frac{{}^*f(x_0 + dx) - f(x_0)}{dx} \approx c \ \textit{ for every } \ 0 \neq dx \approx 0. $$

Or, with the help of the standard part:

$$ f'(x_0) = \mathrm{st}\left(\frac{{}^*f(x_0 + dx) - f(x_0)}{dx}\right) \ \textit{ independent of } 0 \neq dx \approx 0. $$

Every Object of Analysis Can be Transferred to Non-Standard-Analysis

We define a set which contains every object of interest in anlysis:
Let $\mathbb{R}_0 = \mathbb{R}$ and $\mathbb{R}_{n+1} = \mathbb{R}_n \cup \mathfrak{P}(\mathbb{R}_n)$.

$$\widehat{\mathbb{R}} = \bigcup_{n\in\mathbb{N}} \mathbb{R}_n \text{ is called the super structure of } \mathbb{R}$$

$\widehat{\mathbb{R}}$ has the property

$$A \subset \mathbb{R}_n \implies A \in \mathbb{R}_{n+1}.$$

Thus $A \subset \widehat{\mathbb{R}} \implies A \in \widehat{\mathbb{R}}$.

Examples:

- The *ordered pair* of two real numbers is an element of $\widehat{\mathbb{R}}$:

$$(a,b) := \{ \underbrace{\{a\}}_{\in\mathfrak{P}(\mathbb{R})}, \underbrace{\{a,b\}}_{\in\mathfrak{P}(\mathbb{R})} \} \subset \mathfrak{P}(\mathfrak{P}(\mathbb{R})) \subset \mathbb{R}_2 \subset \widehat{\mathbb{R}}$$

- Thus $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} \in \widehat{\mathbb{R}}$
- Any *relation* (like $<$) is an element of $\widehat{\mathbb{R}}$:

$$<:= \{(a, b) \in \mathbb{R}^2 \mid a < b\} \subset \mathbb{R}^2$$

Transfer of Objects

Let $a \in \widehat{\mathbb{R}}$. We define

$$a_{\mathbb{N}} : \mathbb{N} \to \widehat{\mathbb{R}} \text{ via } a_{\mathbb{N}}(n) := a \text{ for all } n \in \mathbb{N}$$

Now the corresponding object for a in non-standard analysis is

$$^*a = \overline{a_{\mathbb{N}}} \text{ where } \overline{a_{\mathbb{N}}} \text{ is the equivalence class of } a_{\mathbb{N}} \text{ defined similar to } \sim$$

Example: There is a set of hypernatural numbers $^*\mathbb{N}$. This set contains infinitely large numbers $N \in {}^*\mathbb{N} \setminus \mathbb{N}$.

Non-Standard Analysis of Sequences

For every sequence $a : \mathbb{N} \longrightarrow \mathbb{R}$ there is the corresponding sequence ${}^*a : {}^*\mathbb{N} \longrightarrow {}^*\mathbb{R}$.

The numbers ${}^*a_N$ can be seen as the “last” numbers in the sequence:

c is the limit of a if ${}^*a_N \approx c$ for every $N \in {}^*\mathbb{N} \setminus \mathbb{N}$.

Or

If a converges, the limit of a is given by $\mathrm{st}({}^*a_N), N \in {}^*\mathbb{N} \setminus \mathbb{N}$.

The set

$$\{\mathrm{st}({}^*a_N) \mid N \in {}^*\mathbb{N} \setminus \mathbb{N}\}$$

is the set of accumulation points of a.

To the Very End: δ-Functions

Let $f \in C(\mathbb{R})$, $f \geq 0$ and $\int f(x)dx = 1$. For every $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ we have

$$\delta(x) := N\, {}^*f(Nx), \ \text{for } x \in {}^*\mathbb{R}$$

is a δ-function, i. e.

$$^*\!\int \delta(x)dx = 1$$

$$^*\!\int_{-\varepsilon}^{\varepsilon} \delta(x)dx \approx 1, \ \text{for infinitesimal } \varepsilon > 0$$