HIGHER MATHEMATICS

Lectures

Part Two

HIGHER MATHEMATICS

Lectures

Part Two

Stefan Wurm

ATICE LLC, Albany NY

First English language ATICE e-Book Edition | ISBN 978-1-951894-14-6.

Contents

Preface

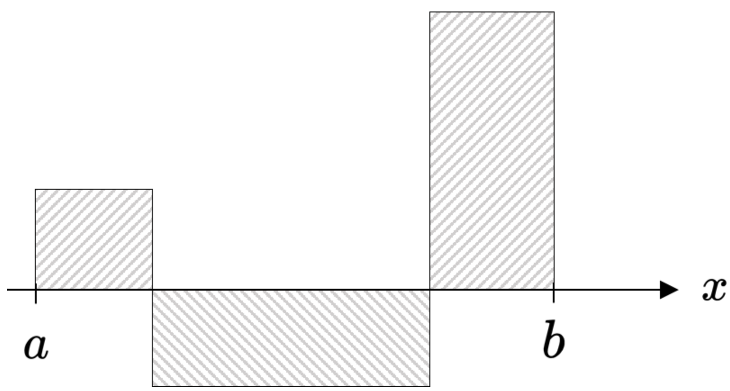
At German universities, lectures on higher mathematics are an integral part of the curriculum in natural and engineering sciences. These lectures aim to provide students with the mathematical foundations for their respective subject areas, typically in the first four semesters. This was also the case for me when a good forty years ago at the beginning of my physics studies I first entered the lecture hall of the Technische Universität München (TUM), the place where Prof. Dr. Armin Leutbecher taught Higher Mathematics. I realize of course that not everyone can or wants to share the same enthusiasm for mathematics. However, I hope that those who are reading these lines will understand what I mean when I say that to me as a student those mathematics courses have been a real source of happiness. Happiness in the sense that back then I always looked forward to each and every one of these lectures. This certainly did not only have to do with the content of the lectures, but at least as much with the way they were delivered by Prof. Leutbecher. Of course, one always expects clarity from a mathematician. But the clarity with which professional mathematicians generally conduct their discussions does not necessarily carry over to how a mathematician might then impart his knowledge to students. Prof. Leutbecher's clarity and style of delivery made his Higher Mathematics lectures an intellectual delight. In addition, I also had the good fortune that the exercises for Prof. Leutbecher lectures were given by Dr. Peter Vachenauer. At the beginning of the 1990s the first edition of a two-volume textbook on Higher Mathematics co-authored by Dr. Vachenauer was published. The exemplary methodology and care with which the lecture materials were studied during my time at TUM in Dr. Vachenauer's tutorial exercises is reflected in this textbook.

A little over a year ago, while tidying up, I stumbled across my transcripts of the Higher Mathematics lectures from the years 1981-1983 and the corresponding exercises. At first I was surprised that these forty-year-old documents were not lost during various moves over four decades, some of them between continents. When I then curiously began to leaf through my rediscovered lecture notes I all of a sudden experienced the same kind of joy which I once felt when I was sitting in the lecture hall, listening spellbound to Prof. Leutbecher's lectures some forty years ago. Although these notes, my transcript of Prof. Leutbecher's lectures, cannot replace a textbook, they do convey the essential content of Higher Mathematics with a vividness that I believe should make them a reading pleasure for students or anyone else seriously interested in mathematics. All too often such lecture notes are riddled with errors, and this was no different here. After reviewing and correcting my notes several times, hopefully the vast majority of them have been corrected. Preserving the clarity and style of Prof. Leutbecher's lectures, as I captured them in my notes more than forty years ago, was something I attached great importance to when revising my notes. Translating those notes from their original German into English added of course another challenge. Quite likely some of the elegance of the German language lectures may have been lost in translation. However, I do hope that the English language version still conveys the essence of the lectures original style and clarity. This volume, **HIGHER MATHEMATICS - Lectures Part Two**, contains the material of the Higher Mathematics II lectures as given by Prof. Leutbecher in the summer semester 1982 at the TUM.

Stefan Wurm

Albany, New York
May 2022

1. Fundamentals of Integral Calculus



An integral for step functions

$\varphi :\ [a,b] \to \mathbb{R}$ is called a step function if a subdivision of $[a,b]$ exists, i.e. a finite set of points $a = x_0 < x_1 < x_2 < \ldots < x_n = b$, such that each of the restrictions $\varphi/\,]x_{k-1}, x_k[$ is constant. Such a subdivision is called compatible with φ. Each refinement (addition of further points) is also compatible with φ!

The integral as a sum over rectangles:

$$\int_a^b \varphi(x)\,\mathrm{d}x \ := \ \sum_{k=1}^{n} \varphi\left(\frac{x_k + x_{k-1}}{2}\right)(x_k - x_{k-1})$$

This definition is independent of the choice of the φ-permissible subdivision because every two subdivisions share a common refinement!

Size comparison between real functions!

$f, g :\ D \to \mathbb{R}$

In case $f \leq g$ holds one says

$$f(x) \;\leq\; g(x) \qquad \forall\, x \in D$$

Elementary properties of the integral for step functions

For every two step functions ψ, φ on $[a, b]$ the following applies:

$$\left.\begin{aligned} \int_a^b (\varphi + \psi)(x)\,\mathrm{d}x \;&=\; \int_a^b \varphi(x)\,\mathrm{d}x + \int_a^b \psi(x)\,\mathrm{d}x \\ \int_a^b (\lambda\,\varphi)(x)\,\mathrm{d}x \;&=\; \lambda \int_a^b \varphi(x)\,\mathrm{d}x \end{aligned}\right\} \qquad \underline{\text{Linearity}}$$

$$\int_a^b \varphi(x)\,\mathrm{d}x \;\leq\; \int_a^b \psi(x)\,\mathrm{d}x \;, \quad \text{if} \quad \varphi \leq \psi \qquad \underline{\text{Monotony}}$$

One has to note that subdivisions of $[a, b]$ exist which are compatible with φ and ψ at the same time!

Every function f that is continuous on $[a, b]$ is even uniformly continuous there, i.e. for every $\epsilon > 0$ exists a $\delta > 0$ with

$$|f(x) - f(x')| \;<\; \epsilon \qquad \text{if} \;\; x, x' \in [a, b] \;\; \text{and} \;\; |x - x'| \;<\; \delta$$

Indirect proof:

What is the negation of the assertion in this theorem? An $\epsilon_0 > 0$ exists such that for every $\delta > 0$, no matter how small, points $x, x' \in [a, b]$ exist with

$$|x - x'| \;<\; \delta \qquad \text{but} \qquad |f(x) - f(x')| \;\geq\; \epsilon_0$$

The thought behind the proof: $\delta = 1/k$ for $k = 1, 2, 3, \ldots$; Then sequences of points (x_k), (x'_k) exist on $[a, b]$ with

$$|x_k - x'_k| \;<\; \frac{1}{k} \qquad (*)$$

$$|f(x_k) - f(x'_k)| \;\geq\; \epsilon_0 \qquad (**)$$

That leads to a contradiction. According to Bolzano-Weierstraß one finds subsequences $x_{n_k} = y_k$, $x'_{n_k} = y'_k$ which converge on $[a, b]$. Let for example

$$y_\infty \;=\; \lim_{k\to\infty} y_k \quad \text{then also because of } (*) \quad \lim_{k\to\infty} y'_k \;=\; y_\infty$$

According to $(**)$

$$\epsilon_0 \;\leq\; |f(y_k) - f(y'_k)| \;\leq\; |f(y_k) - f(y_\infty)| \;+\; |f(y_\infty - f(y'_k))|$$

Because f is continuous in y_∞, the right-hand side converges to 0, so it is not always $\geq \epsilon_0$. Therefore f is, contrary to the assumption, uniformly continuous.

$\square$

<u>Approximation by step functions</u>

Let $f : [a,b] \to \mathbb{R}$ be continuous. Then for every $\epsilon > 0$ there exist step functions φ, ψ with $\varphi \leq f \leq \psi$ and $\psi(x) - \varphi(x) \leq \epsilon \;\; \forall\, x \in [a,b]$

$$\psi(x) \;:=\; f\left(\frac{x_{k-1}+x_k}{2}\right) \;+\; \frac{\epsilon}{2}$$

$$\varphi(x) \;:=\; f\left(\frac{x_{k-1}+x_k}{2}\right) \;-\; \frac{\epsilon}{2} \qquad \text{if } x_{k-1} < x < x_k$$

For continuous $f : [a,b] \to \mathbb{R}$ therefore holds

$$\inf_{\psi \geq f} \int_a^b \psi(x)\,\mathrm{d}x \;=\; \sup_{\varphi \leq f} \int_a^b \varphi(x)\,\mathrm{d}x$$

where φ runs through the step functions $\leq f$ and ψ through the step functions $\geq f$.

Definition

$$\int_a^b f(x)\,\mathrm{d}x \;:=\; \sup_{\varphi \leq f} \int_a^b \varphi(x)\,\mathrm{d}x$$

The identity of infimum of upper sums and supremum of lower sums holds for piecewise continuous functions, i.e. those f for which a subdivision

$$E \;=\; \{a = x_0 < x_1 < x_2 < \ldots < x_n = b\} \qquad (x_{k-1} \;<\; x_k)$$

exists with continuous restrictions $f/\,]x_{k-1}, x_k[\;\; 1 \leq k \leq n$, which can also be extended continuously into the end points.

Elementary integration rules

$$\left.\begin{aligned}\int_a^b (f+g)(x)\,\mathrm{d}x &= \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x \\ \int_a^b (\lambda f)(x)\,\mathrm{d}x &= \lambda\int_a^b f(x)\,\mathrm{d}x\end{aligned}\right\}\quad \underline{\text{Linearity}}$$

$$\int_a^b f(x)\,\mathrm{d}x \le \int_a^b g(x)\,\mathrm{d}x\ ,\quad \text{falls}\quad f\le g \qquad \underline{\text{Monotony}}$$

Integral estimates

a) If $f, g\,[a,b] \to \mathbb{R}$ piecewise continuous, $m \le f(x) \le M \quad \forall x \in [a,b],\ g(x) \ge 0$ $\forall x \in [a,b]$ then

$$m\int_a^b g(x)\,\mathrm{d}x \le \int_a^b f(x)g(x)\,\mathrm{d}x \le M\int_a^b g(x)\,\mathrm{d}x$$

b) If f is also continuous on $[a,b]$, then a $\xi \in [a,b]$ exists with

$$\int_a^b f(x)g(x)\,\mathrm{d}x = f(\xi)t\int_a^b g(x)\,\mathrm{d}x$$

This is the mean value theorem of integral calculus!

The proof of the first two inequalities **a)** uses only monotony and linearity of the integral. For example:

$$\int_a^b f(x)g(x)\,\mathrm{d}x - m\int_a^b g(x)\,\mathrm{d}x =$$

$$= \int_a^b \underbrace{(f(x)-m)g(x)}_{\substack{\ge 0 \\ \text{according to the premise}}}\,\mathrm{d}x \underset{\substack{\downarrow \\ \text{Monotony}}}{\ge} \int_a^b 0\,\mathrm{d}x = 0$$

For part **b)**: $\begin{aligned} m &= \text{minimum } f(x) \ \ x\in[a,b] \\ M &= \text{maximum } f(x) \ \ x\in[a,b]\end{aligned}$

m, M are function values of f according to the maximum and minimum theorem.[1]

[1] Compare with chapter 5, section 3 in HIGHER MATHEMATICS Lectures Part One.

According to **a)** $\displaystyle\int_a^b f(x)g(x)\,\mathrm{d}x$ is a number between

$$m\int_a^b g(x)\,\mathrm{d}x \qquad \text{and} \qquad M\int_a^b g(x)\,\mathrm{d}x$$

hence equal to

$$Z\int_a^b g(x)\,\mathrm{d}x \qquad \text{with} \qquad m \le Z \le M$$

Intermediate value theorem: $Z \;=\; f(\xi)$

□

Remark:

(1) If one chooses $M := \sup|f(x)|\;\; x \in [a,b]$, $m = -M$ and $g(x) = 1\;\;\forall\, x \in [a,b]$, then the standard integral estimate follows

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \;\le\; |b-a|\,M$$

Differentiation and integration

The integral is considered as a function of the upper (or lower) integral limit. To standardize the calculation

$$\int_a^a f(x)\,\mathrm{d}x \;=\; 0 \quad ; \quad \int_b^a f(x)\,\mathrm{d}x \;=\; -\int_a^b f(x)\,\mathrm{d}x \qquad (\text{if } a < b)$$

If f is a continuous real function on the open interval I, then for every triple $a_1, a_2, a_3 \in I$ the following holds:

$$\int_{a_1}^{a_2} f(x)\,\mathrm{d}x + \int_{a_2}^{a_3} f(x)\,\mathrm{d}x \;=\; \int_{a_1}^{a_3} f(x)\,\mathrm{d}x$$

If thus $a = \min(a_1, a_2, a_3)$, $b = \max(a_1, a_2, a_3)$ and c is the third number, then

$$\int_a^b f(x)\,\mathrm{d}x \;+\; \int_a^c f(x)\,\mathrm{d}x \;=\; \int_c^b f(x)\,\mathrm{d}x$$

All six cases follow from this by interchanging integral limits. The standard integral estimate remains valid for the extension.

Existence theorem for primitive functions

If f is a continuous function on the open interval I and $a \in I$, then

$$F(x) \;:=\; \int_a^x f(t)\,\mathrm{d}t$$

defines a primitive function F of f on I, which means that F is differentiable on I with the derivative

$$F'(x) \;=\; f(x) \qquad \forall\, x \in I$$

Remark:

(2) The mean value theorem of differential calculus shows that every two primitive functions of f on the interval I have a constant as the difference.

Proof:

Assume $x \in I$ is fixed and $\delta > 0$ always so small that $[x-\delta, x\delta] \subset I$. Let $\epsilon > 0$ be arbitrary. Then estimate the absolute value of

$$\frac{1}{h}F(x+h) \;-\; \frac{1}{h}F(x) \;-\; f(x)\,,$$

which is the difference between the differential quotients for $F(x)$ and $f(x)$.

$$\left| \frac{1}{h}\left(\int_a^{x+h} f(t)\,\mathrm{d}t \;-\; \int_a^x f(t)\,\mathrm{d}t \right) \;-\; f(x) \right| \;=\; \left| \frac{1}{h} \int_x^{x+h} f(t)\,\mathrm{d}t \;-\; \frac{1}{h} \int_x^{x+h} f(x)\,\mathrm{d}x \right|$$

$$= \;\left|\frac{1}{h}\right| \cdot \left| \int_x^{x+h} (f(t) \;\;-\;\; f(x))\,\mathrm{d}t \right| \;\leq\; \left|\frac{1}{h}\right| \cdot |h| \cdot \sup_{|t-x|\,\leq\,\delta} |f(t) - f(x)|$$

Since f is continuous in x, a $\delta > 0$ exists with

$$\sup_{|t-x|\,\leq\,\delta} |f(t) - f(x)| \;\leq\; \epsilon$$

This shows: $$\lim_{\substack{h\to 0\\ h\neq 0}} \frac{1}{h}\big(F(x+h) - F(x)\big) \;=\; f(x)$$

$\square$

1.1 The Fundamental Theorem of Calculus

Let f be continuous on the open interval I, and let F be any primitive function of f. Then for all $a, b \in I$

$$\boxed{\int_a^b f(x)\,\mathrm{d}x \;=\; F(b) - F(a) \;=:\; F(x)\Big|_a^b}$$

<u>Remark:</u> Significance of the fundamental theorem

A primitive function (F) is known for many f. For such f the calculation of

$$\int_a^b f(x)\,\mathrm{d}x$$

is just as easy as calculating the values $F(a)$, $F(b)$; A short notation for primitive functions of f is

$$\int f(x)\,\mathrm{d}x \qquad \text{(indefinite integral)}$$

<u>Example integrals:</u>

<u>1.</u> $$\int x^\alpha\,\mathrm{d}x = \frac{x^{\alpha+1}}{\alpha+1}\,, \quad \text{if} \quad \alpha \in \mathbb{R} \setminus \{-1\}$$

<u>2.</u> $$\int \frac{\mathrm{d}x}{x} = \ln|x| + c$$

<u>3.</u> $$\int \mathrm{e}^x\,\mathrm{d}x = \mathrm{e}^x + c$$

<u>4.</u> $$\int \sin x\,\mathrm{d}x = -\cos x + c\,; \qquad \int \cos x\,\mathrm{d}x = \sin x + c$$

<u>5.</u> $$\int \frac{\mathrm{d}x}{1+x^2} = \arctan x + c$$

<u>6.</u> $$\int \frac{\mathrm{d}x}{1-x^2} = \arcsin x + c \qquad |x| < 1$$

The main integration methods are obtained from a) the chain rule and b) the product rule.

1.1.1 The Substitution Method

If $\varphi : [a,b] \to \mathbb{R}$ is continuously differentiable and f is continuous on the image interval $\varphi([a,b])$, then

$$\int_{\varphi(a)}^{\varphi(b)} f(x)\,\mathrm{d}x \;=\; \int_a^b f(\varphi(t))\varphi'(t)\,\mathrm{d}t$$

Proof:

Let F be a primitive function of f on the image interval $\varphi([a,b])$. The chain rule says:

$F \circ \varphi$ is the primitive function of the function $t \longmapsto f(\varphi(t))\varphi'(t)$

Evaluation according to the fundamental theorem: ($LS =$ left side , $RS =$ right side)

$$LS \;=\; F(x)\Big|_{\varphi(a)}^{\varphi(b)} \;=\; F(\varphi(b)) \,-\, F(\varphi(a))$$

$$RS \;=\; F\circ\varphi(t)\Big|_a^b \;=\; F\circ\varphi(b) \,-\, F\circ\varphi(a)$$

$\square$

An example: Integration of the logarithmic derivative

For continuously differentiable functions φ without any zeros the following applies

$$\int \frac{\varphi'(t)}{\varphi(t)}\,\mathrm{d}t \;=\; \ln|\varphi(t)| \,+\, c$$

Proof comes by the integral example 2) with $f(x) = \dfrac{1}{x}$.

Example integrals (continued):

7. $$\int \frac{x}{1+x^2}\,\mathrm{d}x \;=\; \frac{1}{2}\ln\big(1+x^2\big) \,+\, c$$

8. $$\int \frac{x}{(1+x^2)^{k+1}}\,\mathrm{d}x \;=\; \frac{-1}{2k(1+x^2)^k} \,+\, c$$

One can try a solution for example with substitution $u = \varphi(x) = 1 + x^2$ and integral example 1) or check directly.

1.1.2 The Method of Partial Integration

For any two continuous differentiable functions $u(x), v(x)$ the following applies

$$\int u(x)v'(x)\,\mathrm{d}x \;=\; u(x)v(x) \;-\; \int u'(x)v(x)\,\mathrm{d}x$$

Proof:

The product rule of differential calculus says that $u'(x)v(x) + u(x)v'(x)$ has the primitive function $u(x)v(x)$. Using the linearity of the integral, the assertion follows.

<u>On the integration of rational functions</u>

The decomposition of partial fractions in the real leads to three types of integrals

$$\int p(x)\,\mathrm{d}x \quad , \quad \int \frac{\mathrm{d}x}{(x-x_0)^k} \quad , \quad \int \frac{Ax+B}{(x^2+ax+b)^k}\,\mathrm{d}x$$

where $p(x)$ denotes a polynomial function. One obtains the first two from integral examples 1) and 2). Furthermore

$$a^2 - 4b \;<\; 0 \;<\; \left(x + \frac{a}{2}\right)^2 - \frac{1}{4}(a^2 - 4b)$$

suggests that the third type can be traced back to

$$\int \frac{x}{(1+x^2)^k}\,\mathrm{d}x \quad , \quad \int \frac{\mathrm{d}x}{(1+x^2)^k}$$

The first class is dealt with by the integral examples 7) and 8). The case $k = 1$ in the final class is integral example 5).

Let $\quad J_k(x) \;=\; \displaystyle\int \frac{\mathrm{d}x}{(1+x^2)^k}$:

$$J_k(x) \;-\; J_{k+1}(x) \;=\; \int \frac{x^2}{(1+x^2)^{k+1}}\,\mathrm{d}x$$

Partial integration with:

$$u \;=\; x \quad , \quad v' \;=\; \frac{x}{(1+x^2)^{k+1}} \quad ; \quad v \;=\; \frac{-1}{2k(1+x^2)^k}$$

$$\int \underset{u}{\underset{\downarrow}{x}} \underbrace{\frac{x}{(1+x^2)^{k+1}}}_{v'} \,\mathrm{d}x \;=\; \frac{-x}{2k(1+x^2)^k} + \int \frac{\mathrm{d}x}{2k(1+x^2)^k}$$

<ins>Example integrals (continued):</ins>

<ins>9.</ins> $$\int \frac{\mathrm{d}x}{(1+x^2)^{k+1}}\,\mathrm{d}x \;=\; \frac{2k-1}{2k}\int \frac{\mathrm{d}x}{(1+x^2)^k} + \frac{x}{2k(1+x^2)^k}$$

1.2 The Improper Integral

Previous restrictions: The integrand $f(x)$ is defined on a closed, bounded interval, furthermore f is bounded there. Occasionally one can relinquish these terms. For example, if $f : [a,b[\to \mathbb{R}$ and continuous, then in case the limit value exists one declares as improper integral

$$\lim_{\beta \nearrow b} \int_a^\beta f(x)\,\mathrm{d}x \;=:\; \int_a^b f(x)\,\mathrm{d}x$$

Note that $b = \infty$ is allowed. If f is continuously extendable on $[a,b]$ for a finite b, then the concept of the improper integral is the same as the previous discussed concept of the "proper" integral. Obvious is also the extension to the case that the integral is likewise defined for $]a,b]$, $]a,b[$

<ins>Example:</ins>

(1) $$\int_1^\infty x^\alpha \,\mathrm{d}x$$

converges to $\dfrac{-1}{\alpha+1}$ if $\alpha < -1$, otherwise the integral is divergent.

Proof: First let $\alpha \neq -1$, then

$$\int_1^\beta x^\alpha\,\mathrm{d}x \;=\; \left.\frac{x^{\alpha+1}}{\alpha+1}\right|_1^\beta \;=\; \frac{-1}{\alpha+1} + \frac{\beta^{\alpha+1}}{\alpha+1}$$

For $\alpha+1<0$: $\displaystyle\lim_{\beta\to\infty}\beta^{\alpha+1} \;=\; \lim_{\beta\to\infty} \mathrm{e}^{(\alpha+1)\ln\beta} \;=\; 0$

If $\alpha + 1 > 0$ then $\lim\limits_{\beta \to \infty} \beta^{\alpha+1} = \infty$, the integral is therefore divergent

$$\int_1^\beta \frac{\mathrm{d}x}{x} = \ln \beta$$

$\square$

Example:

(2) $$\int_0^1 x^\alpha \, \mathrm{d}x$$

converges to $\dfrac{1}{\alpha+1}$, if $\alpha > -1$ and otherwise divergent!

Proof:

$$\int_\epsilon^1 x^\alpha \, \mathrm{d}x = -\int_{1/\epsilon}^1 t^{-\alpha-2} \, \mathrm{d}t \qquad x = \frac{1}{t} \quad , \quad \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{-1}{t^2}$$

$$= \int_1^{1/\epsilon} t^{-\alpha-2} \, \mathrm{d}t$$

The limit value for $\epsilon \searrow 0$ exists exactly then (see example 1) when the exponent < -1, i.e. $\alpha > -1$. If this is the case, the limit value is equal to

$$\frac{-1}{(-\alpha-2)+1} = \frac{1}{\alpha+1}$$

Majorant criterion for improper integrals

Let $g : \,]a,b[\, \to \mathbb{R}$ be continuous and ≥ 0, and the improper integral

$$\int_a^b g(t) \, \mathrm{d}t \quad \text{shall exist.}$$

Furthermore, let $f : \,]a,b[\, \to \mathbb{R}$ be continuous and $|f(t)| \leq g(t) \quad \forall\, t \in \,]a,b[\,$. Then also the improper integral

$$\int_a^b f(t) \, \mathrm{d}t$$

converges. To proof this a (refined) version of the Cauchy criterion is used.

<u>Example:</u>

(3) For every real x $\quad \int\limits_1^\infty t^x \mathrm{e}^{-t}\,\mathrm{d}t \quad$ converges

Proof:

According to example 1 $\quad \int\limits_1^\infty t^{-2}\,\mathrm{d}t \quad$ converges

For every bound $M > 0$, $\quad M\int\limits_1^\infty t^{-2}\,\mathrm{d}t \quad$ also converges.

The exponential series yields $\mathrm{e}^t \geq \dfrac{t^N}{N!}$ for each $N \in \mathbb{N}$, if $t \geq 0$. Therefore $\mathrm{e}^{-t} \leq N!\,t^{-N}$ and thus

$$t^x\mathrm{e}^{-t} \;\leq\; N!\,t^{x-N} \;\leq\; N!\,t^{-2} \quad , \qquad \text{if} \quad N \geq x+2$$

According to the majorant criterion, $\quad \int\limits_1^\infty t^x \mathrm{e}^{-t}\,\mathrm{d}t \quad$ is therefore convergent!

1.2.1 Euler's Gamma Function

is declared for positive x by the (convergent) improper integral

$$\Gamma(x) \;:=\; \int\limits_0^\infty t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t$$

The convergence follows from considering the subintervals

$\int\limits_1^\infty t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t \quad$ together with example 3

and

$\int\limits_0^1 t^{x-1}\mathrm{e}^{-t}\,\mathrm{d}t \quad$ together with example 1

and by using the majorant criterion. The following applies specifically:

$$\Gamma(1) \;=\; 1$$

$$\Gamma(x+1) \;=\; x\,\Gamma(x) \;, \qquad \forall\, x > 0 \qquad \text{(functional equation)}$$

Hence $\quad \Gamma(n+1) \;=\; n!$

$$\int_0^\beta e^{-t}\,dt \;=\; -e^{-t}\Big|_0^\beta \;=\; 1 - e^{-\beta} \;\to\; 1\,, \qquad \text{if} \quad \beta \to \infty$$

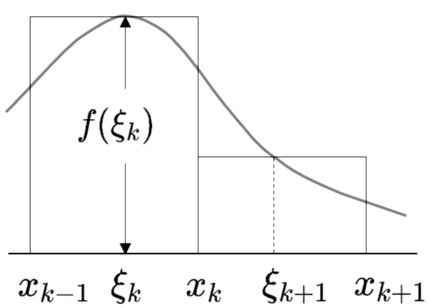
$$\Gamma(x+1) \;\leftarrow\; \int_{(\epsilon}^{\beta} \underbrace{t^x}_{u}\,\underbrace{e^{-t}}_{v'} \;=\; -t^x e^{-t}\Big|_{(\epsilon}^{\beta} + \int_{(\epsilon}^{\beta} x\,t^{x-1}e^{-t}\,dt \;\to\; \Gamma(x)$$

1.3 Riemann Sums and Arc Length

Let $f : [a,b] \to \mathbb{R}$ be piecewise continuous. For each subdivision $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ of the interval $[a,b]$ and for each system of intermediate points $\xi_k \in [x_{k-1}, x_k]\,,\ 1 \le k \le n$

$$S \;=\; \sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1})$$

is called a Riemann sum to the subdivision E for the integral $\int_a^b f(x)\,dx$



$S \;=\; \int_a^b \varrho(t)\,dt$ applies to the special step function

$$\varrho(x) \;=\; f(\xi_k)\,, \quad \text{if} \quad x_{k-1} \le x \le x_k$$

$$\varrho(b) \;=\; f(b) \qquad (1 \le k \le n)$$

The integral $\int_a^b f(t)\,dt$ over the piecewise continuous integrand f is then approximated with arbitrary precision by every Riemann sum having a sufficiently fine subdivision E of $[a,b]$.

Proof (for continuous f):

According to the introductory part of the chapter, f is even uniformly continuous. Let $\epsilon > 0$. Then there exists to $\epsilon' = \dfrac{\epsilon}{b-a}$ a $\delta > 0$ with $|f(x) - f(x')| \le \epsilon'$ for all $x, x' \in [a,b]$ with $|x - x'| \le \delta$. Therefore, for granularity $\max\limits_k |x_k - x_{k-1}| \le \delta$

$$\left| S - \int_a^b f(t)\,\mathrm{d}t \right| \;=\; \left| \int_a^b (\varrho(t) - f(t))\,\mathrm{d}t \right|$$

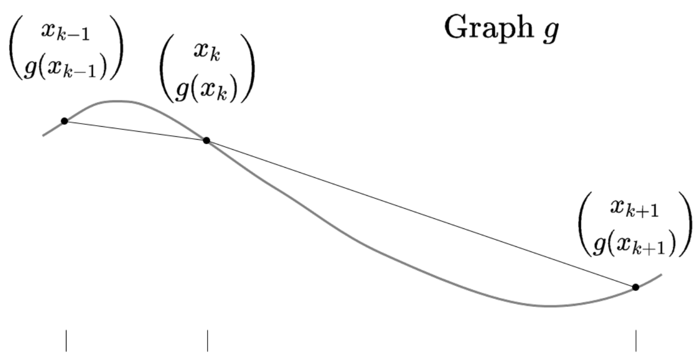
and for that one finds with the standard estimate

$$\leq \; |b-a|\,\epsilon' \;=\; \epsilon$$

$\square$

Let a curve in the plane be given by the (piecewise) continuously differentiable function $g:\,[a,b] \to \mathbb{R}$ in the point set

$$\left\{ \begin{pmatrix} x \\ g(x) \end{pmatrix} ; \quad x \in [a,b] \right\} \quad = \quad \text{Graph } g$$



With a subdivision E of $[a,b]$ neighboring points

$$\begin{pmatrix} x_{k-1} \\ g(x_{k-1}) \end{pmatrix} \quad , \quad \begin{pmatrix} x_k \\ g(x_k) \end{pmatrix}$$

are connected by a segment. The resulting segmented polynomial has the following Euclidean length:

$$S \;=\; \sum_{k=1}^{n} \sqrt{\big(x_k - x_{k-1}^2 + \big(g(x_k) - g(x_{k-1})\big)^2}$$

According to the mean value theorem:

$$g(x_k) \,-\, g(x_{k-1}) \;=\; g'(\xi_k)(x_k - x_{k-1}) \;, \qquad \xi \in [x_{k-1}, x_k]$$

that means:

$$S \;=\; \sum_{k=1}^{n} \sqrt{1 + g'^2(\xi_k)} \cdot (x_k - x_{k-1})$$

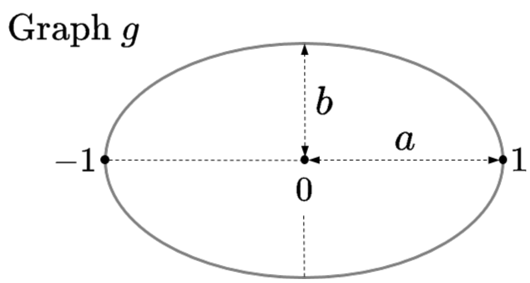
is a Riemann sum with respect to the subdivision E for the integral

$$\int_a^b \sqrt{1+g'^2(t)}\,\mathrm{d}t$$

The lengths of the segmented polynomials for a sufficiently fine subdivision of $[a,b]$ approximate the integral with arbitrary precision. Therefore this integral is the arc length of the curve!

Example:

The ellipse $x^2+\frac{y^2}{b^2}=1$ with semi-major axis $a=1$ and $b=\sqrt{1-\epsilon^2}\ leq 1$ as semi-minor axis.



$$y = g(x) \quad = \quad b\sqrt{1-x^2}$$

$$g'(x) \quad = \quad \frac{-b\,x}{\sqrt{1-x^2}}$$

$$2\int_{-1}^{1} \sqrt{1+g'^2(x)}\,\mathrm{d}x \quad = \quad 2\int_{-1}^{1} \sqrt{\frac{1-\epsilon^2x^2}{1-x^2}}\,\mathrm{d}x$$

This is one of the so-called elliptic integrals. It cannot be evaluated with the elementary functions discussed so far. Moreover, it is an improper integral. The convergence: from the majorant criterion

$$2\int_{-1}^{1} \frac{\mathrm{d}x}{\sqrt{1-x^2}} \quad = \quad \lim_{\epsilon\searrow 0} \arcsin x \Big|_{-1+\epsilon}^{1+\epsilon} \quad = \quad 2\pi$$

This is also the circumference in the special case $\epsilon = 0$, i.e. the circumference of the unit circle.

2. Fourier Series

Periodic processes are described with periodic functions f, f may be real or complex.

So there are $T \neq 0$ with

$$f(x+T) \;=\; f(x)\,, \qquad x \in \mathbb{R}$$

If one replaces $t = \dfrac{2\pi x}{T}$ and considers $g(t) = f(x)$, then $g(t)$ has the period 2 pi. Therefore, in principle, it is sufficient to consider periodic functions with period 2π.

Well-known examples are $\sin x$ and $\cos x$ and in complex notation:

$$t \;\longmapsto\; \mathrm{e}^{it} \;=\; \cos t \,+\, i \sin t$$

$$t \;\longmapsto\; \mathrm{e}^{int} \;=\; \cos nt \,+\, i \sin nt \qquad n \in \mathbb{Z}$$

The last function has in fact the smallest period $\dfrac{2\pi}{n}$ but (as a multiple) 2π is also a period.

For every two functions f_1,f_2 with the period 2π and every two (real or complex) numbers λ_1, λ_2, the "linear combination"

$$f \;:=\; \lambda_1 f_1 \,+\, \lambda_2 f_2 \qquad (\, f(t) \;:=\; \lambda_1 f_1(t) \,+\, \lambda_2 f_2(t) \quad \forall\, t \in \mathbb{R}\,)$$

is a periodic function with period 2π. This results in a wealth of 2π-periodic functions in the "trigonometric polynomials"

$$\begin{aligned} t \;\longmapsto\; & \sum_{n=-N}^{N} c_n \mathrm{e}^{int} \\ &= c_0 + (c_1\mathrm{e}^{it} + c_{-1}\mathrm{e}^{-it}) + \ldots + (c_N\mathrm{e}^{iNt} + c_{-N}\mathrm{e}^{-iNt}) \\ &= c_0 + (c_1 + c_{-1}) \cos t + \ldots + (c_N + c_{-N}) \cos Nt + \\ &\qquad + i\,(c_1 - c_{-1}) \sin t + \ldots + i\,(c_N - c_{-N}) \sin Nt \end{aligned}$$

The theory of Fourier series deals with the question of which 2π-periodic functions can be approximated by trigonometric polynomials.

Basic example:

$$S_N(x) \;=\; \sum_{n=1}^{N} \frac{\sin nx}{n} \;=\; \sin x + \frac{\sin 2x}{2} + \ldots + \frac{\sin Nx}{N}$$

Assertion: For every $x \in \mathbb{R}$ there exists $S(x) = \lim\limits_{N \to \infty} S_N(x)$, viz

$$S(x) \;=\; \begin{cases} 0 & \text{if} \quad x = 0 \\ \dfrac{\pi - x}{2} & \text{if} \quad 0 < x < 2\pi \\ \quad 2\pi\text{-periodic otherwise.} \end{cases}$$

Proof:

Consider: $$\sum_{n=0}^{N} \mathrm{e}^{int} \;=\; \frac{1 - \mathrm{e}^{i(N+1)t}}{1 - e^{it}}$$

To separate real and imaginary part extend with the complex conjugate number!

$$= \frac{1 - \cos\big[(N+1)t\big] - i \sin\big[(N+1)t\big]}{1 - \cos t - i \sin t} \;\Big/\; \times \frac{1 - \cos t + i \sin t}{1 - \cos t + i \sin t}$$

$$= \frac{1}{2 - 2\cos t}\Big[\Big(1 - \cos\big[(N+1)t\big]\Big)(1 - \cos t) + \sin\big[(N+1)t\big] \sin t\Big]$$

$$+ \frac{i}{2 - 2\cos t}\Big[\Big(1 - \cos\big[(N+1)t\big]\Big) \sin t - \sin\big[(N+1)t\big](1 - \cos t)\Big]$$

Calculating the real part by using

$$\sin t \;=\; 2 \sin\left({}^{t}\!/_{2}\right) \cos\left({}^{t}\!/_{2}\right)$$

and

$$1 - \cos t \;=\; 2 \sin^2\left({}^{t}\!/_{2}\right)$$

results in:

$$1 + \sum_{n=1}^{N} \cos nt \;=\; \frac{1}{2} - \frac{\cos\big[(N+1)t\big] \sin\left({}^{t}\!/_{2}\right)}{2 \sin\left({}^{t}\!/_{2}\right)} + \frac{\sin\big[(N+1)t\big] \cos\left({}^{t}\!/_{2}\right)}{2 \sin\left({}^{t}\!/_{2}\right)}$$

$$\sum_{n=1}^{N} \cos nt \;=\; -\frac{1}{2} + \frac{\sin\left(N + {}^1\!/_2\right)t}{2\sin\left({}^t\!/_2\right)}$$

Integration from π to x results in

$$S_N(x) \;=\; \frac{\pi - x}{2} + J_N(x)$$

Calculation of $J_N(x)$ by using partial integration

$$u \;=\; -\frac{\cos\left(N + {}^1\!/_2\right)t}{N + {}^t\!/_2} \qquad , \qquad v \;=\; \frac{1}{2\sin\left({}^t\!/_2\right)}$$

$$J_N(x) \;=\; \int_{\pi}^{x} u'(t)\, v(t)\,\mathrm{d}t \;=\; u(t)\, v(t)\Big|_{\pi}^{x} - \int_{\pi}^{x} u(t)\, v'(t)\,\mathrm{d}t$$

$$= \left[\frac{\cos\left(N + {}^1\!/_2\right)t}{(2N+1)\sin\left({}^t\!/_2\right)}\right]_{\pi}^{x} - \int_{\pi}^{x} \frac{\cos\left[\left(N + {}^1\!/_2\right)t\right]\cos\left({}^t\!/_2\right)}{(2N+1)2\sin^2\left({}^t\!/_2\right)}\,\mathrm{d}t$$

From this follows the estimation:

$$|J_N(x)| \;\leq\; \frac{1}{(2N+1)\sin\left({}^x\!/_2\right)} + \frac{\pi}{(2N+1)2\sin^2\left({}^x\!/_2\right)}$$

It follows

$$\lim_{N\to\infty} J_N(x) \;=\; 0 \qquad \text{for} \quad 0 < x < 2\pi$$

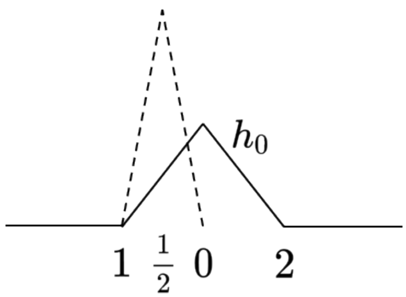
□

With respect to the handling of functions which are being represented by series one considers two general notions of convergence:

1. Uniform convergence of a sequence of (real or respectively complex-valued) functions f.
2. The absolute convergence of an infinite series.

2.1 Uniform Convergence of Function Sequences

Counterexample:



$$h_0(t) = \max(1 - |t-1|, 0)$$

$$h_{n+1}(t) = 2\,h_n(2t)$$

for every real $t \in \mathbb{R}$, $h_n(t)$ is a zero sequence.

Definition of uniform convergence:

Let $(f_n(x))_{n\geq 1}$ be a sequence of functions (with real or complex values) defined for $x \in X$. It is called uniformly convergent to the function $f(x)$ if for each $\epsilon > 0$ an index N_ϵ exists with

$$|f_n(x) - f(x)| < \epsilon \qquad \forall\, n \geq N_\epsilon \quad \forall\, x \in X$$

$(h_n(x))_{n\geq 0}$ converges "pointwise" to the zero function.

$$h_n(2^{-n}) = 2^n$$

The convergence is non-uniform!

Every uniformly convergent sequence of continuous functions $f_n : X \to \mathbb{C}$ has a continuous limiting function f.

Proof:

Let $a \in X$ and let $\eta > 0$ be a tolerance limit. Then $V = \{w \in \mathbb{C}\,;\, |w - f(a)| < \eta\}$ is the η-neighborhood of $fa)$. Let $\epsilon = {}^{\eta}/_{3}$. The triangle inequality gives

$$\begin{aligned} |f(x) - f(a)| &= |f(x) - f_n(x) + f_n(x) - f_n(a) + f_n(a) - f(a)| \\ &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| \\ &\leq \epsilon + |f_n(x) - f_n(a)| + \epsilon \end{aligned}$$

for all $x \in X$ and all $n \geq N_\epsilon$ (according to the premise).

Now $n = n_0 \geq N_\epsilon$ is being fixed and the continuity of f_n is used. In addition, a neighborhood U of a exists such that $|f_{n_0}(x) - f_{n_0}(a)| < \epsilon$ if $x\, inU$. Therefore together for all $x \in U$:

$$|f(x) - f(a)| \;\leq\; \epsilon + \epsilon + \epsilon \;=\; \eta$$

□

Remark:

(1) The sequence of functions $\quad S_n(x) \;=\; \sum_{n=1}^{N} \frac{\sin nx}{n}$

consists of continuous functions and has a discontinuous limiting function $S(x)$. Therefore the convergence is not uniform on all of $\mathbb{R}$. However, the estimate for the limiting function $J_n(x) = S_N(x) - S(x)\,,\ 0 < x < 2\pi$ shows that the restrictions on an interval $[\delta, 2\pi - \delta]$ $(\delta \in\,]0, \pi[\,)$ are always uniform because with the relevant denominator $\sin^2\left({}^x\!/_2\right)$ there

$$\sin^2\left({}^x\!/_2\right) \;\geq\; \sin^2\left({}^\delta\!/_2\right) \;>\; 0$$

For the integral over the limiting function f of a uniformly convergent sequence of continuous functions f_n , the following formula applies

$$\boxed{\int_a^b f(t)\,\mathrm{d}t \;=\; \lim_{n\to\infty} \int_a^b f_n(t)\,\mathrm{d}t}$$

Proof: Another form of the assertion is

$$\lim_{n\to\infty} \int_a^b (f(t) - f_n(t))\,\mathrm{d}t \;=\; 0$$

Now after the standard integral estimation

$$\left| \int_a^b (f(t) - f_n(t))\,\mathrm{d}t \right| \;\leq\; |b - a|\,\epsilon_n$$

where $\quad \epsilon_n \;=\; \sup|f(x) - f_n(x)| \quad a \;\leq\; x \;\leq\; b$

According to the assumption of uniform convergence, the ϵ_n form a null sequence, so the assertion holds.

□

Orthogonality relations

For integer k applies

$$\int_{-\pi}^{\pi} \mathrm{e}^{kit}\,\mathrm{d}t \;=\; \begin{cases} 2\pi & \text{if} \quad k = 0 \\ 0 & \text{if} \quad k \neq 0 \end{cases}$$

Proof:

In the case $k = 0$, the integrand $\mathrm{e}^{0it} = 1$ is the constant function 1. Therefore the assertion is clear.

In the case $k \neq 0$, the assertion follows from the fact that the integrand $\mathrm{e}^{ikt} = \cos kt + i \sin kt$ has the 2π-periodic primitive function:

$$\frac{\mathrm{e}^{ikt}}{ik} \;=\; \frac{\sin kt}{k} - i\,\frac{\cos kt}{k}$$

□

Given be the Fourier series $\displaystyle\sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{int}$

and let it be presupposed that it converges uniformly to the function $f(t)$. Then to the Fourier coefficients applies

$$\boxed{c_k \;=\; \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\mathrm{e}^{-ikt}\,\mathrm{d}t}$$

Remark:

(2) For arbitrary piecewise continuous 2π-periodic functions $f(t)$, a Fourier series belonging to f is declared by the above formula, the “Fourier series” of f.

Proof:

Uniform convergence refers to the sequence $\displaystyle f_N(t) = \sum_{n=-N}^{N} c_n \mathrm{e}^{int}$

Hence $f_N(t)\mathrm{e}^{ikt}$ is uniformly convergent because

$$\left| f_N(t)\mathrm{e}^{ikt} - f(t)\mathrm{e}^{ikt} \right| \;=\; |f_N(t) - f(t)|\,\underbrace{\left|\mathrm{e}^{ikt}\right|}_{=1}$$

Now one can integrate term by term:

$$\begin{aligned}
\int_{-\pi}^{\pi} f(t)\mathrm{e}^{ikt}\,\mathrm{d}t &= \lim_{N\to\infty}\int_{-\pi}^{\pi} f_N(t)\mathrm{e}^{ikt}\,\mathrm{d}t \\
&= \lim_{N\to\infty}\int_{-\pi}^{\pi}\sum_{n=-N}^{N} c_n \underbrace{e^{int}\mathrm{e}^{-ikt}}_{e^{i(n-k)t}}\,\mathrm{d}t \\
&= \lim_{N\to\infty}\sum_{n=-N}^{N} c_n \underbrace{\int_{-\pi}^{\pi} e^{i(n-k)t}\,\mathrm{d}t}_{=0 \text{ except for } n=k} \qquad \text{orthogonality relation} \\
&= c_k\cdot 2\pi
\end{aligned}$$

$\square$

<u>Remark:</u>

(3) The integration over complex-valued continuous functions $g(t)$ with the decomposition of

$$g(t) \;=\; u(t) + iv(t)$$

into real and imaginary parts was used. This means

$$\int_a^b g(t)\,\mathrm{d}t \;=\; \int_a^b u(t)\,\mathrm{d}t + i\int_a^b v(t)\,\mathrm{d}t$$

Obviously then for sums $g = g_1 + g_2$ $(g_1, g_2$ continuous$)$

$$\int_a^b (g_1(t)+g_2(t))\,\mathrm{d}t \;=\; \int_a^b g_1(t)\,\mathrm{d}t + \int_a^b g_2(t)\,\mathrm{d}t$$

If $c = c_1 + ic_2$ is a decomposed complex scalar, then

$$cg(t) \;=\; c_1u(t) - c_2v(t) + i(c_1v(t)+c_2u(t))$$

Hence:

$$\begin{aligned}
\int_a^b cg(t)\,\mathrm{d}t &= \int_a^b (c_1u(t)-c_2v(t))\,\mathrm{d}t + i\int_a^b (c_1v(t)+c_2u(t))\,\mathrm{d}t \\
&= c_1\int_a^b g(t)\,\mathrm{d}t + c_2 i\int_a^b g(t)\,\mathrm{d}t \;=\; c\int_a^b g(t)\,\mathrm{d}t
\end{aligned}$$

To briefly summarize the two calculation rules: The integral over complex-valued continuous integrands is $\mathbb{R}$ linear.

Treatment of the basic example: $$S(x) \;=\; \begin{cases} 0 & \text{if} \quad x = 0 \\ \dfrac{\pi - x}{2} & \text{if} \quad 0 < x < 2\pi \\ & 2\pi\text{-periodic otherwise.} \end{cases}$$

The associated Fourier series has the coefficients

$$c_k(S) \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} S(t)\mathrm{e}^{-ikt}\,\mathrm{d}t \;=\; \frac{1}{2\pi}\int_{0}^{2\pi} \frac{\pi - t}{2}\mathrm{e}^{-ikt}\,\mathrm{d}t$$

($c_0 = 0$ because S is odd)

if $k \neq 0$:

$$c_k(S) \;=\; -\frac{1}{2\pi}\int_{0}^{2\pi} \underset{\substack{\downarrow \\ u}}{\frac{t}{2}}\, \underset{\substack{\downarrow \\ v'}}{\underbrace{\mathrm{e}^{-ikt}}}\,\mathrm{d}t$$

$$= \; -\frac{1}{2\pi}\left[\frac{t}{2}\,\frac{\mathrm{e}^{-ikt}}{-ik}\right]_0^{2\pi} - \underbrace{\frac{1}{2\pi k i}\int_{0}^{2\pi}\frac{1}{2}\mathrm{e}^{-ikt}\,\mathrm{d}t}_{\substack{=0 \\ \text{(orthogonality relation)}}} \;=\; \frac{1}{2\pi i}$$

therefore $\quad c_k \;=\; \dfrac{1}{2\pi i} \quad$ if $k \neq 0$

Summarize for $k > 0$ (also see beginning of chapter)

$$c_k + c_{-k} \;=\; 0 \quad ; \quad i(c_k - c_{-k}) \;=\; \frac{1}{k}$$

<u>Remark:</u>

(4) For continuous functions f with period T

$$\int_{a}^{a+T} f(t)\,\mathrm{d}t \;=\; \int_{b}^{b+T} f(t)\,\mathrm{d}t$$

Because from the decomposition

$$\int_a^{a+T} = \int_a^b + \int_b^{b+T} + \int_{b+T}^{a+T}$$

one obtains the result through substitution in the third integral by exploiting $f(s+T) = f(s)$.

2.1.1 Gibbs' Phenomenon

$S_N(x) = \sum_{n=1}^{N} \frac{\sin nx}{n}$ have the derivative

$$S_N'(x) = \sum_{n=1}^{N} \cos nx = \frac{\sin\left[\left(N+{}^1\!/_2\right)x\right] - \sin\left({}^x\!/_2\right)}{2\sin\left({}^x\!/_2\right)}$$

$$= \frac{2\cos\left[(N+1)\,{}^x\!/_2\right]\sin\left({}^{Nx}\!/_2\right)}{2\sin\left({}^x\!/_2\right)}$$

The smallest root $a_n = \frac{\pi}{N+1} > 0$ belongs to a local maximum of S_N; there the corresponding value is

$$S_N(a_n) = \sum_{n=1}^{N} \frac{1}{n} \sin\left(\frac{n\pi}{N+1}\right)$$

This is a Riemann sum to the integral

$$\int_0^\pi \frac{\sin x}{x}\,\mathrm{d}x$$

for the subdivision of $[0,\pi]$ in $N+1$ subintervals with subpoints $x_i = \frac{i\pi}{N+1}$ $(0 \le i \le N+1)$ and the intermediate points $\xi_i = x_i$. With that

$$\lim_{N\to\infty} S_N(a_n) = \int_0^\pi \frac{\sin t}{t}\,\mathrm{d}t$$

Its value is 18% larger than $\lim_{x\searrow 0} S(x) = \frac{\pi}{2}$.

2.1.2 Differentiation of the Limiting Function

Let the function sequence $(f_n)_{n\geq 0}$ on the interval I converge pointwise to f. All f_n are assumed to be continuously differentiable. Let the sequence $(f_n')_{n\geq 0}$ of derivatives converge uniformly on I to the limiting function g.

Then: f is differentiable on I and the derivative is

$$\boxed{f'(x) = \lim_{n\to\infty} f_n'(x)}$$

Proof:

The limiting function is continuous as a uniform limit of a continuous sequence of functions. Let $x, x+h \in I$ with $h \neq 0$

$$\frac{f(x+h)-f(x)}{h} = \frac{1}{h}\lim_{n\to\infty}\big(f_n(x+h)-f_n(x)\big)$$

$$= \frac{1}{h}\lim_{n\to\infty}\int_x^{x+h} f_n'(t)\,\mathrm{d}t$$

(and since f_n' uniformly convergent to g)

$$= \frac{1}{h}\int_x^{x+h} g(t)\,\mathrm{d}t$$

As in the fundamental theorem of calculus

$$\lim_{\substack{h\to 0\\ h\neq 0}}\frac{f(x+h)-f(x)}{h} = g(x)$$

□

Example:

From $\left|\dfrac{\cos nx}{n^2}\right| \leq \dfrac{1}{n^2}$ and because of $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^2} < \infty$ follows the uniform convergence of the series

$$F(x) = \sum_{n=1}^{\infty}\frac{\cos nx}{n^2}$$

For $\delta \leq x \leq 2\pi - \delta$ with $\delta \in \,]0, \pi[$ the sequence of derivatives of the partial sums is

$$-\sum_{n=1}^{\infty} \frac{\sin nx}{n} \qquad \text{uniformly convergent to} \qquad \frac{x-\pi}{2}$$

According to the last theorem, $F(x)$ is differentiable in each of these intervals, i.e. also in $0 < x < 2\pi$, with the derivative

$$F'(x) \;=\; \frac{x-\pi}{2}$$

It follows

$$F(x) \;=\; \left(\frac{x-\pi}{2}\right)^2 + c\ , \qquad c \ \text{ appropriately constant.}$$

Because both sides are continuous in the interval $0 \leq x \leq 2\pi$, it follows

$$F(0) \;=\; \frac{\pi^2}{4} + c$$

Now uniform convergence of the F defining series is utilized for term-by-term integration

$$\int_0^{2\pi} F(x)\,\mathrm{d}x \;=\; \sum_{n=1}^{\infty} \underbrace{\int_0^{2\pi} \frac{\cos nx}{n}\,\mathrm{d}x}_{\substack{=\ 0 \\ \text{(orthogonality relation)}}} \;=\; 0$$

$$=\; \int_0^{2\pi} \left(\frac{x-\pi}{2} + c\right) \mathrm{d}x \;=\; \left.\frac{(x-\pi)^3}{12}\right|_0^{2\pi} + 2\pi c$$

With that follows: $\qquad 0 \;=\; \dfrac{\pi^3}{6} + 2\pi c \qquad \Rightarrow \qquad c \;=\; -\dfrac{\pi^2}{12}$

$$\boxed{F(x) \;=\; \frac{(x-\pi)^2}{4} - \frac{\pi^2}{12}} \qquad\qquad \boxed{F(0) \;=\; \sum_{n=1}^{\infty} \frac{1}{n^2} \;=\; \frac{\pi^2}{6}}$$

2.2 Absolutely Convergent Series

Definition

A series $\sum\limits_{n=0}^{\infty} a_n$ with real or complex terms a_n is called "absolutely convergent", if the series of absolute values, i.e. $\sum\limits_{n=0}^{\infty} |a_n|$, converges.

Remark:

The latter is the case if and only if the sequence $\left(\sum\limits_{n=0}^{N} |a_n|\right)_{N \geq 0}$ is bounded: monotony criterion. Every absolutely convergent series is convergent.

Proof:

The sequence $S_N = \sum\limits_{n=0}^{N} a_n$ satisfies the Cauchy criterion! Because for $N, p \in \mathbb{R}$

$$|S_{N+p} - S_N| = \left| \sum_{N+1}^{N+p} a_n \right|$$

and with the triangle inequality that is $\leq \sum\limits_{N+1}^{N+p} |a_n|$

$\square$

Comparison criterion

Let $b_n \geq 0$ and $\sum\limits_{n=0}^{\infty} b_n < \infty$. Then for every series $\sum\limits_{n=0}^{\infty} a_n$ absolute convergence applies if $|a_n| \leq b_n \;\forall\, n$. Proof just like before!

The geometric series $\sum\limits_{n=0}^{\infty} M q^n$ with $0 < q < 1$ is for each $M > 0$ convergent.

Quotient criterion

If for the terms a_n with a $q \in \,]0, 1[$ the estimate

$$|a_{n+1}| \leq |a_n|\, q$$

applies for all $n \geq n_0$, then $\sum\limits_{n=0}^{\infty} a_n$ is absolutely convergent, i.e. convergent.

Let $\sum\limits_{n=0}^{\infty} a_n (z - z_0)^n$ be a power series with radius of convergence $\varrho > 0$. For each r with $0 < r < \varrho$ the power series is absolutely uniformly convergent on the disk $|z - z_0| \leq r$.

Proof:

According to the definition of ϱ, there exists a $R :\ r < R < \varrho$ such that the sequence $(a_n R^n)_{n \geq 0}$ is bounded.

Like for example $|a_n| R^n \leq M$. Notation $q := \dfrac{r}{R} \in \,]0, 1[$

If $|z - z_0| \leq r$, then $|a_n (z - z_0)^n| \leq |a_n| r^n \leq |a_n| R^n q^n \leq M\, q^n$

Therefore, $\sum\limits_{n=0}^{\infty} a_n (z - z_0)^n$ is on the one hand absolutely convergent, on the other hand the remainder of this series can be estimated by the series remainder of the geometric series $\sum M q^n$. The latter is independent of the special choice of z. Hence uniform convergence on the disk $|z - z_0| \leq r$ follows.

$\square$

Rearrangement theorem

Let $\sum\limits_{n=0}^{\infty} a_n$ be an absolutely convergent series with real or complex terms. Let $\pi : \mathbb{N}_0 \to \mathbb{N}_0$ be a permutation of the index set $s := \sum\limits_{n=0}^{\infty} a_n$. Then the rearranged series $\sum\limits_{n=0}^{\infty} a_{\pi(n)}$ is also absolutely convergent with limit value s.

Example of a rearrangement:

$$0,\ \underbrace{1,\ 2},\ 4,\ \underbrace{3,\ 6},\ 8,\ \underbrace{5,\ 10},\ 12,\ \underbrace{7,\ 14},\ 16,\ \ldots$$

2.3 Fourier Series Theory

Orthogonality relations: In the "space" V of all real- and complex-valued 2π-periodic and piecewise continuous functions, geometric ideas are to be developed. Just as one recognizes the orthogonality of two vectors $\vec{x}, \vec{y}$ in the plane $\mathbb{R}^2$ from the value of the Euclidean scalar product $\vec{x}\,\vec{y} = 0$, one can also recognize certain minimality properties in V by a scalar product

$$\langle f, g \rangle \;:=\; \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{f(t)} g(t)\,\mathrm{d}t \qquad f, g \in V$$

Basic arithmetic laws

$$\begin{aligned}
\langle f_1 + f_2, g \rangle &= \langle f_1, g \rangle + \langle f_2, g \rangle \\
\langle f, g_1 + g_2 \rangle &= \langle f, g_1 \rangle + \langle f, g_2 \rangle \\
\langle \alpha f, g \rangle &= \overline{\alpha} \langle f, g \rangle \\
\langle f, \beta g \rangle &= \beta \langle f, g \rangle \\
\langle g, f \rangle &= \overline{\langle f, g \rangle} \\
(P) : \langle f, f \rangle &\geq 0 \qquad \text{positivity of the scalar product}
\end{aligned}$$

for all $f,\ f_1,\ f_2,\ g,\ g_1,\ g_2 \subset V$ and all $\alpha,\ \beta \in \mathbb{C}$.

A distance concept in V

(Mean square deviation or also called 2-norm)

$$\|f\| \;:=\; \sqrt{\langle f, f \rangle} \;=\; \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2\,\mathrm{d}x \right)^{1/2}$$

The Cauchy-Schwarz inequality

Always $\boxed{|\langle f, g \rangle| \;\leq\; \|f\|\,\|g\|}$

Proof from (P) with $f+\lambda g$ instead of f

$$0 \;\leq\; \langle f+\lambda g, f+\lambda g\rangle \;=\; \langle f,f\rangle \;+\; \overline{\lambda}\underbrace{\langle g,f\rangle}_{\overline{\langle f,g\rangle}} + \lambda\langle f,g\rangle \;+\; |\lambda|^2\langle g,g\rangle$$

In case $\langle g,g\rangle \neq 0$ choose $\lambda = -\dfrac{\overline{\langle f,g\rangle}}{\langle g,g\rangle}$

$$0 \;\leq\; \|f\|^2 - \frac{|\langle f,g\rangle|^2}{\|g\|^2} \qquad ; \qquad |\langle f,g\rangle|^2 \;\leq\; \|f\|^2\|g\|^2$$

If $\langle g,g\rangle = 0$ but $\langle f,f\rangle \neq 0$ then swap the role of f and g.

Final case: $\langle f,f\rangle \;=\; 0 \;=\; \langle g,g\rangle$

set $\lambda \;=\; -\overline{\langle f,g\rangle}$, then $2|\langle f,g\rangle|^2 \;\leq\; 0$

$\square$

<u>The triangle inequality</u>

$$\|f+g\| \;\leq\; \|f\| + \|g\|$$

Proof with Cauchy-Schwarz inequality (CS-I)

$$\begin{aligned}
\|f+g\|^2 &= \langle f+g, f+g\rangle \\
&= \langle f,f\rangle + \langle g,g\rangle + \langle f,g\rangle + \overline{\langle f,g\rangle} \\
&\leq \langle f,f\rangle + \langle g,g\rangle + 2|\langle f,g\rangle| \\
&\underset{\text{CS-I}}{\overset{\downarrow}{\leq}} \|f\|^2 + \|g\|^2 + 2\|f\|\|g\| \;\leq\; (\|f\|+\|g\|)^2
\end{aligned}$$

$\square$

<u>The basis functions e_k in V</u>

are declared through: $\mathrm{e}_k(x) \;=\; \mathrm{e}^{ikx}, \quad k\in\mathbb{Z}$

$$\langle \mathrm{e}_k, \mathrm{e}_l\rangle \;=\; \frac{1}{2\pi}\int_{-\pi}^{+\pi} \mathrm{e}^{-ikx}\mathrm{e}^{ilx}\,\mathrm{d}x \;=\; \frac{1}{2\pi}\int_{-\pi}^{+\pi} \mathrm{e}^{i(l-k)x}\,\mathrm{d}x$$

$$\delta_{k,l} = \begin{cases} 1 & \text{if} \quad k = l \\ 0 & \text{if} \quad k \neq l \end{cases}$$

Description of the Fourier coefficients for $f \in V$

$$c_k(f) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \mathrm{e}^{-ikx} \, \mathrm{d}x = \langle \mathrm{e}_k, f \rangle$$

The trigonometric polynomials associated with f are

$$f_N = \sum_{k=-N}^{N} c_k(f) \mathrm{e}_k = \sum_{k=-N}^{N} \langle \mathrm{e}_k, f \rangle \mathrm{e}_k$$

<u>Lemma 1:</u> (compare with the Pythagorean theorem)

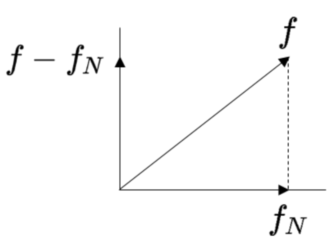
$$\|f\|^2 - \|f_N\|^2 = \|f - f_N\|^2$$

Proof:

1) $c_k(f) = \langle \mathrm{e}_k, f \rangle = \langle \mathrm{e}_k, f_N \rangle \qquad -N \leq k \leq N$

2) $\langle f, f_N \rangle = \left\langle f, \sum_{k=-N}^{N} c_k \mathrm{e}_k \right\rangle = \sum_{k=-N}^{N} c_k \underbrace{\langle f, \mathrm{e}_k \rangle}_{\overline{c_k}} = \sum_{k=-N}^{N} |c_k|^2$

3) $\langle f_N, f_N \rangle = \left\langle \sum_{k=-N}^{N} c_k \mathrm{e}_k , f_N \right\rangle = \sum_{k=-N}^{N} \overline{c_k} \underbrace{\langle \mathrm{e}_k, f_N \rangle}_{c_k} = \sum_{k=-N}^{N} |c_k|^2$



Lemma 1: $\quad \|f - f_N\|^2 = \|f\|^2 - \|f_N\|^2$

$$\langle f, f_N \rangle = \langle f_N, f_N \rangle = \sum_{k=-N}^{N} |c_k|^2$$

$$\begin{aligned} \|f - f_N\|^2 &= \langle f - f_N, f - f_N \rangle \\ &= \langle f, f \rangle - \langle f_N, f \rangle + \langle f, f_N \rangle + \langle f_N, f_N \rangle \\ &= \|f\|^2 - \|f_N\|^2 \end{aligned}$$

Bessel's inequality

$$\sum_{k=-\infty}^{\infty} |c_k(f)|^2 \;\leq\; \|f\|^2 \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} \big|f(x)^2\,\mathrm{d}x\big|$$

Because there the left side is according to lemma $1 \geq 0$, therefore for all N

$$\sum_{k=-N}^{N} |c_k|^2 \;\leq\; \|f\|^2$$

An extremality property

Among all "linear combinations"

$$g \;=\; \sum_{k=-N}^{N} \gamma_k \mathrm{e}_k$$

of the basis function e_k only

$$f_N \sum_{k=-N}^{N} c_k(f)\mathrm{e}_k$$

has minimum distance from f in terms of our norm.

Proof:

$$\langle f_N, f - f_N\rangle \;=\; 0 \quad , \quad \langle \mathrm{e}_k, f - f_N\rangle \;=\; 0 \quad , \quad -N \;\leq\; k \;\leq\; N$$

The compatibility of $\langle\,,\rangle$ with addition (in the first argument) yields $\langle g, f - f_N\rangle \;=\; 0$

Consider:

$$\begin{aligned}
\|f-g\|^2 \;&=\; \langle f-g\,,\, f-g\rangle \\
&=\; \big\langle (f-f_N)+(f_N-g)\,,\,(f-f_N)+(f_N-g)\big\rangle \\
&=\; \langle f-f_N\,,\, f-f_N\rangle + \langle f_N-g\,,\, f_N-g\rangle \\
&=\; \|f-f_N\|^2 \;+\; \|f_N-g\|^2
\end{aligned}$$

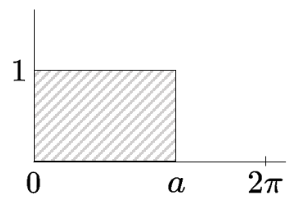
$\|f_N - g\|^2 \;=\; 0$ only if $f_N \;=\; g$

□

Lemma 2: If $f \in V$ is such that the restriction to $[0, 2\pi]$ is a step function, then the following applies to the approximating trigonometric Polynomials: $\lim\limits_{N\to\infty} \|f - f_N\| \;=\; 0$.

Proof:

1)



By assumption, points

$$a_0 = 0 < a_1 < a_2 < \ldots < a_i = 2\pi$$

exist such that $f \,/\,]a_{i-1}, a_i[$ is constant.

On the other hand, the mapping $f \longmapsto (c_k(f))_{k \in \mathbb{Z}}$ is linear in f (compatible with addition and scalar multiplication). It is therefore sufficient to only consider the special function

$$f(x) = \begin{cases} 1 \quad \text{for} \quad 0 \le x \le a \\ 0 \quad \text{for} \quad a \le x < 2\pi \\ \;\; 2\pi\text{-periodic otherwise.} \end{cases}$$

2) For the special f the Fourier coefficients become

$$c_0 = \frac{a}{2\pi} \qquad \text{and for } k \neq 0 \qquad c_k = i\frac{\mathrm{e}^{-ika} - 1}{2\pi k}$$

To be shown (according to lemma 1): The "Parseval equation"

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \|f\|^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2 \, \mathrm{d}x$$

holds for $k \neq 0$: $\quad |c_k|^2 = \dfrac{|\cos ka - i \sin ka - 1|}{2\pi^2 k^2} = \dfrac{2 - 2\cos ka}{2\pi^2 k^2}$

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{a^2}{4\pi^2} + \frac{1}{\pi^2}\sum_{k=1}^{\infty}\frac{1}{k^2} - \frac{1}{\pi^2}\sum_{k=1}^{\infty}\frac{\cos ka}{k^2}$$

with the example of the integration of uniformly convergent sequences of functions

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{a^2}{4\pi^2} + \frac{1}{6} - \frac{1}{\pi^2}\left(\frac{(\pi - a)^2}{4} - \frac{\pi^2}{12}\right)$$

$$= \frac{a}{2\pi} = \frac{1}{2\pi}\int_0^{2\pi} |f(x)|^2 \, \mathrm{d}x$$

2.3.1 The Completeness Theorem

Let $f \in V$, i.e. 2π-periodic and with piecewise continuous restrictions $f\,/\,[0,2\pi]$. Furthermore, let each

$$f_N \;=\; \sum_{n=-N}^{N} c_n \mathrm{e}_n$$

be the f approximating trigonometric polynomial. Then the following applies to the norm of the mean square deviation

$$\lim_{N\to\infty} \|f - f_N\| \;=\; 0$$

In particular, the Parseval equation is always satisfied

$$\|f\|^2 \;=\; \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,\mathrm{d}t \;=\; \sum_{k=-\infty}^{\infty} |c_k(f)|^2$$

Proof: Without restriction f real and $-1 \le f(x) \le 1$. Reason: $f \longmapsto (c_k)\ k \in \mathbb{Z}$ is $\mathbb{R}$ linear.

1) From integration theory one finds for arbitrary $\eta > 0$ step functions φ, ψ on $[0,2\pi]$ with $-1 \le \ varphi(x) \le f(x) \le \psi(x) \le 1$ and

$$\int_0^{2\pi} (\psi(x) - \varphi(x))\,\mathrm{d}x \;\le\; \eta$$

With $\quad g := f - \varphi \quad (0 \le g(x) \le 2) \quad$ holds: $\quad g^2(x) \;\le\; 2\,g(x)$

$$\|g\|^2 \;=\; \frac{1}{2\pi}\int_0^{2\pi} g^2(x)\,\mathrm{d}x \;\le\; \frac{1}{\pi}\int_0^{2\pi} g(x)\,\mathrm{d}x \;\le\; \frac{\eta}{\pi}$$

2) For $\quad f = g + \varphi \quad$ holds $\quad f_N = g_N + \varphi_N$

Lemma 1: $\quad \|g - g_N\|^2 \;\le\; \|g\|^2$

Now $\quad \eta \le \dfrac{\epsilon^2}{4}\ ; \quad \|g - g_N\|^2 \;\le\; \dfrac{\epsilon^2}{4\pi}\ ; \quad \|g - g_N\| \;\le\; \dfrac{\epsilon}{2}$

With lemma 2 holds: $\quad \|\varphi - \varphi_N\| \longmapsto 0 \quad$ for $\quad N \to \infty$.

For $N \ge N_\epsilon$ applies: $\quad \|\varphi - \varphi_N\| \;\le\; \dfrac{\epsilon}{2}$.

Along with the triangle inequality:

$$\begin{aligned}\|f - f_N\| &= \|g - g_N + \varphi - \varphi_N\| \\ &\leq \|g - g_N\| + \|\varphi - \varphi_N\| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\end{aligned}$$

if $N \geq N_\epsilon$

3) According to lemma 1

$$\|f\|^2 = \|f - f_N\|^2 + \|f_N\|^2 \qquad \text{with} \qquad \|f\|^2 = \sum_{k=-N}^{N} |c_k|^2$$

□

Representation theorem

For every continuous function $f \in V$, which is also continuously differentiable, the Fourier series converges to f.

Kernel of proof:

1) Bessel's inequality for the derivative function.

2) Calculation of the Fourier coefficients $c_k(f)$ by means of partial integration from those of the derivative $\gamma_k = c_k(f')$

Result: $c_k = \dfrac{\gamma_k}{ik} \qquad (k \neq 0)$

3) An estamation

$$|a|\,|b| \leq \frac{|a|^2 + |b|^2}{2}$$

gives

$$|c_k|^2 \leq \frac{1}{2}\left(\frac{1}{k^2} + |\gamma_k|^2\right)$$

Absolute and uniform convergence of $\displaystyle\sum_{k=-\infty}^{\infty} c_k f \mathrm{e}^{ikx}$

4) Application of the theorem about uniformly convergent Fourier series!

Example from H. Lebesgue:

$$g(x) = \sum_{n=0}^{\infty} \frac{\sin 2^{n^2} x}{2^n}$$

is a uniformly convergent Fourier series, but the sum $g(x)$ is nowhere differentiable (although g is continuous everywhere).

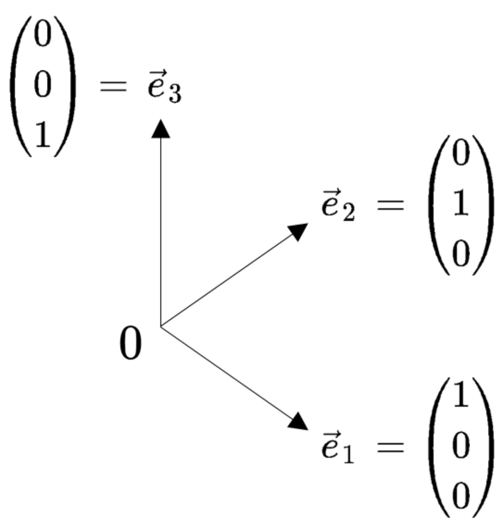
Proof via the difference quotients

$$\begin{aligned}\frac{g(x+h)-g(x)}{h} &= \sum_{n=0}^{\infty} \frac{\sin\left[2^{n^2}(x+h)\right]-\sin\left[2^{n^2}x\right]}{2^n h}\\ &= \sum_{n=0}^{\infty} \frac{\cos\left[2^{n^2}\left(x+{}^{h}\!/_{2}\right)\right]\sin\left[2^{n^2}\left({}^{h}\!/_{2}\right)\right]}{2^n\left({}^{h}\!/_{2}\right)}\end{aligned}$$

Consider: $\quad h_N \;:=\; \pm 2^{-N^2}\cdot\dfrac{\pi}{2}$

Then for $n > N$ all elements of the series are 0. The share of $n = N$, provided the sign is appropriate, far outweighs the sum of the absolute values of all other terms!

3. Three-Dimensional Euclidean Space



After choosing a system of three pairs of perpendicular straight lines through a common point 0, the origin, and after determining an equal scale on each of the three straight lines (axes), each point P of space is assigned a system x_1, x_2, x_3 of three numbers (the coordinates); the first of which denotes the section that the plane through P, which is parallel to the second and third axes, cuts off on the first axis. On the other hand, the second number marks the intersection with the second axis of the plane parallel to the first and third axis through P. Similarly, x_k is the mark on the k-th axis for the point of intersection with the plane parallel to the other axes through P.

One often orients the order of the three axes such that it corresponds with the order of the first three fingers of the right hand!

<u>Example:</u>

Vertices of an axis-parallel cube with edge length 2 and 0 as center.

$$\begin{pmatrix}1\\1\\1\end{pmatrix};\begin{pmatrix}1\\1\\-1\end{pmatrix};\begin{pmatrix}1\\-1\\1\end{pmatrix};\begin{pmatrix}1\\-1\\-1\end{pmatrix};\begin{pmatrix}-1\\1\\1\end{pmatrix};\begin{pmatrix}-1\\1\\-1\end{pmatrix};\begin{pmatrix}-1\\-1\\1\end{pmatrix};\begin{pmatrix}-1\\-1\\-1\end{pmatrix}$$

In addition to defining coordinates in space, the number triplets

$$\vec{a} \; = \; \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

are also used to (determine) describe translations in space. Effect of translation by $\vec{a}$:

$$\vec{x} \; = \; \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto \begin{pmatrix} x_1 + a_1 \\ x_2 + a_2 \\ x_3 + a_3 \end{pmatrix} \; = \; \vec{x} + \vec{a}$$

If one performs the translation by $\vec{b} = (b_1, b_2, b_3)$ after the translation by $\vec{a}$, then the composite mapping is again a translation, namely by the sum vector $\vec{a} + \vec{b}$.

3.1 The Scalar Product

Description of many geometric situations with the Euclidean scalar product (inner product)

$$\vec{x}\,\vec{y} \; := \; x_1 y_1 + x_2 y_2 + x_3 y_3$$

Example:

(1) The distance of the point with the coordinate vector $\vec{x}$ from the origin, also "Euclidean norm" of $\vec{x}$

$$\|\vec{x}\| \; := \; (\vec{x}\,\vec{x})^{1/2} \; = \; \sqrt{x_1^2 + x_2^2 + x_1^3}$$

The Cauchy-Schwarz inequality applies again

$$|\vec{x}\,\vec{y}| \; \leq \; \|\vec{x}\|\,\|\vec{y}\|$$

As a consequence, so does the triangle inequality

$$\|\vec{x} + \vec{y}\| \; \leq \; \|\vec{x}\| + \|\vec{y}\|$$

Like in the Euclidean plane, the geometric meaning of the scalar product is

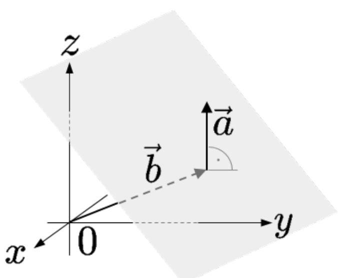
$$\vec{x}\,\vec{y} \; = \; \|\vec{x}\|\,\|\vec{y}\|\cos(\vec{x}, \vec{y})$$

Basic algebraic rules for the scalar product

$$
\begin{aligned}
(\vec{x} + \vec{x})\,\vec{y} &= \vec{x}\,\vec{y} + \vec{x}\,\vec{y} \\
\vec{x}\,(\vec{y} + \vec{y}) &= \vec{x}\,\vec{y} + \vec{x}\,\vec{y} \\
(\lambda\,\vec{x})\,\vec{y} &= \lambda(\vec{x}\,\vec{y}) = \vec{x}\,(\lambda\,\vec{y}) \\
\vec{x}\,\vec{y} &= \vec{y}\,\vec{x} \\
\vec{x}\,\vec{x} &\geq 0\,, \qquad \text{“$=$” only for } \vec{x} = \vec{0} = (0,0,0)
\end{aligned}
$$

The scalar product is defined bilinearly and positively!

Plane equation (implicit with the scalar product)



$$(\vec{x} - \vec{b}) \cdot \vec{a} = 0$$

$\vec{a} \neq 0\,,$ otherwise all $\vec{x}$ are solutions!

$\vec{a}$ “normal vector of the considered plane”.

Another form of the plane equation

$$a_1 x_1 + a_2 x_2 + a_3 x_3 = \beta \qquad \vec{a} \neq 0$$

can be traced back with $\vec{b} = \lambda \vec{a}$

$$\vec{a}\,\vec{b} = \lambda\,\|\vec{a}\|^2 \qquad \text{hence} \qquad \lambda = \frac{\beta}{\|\vec{a}\|^2}$$

On the plane E this vector $\vec{b} = \dfrac{\beta}{\|\vec{a}\|^2}\vec{a}$ is the position vector of the point with minimal distance from $\vec{0}$. Because for $\vec{x} \in E$

$$
\begin{aligned}
\|\vec{x}\|^2 &= \big((\vec{x} - \lambda\,\vec{a}) + \lambda\,\vec{a}\big)\,\big((\vec{x} - \lambda\,\vec{a}) + \lambda\,\vec{a}\big) \\
&= \|\vec{x} - \lambda\vec{a}\|^2 + \lambda^2\|\vec{a}\|^2 \geq \lambda^2\|\vec{a}\|^2
\end{aligned}
$$

Parameter definition of the plane (explicit)

$$\vec{x} \;=\; \vec{b} + \lambda\,\vec{u} + \mu\,\vec{v} \qquad \lambda, \mu \in \mathbb{R}$$

with nonparallel vectors $\vec{u}, \vec{v}$

Example:

(2) Plane through the non-collinear points $\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$

$$\begin{aligned} \vec{x} &= \vec{b}_1 + \lambda\,(\vec{b}_2 - \vec{b}_1) + \mu\,(\vec{b}_3 - \vec{b}_1) \qquad \lambda, \mu \in \mathbb{R} \\ &= \lambda_1\vec{b}_1 + \lambda_2\vec{b}_2 + \lambda_3\vec{b}_3 \qquad (\lambda_1, \lambda_2, \lambda_1 \in \mathbb{R} \;;\quad \lambda_1 + \lambda_2 + \lambda_3 = 1) \end{aligned}$$

(convex combination of $\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$)

Because in Euclidean space straight lines appear as lines of intersection of non-parallel planes, two equations are used for the description of straight lines in scalar product form

$$\begin{aligned} \vec{a}_1\left(\vec{x}_1 - \vec{b}_1\right) &= 0 \qquad\qquad \vec{a}_1, \vec{a}_2 \;\neq\; 0 \\ \vec{a}_2\left(\vec{x}_2 - \vec{b}_2\right) &= 0 \qquad\qquad \vec{a}_1\mathbb{R} \;\neq\; \vec{a}_2\mathbb{R} \end{aligned}$$

Parallelism of these planes is excluded by the forbidden parallelism of their normal vectors. The line of intersection consists of $\vec{x}$ which solve the system of the above two equations.

Also common: parametric form of the straight line (explicit)

$$\vec{x} \;=\; \vec{b} + \lambda\,\vec{u} \qquad \lambda \in \mathbb{R} \qquad \text{where} \qquad \vec{u} \neq 0$$

Now to the transition from the parametric form of a plane to the implicit plane equation.

3.2 The Vector Product

Problem: For vectors $\vec{a}, \vec{b}$ determine a vector that is perpendicular to both.

Solution: Vector product (outer product)

$$\vec{a} \times \vec{b} \;=\; \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}$$

Check: $\vec{a}\,(\vec{a}\times\vec{b}) \;=\; 0 \quad ; \quad \vec{b}\,(\vec{a}\times\vec{b}) \;=\; 0$

$$a_1(a_2b_3 - a_3b_2) \,+\, a_2(a_3b_1 - a_1b_3) \,+\, a_3(a_1b_2 - a_2b_1) \;=\; 0$$

$$b_1(a_2b_3 - a_3b_2) \,+\, b_2(a_3b_1 - a_1b_3) \,+\, b_3(a_1b_2 - a_2b_1) \;=\; 0$$

Examples: $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3 \quad ; \quad \vec{e}_2 \times \vec{e}_3 = \vec{e}_1 \quad ; \quad \vec{e}_3 \times \vec{e}_1 = \vec{e}_2$

Basic algebraic rules of the vector product

$$(\vec{a}+\vec{a}\,')\times\vec{b} \;=\; (\vec{a}\times\vec{b}) \,+\, (\vec{a}\,'\times\vec{b})$$

$$\vec{a}\times(\vec{b}+\vec{b}\,') \;=\; (\vec{a}\times\vec{b}) \,+\, (\vec{a}\times\vec{b}\,')$$

$$(\lambda\,\vec{a})\times\vec{b} \;=\; \lambda\,(\vec{a}\times\vec{b}) \;=\; \vec{a}\times(\lambda\,\vec{b})$$

$$(\vec{a}\times\vec{b}) \;=\; -(\vec{b}\times\vec{a}) \;, \quad \text{skew symmetry}$$

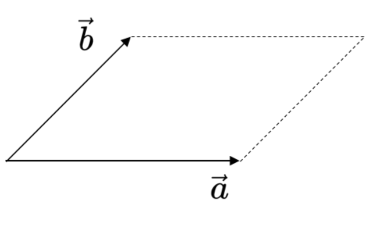
Remark:

The three examples and the rules are enough to calculate all vector products!

$$\left\|\,\vec{a}\times\vec{b}\,\right\|^2 \;=\; (a_2b_3 - a_3b_2)^2 \,+\, (a_3b_1 - a_1b_3)^2 \,+\, (a_1b_2 - a_2b_1)^2$$

$$\left\|\,\vec{a}\times\vec{b}\,\right\|^2 \;=\; \left\|\,\vec{a}\,\right\|^2 \left\|\,\vec{b}\,\right\|^2 \sin^2\!\left(\vec{a},\vec{b}\right)$$

3.2.1 Geometric Interpretation



The norm of $\vec{a}\times\vec{b}$ is at the same time the area of the parallelogram spanned by the vectors $\vec{a}$ and $\vec{b}$. In particular, the vector product vanishes if either one of the two vectors vanishes or if the two vectors come to lie on the same straight line.

The planes $\vec{a}_k(\vec{x}-\vec{b}_k) \;=\; 0 \quad (k = 1,2)$ are not parallel if and only if $\vec{a}_1 \times \vec{a}_2 \neq \vec{0}$ applies to their normal vectors.

If this is the case, and if $\vec{x}_0$ is a point on the line of intersection, then the entire line of intersection is in parametric form

$$\vec{x}_0 \,+\, \mathbb{R}\,(\vec{a}_1 \times \vec{a}_2)$$

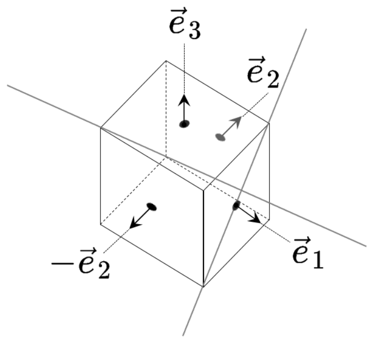
The Großmann identity

For the vector product of three vectors $\vec{a}$, $\vec{b}$, $\vec{c}$ applies

$$\left(\vec{a}\times\vec{b}\right)\vec{c} \;=\; \left(\vec{a}\,\vec{c}\right)\vec{b} - \left(\vec{b}\,\vec{c}\right)\vec{a}$$

Example:

(1)



Consider the straight lines

$$\vec{e}_1 \;=\; \mathbb{R}\left(\vec{e}_2+\vec{e}_3\right)$$

$$-\vec{e}_1 \;=\; \mathbb{R}\left(\vec{e}_2-\vec{e}_3\right)$$

They are skew lines, which means that they do not intersect and are not parallel. Intersections with the plane $x_2 = \text{constant} = c$

$$\begin{pmatrix}1\\0\\0\end{pmatrix} + c\begin{pmatrix}0\\1\\1\end{pmatrix} \;=\; \begin{pmatrix}1\\c\\c\end{pmatrix}$$

or respectively

$$\begin{pmatrix}-1\\0\\0\end{pmatrix} + c\begin{pmatrix}0\\1\\-1\end{pmatrix} \;=\; \begin{pmatrix}1\\c\\-c\end{pmatrix}$$

Connecting line:

$$g_c \;=\; \begin{pmatrix}-1\\c\\-c\end{pmatrix} + \mathbb{R}\begin{pmatrix}2\\0\\2c\end{pmatrix} \;=\; \left\{\begin{pmatrix}-1+2\lambda\\c\\c(-1+2\lambda)\end{pmatrix};\ \lambda\in\mathbb{R}\right\}$$

Set union: $\quad F \;=\; \bigcup_{c\,\in\,\mathbb{R}} g_c$

All points $\vec{x}$ on F satisfy $x_1\,x_2 \;=\; x_3$. Conversely, every solution of this equation lies on a straight line. Every two straight lines of the set are skew lines because for the spanning vectors

$$\begin{pmatrix} 1 \\ 0 \\ c_1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ c_2 \end{pmatrix} \quad \text{the vector product becomes}$$

$$\begin{pmatrix} 1 \\ 0 \\ c_1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ c_1 - c_2 \\ c_0 \end{pmatrix} \neq \vec{0} \quad \text{for different } c_1, c_2$$

Intersections of F with the plane $x_1 = \text{constant} = c$

$$h_c = \left\{ \begin{pmatrix} c \\ \lambda \\ c\lambda \end{pmatrix} = \begin{pmatrix} c \\ 0 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 1 \\ c \end{pmatrix} \right\}$$

This is a second family of straight lines on the surface.

On account of each of the two straight line families on F, F is a “ruled surface”.

Intersections with the planes $x_3 = \text{constant} \neq 0$ yield (in this plane) the points with $x_1\, x_2 = c$, this is a hyperbola.

Intersections with the planes containing the x_3-axis

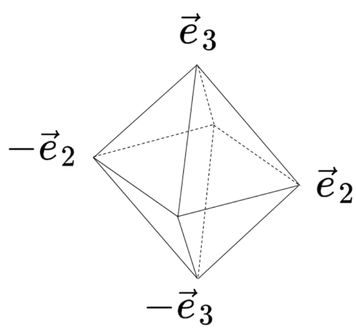
$x_1 = x_2$: Surface points in this plane $x_3 = x_1^2$ form a parabola that opens upwards.

$x_2 = -x_1$: $\Rightarrow \quad x_3 = -x_1^2$; this is a parabola that opens downwards.

In general: $x_2 = \lambda\, x_1$ is a plane containing the x_3-axis and intersects F in the points $\lambda\, x_1^2\ ,= \ x_3$. This is a parabola with the exception $\lambda = 0$.

Example:

(2) Consider 6 points: $\pm\vec{e}_1$, $\pm\vec{e}_2$, $\pm\vec{e}_3$ − the corners of an octahedron.



Determine the inclination angle between the side faces through $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ or respectively $\vec{e}_1$, $\vec{e}_2$, $-\vec{e}_3$ in the planes

$$E_+: \ \vec{e}_1 + \mathbb{R}(\vec{e}_2 - \vec{e}_1) + \mathbb{R}(\vec{e}_3 - \vec{e}_1)$$

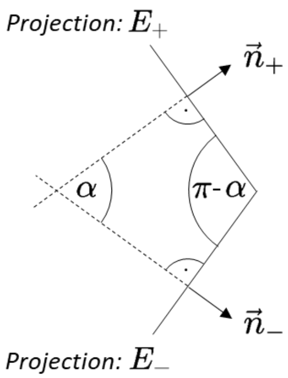
$$E_-: \ \vec{e}_1 + \mathbb{R}(\vec{e}_2 - \vec{e}_1) + \mathbb{R}(-\vec{e}_3 - \vec{e}_1)$$

Normal vectors (outwards with respect to the octahedron)

$$\vec{n}_+ = (\vec{e}_2 - \vec{e}_1) \times (\vec{e}_3 - \vec{e}_1) = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\vec{n}_- = (-\vec{e}_3 - \vec{e}_1) \times (\vec{e}_2 - \vec{e}_1)$$

$$= (\vec{e}_2 - \vec{e}_1) \times (\vec{e}_3 + \vec{e}_1) = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$



$$\cos\alpha = \frac{\vec{n}_+ \, \vec{n}_-}{\|\vec{n}_+\| \, \|\vec{n}_-\|} = \frac{1}{3}$$

$$\alpha \triangleq 70.53°$$

The angle of inclination between both side faces is $180° - \alpha \cong 109.47°$

The tilt angle α occurs between the side faces of the tetrahedron with corners

$$\vec{e}_1, \quad \vec{e}_2, \quad \vec{e}_3, \quad \vec{e}_1 + \vec{e}_2 + \vec{e}_3$$

(Decomposition of a tetrahedron across the edge centers into four tetrahedrons of half the edge length and a central octahedron.)

3.3 The Isometry of Euclidean $\mathbb{R}^3$

The distance between the points $\vec{x}$ and $\vec{y}$ is $\|\vec{x} - \vec{y}\,,\|$, i.e. the norm of the difference vector. Therefore, all distances remain invariant for translations in space by the vector $\vec{a}$. The Euclidean metric is translation invariant. In other words, translation by the vector $\vec{a}$ is always an isometry.

Main question: isometry $f : \mathbb{R}^3 \to \mathbb{R}^3$ with fixed point $\vec{0}$. Where isometry means

$$\|f(\vec{x}\,) - f(\vec{y}\,)\| = \|\,\vec{x} - \vec{y}\,\| \qquad \forall\, \vec{x}, \vec{y} \in \mathbb{R}^3 \qquad \text{(I)}$$

Result: it is a matter of all linear self-mappings of $\mathbb{R}^3$, which leave scalar products invariant, that means

$$f(\vec{x}\,)\, f(\vec{y}\,) = \vec{x}\,\vec{y} \qquad \forall\, \vec{x}, \vec{y} \in \mathbb{R}^3$$

Proof of scalar product fidelity

Consider (I) for $\vec{y} = \vec{0}\,$, because $\vec{0}$ fixed point of f,

$$\begin{aligned} \|f(\vec{x}\,)\| &= \|\vec{x}\| \qquad \forall\, \vec{x} \in \mathbb{R}^3 \\ \|f(\vec{x}\,) - f(\vec{y}\,)\|^2 &= \|f(\vec{x}\,)\|^2 - 2f(\vec{x}\,)f(\vec{y}\,) + \|f(\vec{y}\,)\|^2 \\ &= \|\vec{x}\|^2 - 2f(\vec{x}\,)f(\vec{y}\,) + \|\vec{y}\,\|^2 \qquad \text{(see above)} \end{aligned}$$

Isometry

$$\|\,\vec{x} - \vec{y}\,\|^2 = \|\vec{x}\|^2 - 2\vec{x}\,\vec{y} + \|\vec{y}\,\|^2$$

$\square$

The linearity of f says

$$\begin{aligned} f(\vec{x} + \vec{y}\,) &= f(\vec{x}\,) + f(\vec{y}\,) \\ f(\lambda\,\vec{x}\,) &= \lambda\, f(\vec{x}\,) \qquad \forall\, \vec{x}, \vec{y} \in \mathbb{R}^3 \quad \forall\, \lambda \in \mathbb{R} \end{aligned}$$

This can be calculated using the norm square of the difference between the two sides.

3.3.1 Description of Linear Self-Mappings through Matrices $A = (a_{ik})_{1\,\leq\, i\,,\,k\,\leq\, 3}$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \longmapsto A\vec{x} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

where the element in the k-th row of the vector $A\vec{x}$ is given by

$$a_{k1}x_1 + a_{k2}x_2 + a_{k3}x_3$$

Here, the k-th column of the matrix A contains the image $A\vec{e}m$ of the m-th basis vector ($m = 1, 2, 3$). The matrices for isometries with a fixed point, i.e. the orthogonal transformations, only include matrices A whose column vectors have the length 1 and which are pairwise perpendicular to each other.

$$\begin{aligned} A\vec{e}_m &= \vec{a}_m \quad , \qquad \|\vec{a}_m\| = 1 \quad \text{and} \\ \vec{a}_k\,\vec{a}_m &= 0 \quad , \qquad \text{if} \quad k \neq m \end{aligned}$$

in summary:

$$\vec{a}_i\,\vec{a}_j = \delta_{ij} \qquad \text{for} \quad 1 \leq i, j \leq 3$$

A real 3×3 matrix $A = (\vec{a}_1, \vec{a}_2, \vec{a}_3)$ represents an orthogonal transformation if and only if

$$\vec{a}_k\, \vec{a}_l \;=\; \delta_{kl} \qquad 1 \le k, l \le 3$$

Proof for "only if" follows from $\vec{a}_k$ being the image of $\vec{e}_k$ $(k = 1, 2, 3)$

"if": to be shown that under this condition holds

$$A\vec{x}\, A\vec{y} \;=\; \vec{x}\, \vec{y} \qquad \forall\, \vec{x}, \vec{y} \in \mathbb{R}^3$$

$$A\vec{x}\, A\vec{y} \;=\; \Big(A \sum_{k=1}^{3} x_k \vec{e}_k\Big)\Big(A \sum_{l=1}^{3} y_k \vec{e}_l\Big)$$

(linearity of A)

$$= \Big(\sum_{k=1}^{3} x_k A\vec{e}_k\Big)\Big(\sum_{l=1}^{3} y_k A\vec{e}_l\Big)$$

(bilinearity of scalar products)

$$= \sum_{k,l=1}^{3} x_k y_l \underbrace{(A\vec{e}_k)}_{\vec{a}_k}\underbrace{(A\vec{e}_l)}_{\vec{a}_l} \;=\; \sum_{k,l=1}^{3} x_k y_l \delta_{kl} \;=\; \sum_{k=1}^{3} x_k y_k$$

□

3.4 The Scalar Triple Product

$$D(\vec{a}, \vec{b}, \vec{c}) \;=\; (\vec{a} \times \vec{b})\, \vec{c}$$

Characteristic properties

Detail (1): $D(\vec{a}, \vec{b}, \vec{c})$ is linear in each argument.

Detail (2): If any two arguments in $D(\vec{u}, \vec{v}, \vec{w})$ are equal, then the value of the scalar triple product is zero.

Detail (3): $D(\vec{e}_1, \vec{e}_2, \vec{e}_3) \;=\; 0$

Remark:

From detail (1) and detail (2) follows: If one swaps any two of the three arguments in the scalar triple product, the scalar triple product changes its sign!

Proof:

$$D(\vec{a},\vec{b}+\vec{c},\vec{b}+\vec{c}) \overset{\text{detail (2)}}{\overset{\downarrow}{=}} 0$$

$$\overset{\text{detail (1)}}{\overset{\downarrow}{=}} D(\vec{a},\vec{b},\vec{b}+\vec{c}) + D(\vec{a},\vec{c},\vec{b}+\vec{c})$$

$$\overset{\text{detail (1)}}{\overset{\downarrow}{=}} D(\vec{a},\vec{b},\vec{b}) + D(\vec{a},\vec{b},\vec{c}) + D(\vec{a},\vec{c},\vec{c}) + D(\vec{a},\vec{c},\vec{b})$$

□

3.4.1 Geometric Significance of the Scalar Triple Product

$$D(\vec{a},\vec{b},\vec{c}) = (\vec{a}\times\vec{b})\,\vec{c}$$

follows from the geometric significance of the vector- and of the scalar product: $D(\vec{a},\vec{b},\vec{c})$ is the volume of the parallelepiped spanned by the vectors $\vec{a}$, $\vec{b}$, $\vec{c}$.

$D(\vec{a},\vec{b},\vec{c}) \neq 0$ therefore means that firstly $\vec{a}$ and $\vec{b}$ are not parallel to a straight line and secondly that $\vec{c}$ does not lie in their plane. In other words, $D(\vec{a},\vec{b},\vec{c}) \neq 0$ means that every vector $\vec{x}$ of $\mathbb{R}^3$ is a linear combination of $\vec{a}$, $\vec{b}$ and $\vec{c}$.

The determinant of a 3×3 matrix $A = (\vec{a}_1,\vec{a}_2,\vec{a}_3)$ is defined by the formula

$$\det A = D(\vec{a}_1,\vec{a}_2,\vec{a}_3)$$

In particular, $\det A \neq 0$ means that every vector of $\mathbb{R}^3$ is a linear combination of the columns of A. Hence especially $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$. If $\det A \neq 0$, then $\vec{b}_k \in \mathbb{R}^3$ exist with $A\vec{b}_k = \vec{e}_k$ $(1 \le k \le 3)$. In other words, if $\det A \neq 0$ there is a second matrix B (with columns $\vec{b}_1$, $\vec{b}_2$, $\vec{b}_3$) and $AB = 1_3 = E$, where

$$1_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and where the matrix product is obtained column by column from B.

For 3×3 matrices the determinant product theorem applies:

$$\boxed{\det(AB) = (\det A)(\det B)}$$

Proof:

1) Interpretation of the matrix product

Let $B = (\vec{b}_1, \vec{b}_2, \vec{b}_3)$ be a 3×3 matrix just like A. The matrix product (executed column by column)

$$AB = (A\vec{b}_1, A\vec{b}_2, A\vec{b}_3)$$

stands for that linear self-mapping of $\mathbb{R}^3$ which maps $\vec{e}_1$ onto $A\vec{b}_1$, $\vec{e}_2$ onto $A\vec{b}_2$, and $\vec{e}_3$ onto $A\vec{b}_3$. Therefore $A \cdot B$ describes the mapping with the rule "A to B" (composite mapping).

2) Consider

$$\delta(\vec{b}_1, \vec{b}_2, \vec{b}_3) := \det(AB) = D(A\vec{b}_1, A\vec{b}_2, A\vec{b}_3)$$

δ is linear with D in each argument and also satisfies detail (2). This results in the existence of a $\lambda \in \mathbb{R}$ with

$$\left| \quad \delta(\vec{b}_1, \vec{b}_2, \vec{b}_3) = \lambda\, D(\vec{b}_1, \vec{b}_2, \vec{b}_3) \qquad (*) \quad \right|$$

Specialize:

$$\vec{b}_k = \vec{e}_k \quad ; \quad \det A = \lambda$$

□

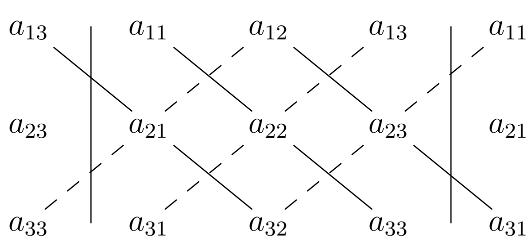
<u>Remark:</u>

Carrying out the calculation $f(*)$ results in the complete expansion of the determinant of

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$

<u>Rule of Sarrus:</u>



The "transposed" of A

$$A^T = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \quad ; \quad \det A^T = \det A \quad ; \quad (AB)^T = B^T A^T$$

Proof of $(AB)^T = B^T A^T$:

$$\sum_{j=1}^{3} a_{ij} b_{jk}$$

is the element in row number k, column number i for $(AB)^T$

$$= \sum_{j=1}^{3} b_{jk} a_{ij}$$

□

For $A, B, C \in M_3(\mathbb{R})$ applies: $(AB)C = A(BC)$

3.5 The Inverse Matrix

If for $A \in M_3(\mathbb{R})$ $\det A \neq 0$ one can (and only then) write each vector of $\mathbb{R}^3$ as a linear combination of the columns $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$. Therefore (as specifically shown for $\vec{e}_1, \vec{e}_2, \vec{e}_3$) a matrix $A^* = (\vec{a}_1^*, \vec{a}_2^*, \vec{a}_3^*)$ exists such that

$$\boxed{AA^* = 1_3}$$

Product theorem: $\det A^* = \dfrac{1}{\det A}$ as $\det 1_3 = 1$; in particular $\det A^* \neq 0$

Therefore $A^{**} \in M_3(\mathbb{R})$ exists with

$$A^* A^{**} = 1_3 \qquad \text{and therefore}$$

$$A = A\,1_3 = A(A^* A^{**}) = (\underbrace{AA^*}_{1_3})A^{**} = A^{**}$$

$$A = A^{**} \quad ; \quad A^* A = 1_3$$

The inverse matrix is uniquely determined. Assumption $AB = 1_3$; multiplication from the left with A^*

$$A^* = A^* 1_3 = (\underbrace{A^* A}_{1_3})B = B$$

3.5.1 The Orthogonal Group O_3 of Euclidean $\mathbb{R}^3$

denotes the set of all orthogonal transformations A of $\mathbb{R}^3$. The associated matrices $A = (\vec{a}_1, \vec{a}_2, \vec{a}_3)$ are characterized by their columns forming pairs of orthogonal unit vectors. In matrix form

$$A^T A \;=\; 1_3$$

$AA^T = 1_3$ is algebraically equivalent because both statements mean $A^{-1} = A^T$.

Geometric meaning of the second relation: The row vectors of A are pairwise orthogonal unit vectors.

With $A, B \in O_3$ also the (matrix) product $AB \in O_3$. Because

$$(AB)^T(AB) \;=\; B^T(A^T(AB)) \;=\; B^T(1_3 B) \;=\; B^T B \;=\; 1_3$$

For all $A \in O_3$ there is also $A^{-1} \in O_3$. Because of $A \in O_3$, $A^{-1} = A^T$ applies.

$$(A^{-1})^T(A^{-1}) \;=\; A^{TT}A^T \;=\; AA^T \;=\; 1_3$$

In particular, $1_3 \in O_3$. The determinant of $A \in O_3$ is by the product theorem for determinants

$$(\det A)^2 \;=\; (\det AA^T) \;=\; 1$$

Hence: $\det A \;=\; +1$ (rotation)
$\det A \;=\; -1$ (improper movement)

The simplest example of an improper movement is provided by

$$A \;=\; -1_3 \;=\; \begin{bmatrix} -1 & & \\ & -1 & \\ & & -1 \end{bmatrix}$$

The images of the unit vectors are $-\vec{e}_1$, $-\vec{e}_2$, $-\vec{e}_3$, i.e. oriented like the first three fingers of the left hand (A with $\det A = -1$ is orientation reversing).

The twelve rotational symmetries of the tetrahedron

Corners: $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$, $\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$, $\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$

180°-rotations around the opposite edge centers $\pm\vec{e}_k$

$$z_1 = \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix} \quad ; \quad z_2 = \begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix} \quad ; \quad z_3 = \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix}$$

$$z_1\,z_2 = \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix} = z_3 = z_2\,z_1$$

$$z_2\,z_3 = \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix} = z_1 = z_3\,z_2$$

$$z_3\,z_1 = \begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix} = z_2 = z_1\,z_3$$

120°-rotation around the axis $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\mathbb{R}$

$$T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad T\vec{e}_1 = \vec{e}_2\,, \quad T\vec{e}_2 = \vec{e}_3\,, \quad T\vec{e}_3 = \vec{e}_1$$

$$T^2 = \begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix} = T^{-1} \quad ; \quad T^3 = 1_3$$

$$z_1\,T = \begin{bmatrix} & & 1 \\ -1 & & \\ & -1 & \end{bmatrix} = T\,z_3$$

$$z_2\,T = \begin{bmatrix} & & -1 \\ 1 & & \\ & -1 & \end{bmatrix} = T\,z_1$$

$$z_3\,T = \begin{bmatrix} & & -1 \\ -1 & & \\ & 1 & \end{bmatrix} = T\,z_2$$

$$z_1\,T^2 \;=\; \begin{bmatrix} & 1 & \\ & & -1 \\ -1 & & \end{bmatrix} \;=\; T^2\,z_2$$

$$z_2\,T^2 \;=\; \begin{bmatrix} & -1 & \\ & & 1 \\ -1 & & \end{bmatrix} \;=\; T^2\,z_3$$

$$z_3\,T^2 \;=\; \begin{bmatrix} & -1 & \\ & & -1 \\ 1 & & \end{bmatrix} \;=\; T^2\,z_1$$

$z_k\,T\,z_k$ has the axis of rotation $\quad T\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}; \quad z_k\begin{bmatrix}1\\1\\1\end{bmatrix}\mathbb{R}$

$$z_1\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}1\\-1\\-1\end{bmatrix} \quad ; \quad z_2\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}-1\\1\\-1\end{bmatrix} \quad ; \quad z_3\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

4. Systems of Linear Equations

The task: for the real orthogonal matrix

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

with $\det A = 1$ a vector $\vec{x} \neq \vec{0}$ is to be determined with $A\vec{x} = \vec{x}$ ($\mathbb{R}\,\vec{x}$ is the "axis of rotation" of A).

A bit more general is the question of "eigenvectors" $\vec{x} \neq 0$ for linear transformations A, that means $A\,\vec{x} = \lambda\ vecx$ for matching λ

$$\text{(E)} \qquad A\,\vec{x} = \lambda\,\vec{x} \quad \Bigg|$$

In (E) all to the left side results in three equations in three indeterminates x_1, x_2, x_3

$$\text{(H)} \quad \left\{ \begin{array}{ccccccc} (a_{11}-\lambda)x_1 & + & a_{12}x_2 & + & a_{13}x_3 & = & 0 \\ a_{12}x_1 & + & (a_{22}-\lambda)x_2 & + & a_{23}x_3 & = & 0 \\ a_{31}x_1 & + & a_{32}x_2 & + & (a_{33}-\lambda)x_3 & = & 0 \end{array} \right.$$

This is a homogeneous linear system of equations with coefficient matrix $A = \lambda\,1_3$.

A homogeneous system of equations $B\,\vec{x} = \vec{0}$ has only a solution $\vec{x} \neq \vec{0}$ if $\det B \neq \vec{0}$. That is so because for $\det B \neq \vec{0}$ the inverse matrix B^{-1} exists to B. It results from its application to $B\,\vec{x} = \vec{0}$

$$B^{-1}(B\,\vec{x}) = (\underbrace{B^{-1}B}_{1_3})\vec{x} = B^{-1}\,\vec{0} = \vec{0} \qquad \Rightarrow \qquad \vec{x} = \vec{0}$$

<u>In particular:</u>

Criterion for the solvability of the eigenvalue equation (E) with a $\vec{x} \neq \vec{0}$.

$$\det(A - \lambda\,1_3) = 0$$

Writing out this condition on the eigenvalue λ results in (see the Rule of Sarrus in the previous chapter)

$$\begin{aligned}\det(A-\lambda\,1_3) &= (a_{11}-\lambda)(a_{22}-\lambda)(a_{33}-\lambda)\\ &\quad + a_{12}\,a_{23}\,a_{31} + a_{13}\,a_{21}\,a_{32} - a_{13}\,a_{31}(a_{22}-\lambda)\\ &\quad - a_{12}\,a_{21}\,(a_{33}-\lambda) - a_{23}\,a_{32}(a_{11}-\lambda) = 0\end{aligned}$$

i.e. ordered according to powers of λ

$$\begin{aligned}0 &= +\lambda^0\det A\\ &\quad - \lambda^1(a_{11}\,a_{22}+a_{22}\,a_{33}+a_{33}\,a_{11}-a_{13}\,a_{31}-a_{12}\,a_{21}-a_{23}\,a_{32})\\ &\quad + \lambda^2(a_{11}+a_{22}+a_{33}) - \lambda^3\end{aligned}$$

$\det(A-\lambda\,1_3)=0$ is called "characteristic" polynomial of A. The existence of the axis of rotation means that $\lambda=1$ is an eigenvalue of the orthogonal matrix $A=(\vec a_1,\vec a_2,\vec a_3)$. $(\det A=1)$ $A^TA=1$ and $\det A=1$ says: For the unit vectors $\vec a_1$, $\vec a_2$, $\vec a_3$ the following formulas apply

$$\begin{aligned}\vec a_1\times\vec a_2 &= \vec a_3\\ \vec a_2\times\vec a_3 &= \vec a_1\\ \vec a_3\times\vec a_1 &= \vec a_2\end{aligned}$$

Hence

$$\begin{aligned}a_{11}\,a_{22} - a_{21}\,a_{12} &= a_{33}\\ a_{22}\,a_{33} - a_{32}\,a_{23} &= a_{11}\\ a_{33}\,a_{11} - a_{13}\,a_{31} &= a_{22}\end{aligned}$$

Inserted into the right-hand side of the above equation $\det(A-\lambda\,1_3)=0$ ordered according to powers of λ, the characteristic polynomial for orthogonal A follows with $\det A=1$

$$\lambda^3 - (a_{11}+a_{22}+a_{33})\lambda^2 + (a_{11}+a_{22}+a_{33})\lambda - 1$$

Obviously $\lambda=1$ is a zero.

If however A orthogonal and $\det A=-1$, then the characteristic polynomial becomes

$$\lambda^3 - (a_{11}+a_{22}+a_{33})\lambda^2 - (a_{11}+a_{22}+a_{33})\lambda + 1$$

It certainly has the eigenvalue $\lambda=-1$ (a zero).

4.1 Solutions for Systems of Linear Equations

$$
\text{(L)} \qquad \begin{array}{ccccccccccc}
a_{11}x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & \ldots & + & a_{1n}x_n & = & b_1 \\
a_{21}x_1 & + & a_{22}x_2 & + & a_{23}x_3 & + & \ldots & + & a_{2n}x_n & = & b_2 \\
\vdots & & & & & & & & & & \vdots \\
a_{m1}x_1 & + & a_{m2}x_2 & + & a_{m3}x_3 & + & \ldots & + & a_{mn}x_n & = & b_m
\end{array}
$$

of m equations in n unknowns x_k $(1 \leq k \leq n)$ with coefficients a_{ik} and right sides b_i, in a previously defined field K (for example $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ etc.).

<u>The two simple special cases:</u>

(1) All $a_{ik} = 0$ $(1 \leq i \leq m\,,\, 1 \leq k \leq n)$. If there is even only one $b_i \neq 0$, then the system of equations is unsolvable. On the other hand, if all $b_i = 0$, then every n-tuple $(x_k)_{1 \leq k \leq n}$ of elements of the field K is a solution of the system of equations.

(2) If $m = 1$ and at least one $a_{ij} \neq 0$, then one can find all solutions if one chooses x_k arbitrarily for all $k \neq j$ and then calculates x_j from the equation.

<u>Elementary transformations:</u>

Under each of the following operations the system of solutions $(x_k)_{1 \leq k \leq n}$ of (L) remains unchanged

(M) multiplication of an equation by a scalar $\neq 0$.

(V) interchanging two equations.

(A) addition of a multiple of one equation to another equation.

4.1.1 The Gaussian Algorithm

reduces through elementary transformations every linear system of equations step by step back to a system that belongs to one of the two simple special cases.

<u>Example:</u>

(1) $(m = n = 3,\ K = \mathbb{Q})$

$$\begin{array}{rcrcrcr} 3x_1 & + & 5x_2 & + & x_3 & = & -4 \\ 2x_1 & + & 4x_2 & + & 5x_3 & = & 9 \\ x_1 & + & 2x_2 & + & 2x_3 & = & 3 \end{array}$$

Tabular form (RS = right side of the equations):

$$\begin{array}{ccc|c} x_1 & x_2 & x_3 & \text{RS} \\ \hline 3 & 5 & 1 & -4 \\ 2 & 4 & 5 & 9 \\ \underline{\underline{1}} & 2 & 2 & 3 \end{array}$$

Forward elimination:

$$\begin{array}{cccc|l} 1 & 2 & 2 & 3 & \text{3. row} \\ 0 & +1 & +5 & +13 & 3 \times \text{3. row} - \text{1. row} \\ 0 & 0 & 1 & 3 & \text{2. row} - 2 \times \text{3. row} \end{array}$$

Backward substitution: $x_3 = 3\,;\ x_2 = -2\,;\ x_1 = 1$

Check:

$$\begin{bmatrix} 3 & 5 & 1 \\ 2 & 4 & 5 \\ 1 & 2 & 2 \end{bmatrix} \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} -4 \\ 9 \\ 3 \end{pmatrix}$$

The system of equations has exactly one solution, namely $(1, -2, 3)$.

(2) $(m = 4,\ n = 5,\ K = \mathbb{R})$

Homogeneous system of equations (i.e. $b_i = 0$) in tabular form:

$$\begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \\ \hline \underline{\underline{1}} & 1 & 0 & 0 & -2 \\ 2 & 0 & 0 & -1 & 0 \\ 1 & -1 & 0 & -1 & 2 \\ 1 & 0 & 0 & 2 & 0 \end{array}$$

$$\begin{array}{ccccc|l} x_1 & x_2 & x_3 & x_4 & x_5 & \\ \hline 1 & 1 & 0 & 0 & -2 & \text{1. row} \\ 0 & -2 & 0 & -1 & 4 & \text{2. row} - 2 \times \text{1. row} \\ 0 & -2 & 0 & -1 & 4 & \text{3. row} - \text{1. row} \\ 0 & -\underline{\underline{1}} & 0 & 2 & 2 & \text{4. row} - \text{1. row} \end{array}$$

x_1	x_2	x_3	x_4	x_5	
1	1	0	0	−2	| 1. row
0	1	0	−2	−2	| 4. row × (−1)
0	0	0	$-\underline{\underline{5}}$	0	| 2. row − 2 × 4. row
0	0	0	−5	0	| 3. row − 2 × 4. row

x_1	x_2	x_3	x_4	x_5		
1	1	0	0	−2		row echelon form
0	1	0	−2	−2		
0	0	0	1	0	| 3. row ÷ (−5)	
0	0	0	0	0	| 3. row − 4. row	

Solution by backward substitution: $x_5 = t$ arbitrary; $x_4 = 0$; $x_3 = s$ arbitrary; $x_2 = 2t$; and $x_1 = 0$.

Check:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & -2 \\ 2 & 0 & 0 & -1 & 0 \\ 1 & -1 & 0 & -1 & 2 \\ 1 & 0 & 0 & 2 & 0 \end{bmatrix} \begin{pmatrix} 0 \\ 2t \\ s \\ 0 \\ t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

The elimination step in the Gaussian algorithm

The system of equations (L) shall not belong to the two simple special cases. Then a first column exists in the coefficient matrix $(a_{ij})_{1 \le i \le m}$ which does not contain all zeros. One chooses a "Pirot element" $a_{i_0 j} \neq 0$. Operation (M): multiplication of the equation number i_0 by the scalar $1/a_{i_0 j}$. Operation (V) brings the hitherto equation with number i_0 into the first line. Then multiple application of operation (A) to make the coefficients in column number j for the remaining equations to 0.
If $(b_i')_{1 \le i \le m}$ denotes the right-hand side column for the row echelon form of the coefficient matrix $A' = (a_{ik}')_{\substack{1 \le i \le m \\ 1 \le k \le n}}$, then one has the following solvability criterion:

(L) is solvable if and only if $b_i' = 0$ if $r + 1 \le i \le m$

Here r is the "rank" of the coefficient matrix, which corresponds to the number of rows which do not exclusively contain zeros. In particular, the solvability criterion is met in that case when r is equal to the number m of equations.

The smallest field has exactly 2 elements: $\mathbb{F}_2 = \{0, L\}$

Calculation rules: (calculation with whole numbers modulo 2)

$$\begin{array}{lcccccc}
\text{Addition:} & 0+0 & = & L+L & = & 0 \\
 & 0+L & = & L+0 & = & L
\end{array}$$

$$\begin{array}{lcccccccc}
\text{Multiplication:} & 0\cdot 0 & = & 0\cdot L & = & L\cdot 0 & = & 0 \\
 & L\cdot L & = & L & & & &
\end{array}$$

Matrix calculation is also performed via $\mathbb{F}_2$.

Example:

(3) Totality of 3×3 matrices: $M_3\mathbb{F}_2$. The number of these matrices is $2^9 = 512$. The number of invertible matrices in $M_3\mathbb{F}_2$ is 168.

Question: Is $\begin{bmatrix} L & 0 & L \\ L & L & 0 \\ L & L & L \end{bmatrix}$ invertible?

System of equations

$$\begin{array}{ccc|l}
x & y & z & \text{RS} \\
L & 0 & L & b_1 \\
\underline{\underline{L}} & L & 0 & b_2 \\
L & L & L & b_3
\end{array}$$

$$\begin{array}{ccc|l}
L & L & 0 & b_2 \\
0 & L & L & b_1+b_2 \\
0 & 0 & L & b_2+b_3
\end{array}$$

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} L \\ 0 \\ 0 \end{bmatrix} \quad ; \quad \begin{bmatrix} 0 \\ L \\ 0 \end{bmatrix} \quad ; \quad \begin{bmatrix} 0 \\ 0 \\ L \end{bmatrix}$$

$$\underline{\underline{1}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} L \\ L \\ 0 \end{bmatrix} \; ; \; \underline{\underline{2}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} L \\ 0 \\ L \end{bmatrix} \; ; \; \underline{\underline{3}} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} L \\ L \\ L \end{bmatrix}$$

Check:

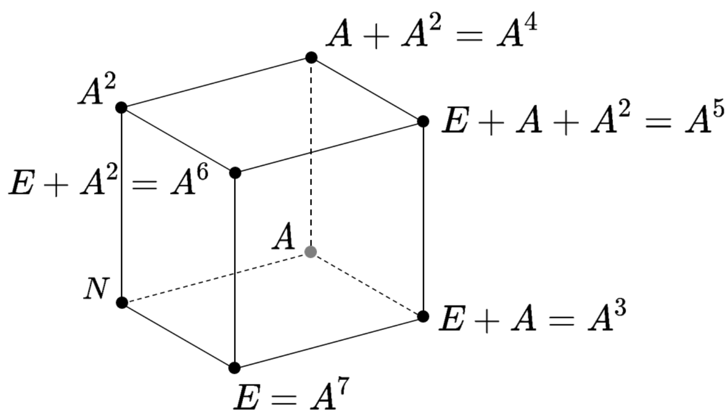
$$\begin{bmatrix} L & 0 & L \\ L & L & 0 \\ L & L & L \end{bmatrix} \begin{bmatrix} L & L & L \\ L & 0 & L \\ 0 & L & L \end{bmatrix} = \begin{bmatrix} L & 0 & 0 \\ 0 & L & 0 \\ 0 & 0 & L \end{bmatrix} = E$$

In $M_3\mathbb{F}_2$ 8 matrices can be selected in such a way that they form a field under addition and multiplication of matrices.

$$N = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \quad E = \begin{bmatrix} L & 0 & 0 \\ 0 & L & 0 \\ 0 & 0 & L \end{bmatrix}; \quad A = \begin{bmatrix} 0 & L & 0 \\ 0 & 0 & L \\ L & L & 0 \end{bmatrix}$$

$$A^2 = \begin{bmatrix} 0 & 0 & L \\ L & L & 0 \\ 0 & L & L \end{bmatrix}; \quad A^3 = \begin{bmatrix} L & L & 0 \\ 0 & L & L \\ L & L & L \end{bmatrix} = E + A$$

$$A^4 = A + A^2; \quad A^5 = E + A + A^2; \quad A^6 = E + A^2; \quad A^7 = E$$



5. Plane and Spatial Curves

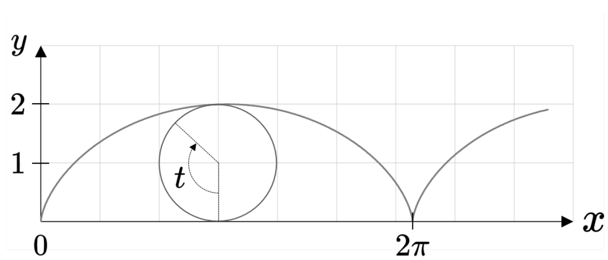
Definition

A curve is understood to be a mapping of an interval I with images in the plane $\mathbb{R}^2$ or $\mathbb{C}$ or in space $\mathbb{R}^3$, whose coordinates are continuously differentiable as a function of $t \in I$ (occasionally: continuous and piecewise continuously differentiable).

Example:

(0) The unit circle line $\quad I = \mathbb{R}\,; \quad \vec{\gamma}(t) = \mathrm{e}^{it}$

(1) The common cycloid

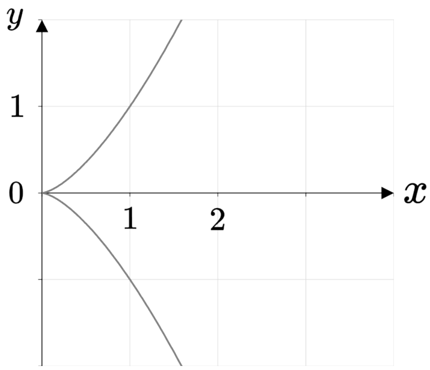


$$I = \mathbb{R}\,;$$

$$\vec{\gamma}(t) = \begin{pmatrix} t \\ 1 \end{pmatrix} - \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$$

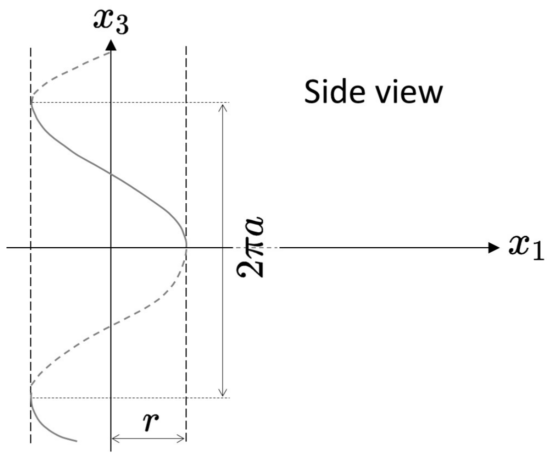
More generally: shortened and lengthened cycloid.

(2) Neil's parabola



$$I = \mathbb{R}\,; \quad \vec{\gamma}(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}$$

(3) The helix



$$I = \mathbb{R}\,; \quad \vec{x}(t) \; = \; \begin{pmatrix} r\cos t \\ r\sin t \\ at \end{pmatrix}$$

The curves assume the existence of the limit value for each $t \in I$

$$\lim_{\substack{h\to 0\\ h\in\mathbb{R}}} \frac{1}{h}\big(\vec{\gamma}(t+h) \;-\; \vec{\gamma}(t)\big) \;=\; \dot{\vec{\gamma}}(t)$$

<u>Geometric interpretation</u>

$\dot{\vec{\gamma}}(t)$ gives a vector in the direction of the curve tangent to the point in “time” t. Its length depends on the throughput speed. $\dot{\vec{\gamma}}(t)$ is the instantaneous velocity in vector form. Tangent: $\vec{\gamma}(t) \,+\, \mathbb{R}\,\dot{\vec{\gamma}}(t)$

Tangents to the curve lines above:

Unit cricle line: $\vec{\gamma}(t) \;=\; \mathrm{e}^{it}$ $\qquad \dot{\vec{\gamma}}(t) \;=\; i\mathrm{e}^{it}$

Cycloid: $\vec{\gamma}(t) \;=\; \begin{pmatrix} t-\sin t \\ 1-\cos t \end{pmatrix}$ $\qquad \dot{\vec{\gamma}}(t) \;=\; \begin{pmatrix} 1-\cos t \\ \sin t \end{pmatrix}$

Neil’s parabola: $\vec{\gamma}(t) \;=\; \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}$ $\qquad \dot{\vec{\gamma}}(t) \;=\; \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix}$

Helix: $\vec{x}(t) \;=\; \begin{pmatrix} r\cos t \\ r\sin t \\ at \end{pmatrix}$ $\qquad \dot{\vec{x}}(t) \;=\; \begin{pmatrix} -r\sin t \\ r\cos t \\ a \end{pmatrix}$

Curve points with $\dot{\vec{\gamma}}(gt) = \vec{0}$ are called “singular curve points”. A curve is called regular if it contains no singular points. The cycloid has the singular set of points $2\pi\mathbb{Z}$. Neil’s parabola has only $t_0 \;=\; 0$ as a singular point. Unit circle line and helix are regular curves.

5.1 Definition of the Curve Length

Convention: The arrows for the vectors are sometimes omitted.

Let $\gamma : [a,b] \to \mathbb{R}^n$ $(n = 2.3)$ be a curve

$$\boxed{l(\gamma) \;:=\; \int\limits_a^b \|\dot\gamma(t)\| \,\mathrm{d}t}$$

Length of the unit circle line:

$$\int\limits_0^{2\pi} \left|i\mathrm{e}^{it}\right| \mathrm{d}t \;=\; 2\pi$$

Length of the cycloid:

$$\int\limits_0^{2\pi} \sqrt{(1-\cos t)^2 + \sin^2 t}\,\mathrm{d}t \;=\; \int\limits_0^{2\pi} \sqrt{2 - 2\cos t}\,\mathrm{d}t$$

$$\text{because} \quad 1 - \cos t \;=\; 2\sin^2\left({}^t/_2\right)$$

$$= \int\limits_0^{2\pi} 2\sin\frac{t}{2}\,\mathrm{d}t \;=\; \left[-4\cos\frac{t}{2}\right]_0^{2\pi} \;=\; 8$$

Length of the helix:

$$\begin{pmatrix} r\cos t \\ r\sin t \\ at \end{pmatrix} \;=\; \vec{x}(t) \quad ; \quad \dot{\vec{x}}(t) \;=\; \begin{pmatrix} -r\sin t \\ r\cos t \\ a \end{pmatrix}$$

$$L \;=\; \int\limits_0^{2\pi} \sqrt{r^2 + a^2}\,\mathrm{d}t \;=\; 2\pi\sqrt{r^2 + a^2}$$

When unrolling the cylinder surface, the helix becomes the hypotenuse of a right-angled triangle with legs of length $2\pi r$ and $2\pi a$.

<u>Parameter transformation</u>

Let $\gamma : [a,b] \to \mathbb{R}^n$ be a curve. A mapping $\varphi : [c,d] \to [a,b]$ between parameters in intervals is called a parameter transformation if $\varphi(c) = a,\;\; varphi(d) = b$; furthermore, φ continuously differentiable with $\dot\varphi(t) > 0 \;\; \forall\, t$

$$\vec{x}(t) = \gamma(\varphi(t))$$

is called reparameterization of the curve γ.

Derivative behavior

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\vec{x}(t) = \begin{pmatrix} \frac{\mathrm{d}}{\mathrm{d}t}\,\gamma_1(\varphi(t)) \\ \frac{\mathrm{d}}{\mathrm{d}t}\,\gamma_2(\varphi(t)) \end{pmatrix} = \begin{pmatrix} \gamma_1'(\varphi(t)) \\ \gamma_1'(\varphi(t)) \end{pmatrix} \dot{\varphi}(t)$$

short form

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\gamma(\varphi(t)) = \gamma'(\varphi(t))\dot{\varphi}(t)$$

Curve length behavior

$$L(\vec{x}) = \int\limits_c^d \big\|\dot{\vec{x}}(t)\big\|\,\mathrm{d}t = \int\limits_c^d \|\gamma'(\varphi(t))\|\dot{\varphi}(t)\,\mathrm{d}t$$

(with transformation formula from section 1.1.1)

$$= \int\limits_{a=\varphi(c)}^{b=\varphi(d)} \|\gamma'(s)\|\,\mathrm{d}s$$

Result: Curve lengths are invariant to parameter transformations.

5.2 The Line Integral over a Vector Field

Let $n = 2, 3\,,\ \vec{h} : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous vector field (continuity means that small perturbations in the argument produce only small changes in value). Let $\gamma : [a, b] \to \mathbb{R}^n$ be a curve. The line integral along $\vec{\gamma}$ over the vector field $\vec{h}$

$$\int_\gamma \vec{h}(\vec{x})\,\mathrm{d}\vec{x} = \int\limits_a^b \vec{h}(\vec{\gamma}(t))\,\dot{\vec{\gamma}}(t)\,\mathrm{d}t$$

Theorem: The line integral along γ over the vector field $\vec{h}$ does not change when γ is reparameterized.

Proof:

Let $\varphi :\ [c, d] \to [a, b]$ be a parameter transformation. Mapping $\vec{\delta}(t) = \vec{\gamma}(\varphi(t))\,.$

Calculate $$\int_\delta \vec{h}(\vec{x})\,\mathrm{d}\vec{x} \;=\; \int\limits_c^d \vec{h}(\vec{\delta}(t))\,\dot{\vec{\delta}}(t)\,\mathrm{d}t$$

$$= \int\limits_c^d \Big[\vec{h}(\vec{\gamma}(\varphi(t)))\,\dot{\vec{\gamma}}(\varphi(t))\Big]\dot{\varphi}(t)\,\mathrm{d}t \;\overset{\substack{\text{Transformation}\\ \text{formula}\\ \downarrow}}{=}\; \int\limits_a^b \vec{h}(\vec{\gamma}(s))\,\dot{\vec{\gamma}}(s)\,\mathrm{d}s$$

<u>An example:</u> $\vec{h}(\vec{x}) \;=\; -\alpha\|\vec{x}\|^{-3}\vec{x}$ (gravitational field) $n=3$; $\vec{x}\neq 0$

<u>Remark:</u> This field cannot be declared continuous in the origin.

A sample particle of mass 1 shall pass through the gravitational field along the curve $\gamma:\,[a,b]\to\mathbb{R}^3$ with $\gamma(t)\neq\ vec0$

$$\int_\gamma \vec{h}(\vec{x})\,\mathrm{d}\vec{x} \;=\; -\alpha\int\limits_a^b \|\gamma(t)\|^{-3}(\gamma(t)\,\dot{\gamma}(t))\,\mathrm{d}t$$

$$= \;\alpha\int\limits_a^b \frac{-\gamma_1(t)\dot{\gamma}_1(t)-\gamma_2(t)\dot{\gamma}_2(t)-\gamma_3(t)\dot{\gamma}_3(t)}{\big[\gamma_1^2(t)+\gamma_2^2(t)+\gamma_3^2(t)\big]^{3/2}}\,\mathrm{d}t$$

$$= \;\alpha\int\limits_a^b \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\|\gamma(t)\|}\right)\mathrm{d}t \;=\; \alpha\left(\frac{1}{\|\gamma(b)\|}-\frac{1}{\|\gamma(a)\|}\right)$$

The work integral in the gravitational field only depends on the start and end point of the curve γ; if these points coincide, the integral vanishes. Newton's law of motion states that the change in the quantity of motion over time $m\,\dot{\vec{x}}$ of a particle of mass m in the force field $\vec{h}(\vec{x})$ of space is proportional to $\vec{h}(\vec{x})$

$$m\ddot{\vec{x}} \;=\; \frac{\mathrm{d}}{\mathrm{d}t}(m\,\dot{\vec{x}}) \;=\; \vec{h}(\vec{x})$$

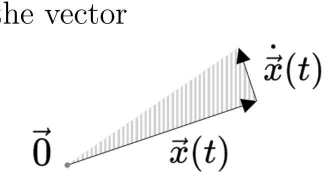
provided the mass is constant in time. Of specific importance: <u>central fields</u> with their center at the coordinate origin

$$\vec{h}(\vec{x}) \;=\; \lambda\,(\|\vec{x}\|)\,\vec{x} \qquad \vec{x}\neq\vec{0}$$

with a continuous function $\lambda:\,]0,\infty[\,\to\mathbb{R}$

<u>Area theorem:</u> For the (free) movement along the curve $\vec{x}(t)$ which follows from the equation of motion of a particle in the central field, the vector

$$\frac{1}{2}\Big(\vec{x}(t)\times\dot{\vec{x}}(t)\Big) \quad \text{is constant.}$$



Proof:

Because of $m\ddot{\vec{x}} = \vec{h}(\vec{x})$, the coordinate functions of $\vec{x}(t)$ are twice continuously differentiable functions of time. therefore $\vec{x}(t) \times \dot{\vec{x}}(t)$ is a vector-valued function of time whose coordinates are continuously differentiable.

The proof: Every coordinate function of $\vec{x}(t) \times \dot{\vec{x}}(t)$ has the derivative 0.

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\vec{x}(t) \times \dot{\vec{x}}(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_2\,\dot{x}_3 - x_3\,\dot{x}_2 \\ x_3\,\dot{x}_1 - x_1\,\dot{x}_3 \\ x_1\,\dot{x}_2 - x_2\,\dot{x}_1 \end{bmatrix} \\
&= \begin{bmatrix} \dot{x}_2\,\dot{x}_3 - \dot{x}_3\,\dot{x}_2 + x_2\,\ddot{x}_3 - x_3\,\ddot{x}_2 \\ \dot{x}_3\,\dot{x}_1 - \dot{x}_1\,\dot{x}_3 + x_3\,\ddot{x}_1 - x_1\,\ddot{x}_3 \\ \dot{x}_1\,\dot{x}_2 - \dot{x}_2\,\dot{x}_1 + x_1\,\ddot{x}_2 - x_2\,\ddot{x}_1 \end{bmatrix} \\
&= \underbrace{\dot{\vec{x}}(t) \times \dot{\vec{x}}(t)}_{=\vec{0}} + \vec{x}(t)\times\ddot{\vec{x}}(t) \quad \text{(because of the equation of motion)} \\
&= \vec{x}(t) \times \left(\frac{1}{m}\vec{h}(\vec{x})\right) = \frac{1}{m}\lambda(\|\vec{x}\|)\underbrace{(\vec{x} \times \vec{x})}_{\vec{0}}
\end{aligned}$$

□

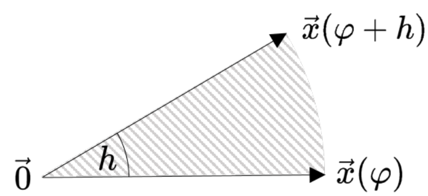
5.3 Polar Coordinates for Plane Curves

Short form: $r = r(\varphi)$

The parameter φ is the angle between the x-axis and the position vector of the curve point ($\vec{x}$); $r = \|\vec{x}\|$. Written out

$$\vec{x}(\varphi) = r(\varphi)\begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} \quad \text{in } \mathbb{R}^2$$

$$\vec{x}(\varphi) = r(\varphi)\mathrm{e}^{i\varphi} \quad \text{in } \mathbb{C}$$



Useful generalization: φ can itself be understood as a function of another parameter t. The vector

$$\vec{x}(\varphi) \times (\vec{x}(\varphi + h) - \vec{x}(\varphi)) = \vec{x}(\varphi) \times \vec{x}(\varphi + h)$$

has as norm twice the area of the triangle with the vertices $\vec{0}$, $\vec{x}(\varphi)$, $\vec{x}(\varphi + h)$. With $h \to 0$ and integration one gets

$$\boxed{\vec{F} \;=\; \frac{1}{2}\int\limits_{\varphi_0}^{\varphi_1} \vec{x}(\varphi) \times \vec{x}\,'(\varphi)\,\mathrm{d}\varphi}$$

in vector form, the area swept over by the “radius vector”

$$\vec{x}(\varphi) \times \vec{x}\,'(\varphi) \;=\; \begin{bmatrix} 0 \\ 0 \\ x_1(\varphi)\,x_2'(\varphi) \;-\; x_2(\varphi)\;\; x_1'(\varphi) \end{bmatrix}$$

Extension for the case that φ is a function of a parameter t.

$$\vec{x}(\varphi) \;=\; r(\varphi)\begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$$

$$\vec{x}\,'(\varphi) \;=\; r\,'(\varphi)\begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix} \;+\; r(\varphi)\begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix}$$

with that, the third coordinate in the vector product becomes

$$x_1\,x_2' \;-\; x_2\,x_1' \;=\; r^2(\varphi)$$

and thus the third component of the vector product becomes

$$\boxed{F \;=\; \frac{1}{2}\int\limits_a^b r^2\dot{\varphi}\,\mathrm{d}t}$$

Example:

(4) The Archimedean spiral: $r \;=\; a\,\varphi \qquad (a > 0) \quad \varphi \geq 0$

In detail:

$$\vec{x}(\varphi) \;=\; \begin{pmatrix} a\,\varphi\cos\varphi \\ a\,\varphi\sin\varphi \end{pmatrix}$$

$$\vec{x}\,'(\varphi) \;=\; a\begin{pmatrix} \cos\varphi - \varphi\sin\varphi \\ \sin\varphi + \varphi\cos\varphi \end{pmatrix}$$

$$\|\vec{x}\,'(\varphi)\|^2 \;=\; a^2(1+\varphi^2)$$

Arc length:

$$s(\phi) \;=\; \int\limits_0^\phi \sqrt{1+\varphi^2} \;=\; \frac{a}{2}\big[\phi(1+\phi^2)^{1/2} + \ln\big(\phi + (1+\phi^2)^{1/2}\big)\big]$$

Swept area:

$$F(\phi) \;=\; \frac{1}{2}\int\limits_0^{\phi} a^2\varphi^2\,\mathrm{d}\varphi \;=\; \frac{a^2}{6}\phi^3$$

If $\phi > 2\pi$, then certain parts of the plane will be swept multiple times.

Remark:

Even in the form $y = f(x)$ with continuously differentiable $f : I \to \mathbb{R}$ one is dealing with a parameterized curve. The parameter is here $x \in I$.

5.4 The Curvature of a Plane Curve

Let γ be an at least twice continuously differentiable curve in $\mathbb{R}^2$ (or $\mathbb{C}$) without singular points. The curvature at the curve point $\gamma(t)$ is defined as the angular velocity of the tangent unit vector relative to the arc length!

Formula for the curvature of the curve at the point $\gamma(t)$:

$$\boxed{\kappa \;=\; \frac{\dot\gamma_1\ddot\gamma_2 - \dot\gamma_2\ddot\gamma_1}{\|\dot\gamma\|^3}}$$

Example:

(0) The circle line around 0 with radius $r > 0$

$$\gamma(t) \;=\; r\mathrm{e}^{it} \qquad \dot\gamma(t) \;=\; ir\mathrm{e}^{it} \qquad \ddot\gamma(t) \;=\; -r\mathrm{e}^{it}$$

$$\kappa \;=\; \frac{r^2[(-\sin t)^2 + (\cos t)^2]}{r^3} \;=\; \frac{1}{r}$$

The curvature of the positively traversed circle line is equal to the reciprocal radius of the circle.

$$\left|\frac{1}{\kappa}\right| \;=\; \text{“radius of curvature”}$$

Invariance of curvature to parameter transformations

Let $\varphi : [c, d] \to [a, b]$ be a parameter transformation that is at least twice continuously differentiable.

Consider: $\quad \delta(t) \;=\; \gamma(\varphi(t))$

Component-wise differentiation!

$$\dot\delta(t) \;=\; \begin{pmatrix}\dot\gamma_1(\varphi(t))\\ \dot\gamma_2(\varphi(t))\end{pmatrix}\dot\varphi(t)$$

$$\ddot\delta(t) \;=\; \ddot\gamma(\varphi(t))\dot\varphi^2(t) \,+\, \dot\gamma(\varphi(t))\ddot\varphi(t)$$

$$\frac{\dot\delta_1\ddot\delta_2-\dot\delta_2\ddot\delta_1}{\|\delta\|^3} \;=\; \frac{\left[\dot\gamma_1(\varphi(t))\,\ddot\gamma_2(\varphi(t)) \,-\, \dot\gamma_2(\varphi(t))\,\ddot\gamma_1(\varphi(t))\right]\dot\varphi^3(t)}{\|\dot\gamma(\varphi(t))\|^3\dot\varphi^3(t)} \,+\, 0$$

$$\frac{\dot\delta_1\ddot\delta_2-\dot\delta_2\ddot\delta_1}{\|\delta\|^3} \;=\; \frac{(\dot\gamma_1\circ\varphi)\,(\ddot\gamma_2\circ\varphi) \,-\, (\dot\gamma_2\circ\varphi)\,(\ddot\gamma_1\circ\varphi)}{\|\dot\gamma\circ\varphi\|^3}$$

Parameterization with respect to the arc length s

Let γ be a regular curve in $\mathbb{R}^n$ $(n=2,3)$ $\gamma:[a,b]\to\mathbb{R}^n$. The arc length

$$s(T) \;=\; \int_{T_0}^{T}\|\dot\gamma(t)\|\,\mathrm{d}t \qquad T\in[a,b]$$

is a differentiable function of T with derivative $\dot s(T)=\|\dot\gamma(T)\|>0$. With $A=s(a)$, $B=s(b)$ there exists for s thus an inverse function $\varphi:\ [A,B]\to[a,b]$ with the derivative

$$\varphi'(s) \;=\; \frac{1}{\dot s(\varphi(s))} \;=\; \frac{1}{\|\dot\gamma(\varphi(s))\|} \qquad (*)$$

Therefore, γ is parameterized by φ:

$$\delta(s) \;=\; \gamma(\varphi(s))$$

Derivative:

$$\delta'(s) \;=\; \dot\gamma(\varphi(s))\,\varphi'(s) \;\overset{\substack{\text{because of } (*)\\ \downarrow}}{=}\; \frac{\dot\gamma(\varphi(s))}{\|\dot\gamma(\varphi(s))\|}$$

The tangent vectors $\delta'(s)$ are unit vectors at every point.

Let now $n=2$. The curvature with parameterization with respect to s is

$$\kappa \;=\; \delta_1'(s)\,\delta_2''(s) \,-\, \delta_2'(s)\,\delta_1''(s)$$

With

$$\delta'(s) \;=\; \begin{pmatrix}\cos\alpha(s)\\ \sin\alpha(s)\end{pmatrix} \quad \text{one gets} \quad \delta''(s) \;=\; \begin{pmatrix}-\sin\alpha(s)\\ \cos\alpha(s)\end{pmatrix}\alpha'(s)$$

and thus

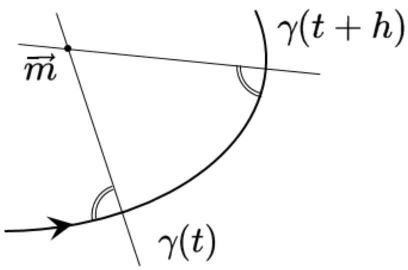
$$\kappa(s) \;=\; \alpha'(s)$$

<u>Back to example 0:</u>

The unit circle line $\gamma(t) \;=\; \mathrm{e}^{it}$ is already parameterized with respect to the arc length:

$$|\dot\gamma(t)| \;=\; |i\mathrm{e}^{it}| \;=1$$

<u>The circle of curvature</u>



Let $\gamma : [a,b] \to \mathbb{R}^2$ be a twice continuously differentiable curve and regular. For sufficiently small $h \neq 0$, one calculates the intersection $\vec{m}$ of the two normal vectors to the tangents in the curve points $\gamma(t)$, $\gamma(t+h)$. One obtains the normal unit vector perpendicular to $\dot\gamma(t)$ or respectively to $\dot\gamma(t+h)$ from the

<u>Point-normal equations:</u>

$$\left.\begin{aligned} \dot\gamma(t)[\gamma(t)-\vec{m}] &= 0 \\ \dot\gamma(t+h)[\gamma(t+h)-\vec{m}] &= 0 \end{aligned}\right\} \quad \text{linear system of equations in } m_1 \text{ and } m_2$$

Elemental transformation

1. $$\dot\gamma(t)\,\vec{m} \;=\; \gamma(t)\,\dot\gamma(t) \;:=\; g(t)$$
   $$\dot\gamma(t+h)\,\vec{m} \;=\; \gamma(t+h)\,\dot\gamma(t+h) \;:=\; g(t+h)$$
2. $$\frac{1}{h}(\dot\gamma(t+h)-\dot\gamma(t)) \;=\; \frac{1}{h}(g(t+h)-g(t))$$

Coefficient matrix

$$A \;=\; \begin{bmatrix} \dot\gamma_1(t) & \dot\gamma_2(t) \\ \frac{1}{h}(\dot\gamma_1(t+h)-\dot\gamma_1(t)) & \frac{1}{h}(\dot\gamma_2(t+h)-\dot\gamma_2(t)) \end{bmatrix}$$

The determinant for $h \to 0$

$$\lim_{h\to 0} \det A \;=\; \dot\gamma_1\ddot\gamma_2 \;-\; \dot\gamma_2\ddot\gamma_1 \;\neq\; 0 \qquad \text{this is now being presupposed!}$$

Hence A is invertible for small h and the point where the normal vectors intersect is

$$\begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = A^{-1} \begin{bmatrix} \gamma(t)\dot{\gamma}(t) \\ \frac{1}{h}(\gamma(t+h)\dot{\gamma}(t+h) - \gamma(t)\dot{\gamma}(t)) \end{bmatrix}$$

Hence a limit position exists for this point of intersection which is again denoted by $\vec{m}$ (for $h \to 0$).

$$\begin{aligned} \vec{m} &= \frac{1}{\dot{\gamma}_1\ddot{\gamma}_2 - \dot{\gamma}_2\ddot{\gamma}_1} \begin{bmatrix} \ddot{\gamma}_2 & -\dot{\gamma}_2 \\ -\ddot{\gamma}_1 & \dot{\gamma}_1 \end{bmatrix} \begin{bmatrix} \gamma\dot{\gamma} \\ \dot{\gamma}\dot{\gamma} + \gamma\ddot{\gamma} \end{bmatrix} \qquad \text{(product rule)} \\ &= \frac{1}{\dot{\gamma}_1\ddot{\gamma}_2 - \dot{\gamma}_2\ddot{\gamma}_1} \begin{bmatrix} +\ddot{\gamma}_2(\gamma\dot{\gamma}) - \dot{\gamma}_2(\dot{\gamma}\dot{\gamma}) - \dot{\gamma}_2(\gamma\ddot{\gamma}) \\ -\ddot{\gamma}_1(\gamma\dot{\gamma}) + \dot{\gamma}_1(\dot{\gamma}\dot{\gamma}) + \dot{\gamma}_1(\gamma\ddot{\gamma}) \end{bmatrix} \\ &= \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} + \frac{\|\dot{\gamma}\dot{\gamma}\|^2}{\dot{\gamma}_1\ddot{\gamma}_2 - \dot{\gamma}_2\ddot{\gamma}_1} \begin{pmatrix} -\dot{\gamma}_2 \\ \dot{\gamma}_1 \end{pmatrix} \end{aligned}$$

<u>Result:</u> The center of curvature curve

$$\boxed{\vec{m} = \vec{\gamma} + \frac{1}{\kappa}\vec{n}} \qquad \text{evolute of } \vec{\gamma}$$

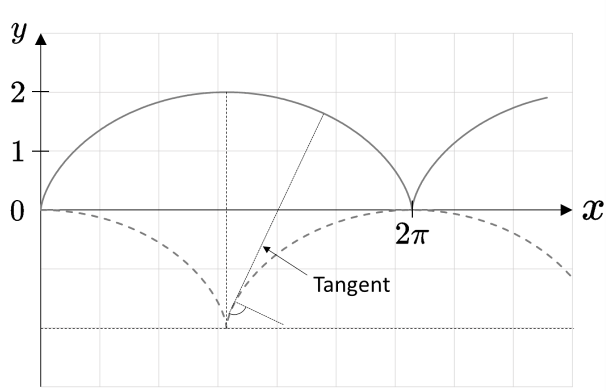
where $\vec{n}$ is the normal unit vector, arisen from the tangent unit vector

$$\frac{1}{\|\dot{\gamma}\|}\begin{pmatrix} \dot{\gamma}_1 \\ \dot{\gamma}_2 \end{pmatrix}$$

by rotating 90° counterclockwise. The circle with center $\vec{m}$ through the curve point $\vec{\gamma}$ has the same curvature as the curve.

<u>Example:</u>

(1) The evolute of the common cycloid



$$\gamma(t) = \begin{pmatrix} t \\ 1 \end{pmatrix} - \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$$

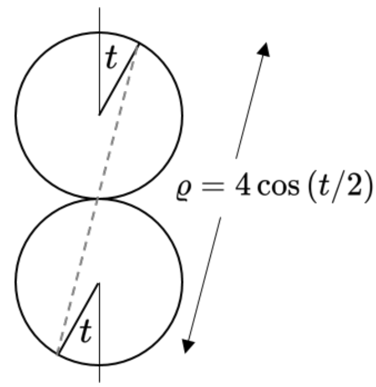
$$\dot{\gamma}(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix}$$

$$\ddot{\gamma}(t) = \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$$

$$\|\dot{\gamma}(t)\|^2 \;=\; (1-\cos t)^2 + \sin^2 t \;=\; 2 - 2\cos t$$

$$\kappa(t) \;=\; \|\dot{\gamma}(t)\|^3 \Big/ \times \text{ third component of } \;\dot{\gamma}\times\ddot{\gamma}$$

$$= \|\dot{\gamma}(t)\|^3(\cos t - 1) \;=\; -\frac{1}{2}\|\dot{\gamma}(t)\|^{-1}$$



$$\vec{m} \;=\; \gamma(t) + \frac{1}{\kappa(t)}\vec{n}(t) \;=\; \gamma(t) - 2\begin{pmatrix} -\dot{\gamma}_2(t) \\ \dot{\gamma}_1(t) \end{pmatrix}$$

$$\vec{m} \;=\; \begin{pmatrix} t \\ -1 \end{pmatrix} + \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$$

<u>General properties of evolutes</u>

Let $\gamma : I \to \mathbb{R}^2$ be a curve parameterized with respect to the arc length s and be at least three times continuously differentiable and regular.

<u>a)</u> $\gamma_1'^2(s) + \gamma_2'^2(s) \;=\; 1$ constant

by differentiation:

$$\gamma_1'(s)\gamma_1''(s) + \gamma_2'(s)\gamma_2''(s) \;=\; 0$$

Geometric meaning: $\gamma'(s)$ is everywhere perpendicular to $\gamma''(s)$.

<u>b)</u> Differentiation of the evolute equation according to s

$$\vec{m}' \;=\; \vec{\gamma}'(s) + \frac{1}{\kappa}\vec{n}'(s) - \frac{\kappa'(s)}{\kappa^2(s)}\vec{n}(s)$$

$$\kappa\vec{\gamma}' + \vec{n}' \;=\; (\gamma_1'\gamma_2'' - \gamma_2'\gamma_1'')\begin{pmatrix} \gamma_1' \\ \gamma_2' \end{pmatrix} + \begin{pmatrix} -\gamma_2' \\ \gamma_1' \end{pmatrix}$$

$$= \begin{bmatrix} (\gamma_1'^2 - 1)\gamma_2'' \;-\; \gamma_1'\gamma_1''\gamma_2' \\ \gamma_1'\gamma_2'\gamma_2'' \;+\; (1-\gamma_2'^2)\gamma_1'' \end{bmatrix}$$

$$= (\gamma_1'\gamma_1'' + \gamma_2'\gamma_2')\begin{pmatrix} -\gamma_2' \\ \gamma_1' \end{pmatrix}$$

$$\overrightarrow{m}' \;=\; -\frac{\kappa'}{\kappa^2}(s)\,\vec{n}(s)$$

Consequences:

In each extremum of the curvature of $\vec{\gamma}$ a singularity of the evolute arises. In the points where the initial curve and its evolute are regular, the tangent to the evolute is also the normal to the initial curve.

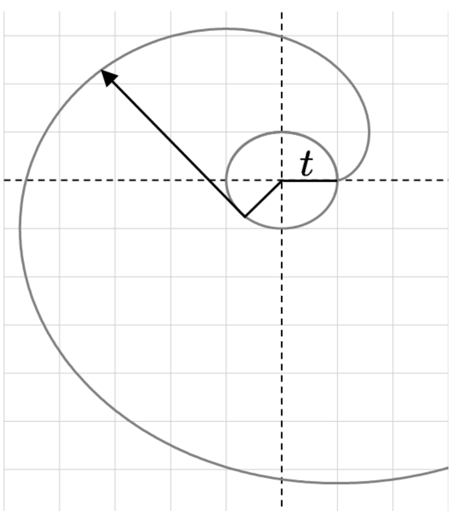
Transition to the norm $\quad \|\vec{m}\,'(s)\| \;=\; \pm\dfrac{\mathrm{d}}{\mathrm{d}s}\,\dfrac{1}{\kappa}$

Expressed geometrically: In regions of strict monotonicity of κ either the sum or the difference between the radius of curvature of γ and the arc length of $\vec{m}$ is constant.

Thread construction of the curve $\vec{\gamma}$ from its evolute

Which curve arises from a given evolute through the thread construction? They are the so-called involutes of a curve!

The involute of a circle:



$$\vec{\gamma}(t) \;=\; \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} + t\begin{pmatrix} \sin t \\ -\cos t \end{pmatrix}; \quad t > 0$$

Check: calculation of the evolute

$$\dot{\gamma}(t) \;=\; t\begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \qquad \text{(Archimedean spiral)}$$

$$\|\dot{\gamma}(t)\| \;=\; t$$

$$\ddot{\gamma}(t) \;=\; \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} + t\begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$$

$$\kappa(t) \;=\; \|\dot{\gamma}(t)\|^{-3}\, t^2\,(\cos^2 t + \sin^2 t) \;=\; \frac{1}{t}$$

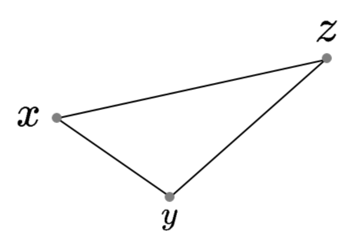
$$\vec{m} \;=\; \vec{\gamma} + \frac{1}{\kappa}\vec{n} \;=\; \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} + t\begin{pmatrix} \sin t \\ -\cos t \end{pmatrix} + t\begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$$

$$\vec{m} \;=\; \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

6. Neighborhoods and Limits

Definition:

A metric space X is a set with a distance function (metric) $(x, y) \longmapsto |x, y| \geq 0$ with the following properties:



(M1) $|x, y| \; = \; 0\,,$ exactly when $x = y$

(M2) $|x, y| \; = \; |y, x|$ (symmetry)

(M3) $|x, z| \; \leq \; |x, y| + |y, z|$ (triangle inequality)

Examples:

(1) $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}^n$ $(n = 2, 3)$ with Euclidean distance

(2) $X \;=\; \{f : [0,1] \to \mathbb{R}\,;\; f \text{ continuous}\}$

$$|f, g| \;\coloneqq\; \sup_{t \,\in\, [0,1]} |f(t) - g(t)|$$

The triangle inequality follows from that for the real numbers (in the context of uniform convergence).

(3) For every metric space X, every subset U with the restriction of the metric from X on U again results in a metric space!

The (open) ϵ-sphere $B_\epsilon(a)$ of radius $\epsilon > 0$ around the point $a \in X$

$$B_\epsilon(a) \;\coloneqq\; \{x \in X;\, |x, a| < \epsilon\}$$

A subset V of X is called a "neighborhood" of a, if at least for one $\epsilon > 0$ applies

$$B_\epsilon(a) \subset V$$

A set U of X is called open if U is a neighborhood of each of its points.

Example:

(4) $B_\epsilon(a)$ is open

For $\quad x \in B_\epsilon(a) \quad , \quad \eta \; = \; |x,a| \; < \; \epsilon$

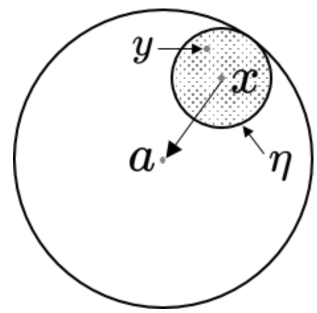
Assertion: $\quad B_{\epsilon-\eta}(x) \subset B_\epsilon(a)$

Let $\;y \in B_{\epsilon-\eta}(x)\,,\;$ that means $\quad |x,y| \; < \; \epsilon - \eta$

With the triangle inequality:

$$|y,a| \;\leq\; |x,y| + |x,a| \;<\; \epsilon - \eta + \eta \;=\; \epsilon$$

hence $y \in B_\epsilon(a) \quad \forall\, y \in B_{\epsilon-\eta}(x)$



Definition of continuity:

Let $f\,:\,X \to Y$ be a mapping (function) between metric spaces. f is called "continuous" in the point $a \in X$ if for each neighborhood V in Y of the pixel $f(a)$, a neighborhood U of the original image point $a \in X$ exists with

$$f(U) \subset V$$

(specifically $V = B_\epsilon(f(a))\,,\; B_d(a) \subset U$)

$$f(B_d(a)) \;\subset\; B_\epsilon(f(a)) \;, \quad \text{that means}$$

for x with $|x,a| < \delta$, $|f(x),f(a)| < \epsilon$. f is called continuous on X, if f is continuous at every point $a \in X$.

Examples:

Every contracting mapping $f:\,X \to X$ is continuous.

Every linear mapping $L:\,\mathbb{R}^n \to \mathbb{R}^m$ is continuous.

Proof:

f "contracting", if a $\vartheta \in \,]0,1[$ exists with

$$|f(x),f(y)| \;\leq\; \vartheta |x,y| \qquad \forall\, x,y \in X$$

L "linear", if

$$\left.\begin{array}{rcl} L(\vec{x}+\vec{y}) & = & L(\vec{x}) \;+\; L(\vec{x}) \\ L(\lambda\vec{x}) & = & \lambda L(\vec{x}) \end{array}\right\} \qquad \begin{array}{l} \forall\,\vec{x},\vec{y}\in\mathbb{R}^n \\ \forall\,\lambda\in\mathbb{R} \end{array}$$

If $\epsilon > 0$ and $|x,a| < \epsilon\,,$ then

$$|f(x),f(a)| \;\le\; \vartheta|x,a| \;\le\; \vartheta\epsilon \;<\; \epsilon$$

that means

$$f(B_\epsilon(a)) \,\subset\, B_\epsilon(f(a)) \qquad \forall\,\epsilon > 0$$

$$\begin{aligned} \Big\| L\Big(\sum_{i=1}^n \lambda_i\vec{e}_i - \underbrace{L(\vec{0})}_{=\vec{0}}\Big)\Big\| \;&=\; \Big\|\Big(\sum_{i=1}^n \lambda_i\vec{e}_i\Big)\Big\| \\ &\le\; \sum_{i=1}^n |\lambda_i|\|L(\vec{e}_i)\| \;\le\; M\sum_{i=1}^n|\lambda_i| \\ &\le\; Mn\Big(\sum_{i=1}^n \lambda_i^2\Big)^{1/2} \quad ; \quad M \coloneqq \max_{1\le i\le n}\|L(\vec{e}_i)\| \end{aligned}$$

$$\|L(\vec{x}) - L(\vec{a})\| \;=\; \|L(\vec{x}-\vec{a})\| \;\le\; Mn\|\vec{x}-\vec{a}\|$$

□

> The composite of two continuous mappings is again continuous.

Let $f : X \to Y\,,\; g : Y \to Z$ be mappings between metric spaces X, Y, Z. If f is continuous in $a \in X$ and g in $b = f(a) \in Y$, then $g \circ f$ is continuous in $a \in X$.

Proof:

Let $c \coloneqq g\,\underbrace{(f(a))}_{b}$ and $W \subset Z$ be a neighborhood of c.

Because g is continuous in b, a neighborhood V of b exists with

$$g(V) \subset W$$

Because f is continuous in a, a neighborhood U of a exists with

$$f(U) \subset V$$

Application of g:

$$g\circ f(U) \;=\; g(f(U)) \;\subset\; g(V) \;\subset\; W$$

with respect to a, U has the property: $\quad (g \circ f)(U) \;\subset\; W.$

$\square$

Example:

(5) $X = Y = \mathbb{C}, \quad Z = \mathbb{R}$

$$f(w) \;=\; w^3 \quad \text{is continuous}$$

$$g(z) \;=\; \mathrm{Re}(z) \quad \text{is continuous (because } \mathbb{R} \text{ linear)}$$

Therefore $g \circ f(z) \;=\; \mathrm{Re}(z^3)$ is also continuous (the graph is a so-called monkey saddle).

Definition:

Let X be a metric space, $a_k \in X$ $(k \in \mathbb{N}_0)$. The sequence $(a_k)_{k \in \mathbb{N}_0}$ is called "convergent" to $a \in X$ $(\lim\limits_{k \to \infty} a_k = a)$ if

$$\lim_{k \to \infty} |a_k, a| \;=\; 0$$

Expressed with neighborhoods!

In every neighborhood U of a exists an index N_U with $a_k \in U$ for all $k \geq N_U$.

Theorem:

If $(a_k)_{k \geq 0}$ is a convergent sequence on the metric space X with limit a and if $f : X \to Y$ is a continuous mapping between metric spaces, then in Y

$$\lim_{k \to \infty} f(a_k) \;=\; f(a)$$

Proof:

Let $V = B_\epsilon(f(a))$ be a neighborhood of $f(a)$. Since f is continuous, there exists a neighborhood $U = B_\delta(a)$ with $f(U) \subset V$

$(a_k)_{k \geq 0}$ converges to a, so an index bound N_U exists with $a_k \in U \;\; \forall\, k \geq N_U$, hence

$$f(a_k) \;\in\; f(U) \;\subset\; V \qquad \forall\, k \geq N_U$$

$\square$

Definition:

A sequence $(c_k)_{k\geq 0}$ on the metric space X is called a "Cauchy sequence" if for every $\epsilon > 0$ a N_ϵ exists with

$$|c_k, c_l| \;\leq\; \epsilon \qquad \forall\, k, l \geq N_\epsilon$$

The space X is called complete if every Cauchy sequence on X converges.

Example:

(6) With the Euclidean metric $\mathbb{R}^n$ is complete ($n = 1, 2, 3$). For $n = 1$ this is the basic insight of analysis. Now let n be arbitrary. Furthermore, let $(c_k)_{k\geq 0}$ be a Cauchy sequence on $\mathbb{R}^n$; and let r be one of the coordinate indices ($1 \leq r \leq n$)

$$|c_{k,r} - c_{l,r}| \;\leq\; \|c_k - c_l\| \quad \text{(inequality for the coordinate differences)}$$

Thus for each r, as a Cauchy sequence the real sequence $({}_{k,r})_{k\geq 0}$ is in $\mathbb{R}$ convergent!

For example, like $\quad \gamma_r \;:=\; \lim\limits_{k\to\infty} c_{k,r}$

$$\text{Assertion:} \qquad \lim_{k\to\infty} c_k \;=\; \gamma \qquad \text{where} \qquad \gamma \;=\; \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{pmatrix}$$

$$\lim_{k\to\infty} \|c_k - \gamma\| \;=\; \lim_{k\to\infty} \sqrt{\sum_{r=1}^{n} (c_{k,r} - \gamma_r)^2} \;=\; 0$$

6.1 Fixed Point Theorem

Let X be a completely metric space, $f : X \to X$ is a contracting mapping. Then: f has exactly one fixed point x_∞ in X (that means $f(x_\infty) = x_\infty$).

Proof:

because f contracted, $\;\vartheta \in\,]0, 1[\;$ exists with $\;|f(x), f(y)| \;\leq\; \vartheta |x, y|$

Assumption: $\;f(x_\infty) \;=\; x_\infty\;$ and $\;f(a) \;=\; a$

$$|x, a| \;=\; |f(x_\infty), f(a)| \;\leq\; \vartheta |x_{\infty,a}|$$

that means $\;(1 - \vartheta)|x_\infty, a| \;\leq\; 0\;$ although $\;\vartheta \;<\; 1$

hence $|x_\infty, a| = 0$, so according to M1, $x_\infty = a$

Existence of the fixed point

Let $a_0 \in X$ be arbitrary. Recursive definition

$$a_{n+1} = f(a_n)$$

$(a_n)_{n\geq 0}$ is a Cauchy sequence:

$$|a_{r+1}, a_r| \leq \vartheta^r |a_0, a_1| \qquad \text{(complete induction)}$$

$$|a_n, a_{n+p}| \leq \sum_{r=n}^{n+p-1} |a_r, a_{r+1}| \qquad \text{(triangle inequality)}$$

$$\leq |a_0, a_1| \sum_{r=n}^{n+p-1} \vartheta^r \leq \underbrace{\frac{|a_0, a_1|\vartheta^n}{1-\vartheta}}_{\to 0} \qquad \text{(geometric series)}$$

Hence, as a Cauchy sequence, $(a_k)_{k\geq 0}$ is convergent: There exists $x_\infty = \lim\limits_{k\to\infty} a_k$

From $a_{k+1} = f(a_k)$ and the continuity of f, follows through transition to the limit

$$x_\infty = \lim_{k\to\infty} a_{k+1} = f\left(\lim_{k\to\infty} a_k\right) = f(x_\infty)$$

□

Definition:

Let $f_k : X \to Y$ be mappings (functions) between metric spaces X, Y ($k \geq 0$). $(f_k)_{k\geq 0}$ is called uniformly convergent, if first for each $x \in X$ the point sequence

$$(f_k(x))_{k\geq 0}$$

converges in Y and secondly (with $f(x) = \lim\limits_{k\to\infty} f_k(x)$ as the limit designation) for every $\epsilon > 0$ an index N_ϵ exists with

$$|f_k(x), f(x)| \leq \epsilon \qquad \forall\, k \geq N_\epsilon \;\; \forall\, x \in X$$

Uniform limits of continuous function sequences are continuous!

Proof:

Double application of the triangle inequality ($x, a \in X$, $N \geq 0$)

$$|f(x), f(a)| \;\leq\; |f(x), f_N(x)| + |f_N(x), f_N(a)| + |f_N(a), f(a)|$$

Let $\epsilon > 0$, choose $N = N_{\epsilon/3}$

$$|f(x), f(a)| \;\leq\; \frac{2}{3}\,\epsilon + |f_N(x), f_N(a)|$$

because f_N is continuous by assumption, $\delta > 0$ exists with $|f_N(x), f_N(a)| < {}^{\epsilon}/_{3}$ if $|x, a| < \delta$.

Together: If $|x, a| < \delta$ then

$$|f(x), f(a)| \;<\; \frac{2}{3}\,\epsilon + \frac{\epsilon}{3} \;=\; \epsilon$$

$\square$

7. Partial and Total Derivative

Mappings (functions) of subsets $U \subset \mathbb{R}^n$ with values in $\mathbb{R}$ are considered with respect to differentiability. Instead of the intervals I in $\mathbb{R}$, the "regions" U in $\mathbb{R}^n$ appear here. These regions are open, nonempty subsets in which two points can be connected by a curve in U.

For functions $f : U \to \mathbb{R}$, the graph as a subset of $\mathbb{R}^{n+1}$ becomes

$$\Gamma_f \ := \ \big\{(x_1, \ldots, x_n, y)\,;\, (x_1, \ldots, x_n) \in U\,,\ \ y = f(x_1, \ldots, x_n)\big\}$$

and the level curves

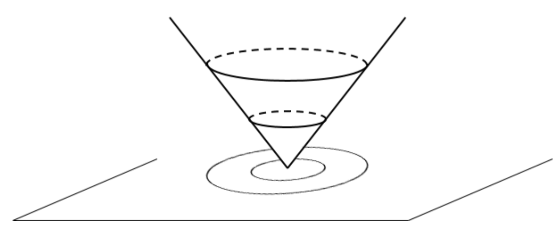
$$\{\vec{x} \in U :\ f(\vec{x}) = c\}\,, \qquad c \ = \ \text{constant}$$

Best illustration for $n = 2$; then the graph f is a kind of mountain range over the set of planes U.

<u>Example:</u>

(0) $n \ = \ 2\,, \quad U \ = \ \mathbb{R}^2$

$$f(\vec{x}) \ = \ \|\vec{x}\|$$



7.1 Definition of the Partial Derivative

Let $f : U \to \mathbb{R}$ be a real function in the region U of $\mathbb{R}^n$. The k-th partial derivative in $\vec{x} \in U$ is declared (in case of existence) as the limit

$$D_k f(\vec{x}) \ = \ \lim_{h \to 0} \frac{f(\ldots x_k + h \ldots) - f(\ldots x_k \ldots)}{h}$$

Other nomenclatures:

$$\frac{\partial f}{\partial x_k}(\vec{x}) \ \ ; \quad f_{x_k}(\vec{x}) \ \ ; \quad f'_{x_k}(\vec{x})$$

If f is differentiable at point $\vec{x}$ in all n variables, then the vector

$$\left(\frac{\partial f}{\partial x_1}(\vec{x}), \ldots\ldots, \frac{\partial f}{\partial x_n}(\vec{x})\right) \;=\; \mathrm{grad} f$$

is called the "gradient" of f in $\vec{x}$. With the symbolic Nabla vector

$$\left(\frac{\partial}{\partial x_1}, \ldots\ldots, \frac{\partial}{\partial x_n}\right) \;=\; \boldsymbol{\nabla}$$

$$\mathrm{grad} f(\vec{x}) \;=\; \boldsymbol{\nabla} f \;=\; \left(\frac{\partial f}{\partial x_1}(\vec{x}), \ldots\ldots, \frac{\partial f}{\partial x_n}(\vec{x})\right)$$

Remarks:

(1) Every partial derivative is nothing more than an ordinary derivative with respect to a real variable. The other variables are regarded as constant!

(2) If f has partial derivatives everywhere on U with respect to $x_1, \ldots, x_n$ then the gradient returns a vector field $\mathrm{grad} f : U \to \mathbb{R}^n$.

Back to example (0):

$$\vec{x} \;\longmapsto\; \|\vec{x}\| \;=\; \Big(\sum_{i=1}^{n} x_i^2\Big)^{1/2}$$

is differentiable on $\quad U \;=\; \mathbb{R}^n \setminus \{\vec{0}\}$

$$\frac{\partial \|\vec{x}\|}{\partial x_k} \;=\; \frac{2x_k}{2\Big(\sum_{i=1}^{n} x_i^2\Big)^{1/2}} \qquad ; \qquad \mathrm{grad}\|\vec{x}\| \;=\; \|\vec{x}\|^{-1}\vec{x}$$

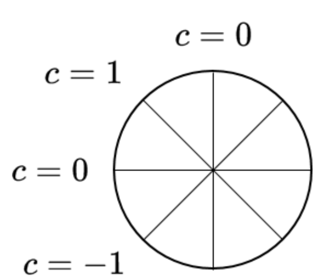
Example:

(1) Cylindroid $\quad n \;=\; 2, \quad U \;=\; \{\vec{x} \in \mathbb{R};\, \|\vec{x}\| < 1\}$

$$z \;=\; f(x,y) \;=\; \begin{cases} \dfrac{2xy}{x^2+y^2} & \text{if not} \quad x = y = 0 \\[2ex] 0 & \text{if} \quad x = y = 0 \end{cases}$$

Level curves:

$\{\vec{x},\, f(\vec{x}) \;=\; c\}$



The partial derivatives

$$\vec{x} \neq \vec{0}: \qquad \frac{\partial f}{\partial x} = \frac{2y(x^2+y^2) - 4x^2y}{(x^2+y^2)^2} = \frac{2y(-x^2+y^2)}{(x^2+y^2)^2}$$

$$\frac{\partial f}{\partial y} = \frac{2x(x^2+y^2) - 4xy^2}{(x^2+y^2)^2} = \frac{2x(x^2-y^2)}{(x^2+y^2)^2}$$

$$\vec{x} = \vec{0}: \qquad \frac{\partial f}{\partial x}(\vec{0}) = 0 \quad ; \quad \frac{\partial f}{\partial y}(\vec{0}) = 0$$

Remark:

(3) Although the function f for the cylindroid has partial derivatives everywhere with respect to both variables, f is discontinuous at the zero point!

Theorem:

Let f be real and partially differentiable with respect to all variables in the region U of $\mathbb{R}^n$. Moreover, let the partial derivatives $\dfrac{\partial f}{\partial x_k}$ be bounded on U for all k. Then f is continuous at every point $\vec{a} \in U$.

Proof: The ordinary mean value theorem

$$f(\vec{a}+\vec{h}) - f(\vec{a}) = f(a_1+h_1, \ldots, a_n+h_n) - f(a_1, \ldots, a_n)$$

becomes with insertions

$$= f(a_1+h_1, \ldots, a_n+h_n) - f(a_1, a_2+h_2, \ldots) + f(a_1, a_2+h_2, \ldots) - \ldots$$

$$+ \ldots - f(a_1, a_2, \ldots, a_n+h_n) + f(a_1, a_2, \ldots, a_n+h_n) - f(a_1, \ldots, a_n)$$

$$= h_1 \frac{\partial f}{\partial x_1}(\vec{a}+\vec{t}_1) + h_2 \frac{\partial f}{\partial x_2}(\vec{a}+\vec{t}_2) + \ldots + h_n \frac{\partial f}{\partial x_n}(\vec{a}+\vec{t}_n)$$

Therefore, the assumption about the $\dfrac{\partial f}{\partial x_k}$ results in

$$\left| f(\vec{a}+\vec{h}) - f(\vec{a}) \right| = \left\| \vec{h} \right\| M\, n$$

In particular, the continuity of f in $\vec{a}$ follows.

$\square$

Remark:

(4) Continuous functions are locally bounded! Because if g is continuous in $\vec{a}$, then there exists a neighborhood U of $\vec{a}$ with

$$|g(\vec{x}) - g(\vec{a})| \leq 1\,, \quad \text{i.e.}$$

$$|g(\vec{x})| \;\leq\; |g(\vec{a})| + 1$$

Consequence:

If, with the assumption of the theorem, each of the partial derivatives $\dfrac{\partial f}{\partial x_k}$ as a function on U is continuous, then f itself is continuous on U.

One calls a vector field $\vec{h} : G \to \mathbb{R}^n$ in the region G of $\mathbb{R}^n$ "conservative" if it is the gradient of a function $-V : G \to \mathbb{R}^n$ (V is called the potential of $\vec{h}$). For continuous conservative vector fields $\vec{h} : G \to \mathbb{R}^n$ and any curves $\gamma : [a, b] \to G$:

$$\int_\gamma \vec{h}(\vec{x})\,\mathrm{d}\vec{x} \;=\; -\int_\gamma \operatorname{grad} V(\vec{x})\,\mathrm{d}\vec{x} \;=\; V(\gamma(a)) - V(\gamma(b))$$

The line integral over the (negative) gradient of a potential V is equal to the potential difference at the end points of the curve!

7.1.1 Generalized Chain Rule

Given is a real function $V : G \to \mathbb{R}^n$ on the region G of $\mathbb{R}^n$, which is partially differentiable in all variables and the partial derivatives are continuous!

Then for every curve $\gamma : [a, b] \to G$ with the composite function

$$F(t) \;:=\; V(\gamma(t))$$

$$\dot{F}(t) \;=\; \operatorname{grad} V(\gamma(t))\,\dot{\gamma}(t) \;=\; \sum_{i=1}^{n} \frac{\partial V}{\partial x_i}\,\dot{\gamma}_i(t)$$

Every continuously partially differentiable function can be linearly approximated. Let f be a continuously partially differentiable function on the region $G \subset \mathbb{R}^n$. Then for $\vec{x}$, $\vec{x} + \vec{y} \in G$

$$f(\vec{x} + \vec{y}) - f(\vec{x}) \;=\; \operatorname{grad} f(\vec{x})\vec{y} + \|\vec{y}\|\, r(\vec{y})$$

$$\text{with} \quad \lim_{\vec{y} \to \vec{0}} r(\vec{y}) \;=\; 0$$

Proof: of the approximation formula

$f(\vec{x} + \vec{y}) - f(\vec{x})$ becomes with insertions

$$= f(x_1 + y_1, \ldots, x_n + y_n) - f(x_1, x_2 + y_1, \ldots) \;\ldots$$

$$\pm \ldots + f(x_1, \ldots, x_{n-1}, x_n + y_n) - f(x_1, \ldots, x_n)$$

with the mean value theorem

$$= \frac{\partial f}{\partial x_1}(\vec{x}+\vec{w}_1)\vec{y}_1 \ + \ \ldots \ + \ \frac{\partial f}{\partial x_n}(\vec{x}+\vec{w}_n)\vec{y}_n$$

Hence the rest has the explicit form

$$\begin{aligned} r(\vec{y}) &= \left(\frac{\partial f}{\partial x_1}(\vec{x}+\vec{w}_1) - \frac{\partial f}{\partial x_1}(\vec{x})\right)\frac{y_1}{\|\vec{y}\|} + \ldots \\ &\quad + \left(\frac{\partial f}{\partial x_n}(\vec{x}+\vec{w}_n) - \frac{\partial f}{\partial x_n}(\vec{x})\right)\frac{y_n}{\|\vec{y}\|} \end{aligned}$$

because of the continuity of the partial derivatives at the point $\vec{x}$ it follows:

$$\lim_{\vec{y}\to\vec{0}} r(\vec{y}) = 0$$

□

Proof: of the chain rule

$$\frac{F(t+h)-F(t)}{h} = \frac{V(\gamma(t+h))-V(\gamma(t))}{h}$$

(From the approximation formula with $\vec{x} = \gamma(t)$ and $\vec{y} = \gamma(t+h)-\gamma(t)$)

$$\begin{aligned} &= \frac{1}{h}\operatorname{grad} V(\gamma(t))\,(\gamma(t+h)-\gamma(t)) + \|\gamma(t+h)-\gamma(t)\|\, r(\gamma(t+h)-\gamma(t)) \\ &= \operatorname{grad} V(\gamma(t))\frac{\gamma(t+h)-\gamma(t)}{h} + \left\|\frac{\gamma(t+h)-\gamma(t)}{h}\right\| r(\gamma(t+h)-\gamma(t)) \end{aligned}$$

From this follows the chain rule.

□

Proof of the statement about the conservative vector fields

$$\int_\gamma -\operatorname{grad} V(\vec{x})\,\mathrm{d}\vec{x} \underset{\substack{\downarrow \\ \text{per definition}}}{=} \int_a^b -\operatorname{grad} V(\gamma(t))\,\dot{\gamma}(t)\,\mathrm{d}t$$

According to the chain rule, the integrand is the derivative with respect to t of the function $-V(\gamma(t))$. Value of the integral therefore:

$$\int_a^b -\operatorname{grad} V(\gamma(t))\,\dot{\gamma}(t)\,\mathrm{d}t = -V(\gamma(b)) + V(\gamma(a))$$

□

<u>Examples:</u> All central fields

$$\vec{h}(\vec{x}) = \lambda(\|\vec{x}\|)\,\vec{x}\ , \qquad \vec{x}\in\mathbb{R}^n,\ \vec{x}\neq\vec{0}$$

are conservative if $\lambda :\ ]a,b[\ \to \mathbb{R}$ is continuous.

Proof: Consider the auxiliary function

$$g(s) \;=\; \int_1^s t\lambda(t)\,\mathrm{d}t \qquad s \in \,]0,\infty[$$

With $V(\vec{x}) \coloneqq -g(\|\vec{x}\|)$ follows

$$\frac{\partial V(\vec{x})}{\partial x_i} \;=\; -g'(\|\vec{x}\|)\,\frac{\partial}{\partial x_i}\|\vec{x}\|$$

and hence with the fundamental theorem of calculus

$$\frac{\partial V(\vec{x})}{\partial x_i} = \; -\|\vec{x}\|\,\lambda(\|\vec{x}\|)\,\frac{x_i}{\|\vec{x}\|}$$

$$\operatorname{grad} V(\vec{x}) \;=\; -\lambda\,(\|\vec{x}\|\,)\,\vec{x}$$

Like $n = 3$, $\vec{h}(\vec{x}) \;=\; -\dfrac{\alpha}{\|\vec{x}\|^3}\,\vec{x}$ (gravitational field, $\lambda(t) \;=\; -\alpha t^{-3}$)

$$g(s) \;=\; \alpha\int_1^2 t^{-2}\,\mathrm{d}t \;=\; \frac{\alpha}{s} + \text{const.}$$

$$V(\vec{x}) \;=\; \frac{-\alpha}{\|\vec{x}\|} \qquad \text{after normalization} \qquad V(\vec{x}\to\infty) \;=\; 0$$

7.2 Definition of the Total Derivative

Conceptual remark regarding the total derivative over linear forms of $\mathbb{R}^n$.

One asks about the linear maps $L:\,\mathbb{R}^n \to \mathbb{R}$ (also called linear forms). They are uniquely determined by the n values $L(\vec{e}_i) \;=\; a_i$ $(1 \le i \le n)$. Then

$$\begin{aligned} L(\vec{x}) \;&=\; L(x_1\vec{e}_1 \,+\, x_2\vec{e}_2 \,+\, \ldots \,+\, x_n\vec{e}_n) \\ &=\; a_1x_1 \,+\, a_2x_2 \,+\, \ldots \,+\, a_nx_n \;=\; \vec{a}\,\vec{x} \end{aligned}$$

Definition:

Let f be a real-valued function on the region G of $\mathbb{R}^n$. f is called "totally differentiable" (linearly approximable) in $\vec{x}\in G$, if with a linear form (linear mapping) $L:\,\mathbb{R}^n\to\mathbb{R}$ the following holds:

$$f(\vec{x}+\vec{h}) \,-\, f(\vec{x}) \;=\; L(\vec{h}) \,+\, \left\|\vec{h}\right\| r(\vec{h}) \qquad \text{where} \qquad \lim_{\vec{h}\to\vec{0}} r(\vec{h}) \;=\; 0$$

Remarks:

(1) If f in $\vec{x}$ totally differentiable, then f in $\vec{x}$ has partial derivatives with respect to all variables: Because then

$$\frac{f(\vec{x}+t\vec{e}_k)-f(\vec{x})}{t} \;=\; \frac{L(t\vec{e}_k)}{t} + \operatorname{sign}(t)\, r(t\vec{e}_k) \;=\; L(\vec{e}_k)$$

(2) If f in $\vec{x} \in G$ totally differentiable, then for every unit vector $\vec{e}$ there exists the "directional derivative"

$$D_{\vec{e}} f(\vec{x}) \;=\; \lim_{t\to 0} \frac{f(\vec{x}+t\vec{e})-f(\vec{x})}{t} \;=\; L(\vec{e})$$

(3) If f is differentiable in all points of G with respect to all variables and the partial derivatives are continuous, then f is differentiable in all points (compare chain rule).

Example:

(1) The sine cone $\;n = 2\,,\quad G \;=\; \{(\vec{x},\vec{y}) \in \mathbb{R}^2;\, x^2+y^2<1\}$

$$z \;=\; f(x,y) \;=\; \left\{ \begin{array}{cl} \dfrac{2xy}{\sqrt{x^2+y^2}} & \text{if} \quad x^2+y^2>0 \\ 0 & \text{if} \quad x=y=0 \end{array} \right\} \;=\; r\sin 2\varphi$$

$$\frac{\partial f}{\partial x}(0,0) \;=\; 0 \quad , \quad \frac{\partial f}{\partial y}(0,0) \;=\; 0 \qquad \text{otherwise}$$

$$\frac{\partial f}{\partial x} \;=\; \frac{2y(x^2+y^2)-2x^2y}{(x^2+y^2)^{3/2}} \;=\; \frac{2y^3}{(x^2+y^2)^{3/2}} \;=\; 2\sin^3\varphi$$

$$\frac{\partial f}{\partial y} \;=\; \frac{2x(x^2+y^2)-2xy^2}{(x^2+y^2)^{3/2}} \;=\; \frac{2x^3}{(x^2+y^2)^{3/2}} \;=\; 2\cos^3\varphi$$

f is not totally differentiable at the origin (reason: the directional derivatives do not always exist)

$$\vec{e} \;=\; (c,s) \qquad (c^2 \,+\, s^2 \;=\; 1)$$

$$\frac{f(ct,st)-f(0,0)}{t} \;=\; \frac{2cst^2}{t|t|} \;=\; 2cs\operatorname{sign} t$$

Therefore, (because of $\operatorname{sign} t$) the limit $\displaystyle \underset{t\neq 0}{t \to 0}$ does not exist!

7.2.1 Geometric Properties of the Total Derivative

1. If f is totally differentiable in $\vec{x}$, then by remark (1) the linear form L is unique, namely

$$L(\vec{h}) \;=\; \operatorname{grad} f(\vec{x})\, \vec{h} \qquad (\vec{h} \longmapsto L(\vec{h})) \qquad \text{because} \quad \frac{\partial f}{\partial x_k} \;=\; L(\vec{e}_k)$$

The gradient points in the direction of maximum growth of f. In particular, the gradient vanishes in local extrema! The level curves ($n = 2$) or level surfaces (if $n = 3$) run perpendicular to the gradient.

2. Examining the graph of f

$$\Gamma_f \;=\; \{\vec{x}\,,\, f(\vec{x})\,,\, \vec{x} \in G\}$$

The curves $\gamma :\, [a, b] \to G$ can be lifted to Γ_f

$$\tilde{\gamma}(t) \;=\; \big(\gamma(t)\,,\, f(\gamma(t))\big) \qquad (n = 2)$$

This curve has the derivative

$$\dot{\tilde{\gamma}}(t) \;=\; \big(\dot{\gamma}_1(t)\,,\, \dot{\gamma}_2(t)\,,\, \operatorname{grad} f(\gamma(t))\, \dot{\gamma}(t)\big)$$

Tangent at a curve point

$$\vec{x}_0 \;=\; (x_0, y_0, z_0) \;=\; \big(\gamma_1(t_0)\,,\, \gamma_2(t_0)\,,\, f(\gamma(t_0))\big)$$

$$\vec{x}_0 + \mathbb{R}\, \dot{\tilde{\gamma}}(t_0) \;=\; \vec{x}_0 + \mathbb{R}\, \big(\dot{\gamma}_1(t_0)\,,\, \dot{\gamma}_2(t_0)\,,\, \operatorname{grad} f\big(\gamma(t_0)\big)\, \dot{\gamma}(t_0)\big)$$

The tangents for each of the curves $\tilde{\gamma}$ running through $\vec{x}_0$ lie on the plane

$$(z - z_0) \;=\; \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \,+\, \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

This is the "tangent plane" at Γ_f at the point $\vec{x}_0$. Normal vector:

$$\vec{n} \;=\; \left(\frac{\partial f}{\partial x}\,,\, \frac{\partial f}{\partial y}\,,\, -1\right)$$

In general $\vec{n} \;=\; \vec{n}(x_0, y_0, z_0)$ is not normalized.

Example:

(2) $n = 2\,,\; G = \mathbb{R}^2\,;$ Let the function $q(x, y)$ be the square of the distance products of the point (x, y) from the fixed points $(-1, 0)$, $(1, 0)$

$$\begin{aligned} q(x,y) &= \big((x+1)^2+y^2\big)\big((x-1)^2+y^2\big) \\ &= \big((x^2+y^2+1)+2x\big)\big((x^2+y^2+1)-2x\big) \\ &= (x^2+y^2+1)^2+4x^2 \\ &= x^4+2x^2y^2+y^4-2x^2+2y^2+1 \end{aligned}$$

Level curves are the Cassini curves. Particularly named among them is the curve $q(x,y)=1$, the lemniscate!

A set of curves orthogonal to the Cassini curves exists in the direction of the gradients, which, in addition to the coordinate axes, contains nothing but hyperbolas with vertical asymptotes and whose left (or respectively right) branches go through $(-1,0)$ (or respectively $(1,0)$).

$$x^2 \;-\; y^2 \;-\; 1 \;\;=\;\; 2cxy$$

Let $x(t)$, $y(t)$ be a parameterization of this curve. Differentiation:

$$2x\dot{x} \;-\; 2y\dot{y} \;\;=\;\; 2c\dot{x}y \;+\; 2cx\dot{y}$$

Elimination of c yields

$$\begin{aligned} \dot{x}y(x^2+y^2+1) &= \dot{y}x(x^2+y^2-1) \\ \operatorname{grad} q &= (4x^3+4xy^2-4x, 4y^3+4x^2y+4y) \\ &= 4(x(x^2+y^2-1), y(x^2+y^2+1)) \end{aligned}$$

8. Higher Derivatives, Taylor Formula and Local Extrema

8.1 The Symmetry of the Second Derivative

Let f be a real function in the region $G \subset \mathbb{R}^n$, which is twice continuously (partially) differentiable (the partial derivatives are thus again continuously partially differentiable). Then the commutation rule

$$D_i(D_k f) \;=\; D_k(D_i f) \qquad (1 \le i, k \le n)$$

applies for differentiation operators!

Proof idea:

1) It is sufficient to look at $n = 2$.

2) For polynomials $x^k y^l$ applies

$$\begin{aligned} \frac{\partial}{\partial y}\frac{\partial}{\partial x}\big(x^k y^l\big) \;&=\; klx^{k-1}y^{l-1} \\ &=\; \frac{\partial}{\partial x}\frac{\partial}{\partial y}\big(x^k y^l\big) \;=\; klx^{k-1}y^{l-1} \end{aligned}$$

3) Generally consider

$$f(x+h, y+h) \,-\, f(x, y+h) \,-\, f(x+h, y) \,+\, f(x, y)$$

Multiple application of the mean value theorem $\;\Rightarrow\;$ assertion.

Example:

$$(1) \qquad f(x,y) \;=\; \left\{ \begin{array}{cl} xy\,\dfrac{x^2-y^2}{x^2+y^2} & \text{if} \quad x^2+y^2 \neq 0 \\ 0 & \text{if} \quad x = y = 0 \end{array} \right\} \;=\; \frac{r^2}{4}\sin 4\varphi$$

Partial derivatives

$$\frac{\partial f}{\partial x}(0,0) \;=\; 0 \quad , \quad \frac{\partial f}{\partial y}(0,0) \;=\; 0 \qquad \text{otherwise}$$

$$\frac{\partial f}{\partial x} \;=\; r^{-4}\big((3x^2y - y^3)(x^2+y^2) \;-\; (x^3y + xy^3)2x\big) \;=\; \frac{x^4y + 4x^2y^3 - y^5}{r^4}$$

$$\frac{\partial f}{\partial y} \;=\; r^{-4}\big((x^3 - 3y^2x)(x^2+y^2) \;-\; (x^3y - xy^3)2y\big) \;=\; \frac{x^5 - 4x^3y^2 - xy^4}{r^4}$$

$$\frac{\partial}{\partial x}\frac{\partial f}{\partial y}(0,0) \;=\; 1 \quad , \quad \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(0,0) \;=\; -1$$

8.1.1 Integrability Criterion for Vector Fields

Let a continuously partially differentiable vector field be given in the region $G \subset \mathbb{R}^n$ (this means that the coordinates v_i shall be continuously differentiable on G)

$$\vec{v} :\; G \to \mathbb{R}^n$$

$\vec{v} = \boldsymbol{\nabla}\, V$ is the gradient of a potential V only if

$$\frac{\partial v_i}{\partial x_k} \;=\; \frac{\partial v_k}{\partial v_i} \;, \qquad 1 \le i,k \le n \qquad (*)$$

Let G be star-shaped. Conversely, if $(*)$ is satisfied and the region G is star-shaped, then there is a potential V on G with $\vec{v} = \operatorname{grad} V$.

Proof:

1) If $\vec{v} = \operatorname{grad} V$, then according to the premise about $\vec{v}$, the potential V is twice continuously partially differentiable. Because of the symmetry of the second derivative

$$\frac{\partial}{\partial x_i}\underbrace{\frac{\partial}{\partial x_k}V}_{v_k} \;=\; \frac{\partial}{\partial x_k}\underbrace{\frac{\partial}{\partial x_i}V}_{v_i}$$

2) Let G be star-shaped and the criterion $(*)$ for $\vec{v}$ fulfilled. Without restriction let $\vec{0}$ be the star point of G, that means that for all $\vec{x} \in G$ the distance between $\vec{x}$ and $\vec{0}$ lies within G.

$$V(\vec{x}) \;=\; \int_0^1 \big(\vec{v}(t\vec{x})\,\vec{x}\big)\,\mathrm{d}t \qquad \text{(integral with parameter } \vec{x} \in G\text{)}$$

With a theorem about integrals with parameters

$$\frac{\partial V}{\partial x_i} \;=\; \int_0^1 \frac{\partial}{\partial x_i}\big(\vec{v}(t\vec{x})\,\vec{x}\big)\,\mathrm{d}t$$

$$= \int_0^1 \frac{\partial}{\partial x_i}\Big(\sum_{k=1}^n v_k(t\vec{x})\,x_k\Big)\,\mathrm{d}t$$

$$= \int_0^1 \Big(v_i(t\vec{x}) + \sum_{k=1}^n t\frac{\partial v_k}{\partial x_i}(t\vec{x})\,x_k\Big)\,\mathrm{d}t \qquad \text{chain rule}$$

$$= \int_0^1 \Big(v_i(t\vec{x}) + \sum_{k=1}^n t\frac{\partial v_i}{\partial x_k}(t\vec{x})\,x_k\Big)\,\mathrm{d}t$$

$$= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t}\big(tv_i(t\vec{x})\big)\,\mathrm{d}t \qquad \text{generalized chain rule}$$

$$= v_i(\vec{x}) - 0 = v_i(\vec{x}) \qquad (1 \le i \le n)$$

□

Remark:

In the case $n = 3$: integrability criterion in the vector field

$$\left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}\right) = \vec{0}$$

Symbolic notation

$$\mathrm{rot}\,\vec{v} = \boldsymbol{\nabla} \times \mathbf{v} = \vec{0}$$

In words: In the region G of $\mathbb{R}^3$, which is star-shaped, a vector field $\vec{v} : G \to \mathbb{R}^3$ (which is continuously partially differentiable) becomes then and only then the gradient of a potential V if $\mathrm{rot}\,\vec{v} = 0$!

Example:

(2) Isolated vortex, $n = 2$ $\quad G = \mathbb{R}^2 \setminus \vec{0}$

$$\vec{v}(x,y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2}\right) = \frac{1}{r}(-\sin\varphi, \cos\varphi)$$

$$\frac{\partial v_1}{\partial y} = \frac{-(x^2+y^2)+2y^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

$$\frac{\partial v_2}{\partial y} = \frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

Calculation of a line integral along the unit curve

$$\gamma(t) = (\cos t, \sin t) \qquad I = [0, 2\pi]$$

The line integral over the vector field along the unit curve is

$$\int_\gamma \vec{v}(\vec{x})\,\mathrm{d}\vec{x} \;=\; \int_0^{2\pi} \vec{v}(\gamma(t))\,\dot{\gamma}(t)\,\mathrm{d}t$$

$$\int_0^{2\pi} \underbrace{(-\sin t, \cos t)(-\sin t, \cos t)}_{\sin^2 t + \cos^2 t = 1}\,\mathrm{d}t \;=\; 2\pi$$

Therefore $\vec{v}$ is not the gradient of a potential in this case, although the criterion $(*)$ is fulfilled. Reason: $\mathbb{R}^2 \setminus \vec{0}$ is not star-shaped.

From the symmetry of the second derivative for twice continuously partially differentiable functions $f : G \to \mathbb{R}^n$ follows:

The “Hesse matrix” of f

$$A \;=\; \left(\frac{\partial^2 f}{\partial x_i \partial x_k}\right)_{1 \le i,k \le n} \;=\; A^T$$

is symmetric in every point $\vec{a}$ of the region G. Under these conditions, a simple version of the Taylor formula applies in $\mathbb{R}^n$.

8.2 A Simple Version of the Taylor Formula in $\mathbb{R}^n$

$$\boxed{f(\vec{a}+\vec{x}) \;=\; f(\vec{a}) \,+\, \operatorname{grad} f(\vec{a})\,\vec{x} \,+\, \frac{1}{2}\,\vec{x}\,A\,\vec{x} \,+\, \varphi_2(\vec{x})}$$

where A is the Hessian matrix of f at the point $\vec{a}$ and where the remainder function φ_2 satisfies the limes designation

$$\lim_{\substack{\vec{x} \to \vec{0} \\ \vec{x} \neq \vec{0}}} \|\vec{x}\|^{-2} \varphi_2(\vec{x}) \;=\; 0$$

Proof by reduction to the one-dimensional case:

$g(t) = f(\vec{a} + t\vec{x})$ is differentiable at least twice as a function of $t \in [0, 1]$. Formation of the derivation and chain rule

$$g'(t) \;=\; \operatorname{grad} f(\vec{a}+t\vec{x})\,\vec{x} \;=\; \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\,(\vec{a}+t\vec{x})\,x_k$$

$$g''(t) \;=\; \sum_{i,k=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_k}\,(\vec{a}+t\vec{x})\,x_i x_k$$

Second-order Taylor formula

$$\begin{aligned} g(1) &= g(0) + g'(0) + \frac{1}{2} g''(\vartheta) \qquad 0 \le \vartheta = \vartheta_{\vec{x}} \le 1 \\ &= g(0) + g'(0) + \frac{1}{2} g''(0) + \frac{1}{2} \left(g''(\vartheta) - g''(0)\right) \end{aligned}$$

Insert

$$f(\vec{a} + \vec{x}) = f(\vec{a}) + \operatorname{grad} f(\vec{a})\, \vec{x} + \frac{1}{2} \vec{x}\, A\, \vec{x} + \varphi_2(\vec{x})$$

where

$$\varphi_2(\vec{x}) = \frac{1}{2} \sum_{i,k=1}^{n} \left(\frac{\partial^2 f}{\partial x_i \partial x_k} (\vec{a} + \vartheta \vec{x}) - \frac{\partial^2 f}{\partial x_i \partial x_k} (\vec{a}) \right) x_i x_k$$

From the continuity of the second derivation follows the remainder estimate.

□

8.2.1 Application of the Simple Version of the Taylor Formula to Stationary Points

Let f be real on $G \subset \mathbb{R}^n$ and twice continuously differentiable. The zeros $\vec{a}$ of $\operatorname{grad} f$ are called stationary points of f. There:

$$f(\vec{a} + \vec{x}) - f(\vec{a}) = \frac{1}{2} \vec{x}\, A\, \vec{x} + \varphi_2(\vec{x})$$

Consequences for the presence of extreme values in $\vec{a}$

1. If the Hessian matrix of f at the point $\vec{a}$ has the property

$$\vec{x}\, A\, \vec{x} > 0 \quad \text{for all } \vec{x} \neq 0$$

(A then is called “positive definite”), then one has an isolated local minimum of f in $\vec{a}$.

2. If for all $\vec{x} \neq 0$ $\vec{x}\, A\, \vec{x} < 0$ (the Hessian matrix is called negative definite in this case) then one has an isolated local maximum of f in $veca$.

3. If for A vectors $\vec{x}_+$, $\vec{x}_-$ exist with $\vec{x}_+ \, , A\, \vec{x}_+ > 0$ and $\vec{x}_-\, A\, \vec{x}_- < 0$, i.e. A is indefinite: then $\vec{a}$ is neither a local minimum nor a local maximum of f, but a saddle point.

4. If none of the cases 1, 2, 3 are present, then the Hessian matrix alone does not provide any information about the type of stationary point $\vec{a}$.

The **proof** of **1**, **2** , **3** follows from the limit statement

$$\lim_{\substack{\vec{x}\to\vec{0}\\ \vec{x}\neq\vec{0}}} \|\vec{x}\|^{-2}\varphi_2(\vec{x}) \;=\; 0$$

Investigation of the symmetric real 2×2 matrices A

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \quad \text{is} \quad \left\{ \begin{array}{r} \text{positive definite} \\ \text{negative definite} \\ \text{indefinite} \end{array} \right\} \quad \text{if} \quad \begin{cases} ac-b^2>0 \;\text{ and }\; a>0 \\ ac-b^2>0 \;\text{ and }\; a<0 \\ ac-b^2<0 \end{cases}$$

If $ac-b^2 \;=\; 0$, then none of these cases is present.

Proof:

$$\vec{x}\,A\,\vec{x} \;=\; \begin{bmatrix} x \\ y \end{bmatrix}\begin{bmatrix} ax+by \\ bx+cy \end{bmatrix} \;=\; ax^2 \,+\, 2bxy \,+\, cy^2$$

1) The case $a\neq 0$

$$\begin{aligned} \vec{x}\,A\,\vec{x} \;&=\; a\left(x^2+\frac{2b}{a}xy+\frac{b^2}{a^2}\,y^2\right) \;+\; \frac{ac-b^2}{a}\,y^2 \\ &=\; a\left(x+\frac{b}{a}\,y\right)^2 \;+\; \frac{ac-b^2}{a}\,y^2 \end{aligned}$$

If $ac-b^2 \;>\; 0$ then both summands have the sign of a and only become 0 at the same time in the case $x=y=0$.
On the other hand, if $ac-b^2 \;<\; 0$, then the two summands (if they are not 0) have different signs. Through an appropriate choice of x and y, the sum can take on any sign.
In the case of $ac-b^2 \;=\; 0$, the first summand determines the sign, provided it is not zero.

2) The case $c\neq 0$: If $c\neq 0$, then factor out c, complete the square, same result.

3) The case $a=c=0$: The result can be read directly from $\vec{x}\,A\,\vec{x} \;=\; 2bxy$!

□

<u>Example:</u> $f(x,y) \;=\; (4x^2+y^2)\mathrm{e}^{-x^2-4y^2}$

f is partially differentiable any number of times for both variables. Determination of stationary points:

$$\frac{\partial f}{\partial x} \;=\; \left(8x-8x^3-2xy^2\right)\mathrm{e}^{-x^2-4y^2}$$

$$\frac{\partial f}{\partial y} = \left(2y - 32x^2y - 8y^3\right)\mathrm{e}^{-x^2-4y^2}$$

Zeros:

$$1.\ \vec{a} = (0\,,\,0) \quad ; \quad 2.\ \vec{a} = (\pm 1\,,\,0) \quad ; \quad 3.\ \vec{a} = (0\,,\,\pm\tfrac{1}{2})$$

Points (x, y) with $x \neq 0 \neq y$ are not stationary here, because otherwise

$$\begin{aligned} 4x^2 + \ \ y^2 &= 4 \\ 16x^2 + 4y^2 &= 1 \end{aligned}$$

This system of linear equations in the indeterminate $s = x^2$ and $t = y^2$ is unsolvable.

$$\frac{\partial^2 f}{\partial x^2} = \left(8 - 24x - 2y^2 - 16x^2 + 16x^4 + 4x^2y^2\right)\mathrm{e}^{-x^2-4y^2}$$

$$\frac{\partial^2 f}{\partial x \partial y} = x\,y \text{ times something}$$

$$\frac{\partial^2 f}{\partial y^2} = \left(2 - 32x^2 - 24y^2 - 16y^2 + 256x^2y^2 + 64y^4\right)\mathrm{e}^{-x^2-4y^2}$$

1. $\vec{a} = (0\,,\,0)$: $\quad A = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$; local minimum

2. $\vec{a} = (\pm 1\,,\,0)$: $\quad A = \begin{bmatrix} -16 & 0 \\ 0 & -30 \end{bmatrix}\mathrm{e}^{-1}$; two maxima (local)

3. $\vec{a} = (0\,,\,\pm\frac{1}{2})$: $\quad A = \begin{bmatrix} 15/2 & 0 \\ 0 & -4 \end{bmatrix}\mathrm{e}^{-1}$; two saddle points

Remarks:

(1) Because the function becomes arbitrarily small outside the large circular disc, but remains positive, $\vec{0}$ is the global minimum and $(\pm 1, 0)$ are the global maxima of f. This follows from a general property of continuous real functions.

(2) On a compactum C every continuous real function always has a maximum and a minimum.

(3) The name of the compactum has been chosen in such a way that the theorem holds.

Examples of compacta:

a) Axis-parallel closed cuboids in $\mathbb{R}^n$

$$\mathbb{Q} = \{\vec{x} \in \mathbb{R}^n\,;\, a_i \leq x \leq b_i\,;\, (1 \leq i \leq n)\}\,, \quad \text{where} \quad a_i \leq b_i$$

b) Closed spheres in $\mathbb{R}^n$

$$A \;=\; \{\vec{x} \in \mathbb{R}^n\,;\, \|\vec{x}-\vec{a}\| \le r\}\,, \quad \vec{a} \in \mathbb{R}^n\,, \quad r > 0$$

c) Every at the same time closed and bounded set $C \subset \mathbb{R}^n$ is compact.

In the given context one calls those sets $A \subset \mathbb{R}^n$ "closed" whose complement $\mathbb{R}^n \setminus A$ is open, i.e. for which $\mathbb{R}^n \setminus A$ contains with each point a neighborhood.

Remark:

(4) Example **c)** is indeed a characterization of all compacta in $\mathbb{R}^n$. It is precisely the sets C in $\mathbb{R}^n$ for which the Bolzano-Weierstraß theorem applies: Every sequence on C has a convergent subsequence with a limit value in C.

Application to the Taylor formula for stationary points.

Let $A = (a_{ik})$ be a symmetric real $n \times n$ matrix.

$$q(\vec{x}) \;=\; \vec{x}\,A\,\vec{x} \;=\; \sum_{i,k=1}^{n} a_{ik}x_ix_k$$

is continuous on $\mathbb{R}^n$. The "unit sphere"

$$S \;=\; \{\vec{e},\, \|\vec{e}\| = 1\}$$

is compact, because it is bounded and closed. q therefore assumes a maximum M_1 and a minimum M_0 on S, like e.g. $q(\vec{e}_+) = M_1$, $q(\vec{e}_-) = M_0$. If A is definite, then $M_0 > 0$ or respectively $M_1 < 0$. However, for indefinite A one has $M_0 < 0 < M_1$.

Now, let $\vec{a}$ be a stationary point of f (declared in $\mathbb{R}^n$)

$$f(\vec{a}+\vec{x}) \;-\; f(\vec{a}) \;=\; \frac{1}{2}\,\vec{x}\,A\,\vec{x} \;+\; \varphi_2(\vec{x})\,, \quad x \ne 0$$

$$f(\vec{a}+\vec{x}) \;-\; f(\vec{a}) \;=\; \|\vec{x}\|^2\left(\frac{1}{2}\,q(\vec{e}) + \frac{\varphi_2(\vec{x})}{\|\vec{x}\|^2}\right), \quad \text{with} \quad \vec{e} \;=\; \frac{\vec{x}}{\|\vec{x}\|}$$

Application to extreme value problems.

Problem statement: Determine the maximum for a continuously partially differentiable function in $\mathbb{R}^n$ on a compactum C.

1. Because f is continuous, such a maximum exists.
2. Of the interior points of C, for which there is a neighborhood in C, only the stationary points are candidates for the maximum.
3. Treat the boundary points separately - maxima with constraints!

9. Implicit Functions and Applications

As an example, the lemniscate, the set of zeros of the function

$$\begin{aligned} F(x,y) &= \big((x+1)^2+y^2\big)\big((x-1)^2+y^2\big) - 1 \\ &= (x^2+y^2+1)^2 - (2x)^2 - 1 \\ &= x^4 + 2x^2y^2 + y^4 - 2x^2 - 2y^2 \end{aligned}$$

$$\begin{aligned} \operatorname{grad} F &= \big(4x^3+4xy^2-4x\,,\, 4x^2y+4y^3+4y\big) \\ &= \big(4x(x^2+y^2-1)\,,\, 4y(x^2+y^2+1)\big) \end{aligned}$$

There is only one stationary point of F on the set of zeros of F, namely $\vec{a}=(0,0)$.
In general, the stationary points of a function F in the plane, which are also in the set of zeros of F, are precisely those points of the graph of F whose tangent plane coincides with the x,y-plane. In the remaining points of the set of zeros of F, these two planes intersect in a straight line.

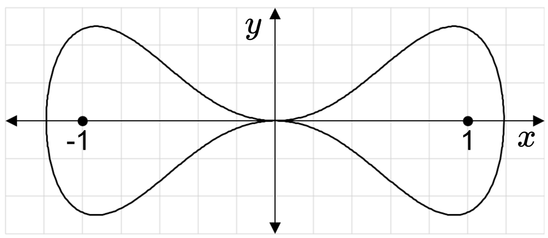
In the points with $\operatorname{grad} F(\vec{x}_0) \neq 0$ the set of zeros can be written as a (local) function, namely

$$y = f(x)\ , \quad \text{if} \quad \frac{\partial F}{\partial y}(\vec{x}_0) \neq 0$$

or respectively

$$x = g(y)\ , \quad \text{if} \quad \frac{\partial F}{\partial x}(\vec{x}_0) \neq 0$$

Explicit calculation is possible in the example.



$$y^2 \;=\; -x^2 - 1 \pm \sqrt{4x^2+1}$$

$$y \;=\; \pm\sqrt{\sqrt{4x^2+1} - x^2 - 1}$$

where $0 < x^2 < 2$; (these are two branches on each x-interval)

$$x^4 - 2x^2(1-y^2) + (1-y^2)^2 \;=\; 1 - 4y^2$$

(completing the square for $F(x,y) = 0$)

$$x^2 \;=\; 1 - y^2 \pm \sqrt{1-4y^2}$$

$$x = \pm\sqrt{1 - y^2 \pm \sqrt{1-4y^2}} \qquad 0 < y^2 < \frac{1}{4}$$

(these are four branches on each interval)

A derivative formula for implicit functions

Let F be a continuously partially differentiable function in the region G of $\mathbb{R}^n$. Furthermore, let $\vec{x}_0 = (x_0, y_0)$ be a root of F with the additional property ${}^{\partial F}\!/\!_{\partial y}\,(\vec{x}_0) \neq 0$. Finally, on an interval I containing x_0, $f : I \to \mathbb{R}$ shall be a continuous function in x_0 with the following three properties

$$f(x_0) \;=\; y_0\,, \quad (x, f(x)) \in G \quad \text{and} \quad F(x, f(x)) \;=\; 0 \quad \forall\, x \in I$$

Then f in x_0 is differentiable with the derivative

$$f'(x_0) \;=\; -\frac{\partial F}{\partial x}(\vec{x}_0) \Big/ \frac{\partial F}{\partial y}(\vec{x}_0)\,, \qquad \text{Short form:} \quad \frac{\mathrm{d}y}{\mathrm{d}x} \;=\; -\frac{F_x}{F_y}$$

Proof:

F is totally differentiable at the point $\vec{x}_0$, hence

$$F(x,y) \;=\; F(x_0,y_0) + \underset{\displaystyle\frac{\partial F}{\partial x}(\vec{x}_0)}{\underset{\downarrow}{a}}\,(x-x_0) + \underset{\displaystyle\frac{\partial F}{\partial y}(\vec{x}_0)}{\underset{\downarrow}{b}}\,(y-y_0) + \varphi(x-x_0, y-y_0)$$

Insert $\quad y \;=\; f(x)$

$$b(f(x) - f(x_0)) \;=\; -a(x-x_0) - \varphi(x-x_0, f(x)-f(x_0)) \qquad (*)$$

A first consequence: the difference quotient

$$\frac{f(x)-f(x_0)}{x-x_0} \quad \text{remains bounded for } x \to x_0$$

Then division of $(*)$ by $b(x-x_0)$

$$\frac{f(x)-f(x_0)}{x-x_0} \;=\; -\frac{a}{b} - \frac{1}{b(x-x_0)}\varphi\big(x-x_0\,,\, f(x)-f(x_0)\big)$$

$$\Rightarrow\;\; f'(x_0) \;=\; -\frac{a}{b}$$

$\square$

9.1 Existence Theorem for Implicit Functions

Let F be real and continuously differentiable on the region $G \subset \mathbb{R}^2$. Furthermore, let $\vec{x}_1 = (x_1, y_1)$ be a root of F with ${}^{\partial F}\!/_{\partial y}\,(\vec{x}_1) \neq 0$. Then a pair of intervals

$$I \;=\;]x_1 - a, x_1 + a[\qquad \text{and} \qquad J \;=\;]y_1 - b, y_1 + b[$$

exists with

$$I \times J \subset G \qquad (I \times J \;=\; \{(x,y);\, x \in I,\, y \in J\})$$

and the following three properties: first

$$\frac{\partial F}{\partial y}\,(\vec{x}) \;\neq\; 0 \qquad \forall\, \vec{x} \in I \times J$$

second, there is a function $f : I \to J$, which is continuous,

$$f(x_1) \;=\; y_1\,, \quad F(x_1, f(x_1)) \;=\; 0 \qquad \forall\, \vec{x} \in I$$

and if thirdly $\vec{x}_0 \in I \times J$ with $F(\vec{x}_0) = 0$ then

$$f(x_0) \;=\; y_0$$

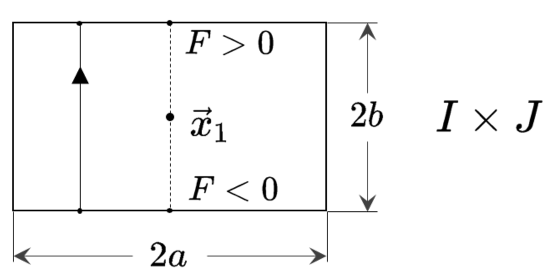
$$\begin{array}{llll} F : G & \to & \mathbb{R} & \text{continuously partially differentiable} \\ & & \cap & \\ & & \mathbb{R}_2 & \vec{x}_1 \in G\;\;, \quad F(\vec{x}_1) \;=\; 0\;\;, \quad \dfrac{\partial F}{\partial x}\,(\vec{x}_1) \;\neq\; 0 \end{array}$$

Then intervals exist

$$I \;=\;]x_1 - a, x_1 + a[$$

$$J \;=\;]y_1 - b, y_1 + b[$$

with $I \times J \;\subset\; G$



$$\frac{\partial F}{\partial y}\,(\vec{x}) \;\neq\; 0\,, \qquad \vec{x} \in I \times J$$

There exists a continuous $f : I \to J$ with

$$f(x_1) = y_1 \ , \qquad F(x, f(x)) = 0 \qquad \forall x \in I$$

If

$$\vec{x}_0 = (x_0, y_0) \in I \times J \quad \text{and} \quad F(x_0, y_0) = 0$$

then

$$y_0 = f(x_0)$$

Proof:

1) Because F and $-F$ have the same set of zeros, it can be assumed

$$\frac{\partial F}{\partial y}(\vec{x}_1) > 0$$

Because $\dfrac{\partial F}{\partial y}$ is continuous, a neighborhood U of $\vec{x}_1$ exists in G with

$$\frac{\partial F}{\partial y}(\vec{x}) > 0 \qquad \forall\, \vec{x} \in U.$$

To this neighborhood $I \times J$ will be restricted. Because

$$y \longmapsto \frac{\partial F}{\partial y}(x, y) \quad \text{is positive,}$$

$$y \longmapsto F(x, y) \quad \text{becomes strictly monotonically increasing}$$

If $b > 0$ only sufficiently small, then

$$F(x_1, y_1 - b) < 0 \quad \text{and} \quad F(x_1, y_1 + b) > 0.$$

Because F is continuous, a (sufficiently small) a exists such that

$$F(x, y_1 - b) < 0\,, \quad F(x, y_1 + b) > 0 \quad \text{if} \quad |x - x_1| < a \qquad (*)$$

If $x_0 \in I = \,]x_1 - a, x_1 + a[$ then consider the function

$$y \longmapsto F(x_0, y)\,, \quad y \in J$$

This function is strictly monotonically increasing and continuous. Because of $(*)$ it has exactly one root: $y_0 \in J$. With that

$$F(x, f(x)) = 0 \qquad \forall\, x \in I.$$

2) Continuity of the implicit function f in $\vec{x}_1$

Instead of starting with b, under **1)** one could also have chosen $b = {}^{\epsilon}/_{2}$. Possibly one would then have to choose an a' to b' smaller than a to b. Just as before, the points $(x, f(x))$ become zeros for the points $x \in \,]x_1 - a', x_1 + a'[$.

For that $|f(x) - f(x_1)| < \epsilon$. Hence f in x_1 is continuous.

3) Continuity of f at the point $x_0 \in I$.

According to **1)** and **2)**, there exists also for the reference point $\vec{x}_0 = (x_0, y_0)$ with $y_0 = f(x_0)$ in a neighborhood a function of x whose values are the second coordinates.

According to **2)**, this unique function is continuous in x_0, on the other hand it must agree with f; Hence f is continuous in x_0.

□

9.2 Local Extrema with Constraints

Let there be given two continuous, partially differentiable real functions F, h in the region G of $\mathbb{R}^n$. Then one asks about (local) extreme values of the restriction of h to the set of zeros of F.

If $\vec{a} \in G$ is a root of F with $\operatorname{grad} F(\vec{a}) \neq 0$, then $\vec{a}$ is a local extremum for the restriction h only if $\lambda \in \mathbb{R}$ (Lagrange multiplier) exists with $\operatorname{grad} h(\vec{a}) = \lambda \operatorname{grad} F(\vec{a})$.

Proof: for $n = 2$

$\operatorname{grad} F(\vec{a}) \neq 0$, $\quad F(\vec{a}) = 0$;

Without limitation let $\dfrac{\partial F}{\partial y}(a_1, a_2) \neq 0$

According to the existence theorem for implicit functions, there exist an open interval I containing a_1 and an interval J containing a_2 with a continuously differentiable function $f : I \to J$ such that

$$f(a_1) = a_2 \,, \qquad F(x, f(x)) = 0 \qquad \forall x \in I$$

$$(x_0, y_0) \in I \times J \quad \text{and} \quad F(x_0, y_0) = 0 \,, \quad \text{then} \quad y_0 = f(x_0)$$

$H(x) := h(x, f(x))$ describes locally (in a neighborhood of $\vec{a}$) the restriction of h to the set of zeros of F. H is differentiable on I; therefore $H(a_1)$ becomes an extremum (local) of H only if $H'(a_1) = 0$.

Determination of the derivative of H according to the chain rule

$$0 = H'(a_1) = \operatorname{grad} h(a_1, a_2)\big(1\,,\ f'(a_1)\big)$$

i.e. according to the derivative formula for the implicit function f

$$0 = \operatorname{grad} h(\vec{a})\left(1\,,\ -\frac{F_x(\vec{a})}{F_y(\vec{a})}\right)$$

The second vector is perpendicular to $\operatorname{grad} F(\vec{a})$. Therefore, the condition above means:

$\operatorname{grad} h(\vec{a})$ is a scalar multiple of $\operatorname{grad} F(\vec{a})$.

$\square$

Example:

(0) Product $h(x, y, z) = x\,y\,z$ is to be examined for extreme values on the closed unit sphere.

a) $\operatorname{grad} h(x, y, z) = (y\,z, x\,z, x\,y)$

The zeros of the gradient lie on the coordinate axes, all these points are saddle points. h is 0 there.

b) The boundary of the unit sphere is described by the constraints:

$$0 = F(x, y, z) = x^2 + y^2 + z^2 - 1$$

$$\operatorname{grad} F(x, y, z) = 2(x, y, z)$$

without stationary points on the boundary of the unit sphere.

Sought after $\vec{a} = (a, b, c)$ with

$$\operatorname{grad} h(\vec{a}) = \lambda\, F(\vec{a}) = 2\lambda\,\vec{a} = (bc, ac, ab)$$

Three equations:

$$2\lambda\, a = bc\,, \quad 2\lambda\, b = ac\,, \quad 2\lambda\, c = ab$$

Outside of the coordinate axes

$$abc = 2\lambda\, a^2 = 2\lambda\, b^2 = 2\lambda\, c^2$$

$$a^2 = b^2 = c^2 = \frac{1}{3}$$

$$a,\, b,\, c \in \left\{\pm\sqrt{\frac{1}{3}}\right\}$$

Values of h at these positions

$$h(a,b,c) \;=\; \pm\sqrt{\frac{1}{27}}$$

The extrema are eight in total, four minima and four maxima. They lie on the edge of the unit sphere, in the corners of the inscribed axis-parallel cube.

Example:

(1) Heron of Alexandria (first century): Area of the plane triangle with side lengths a, b, c

$$A \;=\; \sqrt{s(s-a)(s-b)(s-c)} \quad \text{where} \quad s \;=\; \frac{1}{2}(a+b+c)$$

is half the perimeter of the triangle.

Question: When is the area A at a maximum for a fixed perimeter $2s$?

Exactly when

$$h(a,b,c) \;=\; (s-a)(s-b)(s-c)$$

has a maximum under the constraint $a+b+c \;=\; 2s$.

$$h \;=\; s^3 \;-\; (a+b+c)s^2 \;+\; (ab+bc+ca)s \;-\; abc$$

$$\operatorname{grad} h(a,b,c) \;=\; \begin{pmatrix} -s^2+(b+c)s-ba \\ -s^2+(a+c)s-ac \\ -s^2+(a+b)s-ab \end{pmatrix}$$

$$F(a,b,c) \;=\; a \,+\, b \,+\, c \,-\, 2s \;=\; 0$$

$$\operatorname{grad} F \;=\; (1,1,1)$$

Lagrangian condition $(\operatorname{grad} h(a,b,c) \;=\; \lambda \operatorname{grad} F(a,b,c))$ in difference form:

$$(a-b)s \;=\; (a-b)c \qquad (\text{from } 2^{nd} - 1^{st} \text{ coordinate} = 0)$$

$$(b-c)s \;=\; (b-c)a \qquad (\text{from } 3^{rd} - 2^{nd} \text{ coordinate} = 0)$$

since $0 < a < s$, $0 < b < s$, $0 < c < s$, these two equations can only be satisfied if $a \;=\; b \;=\; c$. Therefore the result is:

$$a \;=\; b \;=\; c \;=\; \frac{2}{3}s$$

$$A_{max} \;=\; \sqrt{s \cdot \frac{s}{3} \cdot \frac{s}{3} \cdot \frac{s}{3}} \;=\; \frac{s^2}{3\sqrt{3}}$$

Example:

(2) Let $A = (a_{ik})$ be a real symmetric $n \times n$ matrix

$$q(\vec{x}) \;=\; \vec{x}\,A\,\vec{x} \;=\; \sum_{i,k=1}^{n} a_{ik}x_i x_k$$

Task: Investigation of the extrema of $q(\vec{x})$ on the unit sphere

$$S \;=\; \{\vec{x} \in \mathbb{R}^n\,,\;\; \|\vec{x}\| = 1\}$$

Constraint: $\quad F(\vec{x}) \;=\; \sum_{i=1}^{n} x_i^2 - 1 \;=\; 0$

$$\operatorname{grad} F(\vec{x}) \;=\; 2\vec{x} \quad \text{has no zeroes on } S$$

$$\begin{aligned}
\operatorname{grad} q(\vec{x}) \;&=\; \left(\frac{\partial}{\partial x_j}\,q(\vec{x})\right)_{1\le j\le n} \;=\; \left(\frac{\partial}{\partial x_j}\sum_{i,k=1}^{n} a_{ik}x_i x_k\right)_{1\le j\le n} \\
&=\; \sum_{i,k=1}^{n} a_{ik}\frac{\partial}{\partial x_j}\,(x_i x_k)_{\,1\le j\le n} \qquad \text{product rule} \\
&=\; \left(\sum_{k=1}^{n} a_{jk}x_k \;+\; \sum_{i=1}^{n} a_{ij}x_i\right)_{1\le j\le n} \qquad \text{symmetry of } A \\
&=\; \left(2\sum_{k=1}^{n} a_{jk}x_k\right)_{1\le j\le n} \;=\; 2\,A\,\vec{x}
\end{aligned}$$

Where does $\operatorname{grad} q(\vec{a}) \;=\; \lambda \operatorname{grad} F(\vec{a}) \quad \vec{a} \in S$ hold? Exactly where

$$A\,\vec{a} \;=\; \lambda\,\vec{a}$$

The extrema of $q(\vec{x})$ on S can therefore be found under the eigenvectors $\vec{a}$ of A with length 1. Because S is compact and because $q(\vec{x})$ is continuous, $q(\vec{x})$ assumes a maximum and a minimum on S. So in particular the symmetric real matrix always has at least one real eigenvalue!

Example:

(3) $b_i \;>\; 0 \quad (1 \le i \le n)\;; \quad$ further let $\; p > 1\,,\;\; (b_i)_{1\le i\le n} \;=\; \vec{b}$

The linear form $L(\vec{x}) \;=\; \vec{x}\,\vec{b}$ is to be examined on the positive cone $P = \{\vec{x}\,,\;\; x_i \ge 0 \;\;\forall\, i\}$ under the constraint $\sum_{i=1}^{n} x_i^p = 1$ for extrema (maxima).

The constraint $F(\vec{x}) = \sum_{i=1}^{n} x_i^p - 1 = 0$ describes a compactum C in P.

First the "inner" of P, that is $\mathring{P} = \{\vec{x} \in \mathbb{R}^n, \; x_i > 0 \;\forall i\}$. There

$$\operatorname{grad} F(\vec{x}) = \left(p\, x_j^{p-1}\right)_{1 \le j \le n}$$

$$\operatorname{grad} L(\vec{x}) = (b_j)_{1 \le j \le n} = \vec{b} \qquad \forall \vec{x} \in \mathbb{R}^n$$

Determine $\vec{a}$ with $F(\vec{a}) = 0$, $\operatorname{grad} L(\vec{a}) = \lambda \operatorname{grad} F(\vec{a})$

Hence $\lambda\, p\, a_j^{p-1} = b_j \qquad (1 \le j \le n)$

Determination of λ from the constraint:

$$\sum_{j=1}^{n} a_j^p = 1 \quad \text{that means} \quad \left(\frac{1}{\lambda p}\right)^{\frac{p}{p-1}} \sum_{j=1}^{n} b_j^{\frac{p}{p-1}} = 1$$

$$\sum_{j=1}^{n} bj^{\,q} = (\lambda\, p)^q \qquad (*)$$

where $q = \dfrac{p}{p-1}$, i.e. $\dfrac{1}{p} + \dfrac{1}{q} = 1$

Determination of the value $L(\vec{a}) = \vec{a}\,\vec{b}$

$$L(\vec{a}) = \sum_{j=1}^{n} a_j b_j = \left(\frac{1}{\lambda p}\right)^{\frac{1}{p-1}} \sum_{j=1}^{n} b_j^{\overbrace{\frac{1}{p-1}+1}^{q}}$$

$$\underset{(*)}{\overset{}{\underset{\downarrow}{=}}} \left(\sum_{j=1}^{n} b_j^q\right)^{1/q} =: M_n$$

Here in $\vec{a}$, L actually has a maximum on C, because on the one hand M_n increases strictly and on the other hand, the pieces of C on the boundary of P (say $x_n = 0$) belong to the analogous $(n-1)$-dimensional problem. Because $n = 1$ is trivial, the maximality of M_n now follows by induction on n. This proves

Hölder's inequality:

$$\sum_{i=1}^{n} a_i b_i \le \left(\sum_{j=1}^{n} a_j^p\right)^{1/p} \left(\sum_{j=1}^{n} b_j^q\right)^{1/q}, \qquad \frac{1}{p} + \frac{1}{q} = 1$$

where the a_i, b_i are all ≥ 0.

The question:

When is a mapping $F : G \to \mathbb{R}^n ; (f_i)_{1 \le i \le n}$, where G denotes a region in $\mathbb{R}^n$ to be called differentiable (linearly approximable) at the point $\vec{a} \in G$?

Definition

F is called at the point $\vec{a} \in G$ totally differentiable (linearly approximable) if a linear mapping

$$L : \mathbb{R}^n \to \mathbb{R}^m$$

exists such that in the formula

$$F(\vec{a} + \vec{h}) \;=\; F(\vec{a}) + L\vec{h} + \varphi(\vec{h})$$

the remainder function φ has the limit property

$$\lim_{\substack{\vec{h} \to \vec{0} \\ \vec{h} \neq \vec{0}}} \|\vec{h}\|^{-1} \varphi(\vec{h}) \;=\; \vec{0}_m$$

Remarks:

(1) If F is differentiable in $\vec{a}$, then all coordinates f_i are partially differentiable with respect to all variables and the linear mapping L is described in matrix form by

$$D\,F(\vec{a}) \;=\; \left(\frac{\partial f_i}{\partial x_k}(\vec{a}) \right)_{\substack{1 \le i \le n \\ 1 \le k \le n}} \qquad \text{“functional matrix or Jacobian matrix” of } F$$

(2) If all f_i on G are continuously partially differentiable, then F is totally differentiable everywhere on G.

A main property of the general notion of differentiability is the general chain rule. Let $F : G \to \mathbb{R}^m$ be defined in the region $G \subset \mathbb{R}^n$, $\vec{a} \in G$ with $F(\vec{a}) = \vec{a}_1$ where $\vec{a}_1$ lies in the domain of definition of another mapping F_1.

$$F_1 : G_1 \in \mathbb{R}^h \;; \quad G_1 \subset \mathbb{R}^m$$

If F is totally differentiable in $\vec{a}$ and F_1 is totally differentiable in $\vec{a}_1$, then the composite mapping $F_1 \circ F$ is totally differentiable at the point $\vec{a}$ with the total derivative

$$D(F_1 \circ F)(\vec{a}) \;=\; D\,F_1(F(\vec{a})) \circ D\,F(\vec{a})$$

9.3 The Problem of Reverse Mapping (Coordinate Transformation)

In the case $m = n$ the functional matrix becomes

$$D\,F(\vec{a}) \;=\; \left(\frac{\partial f_i}{\partial x_k}\,(\vec{a})\right)$$

a square matrix which is invertible if and only if its determinant (the Jakobi determinant of F) is $\neq 0$. The existence theorem for reversibility is to be regarded as a higher-dimensional analogue of the monotony criterion of differential calculus.

<u>Theorem:</u>
Let F be a mapping of a region $G \subset \mathbb{R}^n$ into $\mathbb{R}^n$ which is continuously partially differentiable in all points of G. Furthermore, let $D\,F(\vec{a})$ be invertible, i.e. the Jacobi determinate $\det D\,F(\vec{a}) \neq 0$; then there exist neighborhoods U of $\vec{a}$ and U_1 of $\vec{a}_1 = F(\vec{a})$ such that F is a "bijective" mapping from U to U_1 (that means, $F(U) = U_1$ and every point of U_1 has exactly one F-preimage in U). The thus declared inverse mapping $F_1 : U_1 \to U$ is totally differentiable in $\vec{a}_1 = F(\vec{a})$ ($F_1 \circ F(\vec{x}) = \vec{x} = \mathrm{id}(\vec{x})$).

$$1_n \;=\; D\,(F_1 \circ F)) \;=\; D\,F_1(F(\vec{x})) \circ D\,F(\vec{x})$$

Multiplication by the inverse matrix of $D\,F(\vec{x})$ for $\vec{x} = \vec{a}$:

$$D\,F_1(F(\vec{a})) \;=\; (D\,F(\vec{a}))^{-1}$$

<u>Example:</u>

(1) Polar coordinates in the plane

$$\begin{aligned} x &= r\cos\varphi = f_1(r,\varphi) \\ y &= r\sin\varphi = f_2(r,\varphi) \end{aligned}$$

$$F(r,\varphi) \;=\; \begin{pmatrix} f_1(r,\varphi) \\ f_2(r,\varphi) \end{pmatrix} \qquad \text{defined on } \mathbb{R}^2$$

F does not have a global inverse function as for example F is periodic as a function of φ. The functional matrix:

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial r} & \dfrac{\partial f_1}{\partial \varphi} \\[2ex] \dfrac{\partial f_2}{\partial r} & \dfrac{\partial f_2}{\partial \varphi} \end{bmatrix} = D\,F(r,\varphi) = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}$$

Jacobian determinant

$$\det D\,F(r,\varphi) = r$$

The condition for the local inverse function $\det D\,F \neq 0$ is violated on the line $r = 0$. This straight line only has the zero point as F-image.

<u>Example:</u>

(2) The elliptical coordinates in the plane

$$x = \cosh u \cos v = f_1(u,v)$$
$$y = \sinh u \sin v = f_2(u,v)$$
$$F(u,v) = \begin{pmatrix} f_1(u,v) \\ f_2(u,v) \end{pmatrix} \qquad \text{defined on } \mathbb{R}^2$$
$$D\,F(u,v) = \begin{bmatrix} \sinh u \cos v & -\cosh u \sin v \\ \cosh u \sin v & \sinh u \cos v \end{bmatrix}$$

Jacobian determinant

$$\det D\,F(u,v) = \sinh^2 u \cos^2 v + \cosh^2 u \sin^2 v = \sinh^2 u + \sin^2 v$$

If $n = 0$ and $v = \pi k$, then the Jakobi determinant is 0, but otherwise not!

Image curves $u = \text{const}$ and $v = \text{const}$:

<u>$u = 0$:</u> The image is the line $[-1, 1]$ on the x-axis

<u>$u \neq 0$:</u> $\cosh u =: a\,, \quad |\sinh u| =: b$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \cos^2 v + \sin^2 v = 1$$

This is an ellipse with focal points $\pm(1, 0)$

<u>$v = \text{const}$:</u> If $v = \dfrac{\pi}{2}\,k$, then (pieces of) straight lines result as images; otherwise with $a' = |\cos v|\,,\ b' = |\sin v|$

$$\left(\frac{x}{a'}\right)^2 - \left(\frac{y}{b'}\right)^2 = \cosh^2 u - \sinh^2 u = 1$$

This is the equation of a hyperbola with focal points $\pm(1, 0)$

Example:

(3) Spherical coordinates in space

$$\left.\begin{aligned} x &= R\cos\varphi\cos\vartheta &&= f_1(R,\varphi,\vartheta) \\ y &= R\sin\varphi\cos\vartheta &&= f_2(R,\varphi,\vartheta) \\ z &= R\sin\vartheta &&= f_3(R,\varphi,\vartheta) \end{aligned}\right\} = F(R,\varphi,\vartheta)$$

$$D\,F(R,\varphi,\vartheta) = \begin{bmatrix} x_R & x_\varphi & x_\vartheta \\ y_R & y_\varphi & y_\vartheta \\ z_R & z_\varphi & z_\vartheta \end{bmatrix}$$

$$D\,F(R,\varphi,\vartheta) = \begin{bmatrix} \cos\varphi\cos\vartheta & -R\sin\varphi\cos\vartheta & -R\cos\varphi\sin\vartheta \\ \sin\varphi\cos\vartheta & R\cos\varphi\cos\vartheta & -R\sin\varphi\sin\vartheta \\ \sin\vartheta & 0 & R\cos\vartheta \end{bmatrix}$$

Jacobian determinant

$$\begin{aligned} \det D\,F(R,\varphi,\vartheta) &= R^2\big(\cos^2\varphi\cos^3\vartheta + \sin^2\varphi\cos\vartheta\sin^2\vartheta \\ &\qquad + \cos^2\varphi\cos\vartheta\sin^2\vartheta + \sin^2\varphi\cos^3\vartheta\big) \end{aligned}$$

$$\begin{aligned} \det D\,F(R,\varphi,\vartheta) &= R^2\big(\cos^3\vartheta + \cos\vartheta\sin^2\vartheta\big) \\ &= R^2\cos\vartheta \neq 0 \quad (\text{except } R = 0 \text{ or } \cos\vartheta = 0) \end{aligned}$$

A·T·I·C·E

ATICE LLC, Albany NY

ISBN 978-1-951894-14-6