HIGHER MATHEMATICS

Lectures

Part Three

HIGHER MATHEMATICS

Lectures

Part Three

Stefan Wurm

ATICE LLC, Albany NY

First English language ATICE e-Book Edition | ISBN 978-1-951894-15-3.

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 Three**, contains the material of the Higher Mathematics III lectures as given by Prof. Leutbecher in the winter semester 1982/83 at the TUM.

Stefan Wurm

Albany, New York
May 2022

1. Integrals with Parameters[1]

Theorem: Integrals with parameters over continuous integrands are continuous.

Given: $I = [a,b]\,, \quad J = [c,d]$

$$f\,:\, I \times J \longmapsto \mathbb{R} \ \text{ continuous}$$

Then the function defined by

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

is continuous on J as a function of y.

Remark:

Instead of J one can also make use of a compactum belonging to a higher dimensional space.

Other basic properties of compacta:

Every real continuous function on a compactum is uniformly continuous

$$\begin{aligned}
|F(y) - F(y_0)| \;&=\; \left| \int_a^b (f(x,y) - f(x,y_0))\,\mathrm{d}x \right| \\
&\leq\; \int_a^b |f(x,y) - f(x,y_0)|\,\mathrm{d}x \\
&<\; \frac{\epsilon}{b-a} \qquad \forall\, x\,, \quad \text{if } |y - y_0| < \delta
\end{aligned}$$

[1] See also chapter 1 in HIGHER MATHEMATICS Lectures Part Two.

1.1 Differentiation under the Integral

Under the assumptions above,

$$y \longmapsto f(x,y)$$

shall also be differentiable for all $x \in I = [a,b]$ whose derivative $D_2 f(x,y)$ is continuous as a function of x, y on $I \times J$. Then

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

is differentiable on J (continuous) with derivative

$$F'(y) = \int_a^b D_2 f(x,y)\,\mathrm{d}x$$

Proof sketch:

$$\frac{F(y+h)-F(y)}{h} - \int_a^b D_2 f(x,y)\,\mathrm{d}x$$

$$= \int_a^b \Bigg[\underbrace{\frac{f(x,y+h)-f(x,y)}{h}}_{\substack{\text{Mean value theorem:}\\ D_2 f(x,y+\delta_x h)\,,\ \ 0 \le \delta_x \le 1}} - D_2 f(x,y)\Bigg]\,\mathrm{d}x$$

This formula yields the assertion about $F'(y)$ because of the uniform continuity of $D_2 f(x,y)$ on $I \times J$. The continuity of F' follows from the first theorem applied to $D_2 f(x,y)$ instead of $f(x,y)$.

Example:

$$F(y) = \int_0^\pi \cos(xy)\,\mathrm{d}x = \frac{\sin(\pi y)}{y}$$

is according to the last theorem continuously differentiable. Here, the derivative can easily be calculated directly

$$F'(y) = \frac{\pi y \cos(xy) - \sin(\pi y)}{y^2}$$

1.2 Swapping the Order of Integration in Double Integrals

Let $f : [a,b] \times [c,d] \longmapsto \mathbb{R}$ be continuous. Then the following formula applies:

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

Proof by reduction to the last theorem:

Auxiliary function:

$$F(y) \;=\; \int_a^b \Big(\int_c^y f(x,t)\,\mathrm{d}t \Big)\,\mathrm{d}x$$

This function has the integrand

$$\int_c^y f(x,t)\,\mathrm{d}t$$

According to the fundamental theorem of calculus, it is differentiable with derivative

$$D_2\Big(\int_c^y f(x,t)\,\mathrm{d}t \Big) \;=\; f(x,y)$$

By assumption, this derivative is continuous on $I \times J$. Therefore the last theorem can be applied to F with the result

$$F'(y) \;=\; \int_a^b D_2\Big(\int_c^y f(x,t)\,\mathrm{d}t \Big)\,\mathrm{d}x \;=\; \int_a^b f(x,y)\,\mathrm{d}x$$

Here one observes $F(c) = 0$. Hence

$$\begin{aligned} F(d) \;&=\; \text{right side of the assertion} \\ &=\; F(d) - F(c) \;=\; \int_c^d F'(y)\,\mathrm{d}y \\ &=\; \int_c^d \Big(\int_a^b f(x,y)\,\mathrm{d}x \Big)\,\mathrm{d}y \;=\; \text{left side of the assertion} \end{aligned}$$

$\square$

1.3 Integrals with variable Limits

$$G(y) \;=\; \int\limits_{\varphi(y)}^{\psi(y)} g(x,y)\,\mathrm{d}x$$

can be reduced to the integral types discussed at the beginning with the help of a substitution!

$$\begin{aligned} x \;&=\; \varphi(y) \,+\, (\psi(y)-\varphi(y))\,t \\ &=\; \varphi(y)\,(1-t) \,+\, \psi(y)\,t \;; \qquad 0 \,\le\, t \,\le\, 1 \\ \mathrm{d}x \;&=\; (\psi(y)-\varphi(y))\,\mathrm{d}t \end{aligned}$$

With that

$$G(y) \;=\; (\psi(y)-\varphi(y)) \int\limits_0^1 g\big(\varphi(y)(1-t)+\psi(y)t\,,\,y\big)\,\mathrm{d}t$$

For example, if g is continuously partially differentiable with respect to both variables and φ and ψ are continuously differentiable, then $G(y)$ is continuously differentiable according to the theorem on differentiation under the integral. Again, it is easier to compute the derivative directly

$$\begin{aligned} \frac{G(y+h)-G(y)}{h} \;&=\; \frac{1}{h}\Bigg(\int\limits_{\varphi(y+h)}^{\psi(y+h)} g(x,y+h)\,\mathrm{d}x \;-\; \int\limits_{\varphi(y)}^{\psi(y)} g(x,y)\,\mathrm{d}x\Bigg) \\ &=\; \frac{1}{h}\int\limits_{\psi(y)}^{\psi(y+h)} g(x,y)\,\mathrm{d}x \;-\; \frac{1}{h}\int\limits_{\varphi(y)}^{\varphi(y+h)} g(x,y)\,\mathrm{d}x \\ &\qquad +\; \int\limits_{\varphi(y+h)}^{\psi(y+h)} \frac{g(x,y+h)-g(x,y)}{h}\,\mathrm{d}x \end{aligned}$$

Consideration of the limit case $h \to 0$ yields

$$G'(y) \;=\; \psi'(y)g\big(\psi(y),y\big) \;-\; \varphi'(y)g\big(\varphi(y),y\big) \;+\; \int\limits_{\varphi(y)}^{\psi(y)} D_2 g(x,y)\,\mathrm{d}x$$

1.4 Improper Integrals with Parameters

An improper integral is defined as the limit value for continuous integrands $g(x)$ on half-open intervals $[a, b[$, or respectively $]a, b]$, or on open intervals $]a, b[$ (with special case $b = +\infty$, $a = -\infty$). As for example for $[a, b[$

$$\int_a^b g(x)\,\mathrm{d}x : \quad \lim_{\beta \nearrow b} \int_a^\beta g(x)\,\mathrm{d}x \qquad \text{(compare with HM2)}^2$$

As with series and functions, one requires the concept of "uniform convergence" for improper integrals with parameters:

Definition:

Let $I = [a, b[\ ,\ J = [c, d]$ be intervals and $f : I \times J \longmapsto \mathbb{R}$ continuous.

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

is called uniformly convergent if first $F(y)$ converges for every $y \in J$ and if second for every $\epsilon > 0$ a $\beta_\epsilon \in\,]a, b[$ exists with

$$\left| F(y) - \int_a^\beta f(x, y)\,\mathrm{d}x \right| \ < \ \epsilon \qquad \text{for all} \quad y \in J$$

if $\beta \geq \beta_\epsilon \quad (\beta < b)$.

Remark:

(1) If $I = \,]a, b[$ is open, then the definition has to be correspondingly extended to a. The integration interval can also be broken down by an intermediate point.

Example: The gamma integral

$$\Gamma(x) \ = \ \int_0^\infty \mathrm{e}^{-t} t^{x-1}\,\mathrm{d}t\,, \qquad x > 0$$

is locally uniformly convergent, which means that for all $\delta > 0$, $c > \delta$ the integral is uniformly convergent on $J = [\delta, c]$.

[2] See chapter 1, *Fundamentals of Integral Calculus*, in HIGHER MATHEMATICS Lectures Part Two.

Remark:

(2) Uniform convergence follows from the majorant criterion for improper integrals.[3]

On the interval $]0,1[$ note

$$0 \;<\; \mathrm{e}^{-t}t^{x-1} \;\leq\; t^{\delta-1}\;, \qquad (0<t\leq 1)$$

and on $[1,\infty[$

$$0 \;<\; \mathrm{e}^{-t}t^{x-1} \;\leq\; N!\,t^{-N+c-1}$$

Limit swapping for improper integrals with parameters:

Let $f:\,[a,b[\times[c,d]\longmapsto\mathbb{R}$ be continuous.

(1) If

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

converges uniformly on $J=[c,d]$ then $F(y)$ is a continuous function on J.

(2) Furthermore, if $y\longmapsto f(x,y)$ is differentiable with respect to y and the derivative $D_2 f(x,y)$ as a function of x, y is continuous and

$$\int_a^b D_2 f(x,y)\,\mathrm{d}x \quad \text{converges for every } J,$$

then $F(y)$ is continuously differentiable on J with derivative

$$F'(y) \;=\; \int_a^b D_2 f(x,y)\,\mathrm{d}x$$

Proof:

1) The proof uses a decomposition

$$I=[a,b[\;=\; [a,\beta]\cup[\beta,b[$$

with the estimate from the uniform convergence for $[\beta,b[$ and the (already proven) theorem about the continuity of proper integrals with parameters.

[3]See chapter 1, section *The Improper Integral*, examples (1), (2) and (3), in HIGHER MATHEMATICS Lectures Part Two.

2)
$$\frac{F(y+h)-F(y)}{h} - \int\limits_a^b D_2 f(x,y)\,\mathrm{d}x$$

$$= \int\limits_a^\beta \left(\frac{f(x,y+h)-f(x,y)}{h} D_2 f(x,y)\right)\mathrm{d}x + \int\limits_\beta^b (\ldots)\,\mathrm{d}x$$

As before, one combines the theorem about proper integrals with the estimate of uniform convergence.

□

The integrand of the Γ-function, $\mathrm{e}^{-t}t^{x-1}$, is differentiable in x with the derivative $\mathrm{e}^{-t}\ln(t)\cdot t^{x-1}$. Estimation as with the Γ-integral results in uniform convergence on the intervals $[\delta, c]$, $0 < \delta < c$. Therefore $\Gamma(x)$ is differentiable on $]0,\infty[$ with the derivative

$$\Gamma'(x) = \int\limits_0^\infty \mathrm{e}^{-t}\ln t\, t^{x-1}\,\mathrm{d}t$$

Remark:

Through the application of Hölder's inequality one finds: $\ln\Gamma(x)$ is convex on $]0,\infty[$. With $\Gamma(1) = 1$ and $\Gamma(x+1) = x\Gamma(x)$ one even has a characterization of the gamma function.

1.5 The Laplace Transform

is declared for real (or complex) functions f on $]0,\infty[$ by

$$(\mathscr{L}f)(s) = \int\limits_0^\infty \mathrm{e}^{-st} f(t)\,\mathrm{d}t$$

For example the constant $f = 1$

$$(\mathscr{L}f)(s) = \int\limits_0^\infty \mathrm{e}^{-st}\,\mathrm{d}t = -\frac{1}{s}\,\mathrm{e}^{-st}\Big|_0^\infty = \frac{1}{s}\,, \qquad s > 0$$

For some functions the Laplace transform is constantly divergent, like for $f(t) = \mathrm{e}^{t^2}$, because $\mathrm{e}^{-st}\mathrm{e}^{t^2} = \mathrm{e}^{t(t-s)} \geq \mathrm{e}^{t}$ if $t \geq s+1$

On the other hand, $f(t) := \mathrm{e}^{at}$ returns

$$(\mathscr{L}f)(s) \;=\; \int\limits_0^\infty \mathrm{e}^{-st}\mathrm{e}^{at}\,\mathrm{d}t \;=\; \int\limits_0^\infty \mathrm{e}^{(a-s)t}\,\mathrm{d}t$$

$$= \frac{1}{a-s}\,\mathrm{e}^{(a-s)t}\Big|_0^\infty \;=\; \frac{1}{s-a}\;, \qquad a<s$$

The step function

$$R_a(t) \;:=\; \begin{cases} 1 & \text{in} \quad 0<t<a \\ 0 & \text{in} \quad t>a \end{cases}$$

has convergent Laplace transforms for all s

$$\int\limits_0^\infty \mathrm{e}^{-st}R_a(t)\,\mathrm{d}t \;=\; \int\limits_0^a \mathrm{e}^{-st}\,\mathrm{d}t \;=\; \Big|_0^a \;=\; \frac{1-\mathrm{e}^{-sa}}{s}$$

The starting functions f are all declared on $]0,\infty[$, on the other hand the convergence range of $(\mathscr{L}f)(s)$ in s depends on f.

Important property of the Laplace transform: its linearity, i.e.

$$(\mathscr{L}(f+g))(s) \;=\; (\mathscr{L}f)(s) + (\mathscr{L}g)(s)$$

$$(\mathscr{L}\lambda f)(s) \;=\; \lambda(\mathscr{L}f)(s)\;, \qquad (\lambda \in \mathbb{R} \text{ or } \mathbb{C})$$

Sufficient conditions for the absolute convergence of $(\mathscr{L}f)(s)$:

a) f is continuous on $]0,\infty[$ except for at most finitely many discontinuities.

b) For $t \searrow 0\,,$ $|f|$ does not grow faster than a power of t with exponent > -1, i.e. there exist M_0, t_0, δ with

$$|f(t)| \;\le\; M_0 t^{\delta-1} \quad \text{for} \quad t \le t_0$$

c) f is of exponential growth σ, i.e. there exist positive numbers t_1, M_1 with

$$|f(t)| \;\le\; M_1\mathrm{e}^{\sigma t} \quad \text{for} \quad t \ge t_1$$

Under these conditions, the majorant criterion for improper integrals yields the absolute convergence of the Laplace transform if $\sigma < s$. The estimation is done analogous to the procedure for the gamma integral!

1.5.1 The Laplace Transform of the Derivative

The functions f and f' shall satisfy the properties a), b) and c). Then for the Laplace transform of f'

$$(\mathscr{L} f')(s) \;=\; s(\mathscr{L} f)(s) \;-\; f(0)$$

Proof with partial integration

$$\int_0^T \mathrm{e}^{-st} f'(t)\,\mathrm{d}t \;=\; \underbrace{\mathrm{e}^{-st} f(t)\Big|_0^T}_{\downarrow} + \underbrace{s\int_0^T \mathrm{e}^{-st} f(t)\,\mathrm{d}t}_{\downarrow}$$

$$=\quad -f(0) \quad + \quad s(\mathscr{L} f)(s) \qquad \square$$

<u>Examples:</u>

1) $f(t) \;=\; \sin\omega t\,,\; s>0$

$$(\mathscr{L} f)(s) \;=\; \int_0^\infty \mathrm{e}^{-st}\sin\omega t\,\mathrm{d}t \;=\; -\frac{1}{s}\,\mathrm{e}^{-st}\Big|_0^\infty + \frac{\omega}{s}\int_0^\infty \mathrm{e}^{-st}\cos\omega t\,\mathrm{d}t$$

$$=\; \frac{\omega}{s}\Big[-\frac{1}{s}\,\mathrm{e}^{-st}\cos\omega t\Big|_0^\infty\Big] - \int_0^\infty \frac{\omega^2}{s^2}\sin\omega t\mathrm{e}^{-st}\,\mathrm{d}t$$

$$=\; \frac{\omega}{s^2} - \frac{\omega^2}{s^2}(\mathscr{L} f)(s)$$

$$\Rightarrow\quad (\mathscr{L} f)(s) \;=\; \frac{\omega}{s^2+\omega^2}$$

2) $f(t) \;=\; \cos\omega t\,,\; s>0$

$\cos\omega t \;=\; \dfrac{1}{\omega}\dfrac{\mathrm{d}}{\mathrm{d}t}\sin\omega t\,,$ hence after the last theorem

$$(\mathscr{L}\cos\omega t)(s) \;=\; \frac{s}{\omega}(\mathscr{L}\sin\omega t)(s) \;=\; \frac{s}{s^2+\omega^2}$$

3) $f(t) \;=\; t^\alpha\,,\;\;,\; s>0$

$$(\mathscr{L} f)(s) \;=\; \int_0^\infty \mathrm{e}^{-st} t^\alpha\,\mathrm{d}t \qquad \text{with substitution: } u = st\,,\ \mathrm{d}u = s\mathrm{d}t$$

$$=\; \frac{1}{s^{\alpha+1}}\int_0^\infty \mathrm{e}^{u} u^\alpha\,\mathrm{d}u \;=\; \frac{\Gamma(\alpha+1)}{s^{\alpha+1}}\,,\qquad \alpha > -1$$

2. Multiple Integrals

An introduction to integration theory in two steps

I. Declaration of an integral in the n-dimensional space $\mathbb{R}^n$ for a restricted class of functions $C_0(\mathbb{R}^n)$ by making use of the ordinary integral. Study of characteristic properties.

II. Extension of the function class by approximation of new functions with functions in $C_0(\mathbb{R}^n)$ and transfer of the characteristic properties.

2.1 Integral for a restricted Function Class

The class $C_0(\mathbb{R}^n)$ consists of the continuous real functions $f : \mathbb{R}^n \mapsto \mathbb{R}$ which, outside of a (for f) sufficiently large limited set, only assume the value 0 (in short: continuous functions which vanish outside of a compactum).

Hence, in the case $n = 1$ there are for $f \in C_0(\mathbb{R})$ numbers $a < b$ with

$$f(t) \neq 0 \quad \Rightarrow \quad t \in [a, b]$$

Integral: (formally over the entire $\mathbb{R}^1$)

$$\int\limits_{\mathbb{R}} f(t)\,\mathrm{d}t \;=\; \int\limits_a^b f(t)\,\mathrm{d}t$$

It is independent of the choice of auxiliary numbers a, b.

Important properties:

Linearity:

$$\int (f+g) \;=\; \int f + \int g \;; \qquad \forall\; f, g \in C_0(\mathbb{R}^n)$$

$$\int (\lambda f) \;=\; \lambda \int f \quad ; \qquad \forall\; \lambda \in \mathbb{R}$$

Monotony:

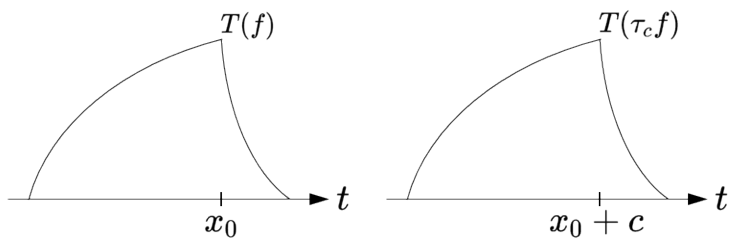
$$f \le g \;\Rightarrow\; \int f \;\le\; \int g \;; \qquad \forall\; f,\, g \,\in\, C_0(\mathbb{R}^n)$$

Translation invariance:

$$f \,\in\, C_0(\mathbb{R}^n)\,, \quad c \in \mathbb{R}^n$$

$$\tau_c(f)(x) \;=\; f(x-c)$$

$$\Downarrow$$



$$\int f \;=\; \int \tau_c f \quad ; \qquad \forall\; f \,\in\, C_0(\mathbb{R}^n)\;, \quad \forall\; c \in \mathbb{R}$$

Now let $n > 1$ and $f \,\in\, C_0(\mathbb{R}^n)$, then there is a $b > 0$ with

$$f(\vec{x}) \;=\; 0 \quad \text{if} \quad \max_{1 \le i \le n} |x_i| \;\ge\; b$$

For arbitrary $(x_2,\, x_3,\, \ldots,\, x_n) \,\in\, \mathbb{R}^{n-1}$

$$g(x_2,\, \ldots,\, x_n) \;:=\; \int_{\mathbb{R}} f(t, x_2,\, \ldots,\, x_n)\,\mathrm{d}t$$

is defined and, according to the theorem in chapter 1 about integrals with parameters with continuous integrands, g is continuous. But if

$$\max_{2 \le i \le n} |x_i| \;\ge\; b \;, \quad \text{then}$$

$$f(t, x_2,\, \ldots,\, x_n) \;=\; 0 \quad \text{for all } t \in \mathbb{R}$$

Hence $g(x_2,\, \ldots,\, x_n) \;=\; 0$ for all $t \in \mathbb{R}$.

Precisely: $\quad g \,\in\, C_0(\mathbb{R}^{n-1})$

Through multiple application

$$\boxed{\int_{\mathbb{R}^n} f(x_1, x_2, \ldots, x_n)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \;=\; \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x_1, x_2, \ldots, x_n)\,\mathrm{d}x_1 \right) \mathrm{d}x_2 \ldots \mathrm{d}x_n}$$

a linear, monotonic and translation-invariant functional on $C_0(\mathbb{R}^n)$ is declared at the same time.

Examples:

(1) Volume of a hemisphere $n = 2\,,\ R > 0$

$$f(x,y) \;=\; \begin{cases} \sqrt{R^2 - x_1^2 - x_2^2} & \text{if} \quad x_1^2 + x_2^2 \le R^2 \\ 0 & \text{if} \quad x_1^2 + x_2^2 > R^2 \end{cases}$$

$$\int\limits_{\mathbb{R}^2} f(x_1,x_2)\,\mathrm{d}x_1\mathrm{d}x_2 \;=\; \int\limits_{\mathbb{R}} \Bigg(\int\limits_{-\sqrt{R^2-x_2^2}}^{\sqrt{R^2-x_2^2}} \sqrt{R^2 - x_1^2 - x_2^2}\,\mathrm{d}x_1 \Bigg)\,\mathrm{d}x_2 \;=\; \frac{2\pi}{3}\,R^3$$

(2) Let $\varphi_1,\,\varphi_2,\,\ldots,\,\varphi_n \in C_0(\mathbb{R})$

With that

$$f(\vec{x}) \;\coloneqq\; \varphi_1(x_1)\cdot\varphi_2(x_2)\cdot\,\ldots\,\cdot\varphi_n(x_n)$$

becomes a function $f \in C_0(\mathbb{R}^n)$

$$\int\limits_{\mathbb{R}^n} f(\vec{x})\,\mathrm{d}x_1\ldots\mathrm{d}x_n \;=\; \Big(\int \varphi_1(x_1)\,\mathrm{d}x_1\Big)\cdot\,\ldots\,\cdot\Big(\int \varphi_n(x_n)\,\mathrm{d}x_n\Big)$$

$$=\; \prod_{m=1}^{n}\int\limits_{\mathbb{R}} \varphi_m(t)\,\mathrm{d}t$$

Remark:

(1) On translation invariance:
If $f \in C_0(\mathbb{R})$ lies outside $[a,b]$ then the shifted function $\tau_c f$ (declared by $\tau_c(f)(x) \coloneqq f(x-c)$) disappears outside of $[a+c, b+c]$. The integral becomes

$$\int\limits_{\mathbb{R}} (\tau_c f)(x)\,\mathrm{d}x \;=\; \int\limits_{a+c}^{b+c} f(x-c)\,\mathrm{d}x\;, \qquad \text{with} \quad x - c = t$$

$$=\; \int\limits_{a}^{b} f(t)\,\mathrm{d}t \;=\; \int\limits_{\mathbb{R}} f(t)\,\mathrm{d}t$$

2.1.1 A Characterization of the Integral

Theorem: If $I : C_0(\mathbb{R}^n) \mapsto \mathbb{R}$ is any monotonic and translation-invariant functional, then there is exactly one $c \geq 0$ with

$$\boxed{I(f) \;=\; c \int\limits_{\mathbb{R}^n} f(\vec{x})\,\mathrm{d}x_1\,\mathrm{d}x_2 \ldots \mathrm{d}x_n} \qquad \text{for all } f \in C_0(\mathbb{R}^n)$$

Remark:

(2) To calculate the constant c it is sufficient to know $I(f_0)$ for a single function $f_0 \in C_0(\mathbb{R}^n)$ with

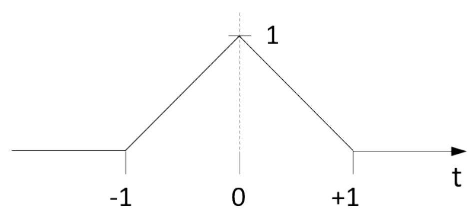
$$\int\limits_{\mathbb{R}^n} f(\vec{x})\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \;\neq\; 0$$

Because then $$c \;=\; \frac{I(f_0)}{\int\limits_{\mathbb{R}^n} f(\vec{x})\,\mathrm{d}x_1 \ldots \mathrm{d}x_n}$$

Proof sketch:

1) The spike function

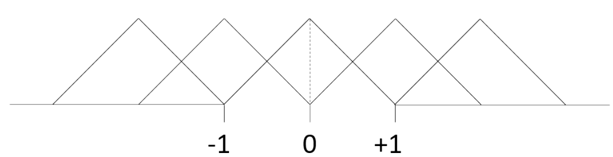
$$\lambda_0(t) \;=\; \max(0, 1-|t|) \quad \text{solved in } C_0(\mathbb{R})$$



$$\int\limits_{\mathbb{R}} \lambda_0(t)\,\mathrm{d}t \;=\; 1 \quad ; \quad \lambda_1(t) \;=\; \lambda_0(2t)$$

$$\sum_{g\in\mathbb{Z}} (\tau_g\lambda_0)(t) \;=\; 1$$

$$\lambda_0 \;=\; \frac{1}{2}\tau_{-\frac{1}{2}}\lambda_1 \,+\, \lambda_1 \,+\, \frac{1}{2}\tau_{+\frac{1}{2}}\lambda_1$$



2) In the n-dimensional case one has with $\lambda_{k+1}(t) \;=\; \lambda_k(2t)$

$$\Lambda_k(\vec{x}) \;=\; \lambda_k(x_1) \cdot \lambda_k(x_2) \cdot \ldots \cdot \lambda_k(x_n)$$

This results in a "partition of the number one"

$$\sum_{g\in 2^{-k}\mathbb{Z}^n} \tau_g\Lambda(\vec{x}) \;=\; 1 \quad \forall\; x$$

3) For $f \in C_0(\mathbb{R}^n)$ one forms the sequences

$$f_k(\vec{x}) \;=\; \sum_{g\,\in\,2^{-k}\mathbb{Z}^n} f(g)(\tau_g\Lambda_k)(\vec{x})$$

Result: f_k converges uniformly to f.

Lemma: Monotonic linear functionals are interchangeable with uniform limit values.

$\square$

2.1.2 Transformation Formula of the Multiple Integral under Linear Mappings

Linear self-mappings A of $\mathbb{R}^n$ have the characteristic properties

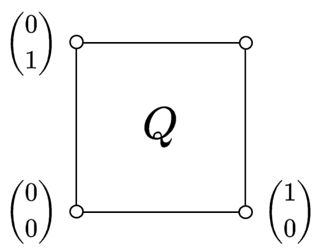
$$\begin{aligned} A(\vec{x}+\vec{y}) \;&=\; A(\vec{x}) + A(\vec{y}) \\ A(\lambda\vec{x}) \;&=\; \lambda(A(\vec{x})) \qquad \forall\ \vec{x},\vec{y} \in \mathbb{R}^n,\ \lambda \in \mathbb{R} \end{aligned}$$

A is invertible if and only if $\det A = 0$.

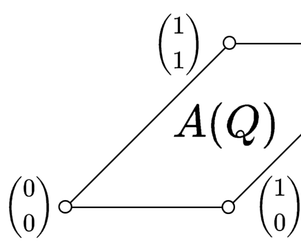
Example:

(3) For $n = 2$ (description of A by matrices)

Consider the square Q

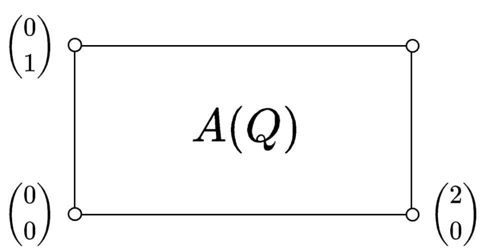


$$A \;=\; \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \;\Rightarrow$$



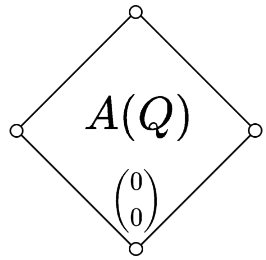
$\det A \;=\; 1$

$$A \;=\; \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \;\Rightarrow$$



$\det A \;=\; 2$

$$\left.\begin{array}{c} A \;=\; \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \\ \text{with} \\ c^2 + s^2 \;=\; 1 \end{array}\right\} \Rightarrow$$



$$\det A \;=\; 1$$

For every invertible linear mapping A of $\mathbb{R}^n$ holds

$$\boxed{\int_{\mathbb{R}^n} f(x_1, x_2, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n \;\coloneqq\; \int_{\mathbb{R}^n} f \circ A(\vec{x}) |\det A|\, \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_n}$$

for all $f \in C_0(\mathbb{R}^n)$

<u>Remark:</u>

(3) $|\det A|$ is an area or respectively a volume distortion factor.

Proof idea:

Reduction to the characterization theorem

$$I_A(f) \;=\; c \int_{\mathbb{R}^n} f \circ A(\vec{x})\, \mathrm{d}x_1\, \mathrm{d}x_2 \ldots \mathrm{d}x_n \qquad \text{for all } f \in C_0(\mathbb{R}^n)$$

This is a linear, monotonic and translation-invariant functional on $C_0(\mathbb{R}^n)$.

Linearity:

$$(f+g)(A\vec{x}) \;=\; f(A\vec{x}) + g(A\vec{x}) \;\Rightarrow$$

$$I_A(f+g) \;=\; I_A(f) + I_A(g)$$

$$I_A(\lambda f) \;=\; \lambda I_A(f)$$

Monotony:

$$f \leq g \;\Rightarrow\; f(A\vec{x}) \leq g(A\vec{x}) \quad \forall\ \vec{x}$$

$$f \circ A \leq g \circ A \;\Rightarrow\; I_A(f) \leq I_A(g)$$

Translation invariance: $\vec{c} \in \mathbb{R}^n$

$$(\tau_{\vec{c}}(f \circ A))(\vec{x}) \;=\; f \circ A(\vec{x} - \vec{c})$$

$$= f(A(\vec{x} - \vec{c})) \qquad (A \text{ is linear})$$

$$= f(A\vec{x} - A\vec{c}) \;=\; (\tau_{A\vec{c}})(A\vec{x})$$

hence $\quad \tau_{\vec{c}}(f \circ A) \;=\; (\tau_{A\vec{c}} f) \circ A$

Result from the characterization theorem

$$I_A(f) \;=\; c_A I(f) \quad \text{with a} \quad c_A \geq 0$$

Since matrix multiplication means image composition, it follows that $c_{AB} \;=\; c_A c_B$. Consequence from the method of elementary transformations: The constant c_A is the absolute value $|\det A|$.

□

2.2 Extension of the Class of Integrable Functions

Let a bounded function $f : \mathbb{R}^n \mapsto \mathbb{R}$ be given. Question: Do approximating functions $\varphi,\, \psi \,\in\, C_0(\mathbb{R}^n)$ exist for every $\epsilon > 0$ with $\varphi \leq f \leq \psi$ and

$$0 \;=\; \int\limits_{\mathbb{R}^n} (\psi - \varphi)(\vec{x})\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \;<\; \epsilon\,?$$

If this is the case, then f is integrable and the integral is

$$\int\limits_{\mathbb{R}^n} f(\vec{x})\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \;=\; \sup_{\varphi} \int\limits_{\mathbb{R}^n} \varphi(\vec{x})\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \;=\; \inf_{\psi} \int\limits_{\mathbb{R}^n} \psi(\vec{x})\,\mathrm{d}x_1 \ldots \mathrm{d}x_n$$

In particular, for partitions B of $\mathbb{R}^n$ an n-dimensional volume is declared using the so-called characteristic function

$$\chi_B(\vec{x}) \;:=\; \begin{cases} 1\,, & \text{if} \quad \vec{x} \in B \\ 0\,, & \text{if} \quad \vec{x} \notin B \end{cases}$$

Definition:

$B \subset \mathbb{R}^n$ is called "measurable" if χ_B is integrable in the above sense and then

$$\mathrm{vol}_n(B) \;:=\; \int\limits_{\mathbb{R}^n} \chi_B(\vec{x})\,\mathrm{d}x_1 \ldots \mathrm{d}x_n$$

Example:

(4) On controlling the abstract volume concept: $n = 3$

$$B = \{x_1, x_2, x_3 \,|\, a_1 \le x_1 \le b_1,\ a_2 \le x_2 \le b_2,\ a_3 \le x_3 \le b_3\}$$

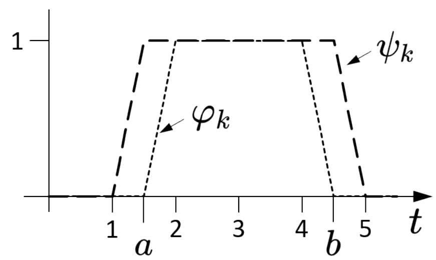
χ_B is not continuous, there are discontinuities on the edge of the cuboid.

Auxiliary functions

$$\varphi_k(t, a, b)\,,\ \psi_k(t, a, b)$$

sketched on the right for:

$$k = 1,\ a = 1.5,\ b = 4.5$$



$$\varphi_k(t, a, b) = \begin{cases} \min\bigl(1, 2^{k-1}(b-a) - \bigl|2^k t - 2^{k-1}(a+b)\bigr|\bigr)\,, & \text{if } t \in [a, b] \\ 0\,, \ \text{if } t \notin [a, b] \end{cases}$$

$$\psi_k(t, a, b) = \begin{cases} \max\bigl(0, 1 + 2^{k-1}(b-a) - \bigl|2^k t - 2^{k-1}(a+b)\bigr|\bigr)\,, & \text{if } t \in [a, b] \\ 1\,, \ \text{if } t \notin [a, b] \end{cases}$$

$$\Phi_k(\vec{x}) := \prod_{m=1}^{3} \varphi_k(x_m, a_m, b_m)$$

$$\Psi_k(\vec{x}) := \prod_{m=1}^{3} \psi_k(x_m, a_m, b_m)$$

With example (2)

$$\int_{\mathbb{R}^3} \Phi_k(\vec{x})\, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 = \prod_{m=1}^{3} \bigl(b_m - a_m - 2^{-k}\bigr)$$

$$\int_{\mathbb{R}^3} \Psi_k(\vec{x})\, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 = \prod_{m=1}^{3} \bigl(b_m - a_m + 2^{-k}\bigr)$$

Hence $(k \to \infty)$

$$\mathrm{vol}_3(B) := \int_{\mathbb{R}^3} \chi_B(\vec{x})\, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 = (b_1 - a_1)(b_2 - a_2)(b_3 - a_3)$$

Remark:

(4) On the characteristic function χ_B

It allows area integrals to be declared by $(B \subset \mathbb{R}^n)$

$$\int_B f(\vec{x})\,\mathrm{d}x_1\ldots\mathrm{d}x_n \;=\; \int_{\mathbb{R}^n} \chi_B(\vec{x})f(\vec{x})\,\mathrm{d}x_1\ldots\mathrm{d}x_n$$

Formal advantage: The geometric complications of B appear later.

(5) On the volume of a cuboid

Instead of integrating over χ_B one could have integrated over every function f which only agrees inside and outside the cuboid B with χ_B, but on the boundary is arbitrary with $f(\vec{x}) \in [0,1]$.

The result with the same auxiliary functions Φ_k, Ψ_k

$$\mathrm{vol}_3(B) \;=\; \int \chi_B \;=\; \int_{\mathbb{R}^3} f(\vec{x})\,\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$$

2.2.1 Reversal of the Order of Integration

$f(x_1,\ldots,x_n)$ is integrable over $\mathbb{R}^n$ and $t = x_k$ is one of the variables and

$$(x_1,\ldots x_{k-1},x_{k+1},\ldots x_n) \;\longmapsto\; f(x_1,\ldots x_{k-1},t,x_{k+1},\ldots x_n)$$

be integrable over $\mathbb{R}^{n-1}$ for every $t \in \mathbb{R}$. Then

$$F(t) \;=\; \int_{\mathbb{R}^{n-1}} f(x_1,\ldots x_{k-1},t,x_{k+1},\ldots x_n)\,\mathrm{d}x_1\ldots\,\mathrm{d}x_n \qquad (\mathrm{d}x_k \text{ missing!})$$

is integrable over $\mathbb{R}$ and it holds

$$\int_{\mathbb{R}} F(t)\,\mathrm{d}t \;=\; \int_{\mathbb{R}^n} f(x_1,\ldots x_n)\,\mathrm{d}x_1\ldots\mathrm{d}x_n$$

Proof:

Simplified notation: $k = 1$

For $\epsilon > 0$ there exists according to the assumption about f : φ, $\psi \in C_0(\mathbb{R}^n)$ with $\varphi \le f \le \psi$

$$0 \;\le\; \int_{\mathbb{R}^n} (\psi-\varphi)(\vec{x})\,\mathrm{d}x_1\ldots\mathrm{d}x_n \;<\; \epsilon$$

$$\Phi(x_1) \;=\; \int_{\mathbb{R}^{n-1}} \varphi(x_1,x_2,\ldots x_n)\,\mathrm{d}x_2\ldots\mathrm{d}x_n$$

$$\Psi(x_1) \;=\; \int_{\mathbb{R}^{n-1}} \psi(x_1,x_2,\ldots x_n)\,\mathrm{d}x_2\ldots\mathrm{d}x_n$$

According to the last chapter, for $\Phi,\,\Psi \in C_0(\mathbb{R}^n):\ \Phi(x_1) \le F(x_1) \le \Psi(x_1)$ and

$$\int\limits_{\mathbb{R}} (\Psi - \Phi)(x_1)\,\mathrm{d}x_1 \;=\; \int\limits_{\mathbb{R}^n} (\psi - \varphi)(x_1, \dots, x_n)\,\mathrm{d}x_1 \dots \mathrm{d}x_n \;<\; \epsilon$$

Hence F is integrable.

□

2.2.2 Application: Cavalieri's Principle

Let B be a set measurable in $\mathbb{R}^3$ and for each $t \in \mathbb{R}$ let the section

$$B_t \;:=\; \{(x,y)\,;\ (x,y,t) \in B\}$$

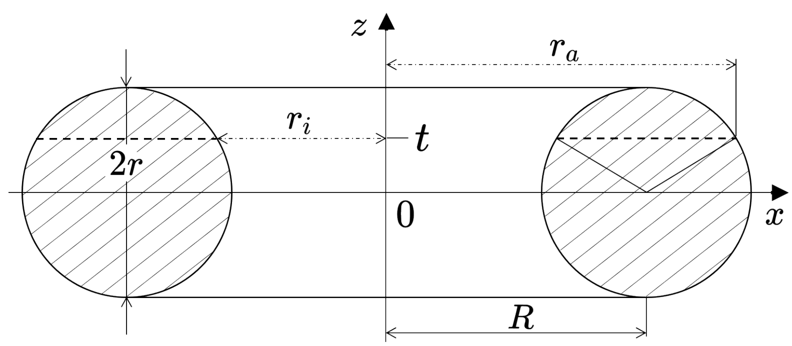
be measurable in the plane. With that

$$\boxed{\mathrm{vol}_3(B) \;=\; \int\limits_{\mathbb{R}} \mathrm{vol}_2(B_t)\,\mathrm{d}t}$$

Example:

(5) The torus volume

$r \in\,]0, R[\,,$ circle with center $(R,0,0)$ in the x,z-plane rotating around the z-axis.



The section B_t becomes empty if $|t| > r$ and an annulus otherwise.

Inner radius r_i : $R - \sqrt{r^2 - t^2}$; Outer radius r_a : $R + \sqrt{r^2 - t^2}$

$$\begin{aligned}
\mathrm{vol}_2(B_t) &= \pi\Big[\big(R + \sqrt{r^2 - t^2}\big)^2 - \big(R - \sqrt{r^2 - t^2}\big)^2\Big] \\
&= 4\pi R\sqrt{r^2 - t^2}
\end{aligned}$$

$$\mathrm{vol}_3(B) \;=\; \int\limits_{-r}^{r} 4\pi R\sqrt{r^2 - t^2}\,\mathrm{d}t$$

$$= \; 4\pi R \left[\frac{t}{2}\sqrt{r^2 - t^2} \; + \; \frac{r^2}{2} \arcsin \frac{t}{r} \right]_{t=-r}^{t=+r}$$

Hence: $\quad \mathrm{vol}_3(B) \; = \; 2\pi^2 R \, r^2$

Example:

(6) The volume of a general cone

Let a measurable set B_0 be given in the x, y-plane, $h > 0$

$$B \; = \; \{(x(1-t)\,,\, y(1-t)\,,\, th)\,; \;\; (x,y) \in B_0\,, \;\; 0 \le t \le 1\}$$

$$\mathrm{vol}_3(B) \; = \; \frac{1}{3}\, h \, \mathrm{vol}_2(B_0)$$

$$B_{th} \; = \; \{(x,y)(1-t)\,; \;\; (x,y) \in B_0\}$$

Transformation formula in $\mathbb{R}^2$

$$\mathrm{vol}_2(B_{th}) \; = \; \mathrm{vol}(B_0)\,(1-t)^2$$

$$\mathrm{vol}_3(B) \; = \; \int\limits_{\mathbb{R}} \mathrm{vol}_2(B_s)\,\mathrm{d}s$$

$$= \; \mathrm{vol}(B_0)\, h \int\limits_0^1 (1 - t^2)\,\mathrm{d}t \; = \; \frac{1}{3}\, h \, \mathrm{vol}(B_0)$$

On the transformation formula in $\mathbb{R}^n$:

Based on the substitution rule in $\mathbb{R}^1$

$$\int\limits_{\varphi(a)}^{\varphi(b)} f(x)\,\mathrm{d}x \; = \; \int\limits_a^b f \circ \varphi(\xi) \varphi'(\xi)\,\mathrm{d}\xi$$

where the parameter transformation φ is strictly monotonically increasing and continuously differentiable.

2.3 Formulation of the General Transformation Formula for Area Integrals

Let ϕ be an invertible and differentiable mapping from an open set $U \subset \mathbb{R}^n$ to an open set $V \subset \mathbb{R}^n$. Then for integrable functions f

$$\boxed{\int\limits_{\phi(U)} f(\vec{x})\,\mathrm{d}x_1\ldots\mathrm{d}x_n \;=\; \int\limits_{U} f\circ\phi(\vec{\xi}\,)\left|\det\frac{\partial\phi_i}{\partial\xi_i}\right|\mathrm{d}\xi_1\ldots\mathrm{d}\xi_n}$$

Note on the **proof** of validity:

First reason for validity: The transformation formula (TF) applies to linear mappings A of $\mathbb{R}^n$.

Second reason: Differentiability means linear approximability. The validity scope can be extended by limit value processes.

Examples:

(0) The case $\phi(\vec{x}) = A\vec{x}$ of linearly invertible mappings of $\mathbb{R}^n$: Because of

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

the Jacobian matrix of the mappings $\vec{x}\longmapsto A\vec{x}$ is equal to the constant matrix A .

(1) Plane polar coordinates

$$\phi(x,y) = (r\cos\varphi\,,\,r\sin\varphi) \qquad 0<r\ ,\ 0<\varphi<2\pi$$

The Jacobian matrix

$$\frac{\partial\phi}{\partial(r,\varphi)} = \begin{bmatrix}\cos\varphi & -r\sin\varphi\\ \sin\varphi & r\cos\varphi\end{bmatrix}$$

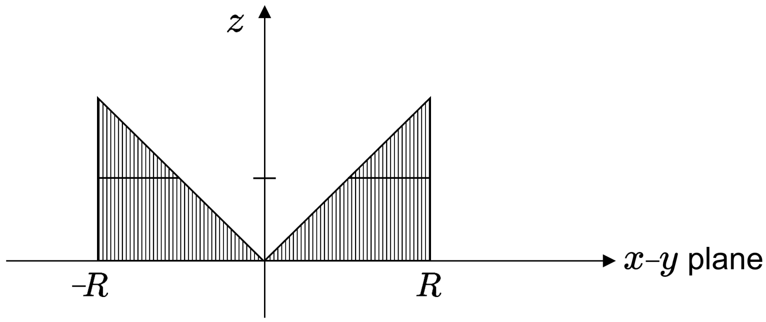
$$\det\frac{\partial\phi}{\partial(r,\varphi)} = r\big(\cos^2\varphi+\sin^2\varphi\big) = r$$

The formula is favorable in cases when the integrand $f(x,y)$ is constant on the circles $x^2+y^2 = r^2$.

For example

$$\int\limits_{x^2+y^2\leq R^2}\sqrt{x^2+y^2}\,\mathrm{d}x\mathrm{d}y \overset{\mathrm{TF}}{=} \int\limits_0^R\int\limits_0^{2\pi}\sqrt{r^2\cos^2\varphi+r^2\sin^2\varphi}\;r\,\mathrm{d}\varphi\mathrm{d}r$$

$$\int\limits_0^R\int\limits_0^{2\pi} r\,r\,\mathrm{d}\varphi\mathrm{d}r = 2\pi\int\limits_0^R r^2\,\mathrm{d}r = \frac{2\pi}{3}R^3$$



Geometric meaning:
The volume of the rest of the body was created from a right circular cylinder with the base area $R^2\pi$ and the height R, from which a circular cone with the same base area and height with the apex at the zero point is removed.

The area of the section B_t of a hemisphere of radius R is

$$\mathrm{vol}_2(B_t) \;=\; \left(\sqrt{R^2 - t^2}\right)^2 \pi \;=\; (R^2 - t^2)\pi$$

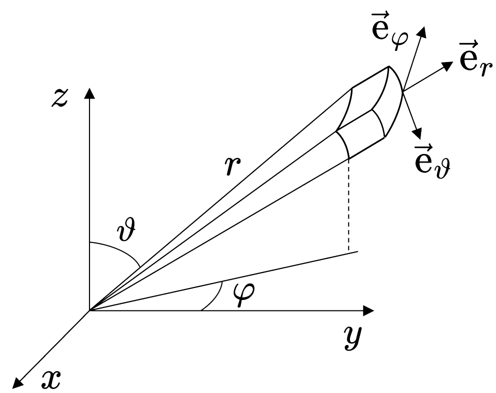
The right side shows that the section of the rest of the body above at height t has the same area as the section B_t of a hemisphere of radius R.

(2) Spherical coordinates

$$x \;=\; r \sin\vartheta \cos\varphi$$

$$y \;=\; r \sin\vartheta \sin\varphi$$

$$z \;=\; r \cos\vartheta$$



The surfaces $r \;=\;$ const are spherical surfaces around the origin.
The surfaces $\vartheta \;=\;$ const are cone surfaces with the vertex at the origin.
The surfaces $\varphi \;=\;$ const are half-planes through the z-axis.

The Jacobian matrix

$$\frac{\partial(x,y,z)}{\partial(r,\varphi,\vartheta)} \;=\; \begin{bmatrix} \sin\vartheta\cos\varphi & r\cos\vartheta\cos\varphi & -r\sin\vartheta\sin\varphi \\ \sin\vartheta\sin\varphi & r\cos\vartheta\sin\varphi & r\sin\vartheta\cos\varphi \\ \cos\vartheta & -r\sin\vartheta & 0 \end{bmatrix}$$

$$\frac{\partial(x,y,z)}{\partial(r,\varphi,\vartheta)} \;=\; r^2 \sin\vartheta$$

3. Integral Theorems in the Plane

Compilation of previous concepts and theorems about curves.

Curve: $\vec{\gamma} \,:\, [a,b] \;\mapsto\; \mathbb{R}$ continuous and piecewise differentiable.

Curve length: $l(\vec{\gamma}) \;:=\; \int\limits_a^b \left\|\dot{\vec{\gamma}}(t)\right\| \mathrm{d}t$ invariant under parameter transformation.

Line integral over a vector field $\vec{v}$ (more precisely: over the tangential component of $\vec{v}$, with $\vec{v}$ along $\vec{\gamma}$ continuous)

$$\int\limits_{\vec{\gamma}} \vec{v}(\vec{x})\,\mathrm{d}\vec{x} \;:=\; \int\limits_a^b \big(\vec{v}(\vec{\gamma}(t))\,\dot{\vec{\gamma}}(t)\big)\,\mathrm{d}t$$

invariant under parameter transformation.

The curve $\vec{\gamma}^*$ opposite to $\vec{\gamma}$

$$\vec{\gamma}^* \;:=\; \vec{\gamma}(a+b-t) \;\;;\quad t \in [a,b]$$

The substitution rule gives

$$\int\limits_{\vec{\gamma}^*} \vec{v}(\vec{x})\,\mathrm{d}\vec{x} \;=\; -\int\limits_{\vec{\gamma}} \vec{v}(\vec{x})\,\mathrm{d}\vec{x}$$

In the plane: vectorial area pieces

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

where $\vec{x}(\varphi)$ is a continuously differentiable function of the angle φ. For the third coordinate

$$\vec{F}_3 \;=\; F \;=\; \frac{1}{2}\int\limits_a^b r^2(\varphi(t))\,\varphi'(t)\,\mathrm{d}t$$

The gradient of a function $f : U \mapsto \mathbb{R}\,,\; U$ region in $\mathbb{R}^n$

$$\operatorname{grad} f \;=\; \left(\frac{\partial t}{\partial x_i}\right)_{1 \le i \le n}$$

The line integral along $\vec{\gamma}$ over the gradient field $\vec{v} = \operatorname{grad} f$ is the difference of the values of f in the endpoints of the curve!

3.1 Integrability Criterion

A continuously differentiable vector field $\vec{v}$ in a star-shaped (generally simply connected) region of $\mathbb{R}^n$ is a gradient field of a potential if and only if

$$D_i v_k \;=\; D_k v_i \qquad 1 \le i, k \le n$$

where D_i denotes the partial derivative with respect to x_i.

In three-dimensional space $\mathbb{R}^3$, written with $\nabla \times \vec{v} = \operatorname{rot} \vec{v}$, this gives the equation

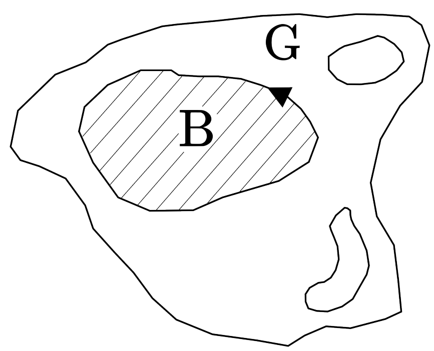
$$\boxed{\operatorname{rot} \vec{v} \;=\; \vec{0}}$$

As an extension of the theorem for line integrals over gradient fields,

Riemann's formula (also Green's formula):

The case $n = 2$:

Let $\vec{v} = (u, v)$ be a continuously differentiable vector field on the region G of the plane. Furthermore, let B be a compactum in G, which is bounded by a simply closed curve $\vec{\gamma}$. Then



$$\boxed{\int\limits_B \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) \mathrm{d}x\mathrm{d}y \;=\; \int\limits_{\vec{\gamma}} \vec{v}(\vec{x})\, \mathrm{d}\vec{x}}$$

where the boundary curve $\vec{\gamma}$ is oriented such that B lies "to the left".

Remark:

(1) Suggestive notation for the right side

$$\int_{\partial B} \vec{v}\,\mathrm{d}\vec{x} \;=\; \oint_{\partial B} \vec{v}\,\mathrm{d}\vec{x}$$

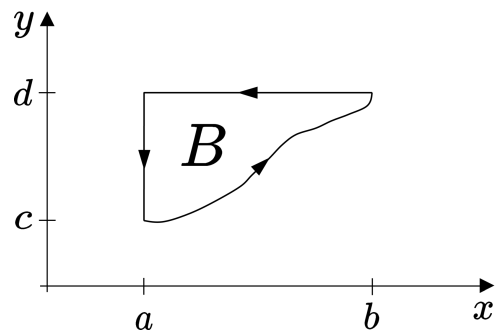
Proof for triangular shaped regions

With a continuously differentiable strictly monotone function

$\delta \;:\; [a,b] \mapsto \mathbb{R}\,,$

for which δ' is without zeros, let

$B \;=\; \{(x,y)\,;\;\; a \le x \le b\,,\;\; \delta(x) \le y \le d\,\}$



Let $\varphi :\; [c,d] \mapsto [a,b]$ denote the inverse function of δ

$$\int_B \Big(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\Big)\,\mathrm{d}x\mathrm{d}y \;=\; \int_B \Big(\frac{\partial v}{\partial x}\,\mathrm{d}x\Big)\,\mathrm{d}y \;-\; \int_B \Big(\frac{\partial u}{\partial y}\,\mathrm{d}y\Big)\,\mathrm{d}x$$

$$=\; \int_c^d \Big(\int_a^{\varphi(y)} \frac{\partial v}{\partial x}\,\mathrm{d}x\Big)\,\mathrm{d}y \;-\; \int_a^b \Big(\int_{\delta(x)}^d \frac{\mathrm{d}u}{\mathrm{d}y}\,\mathrm{d}y\Big)\,\mathrm{d}x$$

$$=\; \int_c^d \big(v(\varphi(y),y) - v(a,y)\big)\,\mathrm{d}x \;-\; \int_a^b \big(u(x,d) - u(x,\delta(x))\big)\,\mathrm{d}x$$

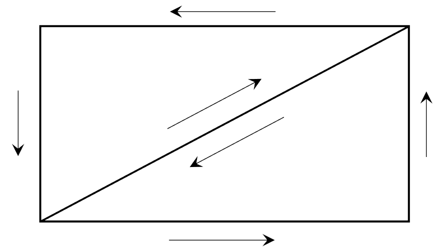
Substitution $\;y = \delta(t)$

$$=\; -\int_c^d v(a,y)\,\mathrm{d}y \;+\; \int_a^b \big(u(t,\delta(t)) + v(t,\delta(t))\,\delta'(t)\big)\,\mathrm{d}t \;-\; \int_a^b u(x,d)\,\mathrm{d}x$$

Considering the special parameterization of the boundary curve of the triangle, this is the assertion.

□

The treatment of general regions can be reduced to our case by decomposing regions into triangles by using internal auxiliary line segments. Note: Integrals over auxiliary line segments are carried out in both possible directions. The corresponding integrals cancel each other out.



If one imagines $\vec{v} = (u, v)$ continued by a coordinate O into three-dimensional space, then the integrand in the area integral represents the third coordinate of the rotation, the Riemann formula is therefore a (plane) version of Stoke's integral theorem

$$\boxed{\int_B \operatorname{rot} \vec{v}(\vec{x})\, \mathrm{d}\vec{F} \;=\; \int_{\partial B} \vec{v}(\vec{x})\, \mathrm{d}x}$$

Special case: $u = -y$, $v = x$. The integrand on the left becomes the constant 2 .

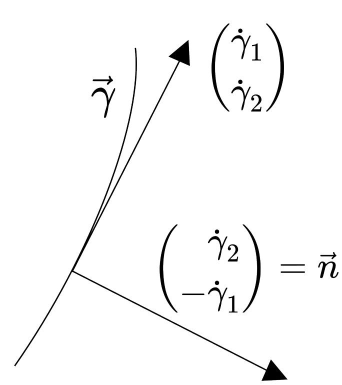
$$\boxed{\mathrm{vol}_2 B \;=\; \frac{1}{2} \int_{\partial B} (-y\, \mathrm{d}x + x\, \mathrm{d}x)}$$

Replace u with $-v$ and v with u in the Riemann formula. Then left in the formula

$$\int_B \underbrace{\Big(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \Big)}_{\operatorname{div} \vec{v}(\vec{x})} \mathrm{d}x\mathrm{d}y$$

In the formula rewrite the integral on the right

$$\begin{aligned}
\int_{\vec{\gamma}} \begin{pmatrix} -v \\ u \end{pmatrix} \mathrm{d}\vec{x} \;&=\; \int_a^b \begin{pmatrix} -v(\gamma_1(t), \gamma_2(t)) \\ u(\gamma_1(t), \gamma_2(t)) \end{pmatrix} \begin{pmatrix} \dot{\gamma}_1(t) \\ \dot{\gamma}_2(t) \end{pmatrix} \mathrm{d}t \\
&= \int_a^b \begin{pmatrix} u(\gamma_1(t), \gamma_2(t)) \\ v(\gamma_1(t), \gamma_2(t)) \end{pmatrix} \begin{pmatrix} \dot{\gamma}_2(t) \\ -\dot{\gamma}_1(t) \end{pmatrix} \mathrm{d}t \\
&= \int_{\vec{\gamma}} \begin{pmatrix} u \\ v \end{pmatrix} \mathrm{d}\vec{n}
\end{aligned}$$



Gauß' theorem in the plane

$$\boxed{\int\limits_B \operatorname{div} \vec{v}(\vec{x})\, \mathrm{d}x\mathrm{d}y \;=\; \int\limits_{\partial B} \vec{v}(\vec{x})\, \mathrm{d}\vec{n}}$$

For the case $B = B_r =$ circular disc of small radius $r > 0$

$$\operatorname{div} \vec{v}(\vec{x}) \;=\; \lim_{r \searrow 0} \frac{1}{\mathrm{vol}_2(B_r)} \int\limits_{\partial B_r} \vec{v}(\vec{x})\, \mathrm{d}\vec{n}$$

3.2 The Cauchy-Riemann Differential Equations and the Cauchy Integral Theorem

In a region $D \subset \mathbb{C}$ let a function with complex values $f : D \mapsto \mathbb{C}$ be given, broken down into real and imaginary parts

$$z \in D \;=\; x + jy \;; \qquad x, y \in \mathbb{R}$$

$$f(z) \;=\; u(x,y) + jv(x,y) \qquad \text{with real functions } u\,,\, v$$

f is called "complex differentiable" in D, if

$$\lim_{\substack{h \to 0 \\ h \in \mathbb{C}^*}} \frac{f(z+h) - f(z)}{h} \qquad \text{exists for all} \quad z \in D$$

Designation of the derivative: $f'(z)$ in analogy to the real!

If f' is furthermore continuous in D, then it is holomorphic in D. Examples of holomorphic functions are provided by every power series

$$\sum_{n=0}^{\infty} a_n (z - z_0)^n$$

in the interior of its region of convergence. Decomposition of f into real and imaginary parts

$$f(z) \;=\; u(x,y) + jv(x,y)$$

$$z \;=\; x + jy \;; \qquad x, y \in \mathbb{R}$$

then the following formulas apply to u and v

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad ; \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}}$$ Cauchy-Riemann differential equations

Proof:

From the premise

$$f'(z) = \lim_{h \to 0} \left[\frac{u(x+h_1, y+h_2) - u(x,y)}{h_1 + j h_2} + j \frac{v(x+h_1, y+h_2) - v(x,y)}{h_1 + j h_2} \right]$$

Special approximation $h \to 0$

$$h_2 = 0 \ , \ h_1 \to 0 \quad ; \quad h_1 = 0 \ , \ h_2 \to 0$$

$$\operatorname{Re} f'(z) = \frac{\partial u}{\partial x} = \frac{\mathrm{d}v}{\mathrm{d}y}$$

$$\operatorname{Im} f'(z) = \frac{\partial v}{\partial x} = -\frac{\mathrm{d}u}{\mathrm{d}y}$$ $\square$

Declaration of a line integral along $\vec{\gamma} : [a,b] \mapsto D$ with continuous complex-valued integrand $f : D \mapsto \mathbb{C}$

$$\int_{\vec{\gamma}} f(z)\,\mathrm{d}z := \int_a^b \big[u(\gamma_1(t), \gamma_2(t)) + j\, v(\gamma_1(t), \gamma_2(t))\big]\big[\dot{\gamma}_1(t) + j\,\dot{\gamma}_2(t)\big]\,\mathrm{d}t$$

Written out

$$\int_{\vec{\gamma}} f(z)\,\mathrm{d}z = \int_a^b \big[u(\vec{\gamma}(t))\dot{\gamma}_1(t) - v(\vec{\gamma}(t))\dot{\gamma}_2(t)\big]\,\mathrm{d}t$$

$$+ j \int_a^b \big[v(\vec{\gamma}(t))\dot{\gamma}_1(t) + u(\vec{\gamma}(t))\dot{\gamma}_2(t)\big]\,\mathrm{d}t$$

$$= \int_{\vec{\gamma}} \begin{pmatrix} u \\ -v \end{pmatrix} \mathrm{d}\vec{x} + j \int_{\vec{\gamma}} \begin{pmatrix} v \\ u \end{pmatrix} \mathrm{d}\vec{x}$$

Estimation of the integrand using the Cauchy-Schwarz inequality

$$|u\dot{\gamma}_1 - v\dot{\gamma}_2| \le \sqrt{u^2+v^2}\sqrt{\dot{\gamma}_1^2 + \dot{\gamma}_2^2}$$

or respectively

$$|v\dot{\gamma}_1 + u\dot{\gamma}_2| \le \sqrt{u^2+v^2}\sqrt{\dot{\gamma}_1^2 + \dot{\gamma}_2^2}$$

Hence

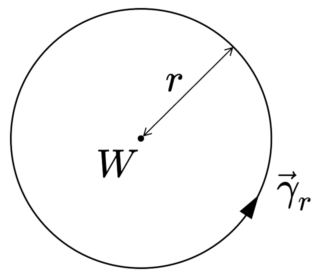
$$\left|\int\limits_{\vec\gamma} f(z)\,\mathrm{d}z\right| \;\le\; 2\cdot\max_{z\,\text{along}\,\vec\gamma}|f(z)|\cdot\int\limits_a^b \left\|\dot{\vec\gamma}\right\|\mathrm{d}t \;\le\; 2\cdot\max_{z\,\text{along}\,\vec\gamma}|f(z)|\cdot l(\vec\gamma)$$

Remark:

(2) Using Riemann sums one gets the estimate without the factor 2 on the right side.

Example:

(1) A line integral in $\mathbb{C}$



$w \in \mathbb{C}\ ,\quad t \in [0, 2\pi]$

$$\begin{aligned}\vec\gamma_r(t) &= w + r\mathrm{e}^{jt}\\ &= w + r(\cos t + j\sin t)\end{aligned}$$

$$\int\limits_{\vec\gamma_r}\frac{\mathrm{d}z}{z-w}\quad;\quad \dot{\vec\gamma}_r(t) = r(-\sin t + j\cos t) = rj\mathrm{e}^{jt}$$

$$\int\limits_{\vec\gamma_r}\frac{\mathrm{d}z}{z-w} = \int\limits_0^{2\pi}\frac{1}{\vec\gamma_r(t)-w}\,\dot{\vec\gamma}_r\,\mathrm{d}t = j\int\limits_0^{2\pi}\frac{r\mathrm{e}^{jt}}{r\mathrm{e}^{jt}}\,\mathrm{d}t = 2\pi j$$

$$\int\limits_{\vec\gamma_r}\frac{\mathrm{d}z}{z-w} = 2\pi j$$

Special case for calculating the winding number of the curve $\vec\gamma_r$ around w.

If u and v are continuously partially differentiable then

$$\left(\operatorname{rot}\begin{pmatrix}u\\-v\end{pmatrix}\right)_3 = -\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}$$

$$\left(\operatorname{rot}\begin{pmatrix}u\\v\end{pmatrix}\right)_3 = \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}$$

Both expressions vanish if $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ of a holomorphic function f.

3.2.1 Cauchy's Integral Theorem

If $f = u + jv$ is holomorphic on the region $D \subset \mathbb{C}$ and if $\vec\gamma : [a,b] \mapsto D$ is a ("closed") curve bounding a compactum $B \subset D$, then

$$\boxed{\int_{\vec{\gamma}} f(z)\,\mathrm{d}z \;=\; 0}$$

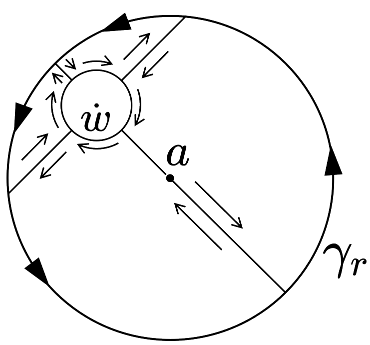
Main application: Cauchy's integral formula

Let f be holomorphic on the region $D \subset \mathbb{C}$. A circular disk $B_r(a)$ of radius $r > 0$ around a shall with its boundary be contained in D. Then

$$\boxed{f(w) \;=\; \frac{1}{2\pi j} \int_{\partial B_r(a)} \frac{f(z)}{z-w}\,\mathrm{d}z \;; \quad \text{if } |w-a| < r}$$

Proof by applying Cauchy's integral theorem to the auxiliary function

$$g(z) \;=\; \begin{cases} \dfrac{f(z)-f(w)}{z-w} & \text{if } w \neq z \\[2ex] f'(w) & \text{if } z = w \end{cases}$$



Because f is holomorphic in D, g becomes a continuous function in D, which is even holomorphic in $D \setminus \{w\}$.

According to the sketch with the auxiliary circle $B_\epsilon(w)$, $0 < \epsilon < (-|w-a| + r)$, integration proceeds over four closed auxiliary curves, each bounding a compactum in the holomorphic region of g.

According to the Cauchy integral theorem (taking into account the cancellation of the integrals along the straight line parts)

$$\int_{\partial B_r(a)} g(z)\,\mathrm{d}z \;=\; \int_{\partial B_\epsilon(w)} g(z)\,\mathrm{d}z \qquad (*)$$

Second step: Show that $\displaystyle\int_{\partial B_r(a)} g(z)\,\mathrm{d}z \;=\; 0$

Hereto consider on the right side of $(*)$ the limit transition $z \searrow 0$.

$|g(z)|$ is bounded on the closed disk $\overline{B_\epsilon(w)}$ by $|f'(w)|+1$.

Application of the integral estimate to the right-hand side (RS) of $(*)$

$$\text{RS} \;=\; 2\cdot \underbrace{l(\gamma_\epsilon(w))}_{2\pi\epsilon}\cdot(|f'(w)|+1)\,; \quad \text{this goes to 0 with } \epsilon \searrow 0$$

Applied to the left side (LS) of $(*)$ follows

$$\Big|\int\limits_{\partial B_r(a)} g(z)\,\mathrm{d}z\Big| \;=\; 0 \quad \text{hence} \quad \int\limits_{\partial B_r(a)} g(z)\,\mathrm{d}z \;=\; 0$$

Third step: Evaluate LS in $(*)$ in a different way

$$\begin{aligned}\text{LS} \;&=\; \int\limits_{\vec{\gamma}_r(a)} \Big(\frac{f(z)}{z-w} - \frac{f(w)}{z-w}\Big)\,\mathrm{d}z \\ &=\; \int\limits_{\vec{\gamma}_r(a)} \frac{f(z)}{z-w}\,\mathrm{d}z \;-\; f(w)\int\limits_{\vec{\gamma}_r(a)} \frac{1}{z-w}\,\mathrm{d}z\end{aligned}$$

With the integral example at the beginning

$$\text{LS} \;=\; \int\limits_{\partial B_r(a)} \frac{f(z)}{z-w}\,\mathrm{d}z \;-\; f(w)\underbrace{\int\limits_{\vec{\gamma}_\epsilon(*)} \frac{1}{z-w}\,\mathrm{d}z}_{2\pi j}$$

With step two

$$\int\limits_{\partial B_r(a)} \frac{f(z)}{z-w}\,\mathrm{d}z \;=\; 2\pi j\, f(w)$$

□

A simple consequence of the Cauchy integral formula:

A holomorphic function can be expanded into a convergent power series around every point a of its holomorphic region. In particular, $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ are at least twice continuously differentiable. Therefore from

$$\frac{\partial u}{\partial x} \;=\; \frac{\partial v}{\partial y} \quad , \quad \frac{\partial u}{\partial y} \;=\; -\frac{\partial v}{\partial x}$$

$$\Delta u \;=\; \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \;=\; \frac{\partial}{\partial x}\Big(\frac{\partial v}{\partial y}\Big) + \frac{\partial}{\partial y}\Big(-\frac{\partial v}{\partial x}\Big) \;=\; 0$$

u (and likewise v) is a solution of the Laplace equation

$$\boxed{\Delta u = 0}$$

that means u (and likewise v) is a harmonic function.

4. Surfaces in Space and Surface Integrals

4.1 Surface Representations

1. In the region $U \subset \mathbb{R}^2$ let f be real-valued and continuously differentiable

$$s \;=\; \Gamma_f \;=\; \{(x_1, x_2, f(x_1, x_2))\,;\;\; (x_1, x_2) \in U\}$$

2. Let G be a region in $\mathbb{R}^3$ and let $F:\, G \mapsto \mathbb{R}$ be continuously differentiable. Contour sets:

$$S_c \;=\; \{\vec{x} \,=\, (x_1, x_2, x_3) \in G\,;\;\; F(\vec{x}) - c \,=\, 0\}$$

From the theorem about implicit functions:[1]

If (at least) one of the partial derivatives $\partial F/\partial x_k(x_1, x_2, x_3)$ in a point $\vec{x} \in S_c \neq 0$, then the coordinate x_k can be written there as a function of the other two coordinates. For example $k = 3$:

For $\vec{a} = (a_1, a_2, a_3)$ shall hold

$$F(a_1, a_2, a_3) \,-\, c \;=\; 0\;; \qquad D_3\, F(\vec{a}) \,\neq\, 0$$

then there exists a neighborhood U of $(a_1, a_2) \in \mathbb{R}^2$ and a continuously differentiable function $f:\, U \mapsto \mathbb{R}^3$ with $f(a_1, a_2) \,=\, a_3$ and for all $(x_1, x_2)\; in \mathbb{R}^2$ holds

$$F(x_1, x_2, f(x_1, x_2)) \,-\, c \;=\; 0$$

Implicit differentiation

[1] See chapter 9 in HIGHER MATHEMATICS Lectures Part Two.

$$\frac{\partial f}{\partial x_1} = -\frac{\dfrac{\partial F}{\partial x_1}\big(x_1, x_2, f(x_1, x_2)\big)}{\dfrac{\partial F}{\partial x_3}\big(x_1, x_2, f(x_1, x_2)\big)}$$

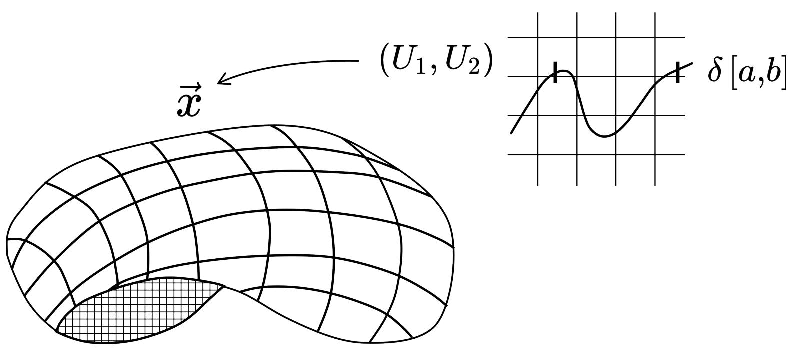
$$\frac{\partial f}{\partial x_2} = -\frac{\dfrac{\partial F}{\partial x_2}\big(x_1, x_2, f(x_1, x_2)\big)}{\dfrac{\partial F}{\partial x_3}\big(x_1, x_2, f(x_1, x_2)\big)}$$

Thus, at least locally, the surface S_c is represented as a graph according to 1.

3. Parameter or map representation

One considers continuously differentiable mappings of plane areas U in $\mathbb{R}^3$ (so-called map representations)

$\vec{x} \,:\, U \mapsto \mathbb{R}^3 \;;\quad S = \vec{x}(u)$



Remark:

(1) The representation as a graph under 1. is a special case. With the parameter lines u_1 = constant or respectively u_2 = constant two sets of curves emerge by lifting onto the surface.

$$t \;\mapsto\; \vec{x}(t, u_2) \quad \text{or respectively} \quad t \;\mapsto\; \vec{x}(u_1, t)$$

Each of these curves is continuously differentiable with derivative

$$\left(\frac{\partial x_1}{\partial u_1}, \frac{\partial x_2}{\partial u_1}, \frac{\partial x_3}{\partial u_1}\right) = \frac{\partial \vec{x}}{\partial u_1} \quad \text{or respectively} \quad \left(\frac{\partial x_1}{\partial u_2}, \frac{\partial x_2}{\partial u_2}, \frac{\partial x_3}{\partial u_2}\right) = \frac{\partial \vec{x}}{\partial u_2}$$

Short form

$$\frac{\partial \vec{x}}{\partial u_1} = \vec{x}_{u_1} \quad ; \quad \frac{\partial \vec{x}}{\partial u_2} = \vec{x}_{u_2}$$

4.1.1 Regularity Condition

$$\frac{\partial \vec{x}}{\partial u_1} \times \frac{\partial \vec{x}}{\partial u_2} \neq 0$$

Points in which the regularity condition holds are called “regular”, the remaining points are called “singular”.

Geometric meaning: The tangential vectors $\vec{x}_{u_1}$ and $\vec{x}_{u_2}$ are not parallel.

Main case: $\vec{x}$ is injective, i.e. different points in U always correspond to different points on S.

Then the (continuously differentiable) curves on S can be obtained by lifting curves $\delta : [a, b] \mapsto U$

$$\vec{\gamma}(t) = \vec{x}(\delta(t)) = \vec{x}(\delta_1(t), \delta_2(t))$$

Chain rule

$$\dot{\gamma}_k(t) = \frac{\partial x_k}{\partial u_1}(\delta(t))\, \dot{\delta}_1(t) + \frac{\partial x_k}{\partial u_2}(\delta(t))\, \dot{\delta}_2(t) \qquad k = 1,\, 2,\, 3$$

In vector form

$$\dot{\vec{\gamma}}(t) = \frac{\partial \vec{x}}{\partial u_1}(\delta(t))\, \dot{\delta}_1(t) + \frac{\partial \vec{x}}{\partial u_2}(\delta(t))\, \dot{\delta}_2(t)$$

Short form

$$\dot{\vec{\gamma}}(t) = \vec{x}_{u_1}\dot{\delta}_1(t) + \vec{x}_{u_2}\dot{\delta}_2(t)$$

4.1.2 Geometric Significance

Each tangent vector at the surface point $\vec{x}(u_1, u_2)$ (that is, any tangent vector of a curve on the surface through that point) is a linear combination of the special tangent vectors $\vec{x}_{u_1}$ and $\vec{x}_{u_2}$. In the regular surface points, the tangential vectors form a plane, the tangential plane.

Examples:

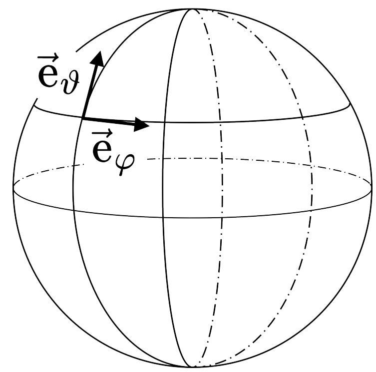
(1) The surface of the sphere (sphere S^2) in geographic coordinates

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

$$\vec{x}_\varphi = \frac{\mathrm{d}\vec{x}}{\mathrm{d}\varphi} = R\begin{pmatrix} -\sin\varphi\cos\vartheta \\ \cos\varphi\cos\vartheta \\ 0 \end{pmatrix}$$

$$\vec{x}_\vartheta = \frac{\mathrm{d}\vec{x}}{\mathrm{d}\vartheta} = R\begin{pmatrix} -\cos\varphi\sin\vartheta \\ -\sin\varphi\sin\vartheta \\ \cos\vartheta \end{pmatrix}$$

$$\vec{x}_\varphi \times \vec{x}_\vartheta = R^2\begin{pmatrix} \cos\varphi\cos^2\vartheta \\ \sin\varphi\cos^2\vartheta \\ \sin\vartheta\cos\vartheta \end{pmatrix}$$



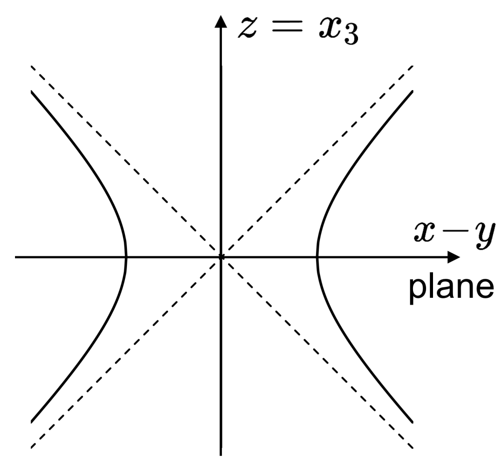
Singular surface points for $\cos\vartheta = 0$. Normalized tangent vectors:

$$\vec{e}_\varphi = \frac{\vec{x}_\varphi}{\|\vec{x}_\varphi\|} \quad ; \quad \vec{e}_\vartheta = \frac{\vec{x}_\vartheta}{\|\vec{x}_\vartheta\|}$$

(2) Hyperboloid of one sheet

$$x_1^2 + x_2^2 - x_3^2 = a^2 \qquad \text{parameterization } \vec{x}(\varphi, z) = ?$$

Cylindrical coordinates



$$\vec{x}(\varphi, z) = \begin{pmatrix} \sqrt{z^2+a^2}\cos\varphi \\ \sqrt{z^2+a^2}\sin\varphi \\ z \end{pmatrix}$$

$$\vec{x}_\varphi = \begin{pmatrix} -\sqrt{z^2+a^2}\sin\varphi \\ \sqrt{z^2+a^2}\cos\varphi \\ 0 \end{pmatrix}$$

$$\vec{x}_z = \begin{pmatrix} \dfrac{z\cos\varphi}{\sqrt{z^2+a^2}} \\ \dfrac{z\sin\varphi}{\sqrt{z^2+a^2}} \\ 1 \end{pmatrix} \qquad \vec{x}_\varphi \times \vec{x}_z = \begin{pmatrix} \sqrt{z^2+a^2}\cos\varphi \\ \sqrt{z^2+a^2}\sin\varphi \\ -z \end{pmatrix}$$

No singular points in the main case $a > 0$. The special case $a = 0$ represents a circular cone surface with apex at the origin, $z = 0$ is now a singular point.

4.2 Area Measurement of Curved Surfaces

is achieved with an integral formula created in analogy to the integral formula for the curve length:

$$l(\vec{\gamma}) \;=\; \int_a^b \|\dot{\vec{\gamma}}\|\,\mathrm{d}t$$

determines the length of the curves $\vec{\gamma}(gt) = \vec{x}(\delta(t))$ on the surface $\vec{x}(u,v)$, which are created through lifting from curves δ in the parameter image. Chain rule

$$\dot{\vec{\gamma}}(t) \;=\; \vec{x}_u(\delta(t))\,\dot{\delta}_1(t) \;+\; \vec{x}_v(\delta(t))\,\dot{\delta}_2(t)$$

$$\|\dot{\vec{\gamma}}(t)\|^2 \;=\; \big(\vec{x}_u\dot{\delta}_1 + \vec{x}_v\dot{\delta}_2\big)\big(\vec{x}_u\dot{\delta}_1 + \vec{x}_v\dot{\delta}_2\big)$$

$$\;=\; \|\vec{x}_u\|^2\dot{\delta}_1^2 \;+\; 2\vec{x}_u\vec{x}_v\dot{\delta}_1\dot{\delta}_2 \;+\; \|\vec{x}_v\|^2\dot{\delta}_2^2$$

Here appear the “fundamental metrics” of surfaces

$$g_{11} = \vec{x}_u\vec{x}_u \quad , \quad g_{12} = g_{21} = \vec{x}_u\vec{x}_v \quad , \quad g_{22} = \vec{x}_v\vec{x}_v$$

$$l(\vec{\gamma}) \;=\; \int_a^b \sqrt{g_{11}\dot{\delta}_1^2 + 2g_{12}\dot{\delta}_1\dot{\delta}_2 + g_{22}\dot{\delta}_2^2}\,\mathrm{d}t$$

The formula for the surface area of the surface S as given by the mapping $\vec{x} : U \mapsto R^3$

$$\boxed{\mathrm{vol}_2(S) \;=\; A(S) \;=\; \int_U \sqrt{g_{11}g_{22} - g_{12}^2}\,\mathrm{d}u\mathrm{d}v}$$

Interpretation of the determinant under the root sign in the integral

$$g_{11}g_{22} - g_{12}^2 \;=\; \|\vec{x}_u\|^2\|\vec{x}_v\|^2 - (\vec{x}_u\vec{x}_v)^2 \;=\; \|\vec{x}_u \times \vec{x}_v\|^2$$

The integral represents the area of the parallelogram spanned by the two tangent vectors $\vec{x_u}$, $\vec{x_v}$.

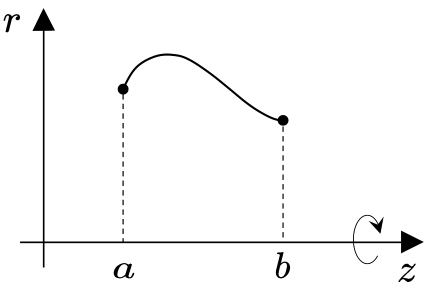
<u>Examples:</u>

(1) The spherical surface

$$\|\vec{x}_\varphi \times \vec{x}_\vartheta\|^2 \;=\; R^4\big(\cos^2\vartheta\,(\cos^2\vartheta + \sin^2\vartheta)\big) \;=\; R^4\cos^2\vartheta$$

$$A(S) \;=\; \int_0^{2\pi}\int_{-\pi/2}^{+\pi/2} R^2\cos\vartheta\,\mathrm{d}\vartheta\mathrm{d}\varphi \;=\; 4\pi R^2$$

(2) Surface area of a surface of revolution



Cylinder coordinates: $r(z) \geq 0$

$$\vec{x}(\varphi,z) \;=\; \big(r(z)\cos\varphi,\, r(z)\sin\varphi,\, z\big)$$

$$\vec{x}_\varphi \;=\; \big(-r(z)\sin\varphi,\, r(z)\cos\varphi,\, 0\big)$$

$$\vec{x}_z \;=\; \big(r'(z)\cos\varphi,\, r'(z)\sin\varphi,\, 1\big)$$

$$\|\vec{x}_\varphi\|^2 \;=\; g_{11} \;=\; r^2(z) \quad , \quad g_{12} \;=\; 0$$

$$\|\vec{x}_z\|^2 \;=\; g_{22} \;=\; r'(z)^2 + 1$$

$$A(S) \;=\; \int_a^b\int_0^{2\pi} r(z)\sqrt{r'(z)^2+1}\,\mathrm{d}\varphi\mathrm{d}z$$

$$\;=\; 2\pi\int_a^b r(z)\sqrt{r'(z)^2+1}\,\mathrm{d}z$$

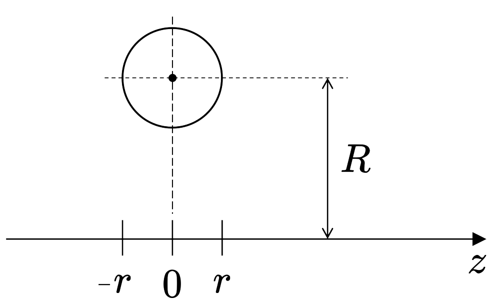
Interpretation:

$l(r) \;=\; \int_a^b r(z)\sqrt{r'(z)^2+1}\,\mathrm{d}z$ is the length of the rotating curve;

$R_0 \;=\; \dfrac{A(S)}{2\pi l(r)}$ is the ordinate of the geometric centroid.

$$\boxed{A(S) \;=\; 2\pi R_0\cdot l(r)}$$ Guldin's first rule

Application: The surface of the torus



For reasons of symmetry

$$R_0 \;=\; R$$

$$A(S) \;=\; 4\pi^2 r\, R$$

4.2.1 Invariance of the Area Content with respect to Orthogonal Transformations A of $\mathbb{R}^3$

Let A be an orthogonal 3×3 matrix, i.e. $A^T A = 1_3$.
Let $S : \vec{x} : U \mapsto \mathbb{R}^3$ be a surface in space.
$\tilde{S} : \vec{y} = A\vec{x}$; fundamental metrics of $\vec{y}$:

$$\vec{y}_u = \frac{\partial}{\partial u} A\vec{x}(u,v) = A\vec{x}_u \quad ; \quad \vec{y}_v = \frac{\partial}{\partial v} A\vec{x}(u,v) = A\vec{x}_v$$

$$\|\vec{y}_u \times \vec{y}_v\|^2 = (\vec{y}_u \times \vec{y}_v)(\vec{y}_u \times \vec{y}_v) = \|\vec{y}_u\|^2 \|\vec{y}_v\|^2 - (\vec{y}_u \vec{y}_v)^2$$

From the scalar product invariance to orthogonal transformations A follows

$$\|\vec{y}_u \times \vec{y}_v\|^2 = \|\vec{x}_u\|^2 \|\vec{x}_v\|^2 - (\vec{x}_u \vec{x}_v)^2 = \|\vec{x}_u \times \vec{x}_v\|^2$$

$$\Rightarrow \quad A(\tilde{S}) = A(S)$$

4.2.2 Invariance of the Area Measure under Parameter Transformations

Let $\vec{x} : U \mapsto \mathbb{R}^3$ be a map of the surface S and $\phi = \phi(\xi, \eta) = (u(\xi,\eta), v(\xi\eta))$ with its inverse map be a continuously differentiable mapping $\phi : V \mapsto U$ (with $\phi(V) = U$).

$$\vec{y}(\xi,\eta) = \vec{x} \circ \phi(\xi,\eta) = \vec{x}(u(\xi,\eta), v(\xi,\eta))$$

is a reparameterization of the surface S.

$$\begin{aligned} \vec{y}_\xi &= \frac{\partial \vec{x}}{\partial u}\big(u(\xi,\eta), v(\xi,\eta)\big)\, u_\xi + \frac{\partial \vec{x}}{\partial v}\big(\ \ldots\ \big)\, v_\xi \\ &= \vec{x}_u u_\xi + \vec{x}_v v_\xi \qquad \text{(chain rule)} \end{aligned}$$

Analogously

$$\begin{aligned} \vec{y}_\eta &= \vec{x}_u u_\eta + \vec{x}_v v_\eta \\ \vec{y}_\xi \times \vec{y}_\eta &= (\vec{x}_u u_\xi + \vec{x}_v v_\xi) \times (\vec{x}_u u_\eta + \vec{x}_v v_\eta) \\ & \qquad \text{(properties of the vector product exploited)} \\ &= (\vec{x}_u \times \vec{x}_v) u_\xi v_\eta + (\vec{x}_v \times \vec{x}_u) v_\xi u_\eta \\ &= (\vec{x}_u \times \vec{x}_v)(u_\xi v_\eta - v_\xi u_\eta) = (\vec{x}_u \times \vec{x}_v) \det \frac{\partial(u,v)}{\partial(\xi,\eta)} \end{aligned}$$

$$\int_V \|\vec{y}_\xi \times \vec{y}_\eta\| \,\mathrm{d}\xi \mathrm{d}\eta \;=\; \int_V \|\vec{x}_u \times \vec{x}_v\| \big(\phi(\xi,\eta)\big) \det \frac{\partial(u,v)}{\partial(\xi,\eta)} \,\mathrm{d}\xi \mathrm{d}\eta$$

(with transformation formula)

$$\boxed{\int_V \|\vec{y}_\xi \times \vec{y}_\eta\| \,\mathrm{d}\xi \mathrm{d}\eta \;=\; \int_{U=\phi(V)} \|\vec{x}_u \times \vec{x}_v\| \,\mathrm{d}u \mathrm{d}v}$$

$$\mathrm{d}S \;=\; \|\vec{x}_u \times \vec{x}_v\| \,\mathrm{d}u \mathrm{d}v$$

is called "surface element" of S with respect to the coordinates (u, v). It was shown: The integral over the surface element supplies a number that is independent of the choice of coordinates!

With the surface element $\mathrm{d}S = \|\vec{x}_u \times \vec{x}_v\| \,\mathrm{d}u\mathrm{d}v$ one can declare for smooth surfaces S an integral for continuous functions f on S by

$$\int f \mathrm{d}s \;:=\; \int_U f(\vec{x}(u,v)) \|\vec{x}_u \times \vec{x}_v\| \,\mathrm{d}u \mathrm{d}v$$

This integral is independent of the parameterization of the surface (proof analogous to the area content, $f = 1$ constant).

Objective: Generalization of the fundamental theorem of calculus.

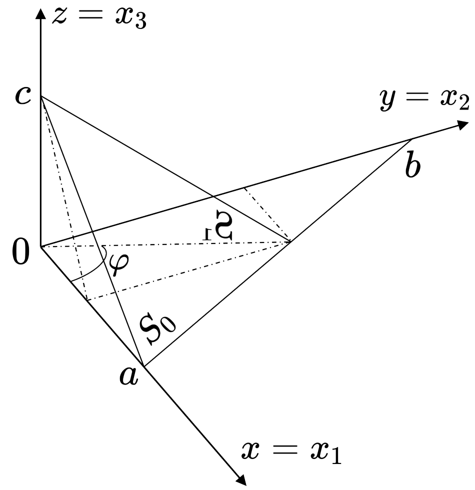
4.3 Gauß' Theorem

Let $\vec{v} : G \mapsto \mathbb{R}^3$ be continuously differentiable in the region $G \subset \mathbb{R}^3$. Let B be a piecewise smooth-bounded compactum in G. Then

$$\boxed{\int_B \operatorname{div} \vec{v}(\vec{x}) \,\mathrm{d}x \mathrm{d}y \mathrm{d}z \;=\; \int_{\partial B = S} \big(\vec{v}(\vec{x})\, \vec{n}\big) \mathrm{d}s}$$

In every point of the surface $S = \partial B$, $\vec{n} = \vec{n}(\vec{x})$ represents the outward-pointing normal vector of length 1 (normal vector field $\vec{n}$).

Proof for polyhedra:



This results from decomposition into simplices (= tetrahedra) T.

$$\big(r\cos\varphi\,,\ r\sin\varphi\,,\ 0\big)$$

$$r \;=\; \frac{a\,b}{a\sin\varphi + b\cos\varphi}$$

1) Description of S_0, the tetrahedron side in the plane through $(a,0,0)$; $(0,b,0)$; $(0,0,c)$

Equation of the plane: $\quad \vec{x}\cdot\left(\frac{1}{a}\,,\,\frac{1}{b}\,,\,\frac{1}{c}\right) \;=\; 1$

S_0 is represented by three maps

$$\vec{x}^{\,k} \;:\; U_k \;\mapsto\; T$$

which are created by reversing the projection from S_0 onto the coordinate planes $x_k = 0$.

Checking the calculation:

$$\left\|\vec{x}^{\,k}_u \times \vec{x}^{\,k}_v\right\| \;=\; a_k\sqrt{a^{-2}+b^{-2}+c^{-2}} \quad \text{with} \quad a_1 = a\,,\ a_2 = b\,,\ a_3 = c$$

2) Description of S_1 by the inversion $\vec{x}(y,z)$ of the projection of S_1 onto the y–z plane and by the inversion $\vec{y}(x,z)$ of the projection of S_1 onto the $x-z$ plane

$$\vec{x}(y,z) \;=\; (y\cot\varphi, y, z) \quad ; \quad \vec{y}(x,z) \;=\; (x, x\tan\varphi, z)$$

$$\vec{x}_y \;=\; (\cot\varphi, 1, 0) \quad ; \quad \vec{x}_z \;=\; (0,0,1)$$

$$\vec{y}_x \;=\; (1, \tan\varphi, 0) \quad ; \quad \vec{y}_z \;=\; (0,0,1)$$

$$\vec{x}_y \times \vec{x}_z \;=\; (1, -\cot\varphi, 0)$$

$$\vec{y}_x \times \vec{y}_z \;=\; (\tan\varphi, -1, 0)$$

$$\begin{aligned} \|\vec{x}_y \times \vec{x}_z\| &= \sqrt{1+\cot^2\varphi} &= \frac{1}{\sin\varphi} \\ \|\vec{y}_x \times \vec{y}_z\| &= \sqrt{\tan^2\varphi + 1} &= \frac{1}{\cos\varphi} \end{aligned}$$

3) Description of the limits of x_k on U_k ($k = 1,\,2,\,3$) and of $x_2 = y$ for the domain of definition V_2 of the map image $\vec{y}$

$U_3:\ 0 \le y \le x\tan\varphi \quad ; \quad \dfrac{x}{a} + \dfrac{y}{b} \le 1$

There: $\quad 0 \le z \le \big(1 - \dfrac{x}{a} - \dfrac{y}{b}\big)c$

$U_1:\ 0 \le y\,,\,z \quad ; \quad \dfrac{y}{r\sin\varphi} + \dfrac{z}{c} \le 1$

There: $\quad \dfrac{y}{\tan\varphi} \le x \le \big(1 - \dfrac{y}{b} - \dfrac{z}{c}\big)a$

$V_2:\ 0 \le x\,,\,z \quad ; \quad \dfrac{x}{r\cos\varphi} + \dfrac{z}{c} \le 1$

$U_2:\ 0 \le x\,,\,z \quad ; \quad 1 - \dfrac{x}{r\cos\varphi} \le \dfrac{z}{c} \le 1 - \dfrac{x}{a}$

There: $\quad 0 \le y \le \big(1 - \dfrac{x}{a} - \dfrac{z}{c}\big)b$

4) Conversion of the space integral into a surface integral

Abbreviations: $\quad \vec{v} = (u,v,w) \quad ; \quad \operatorname{div}\vec{v} = u_x + v_y + w_z$

$$\int_T (u_x + v_y + w_z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z \;=$$

$$\begin{aligned} &\int_{U_1} \Big[u\Big(a\big(1 - \frac{y}{b} - \frac{z}{c}\big), y, z\Big) \;-\; u\big(y\cot\varphi, y, z\big)\Big]\,\mathrm{d}y\mathrm{d}z \\ +&\int_{U_2} \Big[v\Big(x, b\big(1 - \frac{x}{a} - \frac{z}{c}\big), z\Big) \;-\; v\big(x, 0, z\big)\Big]\,\mathrm{d}x\mathrm{d}z \\ +&\int_{V_2} \big[v\big(x, x\tan\varphi, z\big) \;-\; v\big(x, 0, z\big)\big]\,\mathrm{d}x\mathrm{d}z \\ +&\int_{U_3} \Big[w\Big(x, y, c\big(1 - \frac{x}{a} - \frac{y}{b}\big)\Big) \;-\; w\big(x, y, 0\big)\Big]\,\mathrm{d}x\mathrm{d}y \end{aligned}$$

5) Completion of proof

The outward-pointing normal unit vectors

- for the surface S_3 in the $x{-}y$ plane: $\vec{n}_3 = (0,0,-1)\,,$

- for the surface S_2 in the $x{-}z$ plane: $\vec{n}_2 = (0,-1,0)\,,$

- for S_1: $\vec{n}_1 = (-\sin\varphi, \cos\varphi, 0)\,,$

- for S_0: $\vec{n}_0 = (\sqrt{w_0})^{-1}\cdot\left(\frac{1}{a},\frac{1}{b},\frac{1}{c}\right)$ with $w_0 = a^{-2}+b^{-2}+c^{-2}$

Now remember the surface elements. For $U_k'\,\vec{x}^k$

$$\left\|\vec{x}_u^k \times \vec{x}_v^k\right\| = a_k\sqrt{w_0} \quad ; \quad k = 1,\,2,\,3$$

and for $\vec{x}$, $\vec{y}$ for the parameterization of S_1

$$\|\vec{x}_y \times \vec{x}_z\| = \frac{1}{\sin\varphi} \quad ; \quad \|\vec{y}_x \times \vec{y}_z\| = \frac{1}{\cos\varphi}$$

In summary

$$\int_T (u_x + v_y + w_z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z =$$

$$= \int_{S_0} (\vec{v}\,\vec{n}_0)\,\mathrm{d}s + \int_{S_1} (\vec{v}\,\vec{n}_1)\,\mathrm{d}s + \int_{S_2} (\vec{v}\,\vec{n}_2)\,\mathrm{d}s + \int_{S_3} (\vec{v}\,\vec{n}_3)\,\mathrm{d}s$$

$\square$

4.4 Differential Geometric Interpretation of Gauß' Theorem

The scalar product $\vec{v}\cdot\vec{n}$ represents the projection of the vector field onto the (outer) normal of the bounded surface. The right-hand side of Gauß's theorem thus represents the total flow of the vector field $\vec{v}$ through the surface ∂B (possibly consisting of several pieces) of the space part B. With the divergence as the source density, the left-hand side is the integral over the source density of the vector field $\vec{v}$.

Example: Buoyancy (Archimedes principle)

A massive body B is immersed in a liquid of density ϱ. The pressure $\varrho\cdot z\cdot\vec{n}$ acts on the surface at every point (z is negative $\Rightarrow$ pressure directed inwards).

The total force acting on B is

$$K \;=\; \int\limits_{\partial B} \varrho\cdot z\cdot\vec{n}\,\mathrm{d}s\;; \qquad K=(K_1,K_2,K_3) \;\;\text{is a vector}$$

Decomposition of the integrand into its coordinates

$$\vec{n} \;=\; \sum_{i=1}^{3}\big(\vec{n}\,\vec{\mathrm{e}}_i\big)\,\vec{\mathrm{e}}_i \qquad ; \qquad K_i \;=\; \int\limits_{\partial B} \varrho\cdot z\cdot\big(\vec{\mathrm{e}}_i\,\vec{n}\big)\,\mathrm{d}s$$

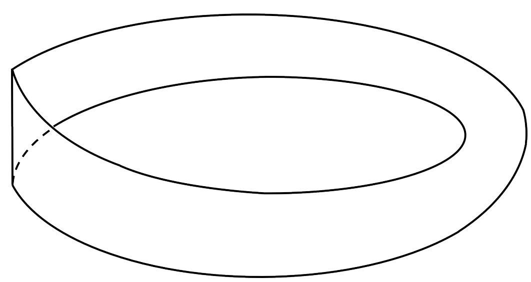
Three vector fields $\quad \vec{v}_i \;=\; \varrho\cdot x_3\cdot\vec{\mathrm{e}}_i \quad (z=x_3)$

For $i \;=\; 1,2$ we obviously have $\operatorname{div}\vec{v}_i \;=\; 0$, from which follows $K_1=K_2=0$. On the other hand $\operatorname{div}\vec{v}_3 \;=\; \varrho$.

$$K_3 \;=\; \int\limits_{B} \varrho\,\mathrm{d}x\mathrm{d}y\mathrm{d}z \;=\; \varrho\cdot\mathrm{vol}_3(B)$$

<u>Orientation:</u>

On the bounding surfaces $S=\partial B$ of smooth-edged compacta $B\subset\mathbb{R}^3$ there exists a continuous normal unit field $\vec{n}(\vec{x})\,,\;\;\vec{x}\in S$. They are "orientable". The "Möbius band" shows the existence of non-orientable surfaces.



- On one-dimensional manifolds (= curves), orientation signifies the sense of direction.
- On two-dimensional manifolds S (= surfaces) orientation signifies the sense of rotation.
- On three-dimensional manifolds, orientation signifies the winding sense.

<u>In common:</u>

On every orientable manifold there always exists a pair of equally entitled orientations. A change in orientation always means a change of sign for integrals over orientable manifolds.

In Euclidean space $\mathbb{R}^n$ the invertible linear self-mappings

$$A : \mathbb{R}^n \longmapsto \mathbb{R}^n$$

have the property $\det A \neq 0$. Change of orientation means $\det A < 0$; orientation fidelity: $\det A > 0$. (The orientation is a global property of the bundle of all tangent planes of the manifolds.)

4.5 Stokes' Integral Theorem

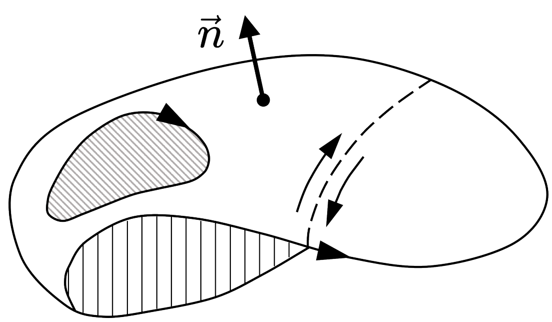
Let $\vec{v}$ be a continuously differentiable vector field in the domain $G \subset \mathbb{R}^3$. In G let S be a smoothly bounded orientable surface with the continuous unit normal field

$$\vec{n} : S \longmapsto \mathbb{R}^3$$

Then

$$\int\limits_S \big(\operatorname{rot} \vec{v}\, \vec{n}\big)\, \mathrm{d}s = \int\limits_{\partial S} \vec{v}\, \mathrm{d}\vec{x}$$

where ∂S carries the orientation induced by $\vec{n}$.



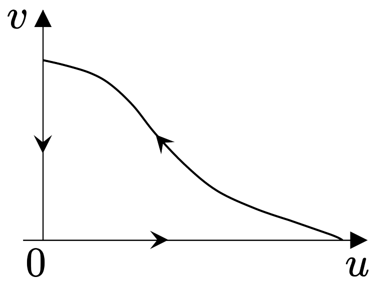
With respect to organizing the **proof**

1. The formula is invariant to reparameterization.

2. The formula is additive with respect to $\vec{v}$. It can be assumed that $\vec{v}$ has only one coordinate $\neq 0$.

3. The formula is additive with respect to the decomposition of the surface S by auxiliary curves, which run from edge to edge.

4. Finally, it is sufficient to discuss triangular surfaces.



$$S: \ \vec{x}(u,v) \ = \ \big(u, v, f(u,v)\big)$$

For example: $\vec{v} \ = \ \big(0, 0, w(x,y,z)\big)$

$B \ = \ U$ parameter range

$$\vec{x}_u \ = \ \left(1\,,\,0\,,\,\frac{\partial f}{\partial u}\right) \quad ; \quad \vec{x}_v \ = \ \left(0\,,\,1\,,\,\frac{\partial f}{\partial v}\right)$$

$$\operatorname{rot}\vec{v} \ = \ \left(\frac{\partial}{\partial x}\,,\,\frac{\partial}{\partial y}\,,\,\frac{\partial}{\partial z}\right)\times\vec{v} \ = \ \left(\frac{\partial w}{\partial y}\,,\,-\frac{\partial w}{\partial x}\,,\,0\right)$$

$$\operatorname{rot}\vec{v}\cdot\vec{n} \ = ? \quad ; \quad \|\vec{x}_u\times\vec{x}_v\| \ = \ \left(-\frac{\partial f}{\partial u}\,,\,-\frac{\partial f}{\partial v}\,,\,1\right)$$

$$\|\vec{x}_u\times\vec{x}_v\|\cdot\operatorname{rot}\vec{v}\cdot\vec{n} \ = \ -\frac{\partial f}{\partial u}\frac{\partial w}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial w}{\partial x}$$

Thus the left-hand side (LS) of Stokes' integral theorem becomes

$$LS \ = \ \int\limits_B \Big(-\frac{\partial w}{\partial y}\frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}\frac{\partial w}{\partial x}\Big)\,\mathrm{d}u\mathrm{d}v$$

Riemann formula for a planar vector field!

$$p(u,v) \ = \ w\big(u,v,f(u,v)\big)\frac{\partial f}{\partial u}$$

$$q(u,v) \ = \ w\big(u,v,f(u,v)\big)\frac{\partial f}{\partial v}$$

$$\frac{\partial p}{\partial v} \ = \ \frac{\partial w}{\partial y}\frac{\partial f}{\partial u} + \frac{\partial w}{\partial z}\frac{\partial f}{\partial v}\frac{\partial f}{\partial u} + w\frac{\partial^2 f}{\partial v\partial u}$$

$$\frac{\partial q}{\partial u} \ = \ \frac{\partial w}{\partial x}\frac{\partial f}{\partial v} + \frac{\partial w}{\partial z}\frac{\partial f}{\partial u}\frac{\partial f}{\partial v} + w\frac{\partial^2 f}{\partial u\partial v}$$

$$LS \ = \ \int\limits_B \big(\frac{\partial q}{\partial u} - \frac{\partial p}{\partial v}\big)\,\partial u\partial v \ = \ \int\limits_{\partial B}(p\,\partial u + q\,\partial v) \ = \ \int\limits_{\partial S}\vec{v}\,\mathrm{d}\vec{x}$$

$\square$

5. Quadratic Matrices and Determinants

Underlying: fields K, like $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$

Considered are "matrix" systems $\big(a_{ik}\big)_{1 \le i\,,\,k \le n}$ with elements $a_{ik} \in K$. Summarized in $M_n(K)$

Matrix sum: $A + B$; Scalar multiples: λA

Definition component wise. With that $M_n(K)$ becomes a vector space (over K). Each matrix $A \in M_n(K)$ maps the space K^n of "columns"

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \vec{x} \quad \text{into itself} \quad A\vec{x} = \left(\sum_{k=1}^{n} a_{ik} x_k\right)_{1 \le i \le n}$$

Calculation rules

$$\left.\begin{aligned} A(\vec{x} + \vec{y}) &= A\vec{x} + A\vec{y} \\ A(\lambda \vec{x}) &= \lambda (A\vec{x}) \end{aligned}\right\} \text{Linearity}$$

Each $A \in M_n(K)$ yields a linear self-mapping of K^n.

$$AB = \left(\sum_{j=1}^{n} a_{ij} b_{jk}\right)_{1 \le i\,,\,k \le n} \qquad \text{matrix multiplication}$$

If $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n$ are the columns of B, then

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

Warning:[1] $AB = BA$ wrong!

[1]Compare the respective warning in chapter 9 of HIGHER MATHEMATICS Lectures Part One.

On the other hand, the associative law holds

$$\boxed{(AB)C \;=\; A(BC)} \qquad \forall\ A, B, C \in M_n(K)$$

Proof:

Consider the element with index i, m on the left side of the equation

$$\sum_{k=1}^{n}\left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)c_{km} \;=\; \sum_{1\le j\,,\,k\le n}(a_{ij}b_{jk})c_{km}$$

and the element with index i, m on the right

$$\sum_{j=1}^{n} a_{ij}\left(\sum_{k=1}^{n} b_{jk}c_{km}\right) \;=\; \sum_{1\le j\,,\,k\le n} a_{ij}(b_{jk}c_{km})$$

Result:

The associative law of multiplication in $M_n(K)$ follows from the law in K.

$\square$

The distributive laws apply analogously

$$\left.\begin{aligned} A(B+C) &= AB + AC \\ (A+B)C &= AC + BC \end{aligned}\right\} \quad \forall\ A, B, C \in M_n(K)$$

Unit element of multiplication

$$1_n \;=\; \left(\delta_{ij}\right) \qquad 1 \le i\,,\, j \le n$$

Recursive definition of the determinants

$$n=1: \quad A = \left(a_{11}\right) \qquad \det_1 A = a_{11}$$

$$n=2: \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad \det_2 A = -a_{21}a_{12} + a_{22}a_{11}$$

$n=3: \quad \left(a_{ik}\right) = A \in M_3(K)$ with the rule of Sarrus[2]

Recursive definition of the determinant for $n \times m$ matrices

$n = 1$: $\quad A = a_{11} \quad ; \quad \det A = a_{11}$

$n > 1$: $\quad A \in M_n(K)$; For each k ($1 \le k \le n$) let $A^{(n,k)} \in M_{n-1}(K)$ arise from A by deleting the nth row and the kth column

[2] See chapter 3 in HIGHER MATHEMATICS Lectures Part Two.

$$\det_n A \;=\; \sum_{k=1}^{n} (-1)^{n+k} a_{nk} \det_{n-1} A^{(n,k)}$$

Example: For any upper triangular matrix

$$A = \begin{pmatrix} a_{11} & & & * \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{pmatrix} \quad \text{holds} \quad \det A = \prod_{i=1}^{n} a_{ii}$$

Proof by induction on n

$n = 1$: Obvious.

$n-1 \mapsto n$: A is a triangular matrix $\Rightarrow$ $a_{nk} = 0$, if $k < n$. Hence

$$\det A \;=\; \sum_{k=1}^{n} (-1)^{n+k} a_{nk} \det_{n-1} A^{(n,k)} \;=\; a_{nn} \det_{n-1} A^{(n,n)}$$

Since $A^{(n,n)}$ is again an upper triangular matrix $\in M_{n-1}(K)$ it follows from the induction premise that

$$\det A \;=\; a_{nn} \prod_{i=1}^{n-1} a_{ii}$$

□

5.1 Characteristic Properties of the Determinant

(1) det : $M_n(K) \mapsto K$ is an alternating multilinear form in the columns of the matrix used $A = (\vec{a}_1, \ldots, \vec{a}_1)$ with normalization $\det_n(1_n) = 1$. These properties, alternating (ALT) and multilinear (ML), are expressed by

$$\det(\ldots \vec{a} \ldots \vec{a} \ldots) \;=\; 0 \qquad \text{(ALT)}$$

$$\left.\begin{aligned} \det(\ldots \vec{a} + \vec{b} \ldots) &= \det(\ldots \vec{a} \ldots) + \det(\ldots \vec{b} \ldots) \\ \det(\ldots \lambda \vec{a} \ldots) &= \lambda \det(\ldots \vec{a} \ldots) \end{aligned}\right\} \quad \text{(ML)}$$

(2) If $D : M_n(K) \mapsto K$ has the properties (ALT) and (ML), then

$$D(\vec{a}_1, \ldots, \vec{a}_n) \;=\; D\underbrace{(\vec{\mathrm{e}}_1, \ldots, \vec{\mathrm{e}}_n)}_{1_n} \det(\vec{a}_1, \ldots, \vec{a}_n)$$

Proof:

1) By induction on n

In the case $n = 1$, (ML) is trivial and (ALT) is correct because it is meaningless.

$n - 1 \to n$:

$$L_k(A) \;=\; (-1)^{n+k} a_{nk} \det_{n-1} A^{(n,k)}$$

is by the induction hypothesis a linear function of each column of A. Therefore the sum of this function, i.e. $\det_n A$, is linear in every column of A. Apparently $\det 1_n \;=\; 1$.

$$D(\ldots \vec{a} \ldots \vec{b} \ldots) \;=\; -D(\ldots \vec{b} \ldots \vec{a} \ldots) \qquad (*)$$

Justification: $\quad D(\ldots \vec{a} + \vec{b} \ldots \vec{a} + \vec{b} \ldots) \;=\; 0 \quad$ because of (ALT)

$$\begin{aligned} 0 &= D(\ldots \vec{a} + \vec{b} \ldots \vec{a} + \vec{b} \ldots) \\ &= D(\ldots \vec{a} \ldots \vec{a} + \vec{b} \ldots) + D(\ldots \vec{b} \ldots \vec{a} + \vec{b} \ldots) \\ &= \underbrace{D(\ldots \vec{a} \ldots \vec{a} \ldots)}_{0} + D(\ldots \vec{a} \ldots \vec{b} \ldots) \\ &\quad + D(\ldots \vec{b} \ldots \vec{a} \ldots) + \underbrace{D(\ldots \vec{b} \ldots \vec{b} \ldots)}_{0} \quad \Rightarrow \quad (*) \end{aligned}$$

2) Assumption: $\quad \vec{a}_s \;=\; \vec{a}_t \quad$ for $\quad s < t$.

Every auxiliary matrix $A^{(n,k)}$ with $k \neq s, t$ also has two equal columns, therefore (by the induction hypothesis) $\det_{n-1} A^{(n,k)} \;=\; 0$.

It remains

$$\det A \;=\; (-1)^{n+s} a_{ns} \det_{n-1} A^{(n,s)} \;+\; (-1)^{n+t} a_{nt} \det_{n-1} A^{(n,t)}$$

because of $\quad \vec{a}_s \;=\; \vec{a}_t \quad$ one has $\quad a_{ns} \;=\; a_{nt}$.

The matrix $A^{(n,s)}$ has (except for the order) the same columns as $A^{(n,t)}$, because of that and because of $(*)$

$$\det A \;=\; 0 \qquad (\text{if } \vec{a}_s \;=\; \vec{a}_t)$$

□

$\det 1_n = 1$, det is an alternating multilinear form of the columns.

5.1.1 Calculation Rules

1. If one adds a multiple of a column of A to another column of A, the determinant does not change.
2. Multiplying a column by λ means multiplying the determinant by λ.
3. Interchanging two columns means going over to the negative of the determinant.

Example:

$$
\begin{aligned}
A &= \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 2 \\ 4 & 1 & 2 & 3 \end{bmatrix} \\
\det A &= \quad \text{(Addition of the first three columns to the fourth column.)} \\
&\overset{1.}{=} \det \begin{bmatrix} 1 & 2 & 3 & 10 \\ 2 & 3 & 4 & 10 \\ 3 & 4 & 1 & 10 \\ 4 & 1 & 2 & 10 \end{bmatrix} \quad \text{(Divide last column by 10 and multiply determinant by 10.)} \\
&\overset{2.}{=} 10 \det \begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 3 & 4 & 1 \\ 3 & 4 & 1 & 1 \\ 4 & 1 & 2 & 1 \end{bmatrix} \quad \text{(Addition of a multiple of the last column to the other columns.)} \\
&= 10 \det \begin{bmatrix} -3 & 1 & 1 & 1 \\ -2 & 2 & 2 & 1 \\ -1 & 3 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = 10 \det \begin{bmatrix} -3 & 1 & 1 \\ -2 & 2 & 2 \\ -1 & 3 & -1 \end{bmatrix} \\
&\overset{1.}{=} 10 \det \begin{bmatrix} -4 & 4 & 0 \\ -4 & 8 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -10 \det \begin{bmatrix} -4 & 4 \\ -4 & 8 \end{bmatrix} \\
&\overset{2.}{=} -10 \cdot 4^2 \det \begin{bmatrix} -1 & 1 \\ -1 & 2 \end{bmatrix} \overset{1.}{=} -160 \det \begin{bmatrix} -1 & -1 \\ -1 & 0 \end{bmatrix} \\
&\overset{3.}{=} 160 \det \begin{bmatrix} -1 & -1 \\ 0 & -1 \end{bmatrix} = 160
\end{aligned}
$$

Characterization of the last theorem (2)

If $D: M_n(K) \mapsto K$ is an alternating multilinear form then it holds that

$$D(A) \;=\; D(1_n)\det A$$

Proof sketch:

Abbreviation: $\delta \coloneqq D(1_n)$; consider: $\tilde{D}(A) \;=\; D(A) - \delta \det A$

With D and det, $\tilde{D}$ itself is again alternating and multilinear, furthermore $\tilde{D}(1_n) = 0$.

If $i_1, i_2, \ldots, i_n$ are indices with $1 \le i_k \le n$ then $\tilde{D}(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_n}) = 0$ because either two columns are equal or the sequence $i_1, i_2, \ldots, i_n$ results from permutation of the sequence $1, 2, \ldots, n$, then

$$\tilde{D}(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_n}) \;=\; \pm\tilde{D}(\underbrace{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n}_{1_n}) \;=\; 0$$

$A \in M_n(K)$, $(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) \;=\; A$, then

$$\vec{a}_k \;=\; \sum_{i_k=1}^{n} a_{i_k} k \vec{e}_k$$

with (ML) : $\tilde{D}(\vec{a}_1, \ldots, \vec{a}_n)$ becomes a linear form of $\tilde{D}(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_n}) \;=\; 0$, explicit

$$\tilde{D}(A) \;=\; \sum_{i_1, i_2, \ldots, i_n = 1}^{n} a_{i_1}, a_{i_2}, \ldots, a_{i_n} \tilde{D}(\vec{e}_{i_1}, \vec{e}_{i_2}, \ldots, \vec{e}_{i_n}) \;=\; 0$$

□

5.2 Product Theorem for Determinants

$$\det AB \;=\; \det A \cdot \det B \qquad A, B \in M_n(K)$$

Proof idea:

Consider $\det AB$ as a function of the columns of B

$$D(B) \;=\; D(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n) \;\coloneqq\; \det AB$$
$$\;=\; \det\Big(A\vec{b}_1, A\vec{b}_2, \ldots, A\vec{b}_n\Big)$$

$D(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n)$ is then an alternating multilinear form of the $\vec{b}_k$. Hence

$$D(B) \;=\; D(1_n \det B)$$

that means

$$\det AB \;=\; \det A \cdot \det B$$

5.2.1 Behavior of the Determinant under Elementary Row Transformations

Elementary row transformation = elementary matrix E multiplication from the left!

Type I: $p \neq q$

$$E = \begin{array}{cc} \begin{array}{cccccc} & & & & q & \\ & & & & \downarrow & \end{array} & \\ \begin{bmatrix} 1 & & & & & \\ & 1 & & & \lambda & \\ & & \ddots & & & \\ & & & \ddots & & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} & \begin{array}{l} \\ \leftarrow p \\ \\ \\ \\ \\ \end{array} \end{array} \qquad \begin{aligned} E &= \left(\delta_{ik} + \lambda\delta_{ip}\delta_{qk}\right)_{1 \leq k,i} \\ \det E &= 1 \end{aligned}$$

EA is created from A by adding a λ-multiple of the q-th column to the p-th row.

Product theorem $\Rightarrow \quad \det EA = \det E \cdot \det A$

$$\det E \overset{\substack{(\mathrm{ML}) \\ \downarrow}}{=} \det 1_n = 1$$

Type II: $p = q$

$$E = \begin{array}{cc} \begin{array}{ccccccc} & & & q & & & \\ & & & \downarrow & & & \end{array} & \\ \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow p \\ \\ \\ \\ \end{array} \end{array} \qquad \begin{aligned} E &= \left(\delta_{ik}(1 + (\lambda - 1)\delta_{ip})\right)_{1 \leq k,i \leq n} \\ \det E &= \lambda \end{aligned}$$

Type III: $p \neq q$

$$E = \begin{array}{cc} \begin{bmatrix} 1 & & & & & & & & & \\ & \ddots & & & & & & & & \\ & & 1 & & & & & & & \\ & & & 0 & & & 1 & & & \\ & & & & 1 & & & & & \\ & & & & & \ddots & & & & \\ & & & & & & 1 & & & \\ & & & 1 & & & & 0 & & \\ & & & & & & & & 1 & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 1 \end{bmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow p \\ \\ \\ \\ \leftarrow q \\ \\ \\ \\ \end{array} \end{array} = \left(\delta i, \tau_{p.q}(k)\right)_{1 \leq i,k \leq n}$$

EA is created from A by swapping the p-th and q-th row in A.

$\tau_{p,q}$ is called “transposition” of p and q

$$\tau_{p,q}(k) \;=\; \left\{ \begin{array}{l} k\,, \;\; \text{if} \;\; k \neq p, q \\ q\,, \;\; \text{if} \;\; k = p \\ p\,, \;\; \text{if} \;\; k = q \end{array} \right\} \qquad \det E \;=\; -1$$

5.2.2 Rules for Row Transformations

Let $A \in M_n(K)$. Addition of a multiple of a row of A to another row does not change the determinant. Multiplying a row by λ means multiplying the determinate by λ. Swapping two lines of A means multiplying the determinant by -1.

Remarks:

(1) The determinant $\det A$ is at the same time an alternating multilinear form of the rows of A. The matrix A^T (transposed matrix) has the columns of A as rows. The function $\det A^T$ is an alternating multilinear form of the columns of A with $\det\bigl(1^T\bigr) = \det 1_n = 1$.
Characterization of det $\Rightarrow$

$$\boxed{\det A^T \;=\; \det A}$$

(2) Every matrix $A \in M_n(K)$ can be written as a product $A = E_1 E_2 \dots E_n$ of elementary matrices.

(3) Let $A \in M_n(K)$ with a “right inverse” $A' \in M_n(K)$ with $AA' = 1_n$.
Product theorem

$$\boxed{\det A \cdot \det A' \;=\; 1}$$

In particular, $\det A \neq 0$. Next objective: If $\det A \neq 0$ then there exists a right inverse A'.

5.2.3 The Adjugate of a Matrix $A \in M_n(K)$

Let $A^{(i,k)} \in M_{n-1}(K)$ be the matrix obtained by deleting the i-th row and k-th column of A. With

$$a^*_{jk} \;=\; (-1)^{j+k} \det A^{(k,j)} \;, \qquad A \;=\; \bigl(a_{ik}\bigr)_{1 \le i\,,\,k \le n}$$

holds

$$\sum_{j=1}^{n} a_{ij} a^*_{jk} \;=\; \delta_{ij} \det A$$

Remark:

(4) In the case $i = k = n$, the last formula is nothing more than the recursive definition of $\det A$. For the case $i = k$ the formula is called "development of $\det A$ after the i-th row".

Proof:

1) In the case $i = k < n$

Let the rows of A be $z_1, z_2, \ldots, z_n$. By swapping rows

$$z'_j \;=\; z_j \qquad 1 \leq j < i$$

$$z'_j \;=\; z_{j+1} \qquad 1 \leq j < n$$

$$z'_n \;=\; z_i$$

The new matrix has the determinant $(-1)^{n-i} \det A$. If one inserts the new matrix into the recursion formula, the assertion follows.

2) In the case $i \neq k$

One replaces in A the k-th row z_k with zi. This creates a matrix A_0 with two equal rows. Hence $\det A_0 = 0 = \delta_{ik} \det A$. The left side is the proven formula for $i = k$ applied to A_0 instead of A.

□

5.2.4 The Inverse Matrix

If $\det A \neq 0$, then the matrix

$$A' \;=\; \left(\frac{a^*_{jk}}{\det A} \right)_{1 \leq j,k \leq n} \;=\; \left(\frac{(-1)^{j+k}}{\det A} \det A^{(k,j)} \right)_{1 \leq j,k \leq n}$$

has the property $AA' \;=\; 1_n$.

In particular $\det A' \neq 0$; Therefore $A'' \in M_n(K)$ exists with $A'A'' \;=\; 1_n$.

Multiplication from the left with A

$$\underbrace{(AA')}_{1_n} A'' \;=\; A(A'A'') \;=\; A \;=\; A''$$

therefore

$$\det A'A \;=\; 1_n$$

Remark:

(5) To calculate the inverse matrix A^{-1} in the case $\det A \neq 0$, the formula for the adjugate is generally too complex. Elementary transformation is the tried and tested approach. If, as row transformations, the elementary matrices $E_1, E_2, \ldots, E_N$ successively turn the matrix A into the identity matrix, i.e. $E_N \ldots E_2 E_1 A \;=\; 1_n$, then

$$E_N \ldots E_2 E_1 \;=\; A^{-1}$$

In words:
If one carries out the elementary row transformations which convert A into the identity matrix in the same order on the identity matrix, one obtains A^{-1}.

Example:

$$A \;=\; \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \qquad\qquad 1_3 \;=\; \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & -5 \\ 3 & 1 & 2 \end{bmatrix} \qquad z_2 \;\mapsto\; z_2 \;-\; 2z_1 \qquad \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -5 \\ 0 & -5 & -7 \end{bmatrix} \qquad z_3 \;\mapsto\; z_3 \;-\; 3z_1 \qquad \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & -5 & -7 \end{bmatrix} \qquad z_2 \;\mapsto\; -z_2 \qquad \begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -3 & 0 & 1 \end{bmatrix}$$

$$18 \cdot \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{array}{l} z_3 \;\mapsto\; 5z_2 \;+\; z_3 \\ z_3 \;\mapsto\; \dfrac{1}{18}\, z_3 \end{array} \qquad \frac{1}{18} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 7 & -5 & 1 \end{bmatrix}$$

$$18 \cdot \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{array}{l} \text{Right: Extract factor} \\ \dfrac{1}{18} \text{ from } z_2 \text{ and } z_3 \end{array} \qquad \frac{1}{18^3} \cdot \begin{bmatrix} 18 & 0 & 0 \\ 36 & -18 & 0 \\ 7 & -5 & 1 \end{bmatrix}$$

$$18 \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \begin{array}{rcl} z_2 & \mapsto & z_2 - 5z_3 \\ z_1 & \mapsto & z_1 - 2z_2 \\ z_1 & \mapsto & z_1 - 3z_3 \end{array} \qquad \frac{1}{18^3} \cdot \begin{bmatrix} -5 & 1 & 7 \\ 1 & 7 & -5 \\ 7 & -5 & 1 \end{bmatrix}$$

$$\Rightarrow \quad A^{-1} = \frac{1}{18^3} \cdot \begin{bmatrix} -5 & 1 & 7 \\ 1 & 7 & -5 \\ 7 & -5 & 1 \end{bmatrix} \quad \text{with } \det A^{-1} = \frac{1}{18}$$

With that $\quad AA^{-1} = 1_n \quad$ and $\quad \det A \cdot \det A^{-1} = 1\,.$

Application to linear systems of equations

Let $\;A \in M_n(K)\,,\;\; \vec{b} \in K^n$

(L) $\qquad A\vec{x} = \vec{b}$

is solvable for every $\vec{b}$ if and only if A is invertible, that means if $\det A \neq 0$. If this is the case, then (L) is uniquely solvable, viz

$$\vec{x} = A^{-1}\vec{b}$$

Proof:

1) If $\det A \neq 0$ then $A^{-1}\vec{b} = \vec{x}$ is a solution of (L). On the other hand, if $\vec{x}$ is any solution of (L) then multiplication by A^{-1} from the left yields

$$A^{-1}\vec{b} = A^{-1}(A\vec{x}) = (A^{-1}A)\vec{x} = \vec{x}$$

2) If (L) has a solution for every $\vec{b}$, then choose $\vec{b} = \vec{e}_k$ and call $\vec{c}_k$ the solution $C = (\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n)$.

Summing-up all systems of equations

$$A(\vec{c}_1, \vec{c}_2, \ldots, \vec{c}_n) = (\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n)$$

in short $\qquad A \cdot C = 1_n$

Hence $\quad \det A \cdot \det C = 1\,, \quad \det A \neq 0$

$\square$

Remark:

(6) It can be shown (compare next chapter) that the “homogeneous system”

(H) $\qquad A\vec{x} = \vec{0}$

has at least one non-trivial solution $\vec{x} > \vec{0}$ if $\det A = 0\,.$

6. Vector Spaces, Linear Self-mappings, Eigenvalues

6.1 Vector Space Axioms

If the set V has an addition

$$\begin{aligned} +V \times V &\mapsto V \\ (\vec{x}, \vec{y}) &\mapsto \vec{x} + \vec{y} \end{aligned}$$

with

$$(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z}) \qquad \text{(associative)}$$

$$\vec{x} + \vec{y} = \vec{y} + \vec{x} \qquad \text{(commutative)}$$

and there exists a "zero element" $\vec{0} \in V$ with

$$\vec{x} + \vec{0} = \vec{x}$$

and there is a negative $-\vec{x}$ for every $\vec{x} \in V$ with

$$\vec{x} + (-\vec{x}) = \vec{0}$$

then V can become a vector space over a field K as the scalar region, if a scalar multiplication exists

$$\begin{aligned} \circ K \times V &\mapsto V \\ (\lambda, \vec{x}) &\mapsto \lambda\vec{x} \end{aligned}$$

with the properties

$$\lambda(\vec{x} + \vec{y}) = \lambda\vec{x} + \lambda\vec{y} \quad ; \quad 1_K\vec{x} = \vec{x}$$

$$\lambda(\mu\vec{x}) = (\lambda\mu)\vec{x}$$

$$(\lambda + \mu)\vec{x} \;=\; \lambda\vec{x} \,+\, \mu\vec{x}$$

for all $\vec{x}, \vec{y} \in V$ and all $\lambda, \mu \in K$.

Examples:

(1) K^n , $K^{\mathbb{N}}$ = set of all sequences with elements in K

(2) $M_n(K)$, $C_{\mathbb{R}}\big([0,1]\big)$: the continuous real functions on the interval $[0,1]$ with $K = \mathbb{R}$

Simple corollary from the vector space axioms

$$\lambda\vec{x} \;=\; \vec{0}_V \quad\Longleftrightarrow\quad \lambda \,=\, 0_K \quad \text{or} \quad \vec{x} \,=\, \vec{0}$$

6.1.1 Linear Combinations

$$\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_r \;\in\; V$$

Linear combinations are

$$\vec{x} \;=\; \sum_{i=1}^{r} \xi_i \vec{a}_i$$

The totality of these linear combinations forms a “subspace” $U = \langle \vec{a}_1, \ldots, \vec{a}_r \rangle$ of V, the linear hull (span) of the a_i . If $U = V$, i.e. if every $\vec{x} \in V$ is a linear combination of the $\vec{a}_i$ $(1 \le i \le r)$, then $\vec{a}_1, \ldots, \vec{a}_r$ is called the generating set of V. If each $\vec{x} \in U$ has only one representation as a linear combination of the $\vec{a}_i$ then the system $\big(\vec{a}_i\big)_{1 \le i \le r}$ is called linearly independent! If not, then the zero vector has a representation

$$\vec{0}_V \;=\; \sum_{i=1}^{r} \xi_i \vec{a}_i \qquad \text{with} \quad \xi_i \;=\; 0_K \quad \text{not for all} \quad i$$

and then $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_r$ are called linearly dependent.

Example:

(3) $K = \mathbb{R}$, $V = \mathbb{R}^3$, $r = 3$

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad ; \quad \vec{a}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad ; \quad \vec{a}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$U \;=\; \langle \vec{a}_1, \vec{a}_2, \vec{a}_3 \rangle \;=\; \mathbb{R}^3$$

because $\vec{e}_1, \vec{e}_2, \vec{e}_3 \in U$. $\vec{a}_1, \vec{a}_2, \vec{a}_3$ is a generating set of $\mathbb{R}^3$ and $\vec{a}_1, \vec{a}_2, \vec{a}_3$ are linearly independent; assumption:

$$\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \xi_1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \xi_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \xi_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

This homogeneous system of linear equations only has the solution: $\xi_1 = 0$, $\xi_2 = 0$, $\xi_3 = 0$.

If the system $\vec{a}_1, \dots, \vec{a}_r$ of vectors in V has the property:

generating set & linear independence

then $\vec{a}_1, \dots, \vec{a}_r$ is called the basis of V.

6.1.2 Steinitz Exchange Lemma

Let a linear independent system $\vec{a}_1, \dots, \vec{a}_r$ and a generating set $\vec{b}_1, \dots, \vec{b}_s$ be given in the vector space V . Then $r \leq s$ applies and $\vec{a}_1, \dots, \vec{a}_r$ can be supplemented by $s-r$ vectors $\vec{a}_j$ $(r+1 \leq j \leq s)$ from the $\vec{b}_1, \dots, \vec{b}_s$ to a further generator set of V . (Proof by induction over $r \geq 0$.)

Corollary:

Every two bases of a vector space V have the same number of elements. This number is called the "dimension" of V .

K^n has dimension $= n$

$M_n(K)$ has (as K-vector space) dimension n^2

The K-vector space of all sequences has no finite dimension, likewise $C_{\mathbb{R}}\big([0,1]\big)$ is infinite-dimensional.

Further consequence of the exchange lemma:

Every system $\vec{a}_1, \dots, \vec{a}_n$ of n linearly independent vectors in K^n is a basis of K^n . In contrast to this, any $n+1$ vectors $\vec{b}_1, \dots, \vec{b}_{n+1}$ in K^n are always linear dependent.

Application to matrices $A \in M_n(K)$

$\det A \neq 0$ if and only if the n columns $\vec{a}_1, \dots, \vec{a}_n$ of A are linearly independent.

Proof:

1) If $\vec{a}_1, \dots, \vec{a}_n$ are linearly independent, then they form a basis of K^n. Then

$$\text{(L)} \qquad A\vec{x} = \vec{b}$$

is always solvable because every $\vec{b} \in K^n$ is a linear combination of $\vec{a}_1, \ldots, \vec{a}_n$. Hence $\det A \neq 0$.

2) On the other hand, if $\vec{a}_1, \ldots, \vec{a}_n$ are linearly dependent, then

$$\vec{0} = \sum_{i=1}^{n} \vec{a}_i \xi_i \quad \text{with not all} \quad \xi_i = 0$$

For example $\xi_k \neq 0$

$$\vec{a}_k = \sum_{i \neq k} \vec{a}_i \left(-\xi_i/\xi_k\right)$$

Then according to (ML) and (ALT)

$$\begin{aligned} \det A &= \det(\vec{a}_1, \ldots, \vec{a}_k, \ldots, \vec{a}_n) \\ &= \det\left(\ldots, \vec{0}, \ldots\right) = 0 \end{aligned}$$

□

6.2 Subspaces and Dimension Formula

A general example for the notion of "subspace"

Let V, W be K-vector spaces. $f : V \mapsto W$ is called "linear mapping" if always

$$\begin{aligned} f(\vec{x} + \vec{y}) &= f(\vec{x}) + f(\vec{y}) \\ f(\lambda \vec{x}) &= \lambda f(\vec{x}) \end{aligned}$$

One denotes as the enquotekernel of f

$$U := \{\vec{x} \in V\,;\ f(\vec{x}) = \vec{0}_W\}$$

(abbreviation $U = \ker f$); U is a subspace.

1. $\vec{0}_V \in U$ because

$$\begin{aligned} f(\vec{0}_V) &= f(\vec{0}_V + \vec{0}_V) = f(\vec{0}_V) + f(\vec{0}_V) \\ \vec{0}_W &= f(\vec{0}_V) \end{aligned}$$

2. $\vec{x}, \vec{y} \in U \Rightarrow \vec{x} + \vec{y} \in U$

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + f(\vec{y}) = \vec{0}_W$$

3. because $\lambda \in K\ ;\ \vec{x} \in U \Rightarrow \lambda \vec{x} \in U$

$$f(\lambda \vec{x}) = \lambda f(\vec{x}) = \vec{0}_W$$

$f(V)$ is called the image of f. $W' = f(V)$ is a subspace of W.

1. $\vec{0}_W \;=\; f(\vec{0}_V) \in W'$

2. $\vec{v}, \vec{w} \in W'$, then $\vec{x}, \vec{y} \in V$ exist with $\vec{v} = f(\vec{x})$, $\vec{w} = f(\vec{y})$

$$\vec{v} + \vec{w} \;=\; f(\vec{x}) + f(\vec{y}) \;=\; f(\vec{x} + \vec{y})$$

3. $\lambda \in K$, $\vec{v} = f(\vec{x}) \in W'$ then

$$\lambda\vec{v} \;=\; \lambda f(\vec{x}) \;=\; f(\lambda\vec{x})$$

Definition:

The dimension of the image $f(V)$ is called "rank" of the linear mapping f, briefly $\operatorname{rk} f$.

Dimension formula

$$\boxed{\dim(\ker f) + \operatorname{rk} f \;=\; \dim V} \qquad \text{if } \dim V = n < \infty$$

Proof:

1) $r := \operatorname{rk} f = \dim V$

Then there is a basis $\vec{b}_1, \ldots, \vec{b}_r$ of $f(V) = W'$. Choose $\vec{a}_i \in V$ with $f(\vec{a}_i) = \vec{b}_i$ $(1 \le i \le r)$

2) $\vec{a}_1, \ldots, \vec{a}_r$ are linearly independent in V, because

$$\vec{0}_V \;=\; \sum_{i=1}^{r} \xi_i f(\vec{a}_i)$$

is a linear combination of $\vec{0}_V$.

Application of f

$$\vec{0}_W \;=\; f\Big(\sum \ldots\Big) \;=\; \sum_{i=1}^{r} \xi_i \underbrace{f(\vec{a}_i)}_{\vec{b}_i}$$

Because the $\vec{b}_1, \ldots, \vec{b}_r$ are linearly independent

$$\Rightarrow \;\xi_i = 0 \quad (1 \le i \le r)$$

3) Supplement the $\vec{a}_i$ $(1 \le i \le r)$ by $\vec{a}'_j$ $(r < i \le n)$ to a basis of V. Then

$$f(\vec{a}'_j) \in W', \quad \text{hence} \quad f(\vec{a}'_j) \;=\; \sum_{i=1}^{r} \lambda_{ij} \vec{b}_i$$

Replace $\vec{a}_j'$ with $\vec{a}_j' - \sum_{i=1}^{r} \lambda_{ij}\vec{a}_i = \vec{a}_j \quad (r < j \le n)$

4) $\vec{a}_1, \ldots, \vec{a}_r, \vec{a}_{r+1}, \ldots, \vec{a}_n$ are again a basis of V and $f(\vec{a}_j) = \vec{0}_W$

With that $\ker f = U = \langle \vec{a}_{r+1}, \ldots, \vec{a}_n \rangle$

$\square$

Applications to matrices and linear equations

Let B be an $m \times n$-matrix with coefficients in K.

$$\text{(L)} \quad B\vec{x} = \vec{y} \quad , \quad \vec{x} \in K^n, \ \vec{y} \in K^m$$

is asking the question if a $\vec{x}$ exists which with a given $\vec{y}$ will satisfy the equation (L).

On the left one has a linear combination of the columns $\vec{b}_1, \ldots, \vec{b}_n$ of B

Note: $B\vec{e}_i = \vec{b}_i$

(L) is solvable if and only if $\vec{y}$ is a linear combination of $\vec{b}_1, \ldots, \vec{b}_n$. In other words, $\vec{y} \in B(K^n)$.

$$\vec{y} \in B(K^n) = W' \subset K^m \qquad \text{if} \qquad \operatorname{rk} B = \operatorname{rk}(B, \vec{y})$$

The homogeneous system of equations

$$\text{(H)} \quad B\vec{x} = \vec{0}_W$$

is the question pertaining to the kernel of the linear mappings belonging to B. With the Gauß algorithm (elementary row transformations) the linear hull (span) of the rows of B does not change. The set U of solution vectors also remains unchanged.

$$\text{row rank}(B) + \dim U = n$$

$$\text{row rank } B = \text{column rank } B$$

6.3 Eigenvalues and Eigenvectors

Let $A \in M_n(K)$. $\lambda \in K$ is called “eigenvalue” of A, if

$$\boxed{A\vec{x} = \lambda\vec{x} \qquad (*)}$$

has a solution $\vec{x} \neq \vec{0}$. If this is the case, then $\vec{x}$ is called “eigenvector” to the eigenvalue λ of A.

Remarks:

(1) $(*)$ is equivalent to

$$(A - \lambda 1_n)\vec{x} \;=\; \vec{0}$$

λ is an eigenvalue if and only if $\det(A - \lambda 1_n) \;=\; 0$.

(2) By developing the determinant

$$P_A(\lambda) \;=\; \det(A - \lambda 1_n)$$

a polynomial of degree n in λ with leading coefficient $(-1)^n$ and the constant term $\det A$ is defined

$$P_A(\lambda) \;=\; (-1)^n \lambda^n \,+\, \ldots \,+\, \det A$$

$P_A(\lambda)$ is the so-called characteristic polynomial of A. Eigenvalues of A are exactly the zeros of $P_A(\lambda)$.

(3) $\lambda = 0$ is an eigenvalue of A if and only if $\det A = 0$.

Examples:

(1) $K = \mathbb{R}$, n odd. According to the intermediate value theorem, $P_A(\lambda)$ has at least one real zero for every $A \in M_n(\mathbb{R})$, so A has at least one eigenvector $\vec{x} \neq \vec{0}$.

(2) $n = 2$, $K = \mathbb{R}$ and

$$A \;=\; \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \qquad c^2 \,+\, s^2 \;=\; 1$$

$$P_A(\lambda) \;=\; \det \begin{bmatrix} c-\lambda & -s \\ s & c-\lambda \end{bmatrix} \;=\; (c-\lambda)(c-\lambda) \,+\, s^2 \;=\; \lambda^2 \,-\, 2c\lambda \,+\, 1$$

Zeros

$$\lambda_\pm \;=\; c \pm \sqrt{c^2 - 1}$$

$$\lambda_\pm \;=\; c \pm js \;\; (= \;\; \mathrm{e}^{\pm j\varphi}) \;\; ; \quad \text{in general, } \lambda_\pm \text{ are not real.}$$

(3) In the case $K = \mathbb{C}$, each polynomial decomposes into a product of linear factors, therefore for each $A \in M_n(\mathbb{C})$

$$P_A(\lambda) \;=\; (-1)^n \prod_{i=1}^{n} (\lambda - \lambda_i)$$

with the roots λ_i of $P_A(\lambda)$. Because of

$$(\lambda - \lambda_1) \dots (\lambda - \lambda_n) \;=\; \lambda^n + \dots + (-1)^n \lambda_1 \lambda_2 \dots \lambda_n$$

$\det A = \lambda_1 \lambda_2 \dots \lambda_n$ is the product of the eigenvalues.

Warning: $P_A(\lambda)$ can have multiple zeros, for example A:

$$A \;=\; \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}$$

has the characteristic polynomial $P_A(\lambda) = (-\lambda)^n$ which has $\lambda = 0$ as the n-fold root.

Diagonalizability of $A \in M_n(K)$ with n different eigenvalues $\lambda_1, \dots, \lambda_n$

If one forms from the eigenvectors $\vec{b}_k$ with $\vec{b}_k \neq \vec{0}$ and $A\vec{b}_k = \lambda_k \vec{b}_k$ $(1 \leq k \leq n)$ the matrix $B = (\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n)$, then one obtains the matrix equation

$$AB \;=\; B \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

Moreover, B is invertible, so $B^{-1}AB = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ is a diagonal matrix with the eigenvalues of A in the main diagonal.

Remark:

(4) If $PA(\lambda)$ has a factorization

$$P_A(\lambda) \;=\; (-1)^n \prod_{k=1}^{n} (\lambda - \lambda_k)$$

with all λ_k different, then A can (but does not have to) be diagonalizable! If $P_A(\lambda)$ is not a product of linear factors, then A is not diagonalizable.

Proof:

Problem: B is invertible, i.e. $\vec{b}_1, \dots, \vec{b}_n$ are linearly independent.

Proof step: eigenvectors $\vec{b}_1, \dots, \vec{b}_r$ for pairwise distinct eigenvalues are linearly independent.

$r = 1$: $\vec{b}_1 = \vec{0}$ by assumption, therefore $\vec{b}_1$ linearly independent.

$r \mapsto r + 1$: $(r < n)$; Assumption:

$$\vec{0} \;=\; \sum_{i=1}^{r+1} \xi_i \vec{b}_i \qquad (*)$$

Apply A

$$\vec{0} \;=\; A(\vec{0}) \;=\; \sum_{i=1}^{r+1} \xi_i \underbrace{A\vec{b}_i}_{\lambda_i \vec{b}_i} \;=\; \sum_{i=1}^{r+1} \xi_i \lambda_i \vec{b}_i$$

Multiplication of $(*)$ by λ_{r+1}

$$\vec{0} \;=\; \sum_{i=1}^{r+1} \xi_i \lambda_{r+1} \vec{b}_i$$

Difference of the two equations above

$$\vec{0} \;=\; \sum_{i=1}^{r} \xi_i (\lambda_i - \lambda_{r+1}) \vec{b}_i$$

Induction hypothesis $\Rightarrow$ $\xi_i = 0 \quad (1 \leq i \leq r)$. From $(*)$ remains

$$\vec{0} \;=\; \xi_{r+1} \vec{b}_{r+1} \;\Rightarrow\; \xi_{r+1} = 0$$

$\square$

Example:

(4) $$A \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in M_2(\mathbb{R})$$

$$P_A(\lambda) \;=\; \det \begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda+1 \end{pmatrix} \;=\; \lambda(\lambda - 1) - 1 \;=\; \lambda^2 - \lambda - 1$$

Zeros $\quad \lambda_\pm \;=\; \frac{1}{2} \pm \frac{1}{2}\sqrt{5}$

$$A \;=\; \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \quad ; \quad \lambda^2 \;=\; \lambda + 1 \qquad \text{(characteristic equation)}$$

$$A \begin{bmatrix} 1 \\ \lambda \end{bmatrix} \;=\; \begin{bmatrix} \lambda \\ \lambda+1 \end{bmatrix} \quad , \quad B \;=\; \begin{bmatrix} 1 & 1 \\ \lambda_+ & \lambda_- \end{bmatrix}$$

$$B^{-1} \;=\; \frac{1}{\lambda_- - \lambda_+} \begin{bmatrix} \lambda_- & -1 \\ -\lambda_+ & 1 \end{bmatrix}$$

$$A \;=\; B \begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix} B^{-1}$$

6.4 Euclidean and Unitary Scalar Products

The Euclidean scalar product in $\mathbb{R}^n$ can be expanded to a “unitary” scalar product in $\mathbb{C}^n$ by

$$\langle \vec{x}, \vec{y} \rangle \; := \; \sum_{k=1}^{n} \overline{x}_k y_k$$

Properties of the unitary scalar product (U)

$$\langle \vec{x}, \vec{y} + \vec{y}' \rangle \; = \; \langle \vec{x}, \vec{y} \rangle \; + \; \langle \vec{x}, \vec{y}' \rangle$$

$$\langle \vec{x}, \lambda\vec{y} \rangle \; = \; \lambda\langle \vec{x}, \vec{y} \rangle$$

$$\langle \vec{y}, \vec{x} \rangle \; = \; \overline{\langle \vec{x}, \vec{y} \rangle}$$

$$\langle \vec{x}, \vec{x} \rangle \; \geq \; 0 \qquad \text{“=” only if } \vec{x} = \vec{0}$$

Consider $\mathbb{C}$-vector spaces V with a “scalar product” $\langle -, -, \rangle$ with the properties (U), unitary vector spaces. (Subordinated is the case of “Euclidean” vector spaces, i.e. $\mathbb{R}$-vector spaces with a (U)-satisfying real-valued scalar product $\langle -, - \rangle$, therefore $\overline{\langle \vec{x}, \vec{y} \rangle} = \langle \vec{x}, \vec{y} \rangle$)

Common notation of unitary $\mathbb{K}$-vector spaces ($\mathbb{K} = \mathbb{R}$, $\mathbb{C}$).

Consequences from (U)

$$\langle \vec{x} + \vec{x}', \vec{y} \rangle \; = \; \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}', \vec{y} \rangle$$

$$\langle \lambda\vec{x}, \vec{y} \rangle \; = \; \overline{\lambda}\langle \vec{x}, \vec{y} \rangle$$

$$|\langle \vec{x}, \vec{y} \rangle|^2 \; \leq \; \langle \vec{x}, \vec{x} \rangle\langle \vec{y}, \vec{y} \rangle \qquad \text{(Cauchy-Schwarz inequality, or short C-S-I)}$$

Norm on V

$$\|\vec{x}\| \; = \; \sqrt{\langle \vec{x}, \vec{x} \rangle}$$

Properties

$$\|\vec{x}\| \; \geq \; 0 \;, \qquad \text{“=” only for } \vec{x} = \vec{0}$$

$$\|\lambda\vec{x}\| \; = \; |\lambda|\|\vec{x}\|$$

$$\|\vec{x} + \vec{y}\| \; \leq \; \|\vec{x}\| + \|\vec{y}\| \qquad \text{(triangle inequality; also follows from C-S-I)}$$

$\vec{x}, \vec{y} \in V$ are called "orthogonal" if $\langle \vec{x}, \vec{y} \rangle = 0$

Note: $\langle \vec{x}, \vec{y} \rangle = 0 \quad \Rightarrow \quad \langle \vec{y}, \vec{x} \rangle = 0$

Lemma 1

If $\vec{a}_1, \dots, \vec{a}_r$ are different from $\vec{0}$ and pairwise orthogonal then they are linearly independent ($\langle \vec{a}_i, \vec{a}_k \rangle = 0$ if $i \neq k$)

Proof:

Assumption $\displaystyle\sum_{k=1}^{r} \xi_k \vec{a}_k = \vec{0}$

Form scalar product with $\vec{a}_i$

$$0 = \langle \vec{a}_i, \vec{0} \rangle = \left\langle \vec{a}_i \, , \sum_{k=1}^{r} \xi_k \vec{a}_k \right\rangle$$

Linearity

$$0 = \sum_{k=1}^{r} \xi_k \underbrace{\langle \vec{a}_i, \vec{a}_k \rangle}_{\substack{=0 \\ \text{for } i \neq k}} = \xi_i \underbrace{\langle \vec{a}_i, \vec{a}_i \rangle}_{>0}$$

Hence $\xi_i = 0 \quad (1 \leq i \leq r)$

□

A system $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r$ of vectors in the unitary $\mathbb{K}$-vector space V is called an orthonormal system, if always $\langle \vec{a}_i, \vec{a}_k \rangle = \delta_{ik}$ $(1 \leq i, k \leq r)$.

6.4.1 Orthonormalization Theorem

Let V be an n-dimensional unitary $\mathbb{K}$-vector space. Every orthonormal system $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_r$ in V can be supplemented to an orthonormal basis $\vec{a}_1, \dots, \vec{a}_n$.

Proof by stepwise enlargement of r .

If $r < n$, then a $\vec{b} \notin \langle \vec{a}_1, \dots, \vec{a}_r \rangle = U$ with $\vec{b} \in V$ exists. Construction of the "orthogonal projection" from $\vec{b}$ onto U

$$\vec{b}_0 = \sum_{k=1}^{r} \langle \vec{a}_k, \vec{b} \rangle \vec{a}_k \qquad \vec{b}_0 \in U \, ;$$

$$\langle \vec{a}_i, \vec{b}_0 \rangle = \left\langle \vec{a}_i \, , \sum_{k=1}^{r} \langle \vec{a}_k, \vec{b} \rangle \vec{a}_k \right\rangle$$

$$= \sum_{k=1}^{r} \langle \vec{a}_i \vec{b} \rangle \underbrace{\langle \vec{a}_i, \vec{a}_k \rangle}_{\delta_{ik}} = \langle \vec{a}_i, \vec{b} \rangle$$

hence $\langle \vec{x}, \vec{b}_0 \rangle = \langle \vec{x}, \vec{b} \rangle$

i.e. $\vec{0} \neq \vec{b} - \vec{b_0}$ is orthogonal to every $\vec{x} \in U$

$$\vec{a}_{r+1} := (\vec{b} - \vec{b}_0) \| \vec{b} - \vec{b}_0 \|^{-1}$$

$\square$

Definition:

Let V be a unitary $\mathbb{K}$-vector space and $A : V \mapsto V$ be a linear self-mapping of V (endomorphism of V).
A is called "self-adjoint" if for all $\vec{x}, \vec{y} \in V$ $\langle \vec{x}, A\vec{y} \rangle = \langle A\vec{x}, \vec{y} \rangle$.

Remark:

(1) If $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ with $n = \dim V$ is an orthonormal basis and

$$A\vec{a}_i = \sum_{k=1}^{n} \alpha_{ik} \vec{a}_k$$

with suitable $\alpha_{ik} \in \mathbb{K}$ then A is self-adjoint if and only if $\overline{\alpha}_{ik} = \alpha_{ki}$ $(1 \leq i, k \leq n)$.
In the case $\mathbb{K} = \mathbb{R}$ this means that the matrix for A is symmetric.
In the case $\mathbb{K} = \mathbb{C}$ the matrix is said to be hermitian.

Lemma 2

A hermetian matrix $A \in M_n(\mathbb{C})$, i.e. $A = A^T$, has nothing but real eigenvalues.

Proof via $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^{n} \overline{x}_i y_i$

According to the fundamental theorem of algebra (about $\mathbb{C}$), A has at least one eigenvalue $\lambda \in \mathbb{C}$. Let $\vec{b}$ be an eigenvector to λ of A

$$\langle \vec{b}, A\vec{b} \rangle = \sum_{i=1}^{n} \overline{b}_i (A\vec{b})_i = \sum_{i=1}^{n} \overline{b}_i \sum_{k=1}^{n} a_{ik} b_k$$

$$= \sum_{i,k=1}^{n} \overline{b}_i \underset{\overline{a}_{ki}}{\underset{\downarrow}{a_{ik}}} b_k = \sum_{i,k=1}^{n} \overline{a_{ki} b_i} \cdot b_k$$

$$= \sum_{k=1}^{n} \left(\overline{A\vec{b}}\right)_k b_k \;=\; \left\langle A\vec{b}, \vec{b}\right\rangle$$

Comparing the left side (LS) and the right side (RS) of the equation gives

$$\left.\begin{array}{l} \text{LS}: \left\langle \vec{b}, A\vec{b}\right\rangle \;=\; \lambda \underbrace{\left\langle \vec{b}, \vec{b}\right\rangle}_{\neq 0} \\ \text{RS}: \left\langle A\vec{b}, \vec{b}\right\rangle \;=\; \overline{\lambda} \overbrace{\left\langle \vec{b}, \vec{b}\right\rangle}^{} \end{array}\right\} \Rightarrow \lambda = \overline{\lambda} \text{ that means } \lambda \text{ must be real}$$

□

Fundamental theorem about self-adjoint endomorphisms

Let A be a self-adjoint endomorphism of the unitary $\mathbb{K}$-vector space V. Then V has an orthonormal basis from eigenvectors of A.

Proof stepwise:

According to lemma 2, there also exists in the case $\mathbb{K} = \mathbb{R}$ an eigenvector to A, such as $\vec{a}_1$, without restriction with norm 1.

Consider $\quad U = \langle \vec{a}_1 \rangle^{\perp} = \{\vec{x} \in V \,;\; \langle \vec{a}_1, \vec{x} \rangle = 0\}$

Main step: A maps U to itself.

Let $\vec{x} \in U$, that means $\langle \vec{a}_1, \vec{x} \rangle = 0$

$$\begin{aligned} \langle \vec{a}_1, A\vec{x} \rangle &= \langle A\vec{a}_1, \vec{x} \rangle \qquad \text{because } A \text{ is self-adjoint} \\ &= \overline{\lambda}_1 \underbrace{\langle \vec{a}_1, \vec{x} \rangle}_{=0} \;=\; 0 \end{aligned}$$

First, because U has a smaller dimension than V $(\dim(n-1))$, and because A on U operates self-adjoint, U has according to the induction hypothesis an orthonormal basis $\vec{a}_2, \vec{a}_3, \ldots, \vec{a}_n$ from eigenvectors of A. It is supplemented with $\vec{a}_1$ to an orthonormal basis of V.

□

Application to $\mathbb{R}$: The principal axis transformation.

6.4.2 The Principal Axis Transformation

For every symmetric matrix $A \in M_n(\mathbb{R})$, $\mathbb{R}^n$ has an orthonormal basis $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n$ consisting of the eigenvectors of A. Expressed with the orthogonal matrix $B = (\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n)$

$$B^T B = 1_n \quad \text{and} \quad B^T A B = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$$

with the (real) eigenvalues of A in the diagonal.

Example:

(1)
$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \in M_3(\mathbb{R})$$

$$P_A(\lambda) = \det(A - \lambda 1_3) = \det \begin{bmatrix} -\lambda & 0 & 1 \\ 0 & -\lambda-1 & 1 \\ 1 & 1 & -\lambda \end{bmatrix}$$
$$= -\lambda^2(\lambda+1) + \lambda + 1 + \lambda = -(\lambda^3 + \lambda^2 - 2\lambda - 1)$$

As a characteristic polynomial of a symmetric 3×3 matrix, this polynomial has 3 real zeros. None is rational.

Zeros are:

$$\lambda' = 2\cos\frac{2\pi}{7} \qquad \lambda'' = 2\cos\frac{4\pi}{7} \qquad \lambda''' = 2\cos\frac{6\pi}{7}$$

Rationale via the seventh (primitive) roots of unity $\mathrm{e}^{2\pi i m/7}$, $m \neq 7$. These are the zeros of

$$\frac{x^7-1}{x-1} = x^6 + x^5 + x^4 + x^3 + x^2 + x + 1$$

$$\zeta^6 + \zeta^5 + \zeta^4 + \zeta^3 + \zeta^2 + \zeta + 1 = 0 \quad / \cdot \zeta^{-3}$$

$$\zeta^3 + \zeta^2 + \zeta + 1 + \zeta^{-1} + \zeta^{-2} + \zeta^{-3} = 0$$

$$\left(\zeta^3 + \zeta^{-3}\right) + \left(\zeta^2 + \zeta^{-2}\right) + \left(\zeta^1 + \zeta^{-1}\right) + 1 = 0$$

$$\left(\zeta + \zeta^{-1}\right)^3 + \left(\zeta + \zeta^{-1}\right)^2 + 2\left(\zeta^1 + \zeta^{-1}\right) - 1 = 0$$

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad ; \quad \lambda^3 + \lambda^2 - 2\lambda - 1 = 0$$

$$A - \lambda 1_3 = \begin{bmatrix} -\lambda & 0 & 1 \\ 0 & -1-\lambda & 1 \\ 1 & 1 & -\lambda \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & -\lambda \\ 0 & 1+\lambda & -1 \\ 0 & \lambda & -\lambda^2+1 \end{bmatrix}$$

From $P_A(\lambda) = 0$ it follows that

$$(\lambda^2-1)(\lambda+1) \,-\, \lambda \;=\; 0$$

hence

$$(\lambda^2-1)(\lambda+1) \;=\; \lambda$$

Now, use this for one more elementary transformation: λ times 2-nd row minus $(\lambda{+}1)$ times 3-rd row yields a new third row (UV = unit vector)

$$\begin{bmatrix} 1 & 1 & -\lambda \\ 0 & 1{+}\lambda & -1 \\ 0 & 0 & 0 \end{bmatrix} \quad\Rightarrow\quad \text{UV:}\quad \vec{b}(\lambda) \;=\; \begin{bmatrix} \lambda^2{+}\lambda{-}1 \\ 1 \\ \lambda{+}1 \end{bmatrix}$$

$$\left.\begin{aligned} \vec{b}_1 \;&=\; \vec{b}(\lambda')\,\left\|\vec{b}(\lambda')\right\|^{-1} \\ \vec{b}_2 \;&=\; \vec{b}(\lambda'')\,\left\|\vec{b}(\lambda'')\right\|^{-1} \\ \vec{b}_3 \;&=\; \vec{b}(\lambda''')\,\left\|\vec{b}(\lambda''')\right\|^{-1} \end{aligned}\right\} \quad B \;=\; (\vec{b}_1, \vec{b}_2, \vec{b}_3)$$

$$B^TB \;=\; 1_3 \quad , \quad B^TAB \;=\; \operatorname{diag}(\lambda', \lambda'', \lambda''')$$

Quadratic form

(in n variables $x_1, \ldots, x_n$ over $\mathbb{R}$) is called any homogeneous polynomial of degree 2 (each term $\neq\; 0$ has total degree 2).

$n=2$: $\quad ax_1^2 \,+\, bx_1x_2 \,+\, cx_2^2$

$n=3$: $\quad ax_1^2 \,+\, bx_1x_2 \,+\, cx_2^2 \,+\, dx_1x_3 \,+\, ex_2x_3 \,+\, fx_3^2$

General case

$$Q(\vec{x}) \;=\; \sum_{i,k=1}^{n} a_{ik}x_ix_k$$

For the pair $i \neq k$ two terms $\;a_{ik}x_ix_k\,,\;\; a_{ki}x_kx_i$

Consequentially, without any restriction, one can write the coefficient matrix A symmetrically

$$Q(\vec{x}) \;=\; \sum_{i,k=1} a_{ik}x_ix_k \;=\; \vec{x}\cdot A\vec{x} \qquad \text{(matrix notation)}$$

Application of principal axis transformation

With an orthogonal matrix B, $(B^T = B^{-1})$ holds $\;B^TAB \;=\; \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$

Coordinate transformation

$$\vec{x} \;=\; B\vec{x}$$

then $\quad \vec{x}^T \;=\; \vec{\xi}^{\,T} B^T$

$$Q(\vec{x}) \;=\; \tilde{Q}(\vec{\xi}) \;=\; Q(B\vec{\xi}) \;=\; \vec{\xi}^{\,T}(B^T A B)\vec{\xi} \;=\; (\xi_1, \ldots, \xi_n)\, A$$

$$A \;=\; \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$$

$$\tilde{Q}(\vec{\xi}) \;=\; \lambda_1 \xi_1^2 \,+\, \lambda_2 \xi_2^2 \,+\, \ldots \,+\, \lambda_n \xi_n^2$$

Example:

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

$$Q(\vec{x}) \;=\; (x_1, x_2, x_3, x_4) \underbrace{\begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 3 & 2 \\ 2 & 3 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix}}_{A} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Because the row sums are constant $(= 6)$, one has

$$\frac{1}{2}\big(\vec{\mathrm{e}}_1 + \vec{\mathrm{e}}_2 + \vec{\mathrm{e}}_3 + \vec{\mathrm{e}}_4\big) \;=\; \vec{b}_1$$

Eigenvector of the eigenvalue 6.

$$\vec{b}_1^{\perp} \;=\; U \;=\; \{\vec{x} \in \mathbb{R}^4\,;\; x_1 + x_2 + x_3 + x_4 = 0\}$$

Guess an orthonormal basis for U

$$\vec{b}_2 \;=\; \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix} \quad ; \quad \vec{b}_3 \;=\; \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \quad ; \quad \vec{b}_4 \;=\; \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

$$A\vec{b}_2 \;=\; \frac{1}{2}\begin{pmatrix} -4 \\ -4 \\ 4 \\ 4 \end{pmatrix} \;=\; -4 \cdot \vec{b}_2 \quad ; \quad A\vec{b}_3 \;=\; \frac{1}{2}\begin{pmatrix} -2 \\ 2 \\ -2 \\ 2 \end{pmatrix} \;=\; -2 \cdot \vec{b}_3$$

$$A\vec{b}_4 \;=\; \frac{1}{2}\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \;=\; 0\cdot\vec{b}_4$$

With $\qquad B \;=\; \dfrac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$ $\qquad$ holds

$$B = B^T = B^{-1} \qquad (B^2 = 1_4)$$

$$B^T\!AB \;=\; \mathrm{diag}\,(6,-4,-2,0)$$

$$\tilde{Q}(\vec{\xi}\,) \;=\; 6\xi_1^2 \,-\, 4\xi_2^2 \,-\, 2\xi_3^2$$

7. Linear Differential Equation with Constant Coefficients

7.1 Differential Equation of Growth and Decay

$$\boxed{\dot{x} \;=\; \alpha \cdot x \qquad (*)} \qquad \alpha \in \mathbb{K} \quad (\mathbb{K} = \mathbb{R} \text{ or } \mathbb{K} = \mathbb{C})$$

Assumption: $f(t)$ be a solution of $(*)$ on the interval I.

Auxiliary function: $g(t) \;=\; f(t)\mathrm{e}^{-\alpha t}$; f, g both differentiable

$$\begin{aligned}
\dot{g}(t) &= \dot{f}(t)\mathrm{e}^{-\alpha t} \;-\; \alpha f(t)\mathrm{e}^{\alpha t} \\
&= \big(\underbrace{\dot{f}(t) - \alpha f(t)}_{\equiv\, 0\;\forall\, t}\big)\mathrm{e}^{\alpha t}
\end{aligned}$$

$$\Rightarrow \quad \dot{g}(t) \;=\; 0 \quad ; \quad g(t) \;=\; \text{const} \;=\; g(0)$$

$$f(t) \;=\; c \cdot \mathrm{e}^{\alpha t}$$

<u>Results:</u>

For each initial value c, $(*)$ has exactly one solution f with $f(0) = c$ (namely $f(t) = c \cdot \mathrm{e}^{\alpha t}$).
With any two solutions f_1, f_2, the linear combinations

$$\varphi \;=\; \beta_1 f_1 \;+\; \beta_2 f_2$$

are also solutions of $(*)$. Every solution is defined on all of $\mathbb{R}$. Any solution other than the zero function is even free of zeros.

7.1.1 Matching Inhomogeneous Differential Equation

$$\boxed{\dot{x} \;=\; \alpha\cdot x + g \qquad (**)}$$

with a $\mathbb{K}$-valued function $g(t)$, defined and continuous on I.

Ansatz to a solution by varying the constants

$$\begin{aligned}\varphi(t) &= c(t)\mathrm{e}^{\alpha t} \\ \dot{\varphi}(t) &= \dot{c}(t)\mathrm{e}^{\alpha t} + \alpha c(t)e^{\alpha t} \qquad \text{and since } \varphi \text{ is the solution} \\ &= \alpha\varphi(t) + g(t)\end{aligned}$$

Hence $\quad \dot{c}(t) = \mathrm{e}^{-\alpha t}g(t)$

This differential equation for c can be solved by "quadrature"

$$c(t) \;=\; \int_0^t \mathrm{e}^{\alpha s}g(s)\,\mathrm{d}s$$

Insertion into $\quad \varphi(t) = c(t)\mathrm{e}^{\alpha t}$

$$\varphi(t) \;=\; \mathrm{e}^{\alpha t}\int_0^t \mathrm{e}^{\alpha s}g(s)\,\mathrm{d}s$$

<u>Results:</u>

By varying the constants in the general solution of $(*)$ one obtains a particular solution φ of $(**)$.
The entirety of all solutions of $(**)$ takes the form

$$\varphi + \psi$$

where ψ runs through all solutions of $(*)$.

7.2 Differential Equation of Damped Oscillation

$$\boxed{\ddot{x} + 2\mu\dot{x} + \omega_0^2 x \;=\; 0 \qquad (*)}$$

$2\mu \geq 0$: damping constant

$\omega_0 > 0$: angular frequency of the undamped oscillation

Exponential ansatz for a solution function φ

$$\varphi(t) = e^{\lambda t} \quad ; \quad \dot{\varphi}(t) = \lambda e^{\lambda t} \quad ; \quad \ddot{\varphi}(t) = \lambda^2 e^{\lambda t}$$

$$(\lambda^2 + 2\mu\lambda + \omega_0^2)e^{\lambda t} = 0$$

It must apply

$$\lambda^2 + 2\mu\lambda + \omega_0^2 = 0 \quad \text{(characteristic equation)}$$

Zeros: $\lambda_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega_0^2}$

Normal case: $\mu^2 \neq \omega_0^2$

Two different solution functions $\varphi_1(t) = e^{\lambda_1 t}$; $\varphi_2(t) = e^{\lambda_2 t}$ follow from the ansatz.

Initial values

$$\begin{pmatrix} \varphi_1(0) \\ \dot{\varphi}_1(0) \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_1 \end{pmatrix} \quad ; \quad \begin{pmatrix} \varphi_2(0) \\ \dot{\varphi}_2(0) \end{pmatrix} = \begin{pmatrix} 1 \\ \lambda_2 \end{pmatrix}$$

form a system of two linearly independent vectors

$$\det \begin{pmatrix} \varphi_1(0) & \varphi_2(0) \\ \dot{\varphi}_1(0) & \dot{\varphi}_2(0) \end{pmatrix} = \lambda_2 - \lambda_1 \neq 0$$

Therefore the initial value problem is always solvable: For given initial values $c_1, c_2 \in \mathbb{K}$ there exists a function φ on $\mathbb{R}$ which satisfies $(*)$ as well as the initial conditions $\varphi(0) = c_1$, $\dot{\varphi}(0) = c_2$.

The solution set of $(*)$ contains the zero function and with two functions φ, ψ also each linear combination of φ and ψ .

Hence for any $\alpha_1, \alpha_2 \in \mathbb{R}$,

$$\varphi(t) = \alpha_1\varphi_1(t) + \alpha_2\varphi_2(t)$$

is also a solution of $(*)$ $\quad \Rightarrow$

$$\dot{\varphi}(t) = \alpha_1\lambda_1\varphi_1(t) + \alpha_2\lambda_2\varphi_2(t)$$

$$\left.\begin{aligned} \varphi(0) &= \alpha_1 + \alpha_2 \\ \dot{\varphi}(0) &= \alpha_1\lambda_1 + \alpha_2\lambda_2 \end{aligned}\right\} \quad \text{system of linear equations in } \alpha_1, \alpha_2$$

with coefficient matrix:

$$\begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \;,\quad \det \neq 0 \quad \text{hence uniquely solvable}$$

The solution space of $(*)$ is a two-dimensional vector space. A solution φ with the initial values $\varphi(0) = \dot{\varphi}(0) = 0$ is sufficient. Hereto in general for two solutions φ, ψ of $(*)$

$$W(\varphi, \psi) \;:=\; \varphi\dot{\psi} - \dot{\varphi}\psi \qquad \text{(Wronksi determinant)}$$

W is differentiable as a function of t:

$$\begin{aligned} \dot{W}(t) &= \varphi(t)\ddot{\psi}(t) - \ddot{\varphi}(t)\psi(t) \\ &= \varphi(t)(-2\mu\dot{\psi}(t) - \omega_0^2\psi(t)) - (-2\mu\dot{\varphi}(t) - \omega_0^2\varphi(t))\psi(t) \\ &= -2\mu(\varphi(t)\dot{\psi}(t) - \dot{\varphi}(t)\psi(t)) \;=\; -2\mu W \end{aligned}$$

W is solution to the differential equation of growth and decay!

Therefore W is either zero-free or the zero function, especially in the case $\varphi(0) = \dot{\varphi}(0) = 0$. For every solution ψ of $(*)$ now holds $W(\varphi, \psi) = 0$ $(\forall t)$, $\psi = \varphi_1$ and $\psi = \varphi_2$

$$\begin{aligned} \varphi(t)\lambda_1 e^{\lambda_1 t} - \dot{\varphi}(t)e^{\lambda_1 t} &= 0 \\ \varphi(t)\lambda_2 e^{\lambda_2 t} - \dot{\varphi}(t)e^{\lambda_2 t} &= 0 \end{aligned}$$

Hence

$$\left.\begin{aligned} \lambda_1\varphi &= \dot{\varphi} \\ \lambda_2\varphi &= \dot{\varphi} \end{aligned}\right\} \;\Rightarrow\; \underbrace{(\lambda_1 - \lambda_2)}_{\neq 0}\varphi(t) \;=\; 0 \quad \forall\, t$$

$$\Rightarrow \;\; \varphi(t) \;=\; 0 \quad \forall\, t$$

<u>The limit case:</u> $\quad \mu^2 \;=\; \omega_0^2$

Characteristic equation $\lambda^2 + 2\mu\lambda + \mu^2 \;=\; 0$ has $\lambda = -\mu$ as double root.

One solution $\quad \varphi_1(t) \;=\; e^{-\mu t} \quad ; \quad$ further solution $\quad \varphi_2(t) \;=\; te^{-\mu t}$

Rationale:

$$\begin{aligned} \dot{\varphi}_2(t) &= -\mu t e^{-\mu t} + e^{-\mu t} \\ \ddot{\varphi}_2(t) &= \mu^2 t e^{-\mu t} - 2\mu e^{-\mu t} \end{aligned}$$

Inserting into $(*)$ with $\omega_0^2 = \mu^2$ yields

$$\ddot{\varphi}_2(t) + 2\mu\dot{\varphi}_2 + \mu^2\varphi_2 = \mathrm{e}^{-\mu t}\big[(\underbrace{\mu^2 - 2\mu + \mu^2}_{=0}) + (\underbrace{-2\mu + 2\mu}_{=0})\big] = 0$$

Initial values

$$\begin{pmatrix}\varphi_1(0)\\ \dot{\varphi}_1(0)\end{pmatrix} = \begin{pmatrix}1\\ -\mu\end{pmatrix} \quad ; \quad \begin{pmatrix}\varphi_2(0)\\ \dot{\varphi}_2(0)\end{pmatrix} = \begin{pmatrix}0\\ 1\end{pmatrix} \quad \Rightarrow$$

initial value problem is (also in the limit case) solvable: $\varphi = \alpha_1\varphi_1 + \alpha_2\varphi_2$

Here too, $W(\varphi, \psi)$ is formed to determine the dimension of the solution pair. Again $\dot{W} = -2\mu W$. Applied to a solution φ with $\varphi(0) = \dot{\varphi}(0) = 0$

$$W(\varphi, \psi)(t) = 0 \quad \text{for all solutions } \psi \text{ of } (*)$$

Substitute $\psi = \varphi_1$ and $\psi = \varphi_2$

$$0 = W(\varphi, \varphi_1)(t) = \varphi(t)\dot{\varphi}_1(t) - \dot{\varphi}(t)\varphi_1(t) = (-\mu\varphi(t) - \dot{\varphi}(t))\mathrm{e}^{-\mu t}$$

$$\Rightarrow \quad \dot{\varphi}(t) = -\mu\varphi(t)$$

$$\begin{aligned} 0 = W(\varphi, \varphi_2)(t) &= \varphi(t)\dot{\varphi}_2(t) - \dot{\varphi}(t)\varphi_2(t) \\ &= \varphi(t)\left(-\mu t\mathrm{e}^{-\mu t} + \mathrm{e}^{-\mu t}\right) - \dot{\varphi}(t)\, t\mathrm{e}^{-\mu t} \\ &= \varphi(t)\left(-\mu t\mathrm{e}^{-\mu t} + \mathrm{e}^{-\mu t}\right) - (-\mu\varphi(t))\, t\mathrm{e}^{-\mu t} = \varphi(t)\mathrm{e}^{-\mu t} \end{aligned}$$

$$\Rightarrow \quad \varphi(t) = 0 \quad \forall\, t$$

<u>Remark</u> on the case of low damping: $\mu^2 < \omega_0^2$

$$\lambda^2 + 2\mu\lambda + \omega_0^2 = 0$$

Roots $-\mu \pm \sqrt{\underbrace{\mu^2 - \omega_0^2}_{-\omega^2 < 0}} = -\mu \pm j\omega$ not real

Solutions

$$\varphi_1(t) = \mathrm{e}^{\lambda_1 t} = \mathrm{e}^{-\mu t}\mathrm{e}^{+j\omega t} = \mathrm{e}^{-\mu t}(\cos\omega t + j\sin\omega t)$$

$$\varphi_2(t) = \mathrm{e}^{\lambda_2 t} = \mathrm{e}^{-\mu t}\mathrm{e}^{-j\omega t} = \mathrm{e}^{-\mu t}(\cos\omega t - j\sin\omega t)$$

Real solutions result from linear combinations

$$\frac{1}{2}\big(\varphi_1(t) + \varphi_2(t)\big) = \mathrm{e}^{-\mu t}\cos\omega t$$

$$\frac{1}{2j}\big(\varphi_1(t) - \varphi_2(t)\big) = \mathrm{e}^{-\mu t}\sin\omega t$$

7.2.1 Matching Inhomogeneous Linear Differential Equation

$$\boxed{\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = g \qquad (**)}$$

is the differential equation of forced oscillation.

Because the left-hand side is a "linear differential operator" for the function $x(t)$, the difference between two solutions of $(**)$ is a solution of $(*)$, in other words: from a single solution φ of $(**)$ one gets all solutions in the form $\varphi + \psi$, where ψ runs through all solutions of $(*)$.

The solution space of $(*)$ $(g \equiv 0)$ is a 2-dimensional vector space. Every basis ψ_1, ψ_2 of the solution space is called a "fundamental system" of $(*)$.

Superposition Principle

If φ is a solution of $\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = g$ and if ψ is a solution of $\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = h$ (where g, h are declared on the same interval) then $\alpha\varphi + \beta\psi$ is solution of

$$\ddot{x} + 2\mu\dot{x} + \omega_0^2 x = \alpha g + \beta h$$

(Because the left side of the differential equation is a linear operator for the solution function.)

<u>Special case:</u> $g(t) = a\mathrm{e}^{\lambda t}$

<u>1)</u> λ not a zero of the characteristic equation

$$x^2 + 2\mu x + \omega_0^2 = 0$$

Ansatz $\varphi(t) = c\mathrm{e}^{\lambda t}$ with $c = c(\lambda)$

$$\dot{\varphi} = c\lambda\mathrm{e}^{\lambda t} \quad ; \quad \ddot{\varphi} = c\lambda^2\mathrm{e}^{\lambda t}$$

$$[\ddot{\varphi} + 2\mu\dot{\varphi} + \omega_0^2\varphi - g] = \mathrm{e}^{\lambda t}[c\lambda^2 + 2\mu c\lambda + \omega_0^2 c - a] = 0$$

if $c = \dfrac{a}{\lambda^2 + 2\mu\lambda + \omega_0^2}$

<u>2)</u> λ is a zero of the characteristic equation

<u>A)</u> $x^2 + 2\mu x + \omega_0^2 = 0$ has two different roots: λ_1, λ_2 $(\mu^2 \neq \omega_0^2)$

$\lambda = \lambda_i$: Ansatz $\varphi(t) = cte^{\lambda_i t}$

$$\dot{\varphi}(t) = ct\lambda_i\mathrm{e}^{\lambda_i t} + c\mathrm{e}^{\lambda_i t}$$

$$\ddot{\varphi}(t) \;=\; ct\lambda_i^2 e^{\lambda_i t} \;+\; 2c\lambda_i e^{\lambda_i t}$$

$$\ddot{\varphi}(t) \;+\; 2\mu\dot{\varphi}(t) \;+\; \omega_0^2\varphi(t) \;-\; g$$

$$= \; e^{\lambda_i t}\Big[ct(\underbrace{\lambda_i^2 \;+\; 2\mu\lambda_i \;+\; \omega_0^2}_{=0}) \;+\; (2c\lambda_i \;+\; 2\mu c \;-\; a)\Big] \;=\; 0$$

if $\;c \;=\; \dfrac{a}{2(\lambda_i + \mu)}$

B) $\omega_0^2 = \mu^2$; $\lambda = -\mu$ is the double root of the characteristic equation

Ansatz: $\quad \varphi(t) \;=\; ct^2 e^{-\mu t}$

$$\dot{\varphi}(t) \;=\; -c\mu t^2 e^{-\mu t} \;+\; 2ct e^{-\mu t}$$

$$\ddot{\varphi}(t) \;=\; c\mu^2 t^2 e^{-\mu t} \;-\; 4c\mu t e^{-\mu t} \;+\; 2c e^{-\mu t}$$

$$\ddot{\varphi}(t) \;+\; 2\mu\dot{\varphi}(t) \;+\; \omega_0^2\varphi(t) \;-\; g$$

$$= \; e^{-\mu t}\Big[ct^2(\underbrace{\mu^2 - 2\mu^2 + \mu^2}_{=0}) \;+\; ct(\underbrace{-4c\mu + 4c\mu}_{=0}) \;+\; (2c - a)\Big] \;=\; 0$$

if $\;c \;=\; \dfrac{a}{2}$

7.2.2 Rewriting the 2nd Order Differential Equation into a System of 1st Order Differential Equations

With $\;y_1 = x \;\;,\;\; y_2 = \dot{x}$

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} \;=\; \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\mu \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \qquad \text{in the homogeneous case}$$

short $\quad \dot{\vec{y}} \;=\; A\vec{y}$

According to what has been proven, there exists a fundamental system ψ_1, ψ_2 with the initial conditions

$$\psi_1(0) \;=\; 1 \quad ; \quad \psi_2(0) \;=\; 0$$

$$\dot{\psi}_1(0) \;=\; 0 \quad ; \quad \dot{\psi}_2(0) \;=\; 1$$

Matrix notation

$$U(t) \;=\; \begin{pmatrix} \psi_1(t) \;,\; \psi_2(t) \\ \dot{\psi}_1(t) \;,\; \dot{\psi}_2(t) \end{pmatrix}$$

In matrix notation: U satisfies the differential equation

$$\dot{U}(t) \;=\; AU(t)$$

with initial condition $U(0) \;=\; 1_2$

Rewriting the general solution of the homogeneous differential equation

$$\vec{y} \;=\; \begin{pmatrix} \psi_1 \\ \dot{\psi}_1 \end{pmatrix} c_1 \;+\; \begin{pmatrix} \psi_2 \\ \dot{\psi}_2 \end{pmatrix} c_2 \;=\; U(t)\vec{c} \quad ; \quad \vec{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

Inhomogeneous linear differential equation $(**)$ in matrix notation

$$\dot{\vec{y}} \;=\; A\vec{y} + \vec{b}$$

with a vector-valued function which is continuous on an interval.

Solution ansatz by varying the constants

$$\vec{\varphi}(t) \;:=\; U(t)\vec{c}(t)$$

$$\dot{\vec{\varphi}}(t) \;=\; \dot{U}(t)\vec{c}(t) \;+\; U(t)\dot{\vec{c}}(t)$$

$$\dot{\vec{\varphi}}(t) \;=\; A\,\underbrace{U(t)\vec{c}(t)}_{\vec{\varphi}(t)} + U(t)\dot{\vec{c}}(t)$$

$$=\; A\vec{\varphi}(t) \;+\; b(t)$$

$$U(t) \;=\; \begin{pmatrix} \psi_1(t) \;,\; \psi_2(t) \\ \dot{\psi}_1(t) \;,\; \dot{\psi}_2(t) \end{pmatrix} \quad \text{has the determinant} \quad \psi_1\dot{\psi}_2 - \dot{\psi}_1\psi_2 \;=\; W(\psi_1, \psi_2)$$

This determinant has no zeros, $U(t)$ is invertible

$$\dot{\vec{c}}(t) \;=\; U^{-1}(t)\vec{b}(t)$$

Solvable by quadrature

$$\vec{c}(t) \;=\; \int_{t_0}^{t} U^{-1}(s)\vec{b}(s)\,\mathrm{d}s$$

$$\vec{\varphi}(t) \;=\; U(t)\int_{t_0}^{t} U^{-1}(s)\vec{b}(s)\,\mathrm{d}s$$

7.3 Systems of 1st Order Linear Differential Equations with Constant Coefficients

$$\boxed{\dot{\vec{y}} \;=\; A\vec{y} \qquad \text{(H)}} \;,\quad A \in M_n(\mathbb{K})$$

Wanted: $\vec{y}\,:\;\mathbb{R} \mapsto \mathbb{K}^n$

Definition:

Explicit 1st order ordinary differential equation

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

where e.g. f continuous in both x, y, $x \in I$, interval in $\mathbb{R}$; initial condition $y(x_0) \;=\; b$

Second look at the differential equation while recalling the fundamental theorem of calculus. The solution function $y(x)$ satisfies

$$y(x) \;-\; y(x_0) \;=\; \int\limits_{x_0}^{x} f(t, y(t))\,\mathrm{d}t$$

in other words, the solution function $y(x)$ is fixed point of the integral operator T_φ

$$\big(T_\varphi\big)(x) \;=\; b \;+\; \int\limits_{x_0}^{x} f(t, \varphi(t))\,\mathrm{d}t$$

7.3.1 Picard-Iteration Ansatz

$y_0(x)$ arbitrary (suitable) with $y_0(x_0) \;=\; b$

$$\boxed{y_{N+1}(x) \;=\; b \;+\; \int\limits_{x_0}^{x} f(t, y_N(t))\,\mathrm{d}t}$$

Iteration both for scalar-valued y, f, b and for vector-valued $\vec{y}, \vec{f}, \vec{b}$.

$$\boxed{\dot{Y} \;=\; AY \qquad \text{(HM)}} \;,\quad A \in M_n(\mathbb{K}) \text{ with}$$

initial condition: $Y(0) = 1_n$ with a solution curve $Y\,:\,\mathbb{R} \mapsto M_n(\mathbb{K})$.

From a solution $U(t)$ of (HM) one obtains the general solution of (H).

Ansatz: Picard iteration

$$
\begin{aligned}
Y_0 &= 1_n \\
Y_1(t) &= 1_n + \int_0^t A \,\mathrm{d}s = 1_n + At \\
Y_2(t) &= 1_n + \int_0^t AY_1(s)\,\mathrm{d}s = 1_n + \int_0^t A(1_n + As)\,\mathrm{d}s \\
&= 1_n + \int_0^t (A + A^2 s)\,\mathrm{d}s = 1_n + At + \frac{A^2}{2}t^2 \\
Y_N(t) &= \sum_{k=0}^{N} \frac{1}{k!}A^k t^k
\end{aligned}
$$

Induction on N: $N \to N+1$

$$
\begin{aligned}
Y_{N+1} &= 1_n + \int_0^t AY_N(s)\,\mathrm{d}s \\
&= 1_n + \sum_{k=0}^{N} \frac{1}{k!}A^{k+1} \int_0^t s^k\,\mathrm{d}s \\
&= 1_n + \sum_{k=0}^{N} \frac{1}{(k+1)!}A^{k+1}t^{k+1}
\end{aligned}
$$

Convergence question for the matrix sequence $\big(Y_N(t)\big)_{N\geq 0}$

7.4 Matrix Norms

Norms in $\mathbb{K}$-vector spaces V are functions $N : V \mapsto \mathbb{R}$ with

N1) $N(\vec{x}) \geq 0$, "=" only for $\vec{x} = \vec{0}$

N2) $N(\lambda\vec{x}) = |\lambda| N(\vec{x})$ homogeneous

N3) $N(\vec{x}+\vec{y}) \leq N(\vec{x}) + N(\vec{y})$

Examples:

(1) in $\mathbb{K}^n$, the Euclidean norm

$$N(\vec{x}) = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{1/2} = \|\vec{x}\|_{\mathrm{e}}$$

(2) in $\mathbb{K}^n$, the maximum norm

$$N(\vec{x}) = \max_{1 \le i \le n} |x_i| = \|\vec{x}\|_{\max}$$

Theorem of the equivalence of norms

In a finite dimensional $\mathbb{K}$-vector space, any two norms have the same Cauchy sequence and V is complete, that means, every Cauchy sequence converges (proof elementary).

In particular, $M_n(\mathbb{K})$ is a finite dimensional vector space. For matrix norms, in addition to N1), N2), N3) one requires:

N4) $N(AB) \le N(A)\,N(B) \quad \forall\, A, B \in M_n(\mathbb{K})$ (submultiplicative)

Examples:

(3) Euclidean norm (Frobenius, Schur)

$$A = \big(a_{ik}\big) \quad , \quad \|A\|_{\mathrm{e}} = \Big(\sum_{i,k=1}^{n} |a_{ik}|^2\Big)^{1/2}$$

(4) Matrix norm for matrices

$$\|A\|_{\max} = n \cdot \max_{1 \le i,k \le n} |a_{ik}|$$

Proof of the submultiplicativity of $\|A\|_{\mathrm{e}}$

$$\|AB\|_{\mathrm{e}} = \Big(\sum_{i,k} \underbrace{\Big(\sum_{l=1}^{n} |a_{il}|^2\Big)}_{s_i} \underbrace{\Big(\sum_{m=1}^{n} |b_{mk}|^2\Big)}_{t_k}\Big)^{1/2}$$

$$= \Big(\sum_{i,k=1}^{n} s_i t_k\Big)^{1/2} = \Big(\sum s_i\Big)^{1/2}\Big(\sum t_k\Big)^{1/2}$$

$$= \|A\|_{\mathrm{e}} \|B\|_{\mathrm{e}}$$

□

Objective: Show that $\big(Y_N(t)\big)_{N \ge 0}$ is a Cauchy sequence in $M_n(\mathbb{K})$ with limit value $\exp(At)$.

7.4.1 Exponential Function on $M_n(\mathbb{K})$

For each $A \in M_n(\mathbb{K})$ the series

$$\exp(A) \;=\; \sum_{k=0}^{\infty} \frac{1}{k!} A^k$$

converges. Furthermore, the functional equation applies to interchangeable matrices A, B (i.e. $AB = BA$)

$$\boxed{\exp(A+B) \;=\; \exp(A)\exp(B)}$$

Proof:

1) The partial sums

$$s_N(A) \;=\; \sum_{k=0}^{N} \frac{1}{k!} A^k \qquad \text{form a Cauchy sequence.}$$

The gist: The estimate is traced back to the Taylor series of the ordinary exponential function.

$$\|s_N(A) - s_{N+p}(A)\| \;=\; \left\| \sum_{k=N+1}^{N+p} \frac{1}{k!} A^k \right\|$$

because of N2) and N3):

$$\leq \sum_{k=N+1}^{N+p} \frac{1}{k!} \|A^k\|$$

because of N4):

$$\leq \sum_{N+1}^{N+p} \frac{1}{k!} \|A\|^k \;\leq\; \frac{\|A\|^{N+1}}{(N+1)!}\, \mathrm{e}^{\|A\|}$$

because of the Lagrangian remainder for the ordinary exponential series!

2) Main idea of the functional equation

$$s_N(A) s_N(B) \,-\, s_N(A+B) \;=\; \Big(\sum_{k=1}^{N} \frac{1}{k!} A^k\Big)\Big(\sum_{r=1}^{N} \frac{1}{r!} B^r\Big)$$

$$=\; \sum_{m=1}^{N} \frac{1}{m!} (A+B)^m$$

Binomial formula

$$\sum_{k=0}^{n} \frac{1}{k!}A^k \frac{1}{(m-k)!}B^{m-k} = \frac{1}{m!}(A+B)^m$$

$$= \sum_{\substack{k,r \le N \\ N \le k+r}} \frac{1}{k!}A^k \frac{1}{r!}B^r \to 0 \quad \text{if} \quad N \to \infty$$

$\square$

Special values of the exponential function for matrices

$$\exp(0_n) = 1_n = \exp(A-A) = \exp(A)\exp(-A)$$

In particular, $\exp(A)$ is invertible for every matrix $A \in M_n(\mathbb{K})$.

$$A = D = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

$$D^k = \operatorname{diag}(\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k) \quad \forall\, k$$

$$\exp(D) = \exp\big(\operatorname{diag}(\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k)\big) = \operatorname{diag}(e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n})$$

If $A \in M_n(\mathbb{K})$ is diagonalizable, i.e

$$S^{-1}AS = D = \operatorname{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

$$(S^{-1}AS)^2 = D^2 = \operatorname{diag}(\lambda_1^2, \lambda_2^2, \dots, \lambda_n^2)$$

$$A = SDS^{-1} \quad ; \quad A^2 = SD^2S^{-1} \quad ; \quad A^k = SD^kS^{-1}$$

$$\Rightarrow \exp(A) = S\exp(D)S^{-1} = S\operatorname{diag}(e^{\lambda_1}, e^{\lambda_2}, \dots, e^{\lambda_n})S^{-1}$$

The Picard iteration of (HM) with start function

$$Y_0(t) = 1_n \quad \text{if} \quad Y_N(t) = \sum_{k=0}^{N} \frac{1}{k!}A^k t^k$$

has a limit value for $N \to \infty$, viz

$$U(t) = \exp(At)$$

$U(t)$ is a solution of the differential equation (HM). Develop

$$\frac{1}{h}\big(U(t+h) - U(t)\big) = \frac{1}{h}\big(\exp(At + Ah) - \exp(At)\big) \quad \text{functional equation}$$

$$= \frac{1}{h}\big(\exp(Ah)\exp(At) - \exp(At)\big)$$

$$= \frac{1}{h}\big(\exp(Ah) - 1_n\big)\exp(At)$$

$$= \frac{1}{h}\big(\exp(Ah) - \exp(Ah)\exp(-Ah)\big)\exp(At)$$
$$= \frac{1}{h}\exp(Ah)\Big(1 - \underbrace{\frac{1}{\exp(Ah)}}_{\substack{\to 0 \\ \text{for } h\to 0}}\Big)\exp(At)$$

Series expansion for the first factor

$$\frac{1}{h}\exp(Ah) = A + \sum_{k=2}^{\infty} \underbrace{\frac{1}{h!}A^k h^{k-1}}_{\text{Rest}}$$

estimate of the remainder

$$\left\|\frac{1}{h!}A^k h^{k-1}\right\| \le \frac{\|A\|^2|h|}{2}\exp\big(\|A\|\cdot|h|\big) \qquad \text{(Lagrange)}$$

The remainder approaches 0 for $h \to 0$, so $U(t)$ is differentiable for all real t with derivative

$$\dot{U}(t) = AU(t)$$

7.5 General 1st Order Linear Differential Equation in $\mathbb{K}^n$

Theorem 1:

$$\dot{\vec{y}} = A\vec{y} \qquad \text{(H)}$$

has for each initial value $\vec{y}(0) = \vec{c}$ exactly one solution, viz

$$\vec{y}(t) = \exp(At)\vec{c}$$

Proof:

1) $\vec{\varphi}(t) = \exp(At)\vec{c}$

$$\frac{1}{h}\big(\vec{\varphi}(t+h) - \vec{\varphi}(t)\big) = \frac{1}{h}\big(\exp(At+Ah) - \exp(At)\big)\vec{c}$$
$$= \underbrace{\frac{1}{h}\big(\exp(Ah) - 1_n\big)}_{\to A \text{ if } h\to 0}\underbrace{\exp(At)\vec{c}}_{\vec{\varphi}(t)}$$

that means: $\vec{\varphi}$ solves (H) with initial value $\vec{\varphi}(0) = 1_n\cdot\vec{c} = \vec{c}$

2) Assumption: $\vec{\psi}(t)$ shall solve (H)

Auxiliary function: $\vec{f}(t) \;=\; \exp(-At)\vec{\psi}(t)$ differentiable with derivative

$$\dot{\vec{f}}(t) \;=\; \Big(\frac{\mathrm{d}}{\mathrm{d}t}\exp(-At)\Big)\vec{\psi}(t) \;+\; \exp(-At)\dot{\vec{\psi}}(t)$$

since $\vec{\psi}(t)$ solution of (H)

$$\dot{\vec{f}}(t) \;=\; \big(\underbrace{-A\exp(-At)+\exp(-At)A}_{0_n}\big)\vec{\psi}(t) \;=\; \vec{0}$$

because $\exp(-At)A \;=\; A\exp(-At)$ as can be seen from the partial sums $s_N(-At)$.

Because $\dot{\vec{f}}(t)=\vec{0}$, $\vec{f}(t)$ is constant equal to $\vec{c}_0$

$$\vec{\psi}(t) \;=\; \exp(At)\vec{f}(t) \;=\; \exp(At)\vec{c}_0$$

□

Remarks:

The solutions $\exp(At)\vec{c}_0$ of (H) form a $\mathbb{K}$-vector space (with functions on $\mathbb{R}$ as elements). Its dimension is that of $\mathbb{K}^n$, i.e. n! Each basis of the solution space is called a fundamental system of solutions. For example, the columns of $\exp(At)$ form a fundamental system.

Theorem 2:

Consider the inhomogeneous linear differential equation. Let $\vec{b}: I \mapsto \mathbb{K}^n$ be a continuous function on the interval I. Then

$$\dot{\vec{y}} \;=\; A\vec{y} + \vec{b} \qquad \text{(L)}$$

has the particular solution

$$\vec{\psi}(t) \;=\; \exp(at)\int_{t_0}^{t}\exp(-As)\vec{b}(s)\,\mathrm{d}s$$

with initial value

$$\vec{\psi}(t_0 \;=\; \vec{0}$$

All remaining solutions are of the form $\vec{\psi}+\vec{\varphi}$, where $\vec{\varphi}$ runs through all solutions of (H)!

Proof step:

The fact that $\vec{\psi}$ is the solution (of (L)) follows directly from the product rule. $\vec{\psi}(t)$ is found by varying the constants in the general solution of (H).

Ansatz: $\vec{\psi}(t) = \exp(At)\vec{c}(t)$, then

$$\dot{\vec{\psi}}(t) = A\vec{\psi}(t) + \exp(At)\dot{\vec{c}}(t) = A\vec{\psi}(t) + \vec{b}(t)$$

That turns into $\dot{\vec{c}}(t) = \exp(-At)\vec{b}(t)$, integration over t_0 , t

$$\vec{c}(t) = \int\limits_{t_0}^{t} \exp(-As)\vec{b}(s)\,\mathrm{d}s$$

7.6 The Jordan Normal Form for Matrices in $M_n(\mathbb{C})$

Let $A \in M_n(\mathbb{C})$ with eigenvalue λ_k and eigenvector $\vec{a}_k$. Then $\vec{\psi}(t) = \mathrm{e}^{\lambda_k t}\vec{a}_k$ is a solution of (H). Because

$$\dot{\vec{\psi}}_k(t) = \lambda_k \mathrm{e}^{\lambda_k t}\vec{a}_k = \mathrm{e}^{\lambda_k t}\lambda_k\vec{a}_k = \mathrm{e}^{\lambda_k t}A\vec{a}_k = A\vec{\psi}_k(t)$$

If A is diagonalizable, i.e. if A has n linearly independent eigenvectors $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, then

$$AS = S\,\mathrm{diag}\,(\lambda_1, \ldots, \lambda_n) \quad ; \quad A = S\,\mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)S^{-1}$$

Hence

$$\exp(At) = S\,\mathrm{diag}\,(\mathrm{e}^{\lambda_1 t}, \mathrm{e}^{\lambda_2 t} \ldots, \mathrm{e}^{\lambda_n t})S^{-1}$$

in particular, the system of columns of $\exp(At)S$ is a fundamental system of (H).

<u>The method in the non-diagonalizable case for A</u>

uses the fundamental theorem of algebra in $\mathbb{C}$. The result is a transformation matrix S (regular, i.e. invertible) with

$$S^{-1}AS = \begin{bmatrix} \boxed{C_1} & & & 0 \\ & \boxed{C_2} & & \\ & & \ddots & \\ 0 & & & \boxed{C_r} \end{bmatrix}$$

where along the main diagonal there are square Jordan blocks C_q of the type

$$C \;=\; C_q \;=\; \left[\begin{array}{ccccc} \lambda & 1 & & & \\ & \lambda & \ddots & 0 & \\ & & \ddots & \ddots & \\ & 0 & & \ddots & 1 \\ & & & & \lambda \end{array}\right] \Big\updownarrow p(q_n)$$

The same λ can occur in different blocks, altogether all eigenvalues of A occur.

$$C \;=\; \lambda 1_p + N_p \quad \text{with} \quad N_p \;=\; \left(\delta_{i+1,k}\right)_{1\le i,k\le p}$$

where $\lambda 1_p$ and N_p are commutable. N_p is nilpotent.

<u>Assertion:</u> $N_p^m \;=\; \left(\delta_{i+m,k}\right)_{1\le i,k\le p}$

Proof by induction

$m \mapsto m+1$

$$N_p^{m+1} \;=\; N_p^m N_p \;=\; \left(\sum_{j=1}^{p} \delta_{i+m,j}\,\delta_{j+1,k}\right) \;=\; 0\;, \quad \text{if} \;\; j \neq i+m$$

$$N_p^{m+1} \;=\; \left(\delta_{i+m+1,k}\right)_{1\le i,k\le p}$$

□

In particular $Np^m = 0$ if $m \ge p$. The functional equation can be applied because $\lambda 1_p$ commutes with any matrix.

$$\exp(\lambda t 1_p) \;=\; \mathrm{e}^{\lambda t} 1_p$$

$$\exp(N_p t) \;=\; \sum_{k=0}^{\infty} \frac{1}{k!} N_p^k t^k \;=\; \sum_{k=0}^{p-1} \frac{1}{k!} N_p^k t^k$$

$$\exp(N_p t) \;=\; \left[\begin{array}{cccccc} 1 & t & \dots & \frac{t^k}{k!} & \cdots & \frac{t^{p-1}}{(p-1)!} \\ & 1 & t & & & \vdots \\ & & \ddots & \ddots & & \frac{t^k}{k!} \\ & & & \ddots & \ddots & \vdots \\ & & & & \ddots & t \\ & & & & & 1 \end{array}\right]$$

$$\exp(\lambda t 1_p + N_p t) \;=\; \exp(\lambda t 1_p) \;+\; \exp(N_p t) \;=\; \underbrace{\mathrm{e}^{\lambda t} \exp(N_p t)}_{\exp(Ct)}$$

$$\exp[(S^{-1}AS)] \;=\; \exp(Jt) \;=\; \begin{bmatrix} \exp(C_1t) & & & 0 \\ & \exp(C_2t) & & \\ & & \ddots & \\ 0 & & & \exp(C_rt) \end{bmatrix}$$

7.7 Higher-order Scalar-valued Linear Differential Equations

$$\boxed{\frac{\mathrm{d}^n x}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1}x}{\mathrm{d}t^{n-1}} + \cdots + a_1\frac{\mathrm{d}x}{\mathrm{d}t} + a_0 x \;=\; 0 \qquad (*)}$$

with constant $a \in \mathbb{K}$.

$$y_{k+1} \;= x^{(k)} \qquad 0 \le k < n$$

From this emerges the system

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \vdots \\ \dot{y}_n \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & 0 & & \ddots & 1 \\ -a_0 & -a_1 & & & -a_{n-1} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Short $\boxed{\dot{\vec{y}} \;=\; A\vec{y}}$

with the Frobenius matrix A (also companion matrix) of the differential equation $(*)$.

The characteristic polynomial of A

$$\det(X1_n - A) \;=\; \begin{bmatrix} X & -1 & & & 0 \\ & \ddots & -1 & & \\ & & \ddots & & \\ 0 & & X & -1 & \\ & & & \ddots & \\ a_0 & \cdots & a_{n-2} & & X + a_{n-1} \end{bmatrix}$$

Development after the first column

$$= X\det\begin{bmatrix} X & -1 & \cdots & \cdots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \vdots \\ \vdots & & X & -1 \\ a_0 & \cdots & a_{n-2} & X+a_{n-1}\end{bmatrix} + (-1)^{n+1}a_0\det\begin{bmatrix}-1 & & 0\\ & \ddots & \\ * & & -1\end{bmatrix}\begin{matrix}\uparrow\\ n-1\\ \downarrow\end{matrix}$$

with induction

$$\det(X1_n - A) \;=\; X^n + \sum_{i=0}^{n-1} a_i X^i \qquad \text{if } A \text{ companion matrix of } (*)\,.$$

<u>Special feature of the companion matrix A</u>

For each eigenvalue λ of A one has $\operatorname{rk}(\lambda 1_n - A) = n-1$. Therefore, the Jordan normal form for λ has only one Jordan block. Rationale:

$$\lambda_n 1_n - A \;=\; \begin{bmatrix} \lambda & -1 & & 0 \\ & \ddots & \ddots & \\ 0 & & \lambda & -1 \\ a_0 & a_1 & a_{n-2} & \lambda+a_{n-1}\end{bmatrix}$$

Adding a λ-multiple of the $(n-k+1)$-th column to the $(n-k)$-th column $(k = 1, 2, \ldots)$ does not change the rank.

$$\operatorname{rk}(\lambda 1_n - A) \;=\; \operatorname{rk}\begin{pmatrix} & 0 & 0 \\ & | & | \\ & -1 & 0 \\ & 0 & -1 \\ & \bigcirc & \lambda+a_{n-1}\end{pmatrix}$$

$$\downarrow$$

$$\overbrace{\lambda^2 + \lambda a_{n-1} + a_{n-2}}$$

$$= \cdots = \operatorname{rk}\begin{bmatrix} 0 & -1 & & 0 \\ | & & \ddots & \\ 0 & & & -1 \\ P_A(\lambda) & * & * & * \end{bmatrix} = n-1$$

Consequence for the solution of the differential equation $(*)$ from the factorization (EV = eigenvalues)

$$P_A(\lambda) \;=\; \prod_{\lambda\,\in\,EV} (X-\lambda)^{\mathrm{e}(\lambda)}$$

Solutions of $(*)$ are the functions

$$t^k \mathrm{e}^{\lambda t} \qquad 0 \,\le\, t \,\le\, \mathrm{e}(\lambda) - \lambda \;\; ; \quad \lambda \in EV$$

if $\lambda_1, \ldots, \lambda_r$ are the different eigenvalues then the functions

$$t^k \mathrm{e}^{\lambda m t} \qquad 0 \,\le\, k \,\le\, \mathrm{e}(\lambda_m) - 1 \;\; , \quad 1 \le m \le r$$

form a basis of the solution space of $(*)$.

7.8 Stability Questions

How do solutions of

$$\dot{\vec{y}} \;=\; A\vec{y} \;\; , \quad A \in M_n(\mathbb{K}) \qquad \text{(H)}$$

behave for $t \to \infty$?

Case 1:

A has (in $\mathbb{C}$) an eigenvalue with $\operatorname{Re}\lambda > 0$. Then (H) has a solution

$$\vec{\varphi}(t) \;=\; \mathrm{e}^{\lambda t}\vec{a} \quad (A\vec{a} \;=\; \lambda\vec{a}) \quad ; \quad \left\|\mathrm{e}^{\lambda t}\vec{a}\right\|_{\mathrm{e}} \;=\; \mathrm{e}^{\operatorname{Re}\lambda t}\|\vec{a}\|_{\mathrm{e}}$$

For this

$$\lim_{t\to\infty} \|\vec{\varphi}(t)\|_{\mathrm{e}} \;=\; \infty$$

For the general case, an estimation of the solutions is obtained from

$$\|\exp(At)\|_{\mathrm{e}}$$

Based on:

$$\|A\vec{x}\|_{\mathrm{e}} \,\le\, \|A\|_{\mathrm{e}}\|\vec{x}\|_{\mathrm{e}}$$

Proof with Cauchy-Schwarz inequality (C-S-I):

$$\begin{aligned}
\Big(\sum_{i=1}^{n}\Big|\sum_{k=1}^{n} a_{ik}x_k\Big|^2\Big)^{1/2} \;&=\; \|A\vec{x}\|_{\mathrm{e}} \\
&\le\; \Big(\sum_{i=1}^{n}\Big(\sum_{k=1}^{n}|a_{ik}|^2\Big)\Big(\sum_{m=1}^{n}|x_m|^2\Big)\Big)^{1/2} \qquad \text{(C-S-I)}
\end{aligned}$$

$$\leq \left(\sum_{i,k=1}^{n} |a_{ik}|^2 \right)^{1/2} \left(\sum_{m=1}^{n} |x_m|^2 \right)^{1/2}$$

$$= \|A\|_{\mathrm{e}} \|\vec{x}\|_{\mathrm{e}}$$

□

Problem with using the Jordan normal form: Estimate of $\|\exp(N_p t)\|$ for $t \to \infty$.

$\|\exp(N_p t)\|^2$ is a polynomial of degree $2p-2$. It grows faster than $\mathrm{e}^{\gamma t}$ for every positive γ.

Result: If $\alpha \in \mathbb{R}$, $\alpha > \operatorname{Re}\lambda$ for every eigenvalue λ of A, then the following estimate applies with a bound $C > 0$:

$$\|\exp(At)\|_{\mathrm{e}} \leq C\mathrm{e}^{\alpha t}$$

<u>Case 2:</u>

If $\operatorname{Re}\lambda < 0$ for each eigenvalue $\lambda \in \mathbb{C}$ of A, then one can choose $\alpha < 0$ and it follows

$$\lim_{t \to \infty} \|\exp(At)\| = 0$$

and then for every solution $\vec{\varphi}$ of (H)

$$\lim_{t \to \infty} \|\vec{\varphi}(t)\|_{\mathrm{e}} = 0 \qquad (*)$$

<u>Case 3:</u>

If always $\operatorname{Re}\lambda \leq 0$ and a $\lambda = \lambda_0$ exists with $\operatorname{Re}\lambda_0 = 0$, i.e. $\lambda_0 = j\omega_0$, then a solution $\vec{\varphi}(t) = \mathrm{e}^{j\omega t}\vec{c}$ exists which does not satisfy $(*)$. Whether all solutions are bounded has to be examined in detail.

Definition:

A polynomial P with real or complex coefficients is called “stable” if all zeros in $\mathbb{C}$ have negative real parts.

Method of polynomial division $P_0 \div P_1$: Euclidean algorithm

$$P_{k-1} = P_k Q_k + P_{k+1} \qquad \text{if the degree } P_k \text{ exists}$$

$$\deg P_k < \deg P_{k-1} \qquad 1 \leq k \leq N \quad \text{with} \quad P_{N+1} = 0$$

7.8.1 Stability Criterion for Real Polynomials

$$P \;=\; a_k X^n \,+\, ak-1X^{N-1} \,+\, \ldots \,+\, a_0 \;; \qquad a_i \in \mathbb{R}\,, \quad a_n \neq 0$$

$$P_0 \;=\; a_n X^n \,+\, a_{n-2}X^{n-2} \,+\, \ldots \quad ; \quad \deg 0 \;=\; -\infty$$

$$P_1 \;=\; a_{n-1}X^{n-1} \,+\, a_{n-3}X^{n-3} \quad ; \quad P \text{ real: } \Rightarrow$$

P is stable if and only if firstly in the division chain for $P_0 \div P_1$ the number of steps is $N = n$ and secondly $Q_k = b_k X$ with $b_k > 0$.

Remark:

If $-\alpha_i$ are the negative zeros of P with $-\mu_k \pm j\nu_k$ as the pairs of conjugate complex zeros of P, then

$$P(x) \;=\; a_n \prod (x+\alpha_i) \prod_k \underbrace{(X+\mu_k+j\nu_k)(X+\mu_k-j\nu_k)}_{X^2+2\mu_k X+\mu_k^2+\nu_k^2}$$

In particular, a stable real polynomial has all coefficients of the same sign. That is not enough:

$$(X+1)(X^2+1) \;=\; X^3 \,+\, X^2 \,+\, X \,+\, 1$$

Examples:

(1) $P \;=\; X^3 \,+\, X^2 \,+\, X \,+\, 1$

$$P_0 \;=\; X^3 \,+\, X \quad ; \quad P_1 \;=\; X^2 \,+\, 1$$

$$P_0 \;=\; \underbrace{(X+1)}_{Q_1}\underbrace{(X^2+1)}_{P_1} \,+\, 0 \qquad \text{not stable}$$

because Q_1 has the wrong shape, also N is too small.

(2) $P \;=\; X^4 \,+\, 5X^3 \,+\, 10X^2 \,+\, 10X \,+\, C$

$$P_0 \;=\; X^4 \,+\, 10X^2 \,+\, C \quad ; \quad P_1 \;=\; 5X^3 \,+\, 10X$$

$$P_0 \;=\; \underbrace{\frac{X}{5}}_{Q_1}\,\underbrace{(5X^3+10X)}_{P_1} \,+\, \underbrace{8X^2+C}_{P_2}$$

$$P_1 \;=\; \underbrace{\frac{5X}{8}}_{Q_2}\underbrace{(8X^2+C)}_{P_2} \,+\, \Big(\underbrace{10-\frac{5C}{8}}_{a}\Big)X \quad ; \quad P_3 \;=\; aX$$

$$P_2 \;=\; 8X^2 + C \;=\; \underbrace{\frac{8X}{a}}_{Q_3}\underbrace{(\,aX\,)}_{P_3} + \underbrace{C}_{P_4} \quad ; \quad \underline{a > 0}$$

$$P_3 \;=\; aX \;=\; \underbrace{\frac{a}{C}X}_{Q_4}\underbrace{C}_{P_4} + \underbrace{0}_{P_5} \quad ; \quad \underline{C > 0}$$

$$a \;=\; 10 - \frac{5C}{8} \;=\; \frac{5(16 - C)}{8}$$

P is stable if and only if $0 < C < 16$.

7.9 Applications of the Laplace transform

$$(\mathscr{L}\varphi)(s) \;=\; \int_0^\infty \mathrm{e}^{-st}\varphi(t)\,\mathrm{d}t$$

Doetsch symbol

$$\varphi(t) \quad \circ\!\!-\!\!\bullet \quad \phi(s) \;=\; \int_0^\infty \mathrm{e}^{-st}\varphi(t)\,\mathrm{d}t \qquad \text{(real part of } s \text{ sufficiently large)}$$

Appropriate convergence conditions assumed:[1]

$$\boxed{(\mathscr{L}\varphi^{(m+1)})(s) \;=\; -\varphi^{(m)}(0) + s(\mathscr{L}\varphi^{m})(s)}$$

$$\varphi^{(m+1)}(t) \;\circ\!\!-\!\!\bullet\; -\varphi^{(m)}(0) + s(\mathscr{L}\varphi^{(m)})(s)$$

Proof by partial integration

$$\int_0^\infty \underbrace{\mathrm{e}^{-st}}_{u}\underbrace{\varphi^{(m+1)}(t)}_{v'}\,\mathrm{d}t \;=\; \mathrm{e}^{-st}\varphi^{(m)}\Big|_0^\infty + s\int_0^\infty \mathrm{e}^{-st}\varphi^{(m)}(t)\,\mathrm{d}t$$

Elementary by complete induction on m

$$(\mathscr{L}\varphi^{(m)})(s) \;=\; -\sum_{k=0}^{m-1}\varphi^{(m-k-1)}(0)s^k + s^m(\mathscr{L}\varphi)(s)$$

□

For example, consider the inhomogeneous scalar differential equation

[1] See chapter 1.

$$\sum_{m=0}^{n} a_m x^{(m)}(t) \;=\; g(t) \qquad \text{(L)}$$

Under suitable convergence conditions g has a $\mathscr{L}$-transform $(\mathscr{L}g)(s)$.

If φ is a solution of (L), then the linearity of the $\mathscr{L}$ transform yields

$$\sum_{m=0}^{n} a_m (\mathscr{L}\varphi^{(m)})(s) \;=\; (\mathscr{L}g)(s)$$

Now let $a_n = 1$. Then the characteristic polynomial of the associated Frobenius matrix is

$$P(X) \;=\; X^n + \sum_{m=0}^{n-1} a_m X^m$$

$$\Big(S^n + \sum_{m=0}^{n-1} a_m S^m\Big)(\mathscr{L}\varphi)(s) - \sum_{m=1}^{n} a_m \sum_{k=0}^{m-1} \varphi^{(m-1-k)}(0) s^k \;=\; (\mathscr{L}g)(s)$$

$$(\mathscr{L}\varphi)(s) \;=\; \frac{1}{P(S)}\Big(\sum_{k=0}^{n-1}\Big(\sum_{m=k+1}^{n} a_m \varphi^{(m-1-k)}(0)\Big)s^k + (\mathscr{L}g)(s)\Big)$$

Solving after the original image φ under the Laplace transform becomes easy if the $\mathscr{L}$-transform of g is a rational function in s. Then on the right there is a rational function of $s \to$ partial fraction decomposition.

Examples:

(1) $\dddot{y} + 2\dot{y} \;=\; \cos t$

Let $\varphi(t)$ be the solution of the equation with initial conditions:

$$\varphi(0) \;=\; \dot{\varphi}(0) \;=\; \ddot{\varphi}(0) \;=\; 0$$

$$\cos t \;=\; \frac{1}{2}e^{jt} + \frac{1}{2}e^{-jt} \quad \circ\!\!-\!\!\bullet \quad \frac{s}{s^2+1} \;=\; (\mathscr{L}g)(s)$$

$$\sin t \;=\; \frac{1}{2j}e^{jt} - \frac{1}{2j}e^{-jt} \quad \circ\!\!-\!\!\bullet \quad \frac{j/2}{s-j} - \frac{j/2}{s+j} \;=\; \frac{1}{s^2+1}$$

$$e^{\lambda t} \quad \circ\!\!-\!\!\bullet \quad \int_0^\infty e^{-(s-\lambda)t}\,dt \;=\; \frac{1}{s-\lambda}$$

Characteristic polynomial of the differential equation: $\quad P(X) \;=\; X^3 + 2X$

$$\varphi(t) \quad \circ\!\!-\!\!\bullet \quad \phi(s) \;=\; (\mathscr{L}\varphi)(s)$$

With the initial conditions $\varphi(0) = \dot{\varphi}(0) = \ddot{\varphi}(0) = 0$ follows

$$P(s)\cdot(\mathscr{L}\varphi)(s) \;=\; (\mathscr{L}g)(s)$$

Insert:

$$
\begin{aligned}
(s^3+2s)\cdot\phi(s) &= \frac{s}{s^2+1}\\
\phi(s) &= \frac{1}{(s^2+1)(s^2+2)} = \frac{1}{s^2+1} - \frac{1}{s^2+2}\\
&= \frac{j/2}{s-j} - \frac{j/2}{s+j} - \frac{j/2\sqrt{2}}{s-\sqrt{2}j} + \frac{j/2\sqrt{2}}{s+\sqrt{2}j}\\
\Rightarrow\ \varphi(t) &= \sin t - \frac{\sin\sqrt{2}t}{\sqrt{2}}
\end{aligned}
$$

Still open: The inverse of the Laplace transform in general.

(2) $\dot{\vec{y}} = A\vec{y} + \vec{b}$ (L)

where the inhomogeneity $\vec{b}(t)$ has the Laplace transform $\vec{B}(s)$.

Initial condition: $\vec{\varphi}(0) = \vec{c}$

According to theory, the solution is

$$\vec{\varphi}(t) = \exp(At)\vec{c} + \int_0^t \exp\big(A(t-s)\big)\vec{b}(s)\,\mathrm{d}s$$

$\mathscr{L}$-transform applied to (L)

$$s\vec{\phi}(s) - \vec{c} = A\vec{\phi}(s) + \vec{B}(s) \qquad \text{where} \qquad \vec{\varphi}(t) \circ\!\!-\!\!\bullet\ \vec{\phi}(s)$$

$$(s1_n - A)\vec{\phi}(s) = \vec{c} + \vec{B}(s)$$

$$\vec{\phi}(s) = (s1_n - A)^{-1}(\vec{c} + \vec{B}(s))$$

Instead of integration in the first explicit solution for $\vec{\varphi}(t)$, one encounters here a three-part task.

1) $\vec{b}(t) \circ\!\!-\!\!\bullet\ \vec{B}(s)$

2) Inversion of $(s1_n - A)$

3) Inverse transformation of $\vec{\phi}(s)$

(3) The scalar-valued homogeneous linear differential equation of n-th order with constant coefficients

$$x^{(n)} + \sum_{i=0}^{n-1} a_i x^{(i)} = 0 \qquad \text{(H)}$$

Characteristic polynomial of the companion matrix (Frobenius)

$$A = \begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & \mathbf{0} & \\ & & \ddots & \ddots & \\ & \mathbf{0} & & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{bmatrix}$$

$$P_A(X) = X^n + \sum_{i=0}^{n-1} a_i X^i$$

Let

$$P_A(X) = \prod_{q=1}^{r} (X - \lambda_q)^{\mathrm{e}_q}$$

be the decomposition of P_A into linear factors (over $\mathbb{C}$).

Then the solutions of (H) are:

$$t^k \mathrm{e}^{\lambda_q t} = \varphi_{k,q}(t) \qquad (0 \le k \le \mathrm{e}_{q-1} \ , \quad 1 \le q \le r)$$

Proof of linear independence with Laplace transform:

1) $t^k \mathrm{e}^{\lambda t} \quad \circ\!\!-\!\!\bullet \quad \dfrac{k!}{(s-\lambda)^{k+1}}$

Proof by induction on k

$k = 0$: (see earlier)

$k \to k+1$:

$$\int_0^\infty \mathrm{e}^{-st} t^{k+1} \mathrm{e}^{\lambda t} \mathrm{d}t = \int_0^\infty \underbrace{t^{k+1}}_{u} \underbrace{\mathrm{e}^{-(s-\lambda)t}}_{v'} \mathrm{d}t \quad \text{(partial integration)}$$

$$= \left. \frac{-t^{k+1}\mathrm{e}^{-(s-\lambda)t}}{s-\lambda} \right|_0^\infty + \frac{k+1}{s-\lambda} \underbrace{\int_0^\infty t^k \mathrm{e}^{-(s-\lambda)t} \, \mathrm{d}t}_{\mathscr{L}(t \to t^k \mathrm{e}^{\lambda t})}$$

(Convergence condition for the integral: $\operatorname{Re} s > \operatorname{Re} \lambda$)

$$= \frac{k+1}{s-\lambda} \frac{k!}{(s-\lambda)^{k+1}} = \frac{(k+1)!}{(s-\lambda)^{k+2}}$$

2) Let $\displaystyle \sum_{q=1}^{r} \sum_{k=0}^{\mathrm{e}_q - 1} \alpha_{k,q} \varphi_{k,q}(t) = 0$

$\mathscr{L}$-transform

$$\begin{aligned} 0 &= \mathscr{L}(0) = \sum_{q,k} \alpha_{k,q}\big(\mathscr{L}\varphi_{k,q}\big)(s) \\ &= \sum_{q,k} \frac{\alpha_{k,q} k!}{(s-\lambda_q)^{k+1}} \quad , \quad \text{if } \operatorname{Re} s > \operatorname{Re}\lambda \ \forall\, q \end{aligned}$$

Therefore even for all $s \in \mathbb{C}$

$$\sum_{q,k} \frac{\alpha_{k,q} k!}{(s-\lambda_q)^{k+1}} = 0 \qquad \Rightarrow \qquad \alpha_{k,q} = 0\,,\ \forall\, k \wedge \forall\, q$$

$\square$

8. Existence and Uniqueness Theorem for Explicit Ordinary Differential Equations

Let D be a region in $\mathbb{R}^2$ and let $f:\, D \mapsto \mathbb{R}$ be continuous.

Explicit differential equation of 1-st order:

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

Main objective: Find curves $\varphi(x)$, differentiable with

$$\dot{\varphi}(x) \,=\, f(x,\varphi(x))$$

Direct interpretation: field of line elements on D $=$ direction field.

Examples:

(1) $y' \,=\, f(x,y) \,=\, \dfrac{y}{x} \quad ; \quad D = \,]0,\infty[\, \times \mathbb{R}$

Level lines of f are called isoclines, here they are rays through the origin.

(2) $y' \,=\, f(x,y) \,=\, -\dfrac{x}{y} \quad ; \quad D = \mathbb{R} \times \,]0,\infty[$

The isoclines are again rays through the origin.

(3) $y' \,=\, f(x,y) \,=\, x^2 + y^2$

The isoclines are circles around the origin (zero point).

Consideration of the directional field (often) gives an approximate solution (idea about the shape of the solution curves). Extension to the construction of approximate solutions (approximate solutions).

Tangent method (Euler)

Start $(x_0, u_0) \in D$, increment $h \neq 0$,

$$x_{i+1} = x_i + h \quad , \quad u_{i+1} = u_i + hf(x_i, u_i)$$

Example:

(4) $y' = y$, $x_0 = 0$, $u_0 = 1$

Increment: $h = \frac{x}{n}$, (n = number of steps)

$$u_1 = 1 + \frac{x}{n} \cdot 1 = 1 + \frac{x}{n}$$

$$u_2 = 1 + \frac{x}{n} + \frac{x}{n}\left(1 + \frac{x}{n}\right) = \left(1 + \frac{x}{n}\right)^2$$

$$u_k = \left(1 + \frac{x}{n}\right)^k$$

$$u_{k+1} = \left(1 + \frac{x}{n}\right)^k + \frac{x}{n}\left(1 + \frac{x}{n}\right)^k = \left(1 + \frac{x}{n}\right)^{k+1}$$

$$u_n = \left(1 + \frac{x}{n}\right)^n$$

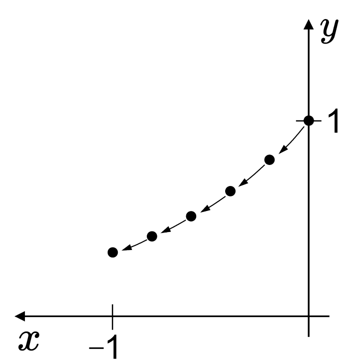
Remark:

(1) Original definition of the solution of $\exp(x)$:

$$\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$$

Sketched on the right: the steps for

$$n = 5 \; ; \; x = -1$$



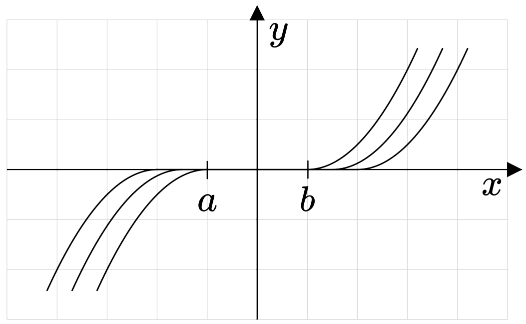
8.1 The Lipschitz Condition

Even if f is continuous, it is possible that under certain circumstances the initial value task, the so-called initial value problem (IVP), can have several solutions.

Example:

(5) $y' = 2|y|^{1/2}$

Level lines are parallels to the x-axis.



Solutions going through $(0,0)$

1) $\varphi(x) = 0$
2) $\varphi(x) = (x-a)^2 \quad \text{for} \quad x \geq a$
3) $\varphi(x) = -(x-b)^2 \quad \text{for} \quad x \leq b < 0$

Reason for the occurrence of several solutions! At $y = 0$ the directional field changes too much.

8.1.1 Definition of the Lipschitz Condition

Let f be real and continuous on the region D. K be a subset of D. f suffices on K a Lipschitz condition with L-constant $L > 0$, if for all (x, y_1), $(x, y_2) \in K$

$$\boxed{|f(x,y_1) - f(x,y_2)| \leq L|y_1 - y_2|}$$

Remark:

(2) For example, if f is continuous and partially differentiable with respect to y, then every point $(x_0, y_0) \in D$ has a neighborhood K, on which, with matching $L > 0$, the Lipschitz condition holds.

Lemma (Gronwall's lemma)

Let non-negative continuous functions φ, ψ be given on $[0, a]$ and a bound $b \geq 0$ such that

$$\varphi(t) \leq b + \int_0^t \varphi(s)\psi(s)\,\mathrm{d}s \,, \quad t \in [0,a] \qquad (*)$$

Then there applies

$$\boxed{\varphi(t) \leq b \cdot \exp\Big(\int_0^t \psi(s)\,\mathrm{d}s\Big)}$$

Proof:

1) If $b = 0$ then the estimate for $b > 0$ is even more valid. Then, via the transition $b \searrow 0$, the statement for $b > 0$ will also yield the assertion here.

2) $b > 0$: The premise says

$$\frac{\varphi(t)\psi(t)}{b + \int\limits_0^t \varphi(s)\psi(s)\,\mathrm{d}s} \;\leq\; \psi(t)$$

(The numerator contains the derivative of the denominator function)

Integration from 0 to t

$$\ln\Big(b + \int\limits_0^t \varphi(s)\psi(s)\,\mathrm{d}s\Big) - \ln b \;\leq\; \int\limits_0^t \psi(s)\,\mathrm{d}s$$

$$\Rightarrow\; b + \int\limits_0^t \varphi(s)\psi(s)\,\mathrm{d}s \;\leq\; b \cdot \exp\Big(\int\limits_0^t \psi(s)\,\mathrm{d}s\Big)$$

with the premise $(*)$

$$\varphi(t) \;\leq\; b \cdot \exp\Big(\int\limits_0^t \psi(s)\,\mathrm{d}s\Big)$$

$\square$

8.1.2 Dependence of Solutions on Initial Values

Let $f : D \mapsto \mathbb{R}$ be continuous, furthermore let a Lipschitz condition with L constant $L > 0$ hold on the part $K \subset D$. If u, v are now two solutions of the differential equation $y' = f(x, y)$ defined on the interval $J \ni x_0$ and it always holds that $(x, u(x)) \in K$, $(x, v(x)) \in K$ then

$$\boxed{\big|u(x) - v(x)\big| \;=\; \big|u(x_0) - v(x_0)\big|\mathrm{e}^{L|x-x_0|}}$$

Proof with Gronwall's lemma:

$$\big|u(x) - v(x)\big| \;=\; \Big|u(x_0 - v(x_0) + \int\limits_{x_0}^{x} \big(f(t, u(t)) - f(t, v(t))\big)\,\mathrm{d}t\Big|$$

$$\text{if} \quad x \geq x_0$$

$$\leq \ \Big|\underbrace{u(x_0) - v(x_0)}_{b}\Big| + \int\limits_{x_0}^{x} \underbrace{\big|f(t, u(t)) - f(t, v(t))\big|}_{\leq \ \underset{\substack{\downarrow \\ \psi}}{L}|u(t) - v(t)|} \,\mathrm{d}t$$

Lemma with $t = x - x_0 \quad \Rightarrow$

$$|u(x) - v(x)| \ = \ |u(x_0 - v(x_0)|\mathrm{e}^{L|x - x_0|}$$

for $x < x_0$ set $t = x_0 - x$.

$\square$

> **Uniqueness Theorem**
>
> Let f be continuous on D, furthermore let f satisfy the Lipschitz condition locally on D. If $\varphi : I \mapsto \mathbb{R}$, $\psi : J \mapsto \mathbb{R}$ are two solutions of $y' = f(x, y)$ and $x_0 \in I \cap J$ defined on intervals, then φ and ψ are equal on $I \cap J$ if $\varphi(x_0) = \psi(x_0)$.

8.2 Picard-Lindelöf Existence Theorem

Let f be a continuous function in the region $D \subset \mathbb{R}^2$, which also locally satisfies the Lipschitz condition. Then there exists for every initial value $(x_0, y_0) \in D$ an $a > 0$ and on $I = [x_0 - a, x_0 + a]$ a solution φ of the differential equation $y' = f(x, y)$ with $\varphi(x_0) = y_0$.

Remark:

(1) If, moreover, f is continuously partially differentiable with respect to y, then local Lipschitz conditions apply. According to the uniqueness theorem, the solution φ with $\varphi(x_0) = y_0$ is complete.

Proof:

According to the premise there is a square

$$Q \ = \ \{(x, y) \,;\ |x - x_0| \leq r \,,\ \ |y - y_0| \leq r\}$$

in D, on which f has a Lipschitz constant L. Because f is continuous, a bound $M \geq 0$ exists with

$$|f(x, y)| \ \leq \ M \qquad \forall\, (x, y) \in Q$$

$$a \ \coloneqq \ \min\!\big(r \,,\ {}^{r}\!/_{M}\big)$$

Picard iteration

$$\begin{aligned}\varphi_0(x) &= 0\\ \varphi_{N+1}(x) &= y_0 + \int\limits_{x_0}^{x} f(t,\varphi_N(t))\,\mathrm{d}t\end{aligned}$$

Proof program: show that on $I = [x_0 - a, x_0 + a]$ the function sequence $\varphi_N(x)$ is uniformly convergent.

1) For all N and all $x \in I$ we have $(x, \varphi_N(x)) \in Q$ i.e. $|\varphi_N(x) - y_0| \leq r$.

Induction on N. $N = 0$ is trivial. $N \to N+1$:

$$|\varphi_{N+1}(x) - y_0| = \left| \int\limits_{x_0}^{x} f(t,\varphi_N(t))\,\mathrm{d}t \right|$$

Standard integral estimate: $\leq |x - x_0|M$

By choice of a: $\leq r$

2) $|\varphi_{N+1}(x) - \varphi_N(x)| \leq \dfrac{ML^N|x-x_0|^{N+1}}{(N+1)!}$

Induction on N.

$N = 0$:

$$|\varphi_1(x) - \varphi_0(x)| = \left| \int\limits_{x_0}^{x} f(t,y_0)\,\mathrm{d}t \right|$$

$N - 1 \to N$:

$$\begin{aligned}|\varphi_{N+1}(x) - \varphi_N(x)| &= \left| \int\limits_{x}^{x_0} \Big[f(t,\varphi_N(t)) - f(t,\varphi_{N-1}(t))\Big]\,\mathrm{d}t \right|\\ &= \left| \int\limits_{x_0}^{x} \Big|f(t,\varphi_N(t)) - f(t,\varphi_{N-1}(t))\Big|\,\mathrm{d}t \right|\end{aligned}$$

Lipschitz condition

$$\leq \left| \int\limits_{x_0}^{x} L\Big|\varphi_N(t) - \varphi_{N-1}(t)\Big|\mathrm{d}t \right|$$

Induction premise for $N - 1$ instead of N

$$\leq \left| \int_{x_0}^{x} L^N M \frac{|t-x_0|^N}{N!} \mathrm{d}t \right| \leq \frac{ML^N|x-x_0|^{N+1}}{(N+1)!}$$

3) Uniform convergence

$$\begin{aligned} |\varphi_{N+p}(x) - \varphi_N(x)| &= \left| \sum_{n=N}^{N+p-1} \big(\varphi_{N+1}(x) - \varphi_N(x)\big) \right| \\ &\leq \sum_{n=N}^{N+p-1} |\varphi_{N+1}(x) - \varphi_N(x)| \overset{\substack{\mathbf{2)} \\ \downarrow}}{\leq} \sum_{n=N}^{N+p-1} \frac{ML^n|x-x_0|^{n+1}}{(n+1)!} \\ &\leq \sum_{n=N}^{\infty} \frac{ML^n a^{n+1}}{(n+1)!} \end{aligned}$$

Estimate this section of an e^x series with Lagrange

$$\leq \frac{ML^N a^{N+1}}{(N+1)!} \,\mathrm{e}^{aL} \;\to\; 0$$

for $N \to \infty$ independent of x

4) Because of uniform convergence the limit function

$$\varphi(x) = \lim_{N\to\infty} \varphi_N(x)$$

is continuous on I.

Moreover, $f(t, \varphi_N(t))$ converges to $f(t, \varphi(t))$ uniformly on I. Hence the recursion formula gives

$$\varphi(x) = \lim_{N\to\infty} \varphi_N(x) = y_0 + \int_{x_0}^{x} f(t, \varphi(t))\,\mathrm{d}t$$

hence $\varphi(x)$ is a differentiable function of the upper limit of integration x with derivative $\varphi'(x) = f(x, \varphi(x))$. Finally $\varphi(x_0) = y_0$.

$\square$

Remark:

(2) If f is continuous but does not satisfy local Lipschitz conditions, then solutions to the initial value problem need not be unique. Nevertheless, according to Peano, there is at least one solution to the initial value problem.

8.2.1 The Runge-Kutta Method

provides approximate solutions $u(x)$ of the initial value problem $y' = f(x, y)$ in the form of the vertices (x_i, u_i) of a polygonal chain. The increment $h \neq 0$, constant, $x_{i+1} = x_i + h$. The forward slope of the polygonal chain (to the right for $h > 0$, to the left for $h < 0$) is chosen as the weighted average of four function values of f near the reference point (x_i, u_i)

$$\begin{aligned} u_{i+1} &= u_i + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\ k_1 &= hf(x_i, u_i) \\ k_2 &= hf(x_i + \frac{h}{2}, u_i + \frac{k_1}{2}) \\ k_3 &= hf(x_i + \frac{h}{2}, u_i + \frac{k_2}{2}) \\ k_4 &= hf(x_i + h, u_i + k_3) \end{aligned}$$

Remark:

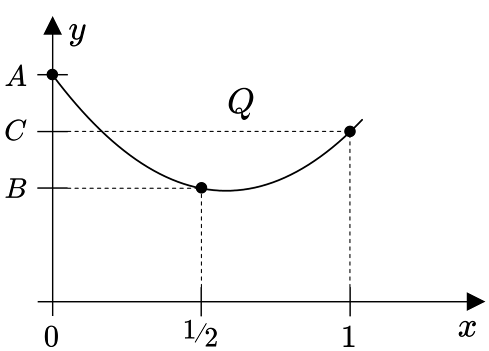
(3) For the special case that the right-hand side of the differential equation only depends on x, i.e. $y' = f(x)$, an approximate solution also means an approximate quadrature of $f(x)$. The resulting formula becomes

$$\int\limits_{x_0}^{x} f(t)\,\mathrm{d}t = \frac{x - x_0}{6n}\Big(f(x_0) + \sum_{i=0}^{n-1}\Big(4f(x_1 + \frac{h}{2}) + 2f(x_{i+1})\Big) - f(x_n)\Big)$$

8.2.2 Simpson's Rule

The single step is known as Kepler's barrel rule

$$\int\limits_{a}^{b} f(t)\,\mathrm{d}t = \frac{b-a}{6}\Big(f(a) + 4f\Big(\frac{a+b}{2}\Big) + f(b)\Big)$$



Geometric interpretation: Find a polynomial $Q(x)$ with $\deg Q \leq 2$.

$$\left.\begin{aligned} Q(0) &= A \\ Q(1/2) &= B \\ Q(1) &= C \end{aligned}\right|$$

(Drawn in the sketch on the left for $A = 0.8$, $B = 0.4$, $C = 0.4$ with the formula for $Q(x)$ given below.)

$$Q(x) \;=\; A \,+\, 2x(B-A) + x(2x-1)(C+A-2B)$$

For this

$$\int_0^1 Q(x)\,\mathrm{d}x \;=\; B \,+\, (C+A-2B)\int_0^1 (2x^2-x)\,\mathrm{d}x$$

that means:

$$\int_0^1 Q(x)\,\mathrm{d}x \;=\; \frac{1}{6}(A+4B+C)$$

The polynomial $P(t)$ of degree ≤ 2 with

$$P(a) \;=\; A \quad ; \quad P\Big(\frac{a+b}{2}\Big) \;=\; B \quad ; \quad P(b) \;=\; C$$

is

$$P(t) \;=\; Q\Big(\frac{t-a}{b-a}\Big)$$

$$\int_a^b P(t)\,\mathrm{d}t \;=\; (b-a)\int_0^1 Q(x)\,\mathrm{d}x \;=\; \frac{b-a}{6}\,(A+4B+C)$$

Make specific with the help of the function to be integrated, i.e. f

$$A \;=\; f(a) \quad ; \quad B \;=\; f\Big(\frac{a+b}{2}\Big) \quad ; \quad C \;=\; f(b)$$

8.2.3 The Power Series Ansatz

for determining a solution to the initial value problem of the differential equation $y' = f(x,y)$ through

$$\varphi(x) \;=\; \sum_{n=0}^{\infty} a_n(x-x_0)^n \quad ; \quad a_0 = \varphi(x_0) = y_0$$

can also be used in 2-nd (n-th) order differential equations.

Example: $\quad y'' \,+\, \dfrac{1}{x}y' \,+\, y \;=\; 0 \qquad$ (Bessel differential equation)

$x_0 = 0\;,\quad y_0 = 1$

$$\varphi(x) \;=\; \sum_{n=0}^{\infty} a_n x^n \quad , \quad \varphi'(x) \;=\; \sum_{n=1}^{\infty} n a_n x^{n-1}$$

$$x\varphi''(x) \;=\; \sum_{n=2}^{\infty} n(n-1)a_n x^{n-1} \quad , \quad x\varphi(x) \;=\; \sum_{n=2}^{\infty} a_{n-2} x^{n-1}$$

With that

$$0 = x\varphi'' + \varphi' + x\varphi = a_1 + \sum_{n=2}^{\infty}(n^2 a_n + a_{n-2})x^{n-1}$$

$$a_1 = 0 \quad , \quad n^2 a_n + a_{n-2} = 0$$

Coefficient comparison

$$a_{2k-1} = 0 \quad , \quad a_{2k} = \frac{-a_{2k-2}}{(2k)^2} \quad , \quad a_{2k} = \frac{(-1)^k}{k!2^{2k}} \qquad \text{(complete induction)}$$

$$\varphi(x) = J_0(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n!)^2} \left(\frac{x}{2}\right)^{2n}$$

8.3 Ordinary Differential Equations of n-th Order

$$y^{(n)} = F(x, y, y', \ldots, y^{(n-1)})$$

reduced to a system of n ordinary differential equations of 1-st order.

$$y_k := y^{(k-1)} \qquad (1 \le k \le n)$$

$$\begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_{n-1}' \\ y_n' \end{pmatrix} = \begin{pmatrix} y_2 \\ y_3 \\ \vdots \\ y_n \\ F(x, y_1, y_2, \ldots, y_n) \end{pmatrix}$$

Short form: $\vec{y}\,' = \vec{f}(x, \vec{y})$

Let a region D be given in $\mathbb{R} \times \mathbb{R}^n$ and a continuous function $\vec{f} : D \mapsto \mathbb{R}^n$ for dealing with $\vec{y}\,' = \vec{f}(x, \vec{y})$.

The Lipschitz condition for $\vec{f}$ is satisfied on the subset $K \subset D$ with L-constant $L > 0$ if

$$\boxed{\left\|\vec{f}(x, \vec{y}_1) - \vec{f}(x, \vec{y}_2)\right\| \le L\|\vec{y}_1 - \vec{y}_2\|}$$

for all $(x, \vec{y}_1)$, $(x, \vec{y}_2) \in K$. Herein L depends on the previously chosen norm in $\mathbb{R}^n$.

Remarkable: If local Lipschitz conditions do apply for $\vec{f}$ then one can transfer from the one-dimensional case:

1) The theorem about the dependence on initial conditions
2) The uniqueness theorem
3) The existence theorem

The existence theorem

On the region $D \subset \mathbb{R} \times \mathbb{R}^n$ let $\vec{f}: D \mapsto \mathbb{R}^n$ be continuous and satisfy local Lipschitz conditions! Then for every starting point $(x_0, \vec{y}_0) \in D$ exist an $a > 0$ and a solution $\vec{\varphi}: [x_0 - a, x_0 + a]$ of the differential equation $\vec{y}\,' = \vec{f}(x, \vec{y})$ with $\vec{\varphi}(x_0) = y_0$.

8.3.1 Systems of 1-st Order Linear Differential Equations with Non-constant Coefficients

Given on the interval $I \subset \mathbb{R}$ a continuous function $A : I \mapsto M_n(\mathbb{R})$, $vecb : I \mapsto \mathbb{R}^n$

$$\boxed{\vec{y}\,' = A(x)\vec{y} + \vec{b}(x) \qquad \text{(L)}}$$

In the previous notation, $\vec{f}(x, \vec{y}) = A(x)\vec{y} + \vec{b}(x)$.

f satisfies local Lipschitz conditions. Because $A(x)$ has continuous coefficients on I, it holds for every compact subinterval $J \subset I$

$$\|A(x)\|_{\mathrm{e}} \leq L \qquad \forall\, x \in J$$

therefore

$$\begin{aligned} \left\|\vec{f}(x, \vec{y}_1) - \vec{f}(x, \vec{y}_2)\right\|_{\mathrm{e}} &= \|A(x)\vec{y}_1 - A(x)\vec{y}_2\|_{\mathrm{e}} \\ &= \|A(x)(\vec{y}_1 - \vec{y}_2)\|_{\mathrm{e}} \leq \|A(x)\|_{\mathrm{e}}\|\vec{y}_1 - \vec{y}_2\|_{\mathrm{e}} \\ &\leq L\|\vec{y}_1 - \vec{y}_2\|_{\mathrm{e}} \qquad \forall\, x \in J \end{aligned}$$

Favorable: First the matrix differential equation

$$Y' = A(x)Y \qquad \text{(HM)}; \qquad Y : I \mapsto M_n(\mathbb{R})$$

Result of existence and uniqueness theorem:

For each $x_0 \in I$ exists a (through the initial conditions $U(x_0, x_0) = 1_n$) uniquely determined and on all of I declared matrix function $U(x, x_0)$ which solves (HM), i.e. $U'(x) = A(x)U(x)$.

Remark:

(1) $U(x, x_0)$with constant coefficients of A corresponds to the function $\exp(A(x - x_0))$.

Moreover: $V(x) = U(x, x_0)U(x_0, x_1)$ is satisfied

$$V'(x) = U'(x, x_0)U(x_0, x_1) = A(x)V(x)$$

Therefore, according to the uniqueness theorem, V is the solution of (HM) with initial condition

$$V(x_0) = U(x_0, x_1)$$

$U(x, x_1)$ also has this property

$$U(x, x_0)U(x_0, x_1) = U(x, x_1)$$

In particular (setting $x = x_1$)

$$U(x_1, x_0)U(x_0, x_1) = 1_n$$

therefore $U(x, x_0)$ is invertible at every position x_1.

With that, first consider the homogeneous differential equation:

$$\vec{y}\,' = A(x)\vec{y} \qquad \text{(H)}$$

For the totality of the solutions one gets:

$$\vec{\varphi}(x) = U(x, x_0)\vec{c}\ , \quad \vec{c} \in \mathbb{R}^n; \qquad \vec{c} = \vec{\varphi}(x_0)$$

To solve the inhomogeneous differential equation (L): variation of the constants

$$\begin{aligned}
\vec{\psi}(x) &= U(x, x_0)\vec{c}(x) \\
\vec{\psi}\,'(x) &= U'(x, x_0)\vec{c}(x) + U(x, x_0)\vec{c}\,'(x) \\
&\overset{\overset{\text{(HM)}}{\downarrow}}{=} A(x)U(x, x_0)\vec{c}(x) + U(x, x_0)\vec{c}\,'(x)
\end{aligned}$$

since $\vec{\psi}$ is solution of (L)

$$\vec{\psi} = A(x)\vec{\psi}(x) + \vec{b}(x) \qquad \Rightarrow$$

$$U(x, x_0)\vec{c}\,'(x) = \vec{b}(x)$$

$$\vec{c}\,'(x) \;=\; U^{-1}(x,x_0)\vec{b}(x)$$

Differential equation can be solved (for $\vec{c}$) by integration.

$$\vec{c}(x) \;=\; \int\limits_{x_0}^{x} U^{-1}(t,x_0)\vec{b}(t)\,\mathrm{d}t$$

With that

$$\vec{\psi}(x) \;=\; U(x,x_0)\int\limits_{x_0}^{x} U^{-1}(t,x_0)\vec{b}(t)\,\mathrm{d}t$$

becomes a particular solution of (L).

The totality of all solutions of (L) is of the form

$$\vec{\psi}(x) + \vec{\varphi}(x)$$

where $\vec{\varphi}(x)$ runs through the entirety of the solutions of (HM).

Remark:

(2) In concrete tasks, the ideal function $U(x,x_0)$ for the theory is only found at the end.

8.3.2 The Case of Minimum Dimension $n = 1$

$$y' \;=\; a(x)y + b(x) \qquad \text{(L)} \quad ; \qquad y' = a(x)y \qquad \text{(H)} = \text{(HM)}$$

where a, b are continuous on the interval I.

$$U(x,x_0) \;=\; \exp\Big(\int\limits_{x_0}^{x} a(t)\,\mathrm{d}t\Big)$$

$$\psi(x) \;=\; c(x)\exp\Big(\int\limits_{x_0}^{x} a(t)\,\mathrm{d}t\Big)$$

$$\begin{aligned}\psi'(x) \;&=\; c'(x)\exp\Big(\int\limits_{x_0}^{x} a(t)\,\mathrm{d}t\Big) + a(x)\psi(x)\\ &=\; a(x)\psi(x) + b(x)\end{aligned}$$

$$c'(x) \;= \exp\Big(-\int\limits_{x_0}^{x} a(t)\,\mathrm{d}t\Big)b(x)$$

Example: $y' = 2xy + x^3$ (L) ; $y' = 2xy$ (H)

Solutions (guessed): $\varphi(x) = c\mathrm{e}^{x^2}$

Variation of the constant

$$\begin{aligned}
\psi(x) &= c(x)\mathrm{e}^{x^2} \\
\psi'(x) &= c'(x)\mathrm{e}^{x^2} + 2x\psi(x) = x^3 + 2x\psi(x) \\
c'(x) &= \mathrm{e}^{-x^2}x^3 \qquad ; \qquad x_0 = 0 \\
c(x) &= \int_0^x t^3\mathrm{e}^{-t^2}\mathrm{d}t \qquad (\text{Substitution: } s = t^2;\ \mathrm{d}s = 2t\,\mathrm{d}t) \\
&= \frac{1}{2}\int_0^{x^2} s\mathrm{e}^{-s}\mathrm{d}s = -\frac{s}{2}\mathrm{e}^{-s}\Big|_0^{x^2} + \frac{1}{2}\int_0^{x^2} \mathrm{e}^{-s}\,\mathrm{d}s \\
&= -\frac{1}{2}x^2\mathrm{e}^{-x^2} - \frac{1}{2}\mathrm{e}^{-x^2} + \frac{1}{2} \\
&= \frac{1}{2} - \frac{1}{2}(x^2+1)\mathrm{e}^{-x^2} \\
\psi(x) &= \frac{1}{2}\mathrm{e}^{-x^2} - \frac{1}{2}(x^2+1)
\end{aligned}$$

General solution:

$$\psi(x) = c\mathrm{e}^{x^2} - \frac{1}{2}(x^2+1) \ , \quad c \in \mathbb{R}$$

A·T·I·C·E

ATICE LLC, Albany NY

ISBN 978-1-951894-15-3