

Mathematik für Naturwissenschaftler*innen*

Stefan Keppeler Universität Tübingen, Fachbereich Mathematik
 Auf der Morgenstelle 10
 72076 Tübingen, Germany

Büro: C4P31

Email-Adresse: stefan.keppeler@uni-tuebingen.de

Inhaltsverzeichnis

1 Prolog: $e^{i\pi} = -1$	4
1.1 π	4
1.2 e	4
1.3 i	5
1.4 Bemerkungen zur Notation	7
2 Vollständige Induktion	8
2.1 Geometrische Summe	8
2.2 Kleiner Gauß	9
2.3 Prinzip der vollständigen Induktion	10
3 Binomische Formel	12
4 Funktionen einer reellen Variablen	16
4.1 Folgenre Grenzwerte	17
4.2 Funktionsgrenzwerte	18
4.3 Stetigkeit	20
4.4 Asymptoten	22
4.5 Differentiation	23
4.5.1 Alternative Definition mit Klein-o	26
4.6 Die Exponentialfunktion	29
4.7 Umkehrfunktionen	33
4.8 Der Logarithmus	36
4.9 Weitere elementare Funktionen	37

*Dieses Skript ist **kein** Lehrbuch sondern das Vorbereitungsmaterial des Dozenten. Es sollte daher eher als ausführliche Formelsammlung betrachtet werden. Das Studium des Skripts kann den Besuch der Vorlesung nicht ersetzen – möglicherweise ist es aber beim Nachbereiten der Vorlesung, d.h. beim Bearbeiten der **Übungsaufgaben**, hilfreich.

Für Hinweise auf (sicherlich vorhandene) Fehler jeglicher Art sowie für sonstige Anregungen bin ich dankbar.
Stefan Keppeler

4.10	Trigonometrische Funktionen	38
4.11	Potenzreihen	41
4.12	Kurvendiskussion	48
5	Vektorrechnung	53
5.1	Vektorräume	53
5.2	Lineare Unabhängigkeit	56
5.3	Lineare Gleichungssysteme und allgemeiner Gauß-Algorithmus	60
5.4	Unterräume, Dimension und Basis	62
5.5	Skalarprodukt und Norm	65
5.6	Kreuzprodukt und Spatprodukt im \mathbb{R}^3	70
5.7	Geraden und Ebenen	72
5.8	Kurven und Spezielle Koordinatensysteme	74
6	Matrizen und Determinanten	77
6.1	Matrizen	77
6.2	Determinanten	82
7	Komplexe Zahlen	87
7.1	Komplexe e-Funktion	88
7.2	\mathbb{C}^n , $\mathbb{C}^{m \times n}$ und Verwandtes	91
8	Integration	93
8.1	Uneigentliche Integrale	96
8.2	Integrationstechniken	97
8.3	Riemannsche Zwischensummen	104
9	Differentialgleichungen (DGLn)	107
9.1	DGLn erster Ordnung mit getrennten Veränderlichen	107
9.2	Lineare DGLn erster Ordnung	109
9.3	Lineare DGLn zweiter Ordnung mit konstanten Koeffizienten	111
9.4	Existenz und Eindeutigkeit für DGLn 1. Ordnung	116
10	Hauptachsentransformation	118
10.1	Eigenwerte und Eigenvektoren	118
10.2	Einige Begriffe	120
10.3	Diagonalisierbarkeit	121
10.3.1	Praktische Durchführung der HAT	124
10.4	Quadratische Formen	125
11	Differentiation in mehreren Variablen	131
11.1	Punktmengen und Folgen im \mathbb{R}^n	131
11.2	Stetige Funktionen mehrerer Variabler	132
11.3	Partielle, Richtungs- und totale Ableitung	133

11.4	Höhere Ableitungen und Satz von Taylor	138
11.5	Lokale Minima und Maxima	142
11.6	Vektorwertige Funktionen	144
11.7	Implizit definierte Funktionen	146
11.8	Extremwerte unter Nebenbedingungen	150
9	Nachtrag: Differentialgleichungen (DGLn)	154
9.5	Existenz und Eindeutigkeit für Systeme und DGLn beliebiger Ordnung . .	154
12	Integration im \mathbb{R}^n	156
12.1	Bereichs- und iterierte Integrale	156
12.2	Oberflächenintegrale	162
12.3	Integralsätze	166
13	Wahrscheinlichkeitsrechnung & Statistik	170
13.1	Axiomatischer Aufbau	170
13.2	Kombinatorik	174
13.2.1	Laplacesche Wahrscheinlichkeitsräume	174
13.2.2	Permutationen	174
13.2.3	Urnenmodelle	175
13.3	Bedingte Wahrscheinlichkeiten	176
13.4	Zufallsvariable & Verteilungsfunktionen	178
13.4.1	Kenngößen von Zufallsvariablen	182
13.5	Exkurs: Beschreibende Statistik	189
13.6	Grenzwertsätze	191
13.7	Schließende Statistik: Hypothesentests	192
13.7.1	Der t-Test	194

1 Prolog: $e^{i\pi} = -1$

In diesem Vorspann werden wir – zunächst noch nicht mit der mathematischen Strenge, die wir uns später zu eigen machen wollen – den Beweis der Gleichung $e^{i\pi} = -1$ skizzieren. In der ZEIT vom 6. Juni 2007 (www.zeit.de/2007/24/N-Eulersche-Zahl) wurde diese Gleichung, umgestellt zu $e^{i\pi} + 1 = 0$, zur *schönsten mathematischen Gleichung schlechthin* gekürt.

1.1 π

Definiert am Kreis durch

$$\pi := \frac{\text{Umfang}}{\text{Durchmesser}} \approx 3,141\dots \quad (1.1)$$

damit auch

$$\pi = \text{Halber Umfang des Einheitskreises (EK)},$$

und als Winkel $\pi \hat{=} 180^\circ$.

1.2 e

Wir verzinsen ein Guthaben, sagen wir 100€, jährlich mit einem Zinssatz z , z.B. $z = 2,5\% = 0,025$.

Das Vermögen nach einem Jahr erhalten wir durch Multiplikation mit dem Faktor $1 + z$, z.B. $100\text{€} \cdot 1,025 = 102,50\text{€}$.

Bei halbjährlicher Verzinsung mit Zinssatz $\frac{z}{2}$ ist Faktor nach einem Jahr

$$\left(1 + \frac{z}{2}\right) \left(1 + \frac{z}{2}\right) = \left(1 + \frac{z}{2}\right)^2 = 1 + z + \left(\frac{z}{2}\right)^2 > 1 + z, \quad (1.2)$$

also attraktiver, wegen des Zinseszins'.

Analog:

- Monatliche Verzinsung mit Zinssatz $\frac{z}{12}$, Faktor: $\left(1 + \frac{z}{12}\right)^{12}$
- Tägliche Verzinsung mit Zinssatz $\frac{z}{365}$, Faktor: $\left(1 + \frac{z}{365}\right)^{365}$
- ... wird immer mehr. Wie groß maximal?

Stetige Verzinsung: Kapital wächst innerhalb eines Jahres um den Faktor

$$\lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n \quad (1.3)$$

Man zeigt (haben Sie in bereits der Schule gesehen, analysieren wir aber auch später nochmals genau)

$$\lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n = \exp(z) = e^z \quad \text{und insbesondere} \quad \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e \approx 2,718\dots \quad (1.4)$$

1.3 i

Wir definieren: $i := \sqrt{-1}$.

Behauptung: $(2,718\dots)^{3,141\dots\sqrt{-1}} = -1$

1. Wie rechnen wir mit i?

- Bilde Zahlen $x + iy$ mit $x, y \in \mathbb{R}$,
- schleppe i überall mit,
- und ersetze i^2 durch -1 ,

z.B. Multiplikation: $(1 + 2i) \cdot (3 - i) = 3 - i + 6i - 2i^2 = 5 + 5i$.

2. Was bedeutet das anschaulich?

Betrachte $z = x + iy$ als Punkt in der (komplexen) Ebene.

In Polardarstellung:²

$$r := \sqrt{x^2 + y^2} \quad (\text{Abstand zum Ursprung})$$

$$\varphi := \sphericalangle \text{ zur } x\text{-Achse}$$

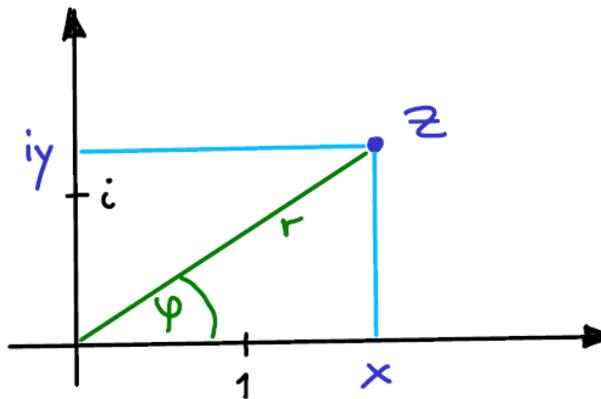
(1.5)

und umgekehrt

$$x = r \cos \varphi$$

$$y = r \sin \varphi$$

(1.6)



Nochmal Multiplikation,

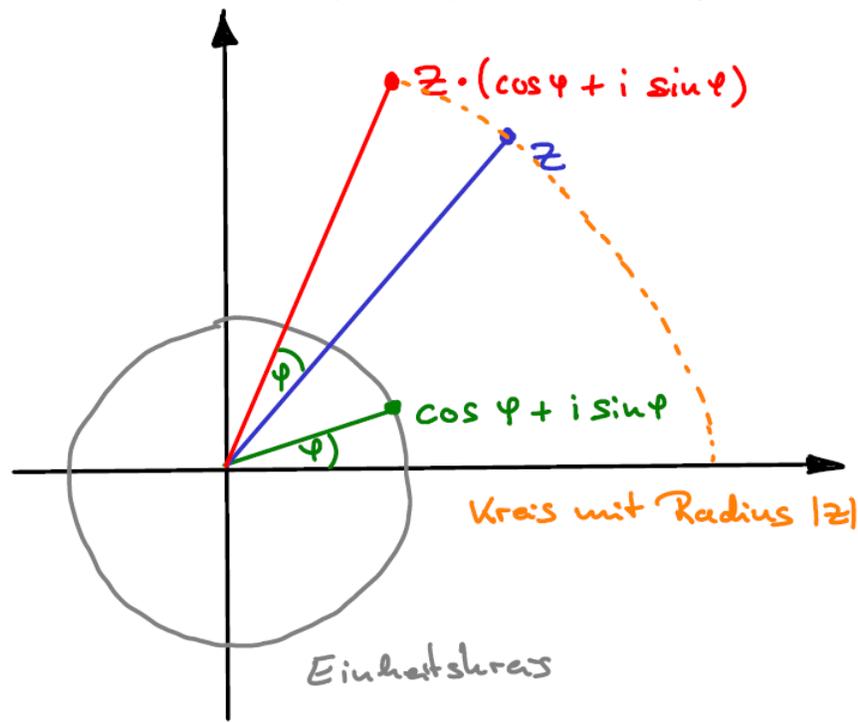
$$\begin{aligned} z_1 \cdot z_2 &= r_1(\cos \varphi_1 + i \sin \varphi_1) \cdot r_2(\cos \varphi_2 + i \sin \varphi_2) \\ &= r_1 r_2 (\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2 + i(\cos \varphi_1 \sin \varphi_2 + \sin \varphi_1 \cos \varphi_2)), \end{aligned} \quad (1.7)$$

mit Additionstheoremen (Trigonometrie, bekannt aus der Schule, kommt aber auch nochmal später),

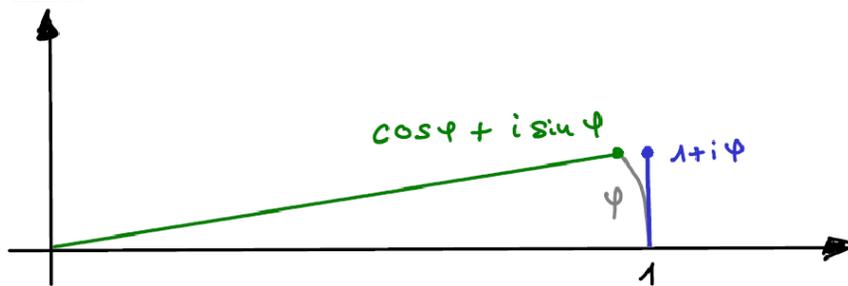
$$z_1 \cdot z_2 = r_1 r_2 (\cos(\varphi_1 + \varphi_2) + i \sin(\varphi_1 + \varphi_2)). \quad (1.8)$$

²Wir nennen r den Betrag und φ das Argument von z und schreiben dafür: $r = |z|$ und $\varphi = \arg(z)$.

Speziell $r = 1$: Multiplikation mit $\cos \varphi + i \sin \varphi$ bewirkt Drehung um Winkel φ .



3. Kleine Winkel



$$\cos \varphi + i \sin \varphi \approx 1 + i\varphi \quad \text{für kleine } \varphi \text{ (siehe Skizze)} \quad (1.9)$$

Je kleiner φ , desto besser ist diese Näherung.

4. Setze endliche Drehung zusammen

$$\cos \varphi + i \sin \varphi = \lim_{n \rightarrow \infty} \left(\underbrace{\cos \frac{\varphi}{n} + i \sin \frac{\varphi}{n}}_{\approx 1 + i \frac{\varphi}{n}, \text{ wenn } n \text{ groß}} \right)^n = \lim_{(*)} \left(1 + \frac{i\varphi}{n} \right)^n = e^{i\varphi} \quad (1.10)$$

(*) Hier sind eine Näherung und ein Limes involviert, d.h. wir müssen genauer untersuchen, ob es erlaubt war, in der Klammer die Näherung zu verwenden. Dazu müssen wir zunächst den bei der Näherung gemachten Fehler abschätzen – das wird später eine wichtiges Thema.

Mit $\varphi = \pi$ ($\cos \pi = -1$, $\sin \pi = 0$) folgt die Behauptung. ☺

1.4 Bemerkungen zur Notation

Zahlenmengen

- \mathbb{N} die natürlichen Zahlen $\{1, 2, 3, \dots\}$,
- $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$ die natürlichen Zahlen inkl. der Null,
- $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ die ganzen Zahlen,
- $\mathbb{Q} = \{\frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0\}$ die rationalen Zahlen – Brüche bzw. Zahlen mit endlicher oder periodischer Dezimalbruchentwicklung,
- \mathbb{R} die reellen Zahlen – enthalten zusätzlich zu \mathbb{Q} Zahlen wie $\sqrt{2} \approx 1,414\dots$, $\pi \approx 3,141\dots$ oder $e \approx 2,718\dots$, die nicht als Bruch geschrieben werden können.
- \mathbb{C} die komplexen Zahlen, sind von der Form $\mathbb{C} \ni z = x + iy$ mit $x, y \in \mathbb{R}$.

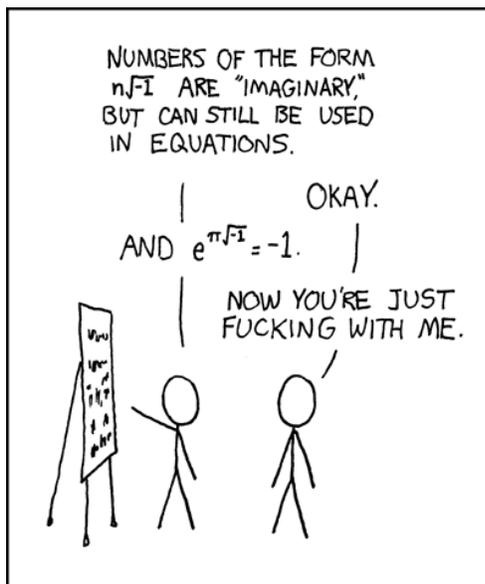
Bemerkung: $\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$

Griechische Buchstaben

A	α	Alpha
B	β	Beta
Γ	γ	Gamma
Δ	δ	Delta
E	ϵ, ε	Epsilon
Z	ζ	Zeta
H	η	Eta
Θ	θ, ϑ	Theta

I	ι	Iota
K	κ, \varkappa	Kappa
Λ	λ	Lambda
M	μ	My
N	ν	Ny
Ξ	ξ	Xi
O	o	Omikron
Π	π	Pi

P	ρ, ϱ	Rho
Σ	σ	Sigma
T	τ	Tau
Υ	υ	Ypsilon
Φ	ϕ, φ	Phi
X	χ	Chi
Ψ	ψ	Psi
Ω	ω	Omega



<http://xkcd.com/179/>

2 Vollständige Induktion

2.1 Geometrische Summe

$$s_n := 1 + x + x^2 + \dots + x^n \quad , \quad x \in \mathbb{R} \text{ (oder } x \in \mathbb{C}) \quad (2.1)$$

$$= \sum_{k=0}^n x^k \quad (x^0 = 1) \quad (2.2)$$

Behauptung: ($n \in \mathbb{N}_0$)

$$s_n = \begin{cases} \frac{x^{n+1} - 1}{x - 1} & , \quad x \neq 1 \\ n + 1 & , \quad x = 1 \end{cases} \quad (2.3)$$

1. Beweis: (durch Induktion)

(i) Induktionsanfang: $n = 0$

$$s_0 = \sum_{k=0}^0 x^0 = 1 \quad (\text{wie in Definition}) \quad (2.4)$$

$$= \begin{cases} \frac{x^{0+1} - 1}{x - 1} = 1 & , \quad x \neq 1 \\ 0 + 1 = 1 & , \quad x = 1 \end{cases} \quad (\text{wie in Behauptung}) \quad (2.5)$$

(ii) Induktionsschritt / $n \rightarrow n + 1$:

Behauptung sei für (ein festens) n bewiesen;

zu zeigen: Behauptung gilt auch für $n + 1$

$$\begin{aligned} s_{n+1} &= 1 + x + \dots + x^n + x^{n+1} \\ &= s_n + x^{n+1} \\ &\stackrel{\text{i.V.}}{=} \begin{cases} \frac{x^{n+1} - 1}{x - 1} + x^{n+1} & , \quad x \neq 1 \\ n + 1 + 1 & , \quad x = 1 \end{cases} \\ &= \begin{cases} \frac{x^{n+2} - 1}{x - 1} & , \quad x \neq 1 \\ n + 2 & , \quad x = 1 \end{cases} \end{aligned} \quad (2.6)$$

□

2. Beweis: (konstruktiv)

für $x = 1$: $s_n = \sum_{k=0}^n 1 = n + 1$

für $x \neq 1$:

$$s_n = 1 + x + \dots + x^n \quad | \cdot x \quad (2.7)$$

$$xs_n = x + x^2 + \dots + x^{n+1} \quad (2.8)$$

subtrahiere erste Zeile von der zweiten

$$s_n(x - 1) = x^{n+1} - 1 \quad \Leftrightarrow_{x \neq 1} \quad \frac{x^{n+1} - 1}{x - 1} \quad (2.9)$$

Nochmal mit Summenzeichen:

□

$$\begin{aligned} (x - 1)s_n &= xs_n - s_n = x \sum_{\nu=0}^n x^\nu - \sum_{\nu=0}^n x^\nu = \sum_{\nu=0}^n x^{\nu+1} - \sum_{\nu=0}^n x^\nu = \sum_{\nu=1}^{n+1} x^\nu - \sum_{\nu=0}^n x^\nu \\ &= x^{n+1} + \sum_{\nu=1}^n x^\nu - \left(x^0 + \sum_{\nu=1}^n x^\nu \right) = x^{n+1} - 1 \end{aligned} \quad (2.10)$$

2.2 Kleiner Gauß

Summe der ersten n ganzen Zahlen

$$\begin{aligned} s_n &:= 1 + 2 + \dots + n \\ &= \sum_{k=1}^n k \quad \left(= \sum_{k=0}^n k \right) \end{aligned} \quad (2.11)$$

Behauptung:

$$s_n = \frac{n(n+1)}{2} \quad (2.12)$$

1. Beweis: (konstruktiv)

$$\begin{array}{r} s_n = 1 + 2 + \dots + n \\ s_n = n + (n-1) + \dots + 1 \\ \hline 2s_n = n(n+1) \end{array} \quad (2.13)$$

$$\Rightarrow s_n = \frac{n(n+1)}{2} \quad \square$$

2. Beweis: (durch vollständige Induktion)

(i) Induktionsanfang: $n = 1$

$$s_1 = \sum_{k=1}^1 k = 1 \quad (\text{Definition}) \quad (2.14)$$

$$= \frac{1 \cdot 2}{2} = 1 \quad (\text{Behauptung}) \quad (2.15)$$

(ii) $n \rightarrow n + 1$:

$$s_{n+1} = s_n + (n + 1) = \frac{n(n + 1)}{2} + (n + 1) = \frac{(n + 2)(n + 1)}{2} \quad (2.16)$$

□

2.3 Prinzip der vollständigen Induktion

Sei $A(n)$ eine Aussage für ganze Zahlen ($n \in \mathbb{Z}$).

Gilt:

(i) $A(n_0)$ ist wahr, für ein $n_0 \in \mathbb{Z}$ (Induktionsanfang), und

(ii) aus $A(n)$ wahr folgt $A(n + 1)$ wahr $\forall n \geq n_0$

so folgt: $A(n)$ ist wahr $\forall n \geq n_0$.

Vorsicht: Nie den Induktionsanfang vergessen!

Sonst lassen sich viele falsche Dinge beweisen, z.B.

Behauptung (falsch!): $x^n < \frac{1}{2} \forall 0 < x < 1$

$n \rightarrow n + 1$:

$$x^{n+1} = x \cdot x^n \underset{\text{i.V.}}{<} x \cdot \frac{1}{2} \underset{0 < x < 1}{<} \frac{1}{2} \quad (2.17)$$

Aber wir haben keinen Induktionsanfang!

Beispiel: Für $n \geq 4$ gilt: $n^2 \leq 2^n$

Induktionsanfang, $n = 4$:

$$16 = 4^2 \leq 2^4 = 16 \quad (2.18)$$

$n \rightarrow n + 1$:

$$2^{n+1} = 2 \cdot 2^n \underset{\text{i.V.}}{\geq} 2n^2 = n^2 + n^2 \underset{n \geq 4}{\geq} n^2 + 4n = n^2 + 2n + 2n \geq n^2 + 2n + 1 = (n + 1)^2 \quad (2.19)$$

□

Notation: Summenzeichen

$a_n \in \mathbb{R}$ (oder $\in \mathbb{C}$)

$$\sum_{k=n}^m a_k := \begin{cases} a_n + a_{n+1} + \dots + a_m & , \quad m \geq n \\ 0 & , \quad m < n \end{cases} \quad (2.20)$$

Indexverschiebung:

$$\begin{aligned}\sum_{k=n}^m a_k &= \sum_{k=0}^{m-n} a_{k+n} \\ &= \sum_{l=n+1}^{m+1} a_{l-1}\end{aligned}\tag{2.21}$$

Merke: Was man an den Grenzen abzieht, muss man am Index der Summanden addieren (und umgekehrt).

Rechenregeln:

$$\sum_{k=0}^n a_k + \sum_{l=0}^n b_l = \sum_{k=0}^n (a_k + b_k)\tag{2.22}$$

$$c \sum_{k=0}^n a_k = \sum_{k=0}^n (ca_k)\tag{2.23}$$

$$\sum_{k=0}^n \left(\sum_{l=0}^m a_{kl} \right) = \sum_{l=0}^m \left(\sum_{k=0}^n a_{kl} \right) \quad \left(=: \sum_{k=0}^n \sum_{l=0}^m a_{kl} \right)\tag{2.24}$$

Was ist hier falsch?

Behauptung: Alle einfarbigen Pferde haben die gleiche Farbe.

Beweis durch vollständige Induktion:

Induktionsanfang:

Ein einfarbiges Pferd hat offensichtlich die gleiche Farbe (wie es selbst)

$n \rightarrow n + 1$:

Wir betrachten eine Menge von $n + 1$ Pferden. Nehmen wir ein Pferd weg, so verbleiben n Pferde. Nach Induktionsvoraussetzung haben diese alle die gleiche Farbe. Nun nehmen wir ein anderes Pferd weg und stellen das zuvor weggenommene wieder dazu. Wieder haben die verbliebenen n Pferde nach Induktionsvoraussetzung die gleiche Farbe. Den beiden jeweils weggenommenen Pferden bleibt nichts anderes übrig, als auch die gleiche Farbe zu haben.

□ (bzw. eben nicht!)

3 Binomische Formel

Ziel:

$$(a + b)^n = \sum_{\nu=0}^n c_{n\nu} a^\nu b^{n-\nu} \quad (3.1)$$

Hintergrund: Ausmultiplizieren

$$(a + b)^n = \underbrace{(a + b)(a + b) \cdots (a + b)}_{n \text{ mal}} \quad (3.2)$$

Was sind die Koeffizienten $c_{n\nu}$?

Definition: (Fakultät)

Für $n \in \mathbb{N}_0$ sei

$$n! := \begin{cases} 1 & , \quad n = 0 \\ 1 \cdot 2 \cdots n = \prod_{\nu=1}^n \nu & , \quad n \in \mathbb{N} \end{cases} \quad (3.3)$$

Notation: Produktzeichen

$$\prod_{\nu=n}^m a_\nu = \begin{cases} a_n \cdot a_{n+1} \cdots a_m & , \quad m \geq n \\ 1 & , \quad m < n \end{cases} \quad (3.4)$$

Definition: (Binomialkoeffizient)

Für $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}$ sei

$$\binom{\alpha}{k} := \frac{\alpha(\alpha - 1) \cdots (\alpha - k + 1)}{k!} \quad \text{sowie} \quad \binom{\alpha}{0} := 1. \quad (3.5)$$

Besser mit Produktzeichen:

$$\binom{\alpha}{k} = \frac{1}{k!} \prod_{\nu=0}^{k-1} (\alpha - \nu). \quad (3.6)$$

Besonders wichtig für ganze Zahlen, also $n, k \in \mathbb{N}_0$,

$$\binom{n}{k} := \frac{n(n - 1) \cdots (n - k + 1)}{k!} \quad \text{sowie} \quad \binom{n}{0} := 1. \quad (3.7)$$

Eigenschaften: $n, k \in \mathbb{N}_0$

(i)

$$\binom{n}{n} = 1 = \binom{n}{0} \quad (3.8)$$

(ii)

$$\binom{n}{k} = 0, \quad \text{für } k > n \quad (3.9)$$

(iii)

$$\begin{aligned}\binom{n}{k} &= \frac{n(n-1)\cdots(n-k+1)}{k!} \cdot \frac{(n-k)(n-k-1)\cdots 1}{(n-k)(n-k-1)\cdots 1} \\ &= \frac{n!}{k!(n-k)!} \\ &= \binom{n}{n-k}\end{aligned}\tag{3.10}$$

(iv) Funktionalgleichung

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}\tag{3.11}$$

Beweis:

$$\begin{aligned}\binom{n}{k-1} + \binom{n}{k} &= \frac{k}{k} \frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!} \frac{(n-k+1)}{(n-k+1)} \\ &= \frac{n!}{k!(n-k+1)!} \underbrace{(k + (n-k+1))}_{=n+1} \\ &= \frac{(n+1)!}{k!(n+1-k)!} \\ &= \binom{n+1}{k}\end{aligned}\tag{3.12}$$

□

Bemerkung: Aufbau des Pascalschen Dreiecks aus (i) und (iv).

Satz 1. (Binomi)

$\forall a, b \in \mathbb{R}$ (oder $\in \mathbb{C}$) und $\forall n \in \mathbb{N}_0$ gilt

$$(a+b)^n = \sum_{\nu=0}^n \binom{n}{\nu} a^\nu b^{n-\nu}.\tag{3.13}$$

Beweis: (mit vollständiger Induktion)

(i) Induktionsanfang, $n = 0$:

links:

$$(a+b)^0 = 1\tag{3.14}$$

rechts:

$$\sum_{\nu=0}^0 \binom{0}{\nu} a^\nu b^{0-\nu} = \binom{0}{0} a^0 b^0 = 1 \cdot 1 \cdot 1 = 1\tag{3.15}$$

...oder auch $n = 1$:

linke Seite:

$$(a+b)^1 = a+b\tag{3.16}$$

rechte Seite:

$$\sum_{\nu=0}^1 \binom{1}{\nu} a^\nu b^{1-\nu} = \binom{1}{0} a^0 b^{1-0} + \binom{1}{1} a^1 b^{1-1} = b + a \quad (3.17)$$

(ii) $n \rightarrow n + 1$:

$$\begin{aligned} (a+b)^{n+1} &= (a+b)(a+b)^n \\ &\stackrel{\text{i.V.}}{=} (a+b) \sum_{\nu=0}^n \binom{n}{\nu} a^\nu b^{n-\nu} \\ &= \sum_{\nu=0}^n \binom{n}{\nu} a^{\nu+1} b^{n-\nu} + \sum_{\nu=0}^n \binom{n}{\nu} a^\nu b^{n+1-\nu} \\ &= \sum_{\nu=1}^{n+1} \binom{n}{\nu-1} a^\nu b^{n-(\nu-1)} + \sum_{\nu=0}^n \binom{n}{\nu} a^\nu b^{n+1-\nu} \\ &= a^{n+1} + \sum_{\nu=1}^n \binom{n}{\nu-1} a^\nu b^{n+1-\nu} + \sum_{\nu=1}^n \binom{n}{\nu} a^\nu b^{n+1-\nu} + b^{n+1} \\ &= a^{n+1} + \sum_{\nu=1}^n \underbrace{\left[\binom{n}{\nu-1} + \binom{n}{\nu} \right]}_{= \binom{n+1}{\nu}} a^\nu b^{n+1-\nu} + b^{n+1} \\ &= \sum_{\nu=0}^{n+1} \binom{n+1}{\nu} a^\nu b^{n+1-\nu} \end{aligned} \quad (3.18)$$

□

Beispiele:

1.

$$(a+b)^2 = a^2 + 2ab + b^2 \quad (3.19)$$

2.

$$(x-y)^3 = x^3 - 3x^2y + 3xy^2 - y^3 \quad (3.20)$$

3.

$$(1+1)^n = 2^n = \sum_{\nu=0}^n \binom{n}{\nu} \quad (3.21)$$

Bemerkung: Stirlingsche Formel (ohne Beweis)

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n, \quad n \rightarrow \infty \quad (3.22)$$

Asymptotisches Verhalten für große n , d.h.

$$\lim_{n \rightarrow \infty} \frac{n!}{\sqrt{2\pi n} \left(\frac{n}{e}\right)^n} = 1 \quad (3.23)$$

Merke: Die Fakultät wächst also im Wesentlichen wie n^n .

Satz 2. (Bernoullische Ungleichung)

$\forall n \in \mathbb{N}_0$ und $\forall x \in \mathbb{R}, x \geq -1$, gilt

$$(1+x)^n \geq 1+nx \tag{3.24}$$

Beweis: (vollständige Induktion)

(i) $n = 0, n = 1$ klar

(ii) $n \rightarrow n + 1$:

$$\begin{aligned} (1+x)^{n+1} &= (1+x)(1+x)^n \\ &\stackrel{\text{i.V.}}{\geq} (1+x)(1+nx) = 1+nx+x+\underbrace{nx^2}_{\geq 0} \\ &\geq 1+(n+1)x \end{aligned} \tag{3.25}$$

(In der ersten Zeile sind beide Faktoren ≥ 0 , da $x \geq -1$.) □

4 Funktionen einer reellen Variablen

Intervalle $I \subset \mathbb{R}$

abgeschlossenes Intervall: $I = [a, b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$

offenes Intervall: $I = (a, b) = \{x \in \mathbb{R} \mid a < x < b\}$

oder gemischt, z.B. $1 \in (0, 1]$ aber $0 \notin (0, 1]$.

Bemerkungen: $\mathbb{R} = (-\infty, \infty)$, $\mathbb{R} \setminus \{0\} = (-\infty, 0) \cup (0, \infty)$

Funktionen

$D \subseteq \mathbb{R}$ (d.h. Teilmenge oder gleich \mathbb{R} , z.B. ein Intervall)

$$\begin{aligned} f : D &\rightarrow \mathbb{R} \\ x &\mapsto f(x) \end{aligned} \tag{4.1}$$

Jedem $x \in D$ wird genau ein $f(x) \in \mathbb{R}$ zugeordnet.

D heißt Definitionsbereich von f .

Das Bild von f ist die Menge

$$f(D) = \{y \in \mathbb{R} \mid \exists x \in D, \text{ so dass } f(x) = y\}. \tag{4.2}$$

Komposition (bzw. Verkettung) von Funktionen

$$\begin{aligned} f : D &\rightarrow W \subset \mathbb{R} \\ x &\mapsto f(x) \end{aligned} \tag{4.3}$$

$$\begin{aligned} g : W &\rightarrow \mathbb{R} \\ y &\mapsto g(y) \end{aligned}$$

Dann existiert auch eine Abbildung

$$\begin{aligned} g \circ f : D &\rightarrow \mathbb{R} \\ x &\mapsto (g \circ f)(x) := g(f(x)) \end{aligned} \tag{4.4}$$

Beispiele:

1.

$$f : \mathbb{R} \rightarrow \mathbb{R} \tag{4.5}$$

$$x \mapsto x^5 + 1$$

$$g : \mathbb{R} \rightarrow \mathbb{R} \tag{4.5}$$

$$x \mapsto x^3 - 2 \tag{4.6}$$

(oder kurz: $f(x) = x^5 + 1$ und $g(x) = x^3 - 2$.)

Sowohl $f \circ g$ als auch $g \circ f$ existieren.

$$\begin{aligned} f \circ g : x &\mapsto f(x^3 - 2) = (x^3 - 2)^5 + 1 \\ g \circ f : x &\mapsto g(x^5 + 1) = (x^5 + 1)^3 - 2 \end{aligned} \tag{4.7}$$

Bemerkung: I.A. ist $f \circ g \neq g \circ f$ (hier z.B. Ausrechnen mit Binomi)

2.

$$f : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+ \qquad g : \mathbb{R} \rightarrow \mathbb{R} \qquad (4.8)$$

$$x \mapsto \sqrt{x} \qquad x \mapsto x^2 - 1 \qquad (4.9)$$

$f \circ g$ existiert und ist definiert durch³

$$\mathbb{R}_0^+ \ni x \mapsto g(\sqrt{x}) = x - 1 \qquad (4.10)$$

$f \circ g$ existiert dagegen nicht (zumindest nicht für alle x im Definitionsbereich von g).

4.1 Folgengrenzwerte

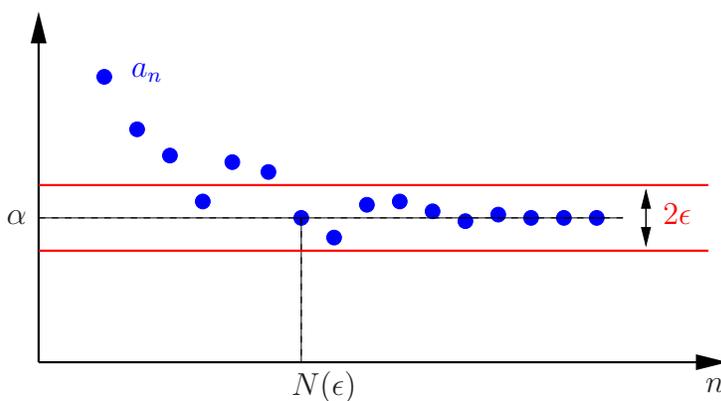
Eine (unendliche) **Folge** (a_n) reeller (oder auch komplexer) Zahlen ist eine Abbildung von \mathbb{N} (oder \mathbb{N}_0) nach \mathbb{R} , d.h. wir ordnen jedem $n \in \mathbb{N}$ ein $a_n \in \mathbb{R}$ (oder \mathbb{C}) zu.

Definition: (Grenzwert)

Eine (unendliche) Folge (a_n) konvergiert gegen den Grenzwert α , wenn gilt:

Für jedes $\epsilon > 0 \exists$ ein $N(\epsilon)$, so dass $|a_n - \alpha| < \epsilon \forall n \geq N(\epsilon)$. Man schreibt dann

$$\lim_{n \rightarrow \infty} a_n = \alpha. \qquad (4.11)$$



Beispiele:

1. $(a_n) = 1, \frac{1}{2}, \frac{1}{3}, \dots$ d.h. $a_n = \frac{1}{n}$, $n \in \mathbb{N}$ konvergiert gegen Null,

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \frac{1}{n} = 0. \qquad (4.12)$$

Bemerkung: Ist der Grenzwert 0, so sprechen wir von einer **Nullfolge**.

³ $\mathbb{R}^+ := \{x \in \mathbb{R} \mid x > 0\} = (0, \infty)$, $\mathbb{R}_0^+ := \{x \in \mathbb{R} \mid x \geq 0\} = [0, \infty)$, analog \mathbb{R}^- , \mathbb{R}_0^-

2. Die Folge $1, -1, 1, -1, 1, \dots$, d.h. $a_n = (-1)^n$, $n \in \mathbb{N}_0$ konvergiert nicht.
3. Die Folge (a_n) mit $a_n = n^2$ konvergiert nicht (divergiert). Hier sagen wir die Folge strebt nach Unendlich und schreiben

$$\lim_{n \rightarrow \infty} n^2 = \infty. \quad (4.13)$$

wobei: $\lim_{n \rightarrow \infty} a_n = \infty \stackrel{\text{Def.}}{\iff} \forall K > 0 \exists N(K), \text{ so dass } a_n > K \forall n \geq N(K)$

4. $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$ (später)
5. $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e \approx 2,718$ (später)

Rechenregeln: Sei $\lim_{n \rightarrow \infty} a_n = \alpha$ und $\lim_{n \rightarrow \infty} b_n = \beta$ (" $\neq \infty$ ", also konvergent) dann gilt:

1. $\lim_{n \rightarrow \infty} (a_n + b_n) = \alpha + \beta$
2. $\lim_{n \rightarrow \infty} (a_n b_n) = \alpha \beta$
3. $\lim_{n \rightarrow \infty} \frac{1}{a_n} = \frac{1}{\alpha}$ falls $\alpha \neq 0$

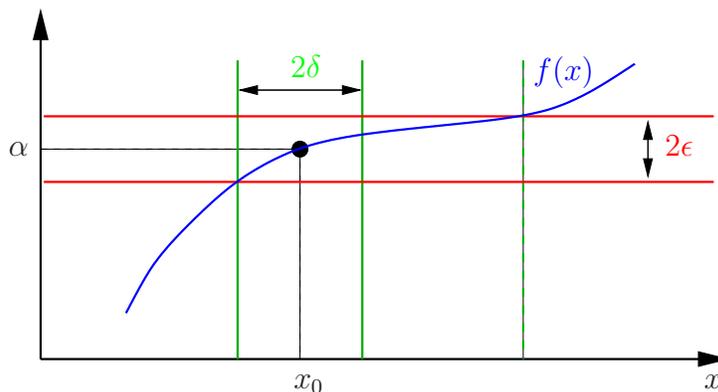
Merke: *Auseinanderziehen erlaubt, falls alles konvergent*

4.2 Funktionsgrenzwerte

Definition: (Grenzwert)⁴

Eine Funktion f hat an der Stelle x_0 den Grenzwert α , wenn gilt: Für jedes $\epsilon > 0 \exists$ ein $\delta(\epsilon) > 0$, so dass $|f(x) - \alpha| < \epsilon \forall x \in D$ mit $x \neq x_0$ und $|x - x_0| < \delta(\epsilon)$. Man schreibt dann

$$\lim_{x \rightarrow x_0} f(x) = \alpha. \quad (4.14)$$



Bemerkung: Ähnlich schreiben wir $\lim_{x \rightarrow x_0} f(x) = \infty$ wenn gilt:

Für jedes $K < \infty \exists$ ein $\delta(K) > 0$, so dass $f(x) > K \forall x \neq x_0$ mit $|x - x_0| < \delta(K)$.

⁴Wird nichts anderes gesagt, so gilt für den Rest des Kapitels stets $f : D \rightarrow \mathbb{R}$, wobei $D \subseteq \mathbb{R}$.

Beispiele:

1. $\lim_{x \rightarrow 2} x^2 = 4$

2. $\lim_{x \rightarrow 0} \frac{1}{x^2} = \infty$, denn

$$\frac{1}{x^2} > K \Leftrightarrow x^2 < \frac{1}{K} \Leftrightarrow |x| < \frac{1}{\sqrt{K}} \quad (4.15)$$

d.h. $\delta(K) = \frac{1}{\sqrt{K}}$ tut's.

3. $\lim_{x \rightarrow 1} \frac{x^2 - 1}{x - 1} = 2$, denn $\frac{x^2 - 1}{x - 1} = \frac{(x-1)(x+1)}{x-1}$ (für $x \neq 1$)

4. $f(x) = \frac{1}{x}$ besitzt bei $x = 0$ keinen Grenzwert, denn

$$\begin{aligned} f(x) > 1 & \quad \forall \quad 0 < x < 1 & \quad \text{und} \\ f(x) < -1 & \quad \forall \quad -1 < x < 0. \end{aligned} \quad (4.16)$$

Definition: (Rechtsseitiger Grenzwert)

Eine Funktion f hat an der Stelle x_0 den rechtsseitigen Grenzwert α , wenn gilt: Für jedes $\epsilon > 0 \exists$ ein $\delta(\epsilon) > 0$, so dass $|f(x) - \alpha| < \epsilon \forall x > x_0$ mit $|x - x_0| < \delta(\epsilon)$. Man schreibt dann

$$\lim_{x \rightarrow x_0+} f(x) = \alpha. \quad (4.17)$$

Analog definiert man den linksseitigen Grenzwert mit $x < x_0$.Zurück zu $f(x) = \frac{1}{x}$:

$$\lim_{x \rightarrow 0+} \frac{1}{x} = +\infty, \quad \lim_{x \rightarrow 0-} \frac{1}{x} = -\infty. \quad (4.18)$$

Bemerkungen:1. Gilt $\lim_{x \rightarrow x_0+} f(x) = \lim_{x \rightarrow x_0-} f(x)$ so existiert der Grenzwert (und ist gleich den Beiden).

2. Man schreibt auch

$$\lim_{x \rightarrow \infty} f(x) = \alpha \quad \text{oder} \quad \lim_{x \rightarrow -\infty} f(x) = \beta, \quad (4.19)$$

obwohl das eher wie ein links- bzw. rechtsseitiger Grenzwert aussieht.

Genauer:

$$\lim_{x \rightarrow \infty} f(x) = \alpha \quad \stackrel{\text{Def.}}{\iff} \quad \text{für jedes } \epsilon > 0 \exists K(\epsilon) < \infty, \text{ so dass } |f(x) - \alpha| < \epsilon \forall x > K(\epsilon).$$

Beispiele:

1.

$$\lim_{x \rightarrow \infty} \frac{1}{x} = 0 \quad \text{oder} \quad \lim_{x \rightarrow \infty} \frac{3x + 2}{4x^2} = 0 \quad (\text{Zählergrad kleiner Nennergrad}) \quad (4.20)$$

2.

$$\lim_{x \rightarrow \infty} \frac{x^2 + 1}{x - 3} = \infty \quad \text{oder} \quad \lim_{x \rightarrow -\infty} \frac{x^2 + 1}{x - 3} = -\infty \quad (\text{Zählergrad größer Nennergrad}) \quad (4.21)$$

3.

$$\lim_{x \rightarrow \infty} \frac{x + 5}{x - 4} = 1 \quad \text{oder} \quad \lim_{x \rightarrow \infty} \frac{1 + 2x^2}{2x + 3x^2} = \frac{2}{3} \quad (\text{Zählergrad gleich Nennergrad}) \quad (4.22)$$

Rechenregeln: Sei $\lim_{x \rightarrow x_0} f(x) = \alpha$ und $\lim_{x \rightarrow x_0} g(x) = \beta$ (wobei $|\alpha|, |\beta| < \infty$) dann gilt:

1. $\lim_{x \rightarrow x_0} (f(x) + g(x)) = \alpha + \beta$,
2. $\lim_{x \rightarrow x_0} (f(x)g(x)) = \alpha\beta$,
3. $\lim_{x \rightarrow x_0} \frac{1}{f(x)} = \frac{1}{\alpha}$, falls $\alpha \neq 0$.

4.3 Stetigkeit

Definition: (Stetigkeit)

Eine Funktion f ist stetig an der Stelle x_0 , wenn gilt: Für jedes $\epsilon > 0 \exists$ ein $\delta > 0$, so dass $|f(x) - f(x_0)| < \epsilon \forall x \neq x_0$ mit $|x - x_0| < \delta$.

Wir fordern also dreierlei:

1. $f(x)$ ist an der Stelle $x = x_0$ definiert.
2. Der Grenzwert an der Stelle x_0 existiert.
3. Grenzwert und Funktionswert stimmen überein.

$$\text{Kurz: } f \text{ stetig in } x_0 \quad \Leftrightarrow \quad \lim_{x \rightarrow x_0} f(x) = f(x_0)$$

Weiter sagt man f ist stetig auf einem Intervall $I \subseteq \mathbb{R}$ wenn f in jedem Punkt $x_0 \in I$ stetig ist.

Beispiele:

1. Monome & Polynome sind überall stetig, rationale Funktionen überall dort, wo der Nenner nicht Null ist.

2.

$$f(x) = \frac{1}{x^2} \quad (4.23)$$

Untersuche Stetigkeit bei $x_0 = 0$:

$f(0)$ existiert nicht \Rightarrow nicht stetig in 0 (Singularität, Polstelle)

3.

$$f(x) = \text{sgn}(x) = \begin{cases} 1 & , x > 0 \\ 0 & , x = 0 \\ -1 & , x < 0 \end{cases} \quad (4.24)$$

Untersuche Stetigkeit bei $x_0 = 0$:

$f(0) = 0$ existiert

$\lim_{x \rightarrow 0} f(x)$ existiert nicht, denn

$$\lim_{x \rightarrow 0^+} f(x) = 1 \quad \text{aber} \quad \lim_{x \rightarrow 0^-} f(x) = -1 \quad (4.25)$$

\Rightarrow nicht stetig in 0 (Sprungstelle)

4.

$$f(x) = \begin{cases} 1 & , \quad x \leq 2 \\ x & , \quad x > 2 \end{cases} \quad (4.26)$$

Untersuche Stetigkeit bei $x_0 = 2$:

$f(2) = 1$ existiert,

$\lim_{x \rightarrow 2^-} f(x) = 1 = f(2)$,

aber $\lim_{x \rightarrow 2^+} f(x) = 2$

\Rightarrow nicht stetig in 2 (Sprungstelle)

Merke: *Eine stetige Funktion kann man ohne Absetzen zeichnen.*

5.

$$f(x) = \begin{cases} 1 & , \quad x \neq 1 \\ 3 & , \quad x = 1 \end{cases} \quad (4.27)$$

Untersuche Stetigkeit bei $x_0 = 1$:

$f(1) = 3$ existiert,

aber $\lim_{x \rightarrow 1} f(x) = 1$

\Rightarrow nicht stetig in 1

6.

$$f(x) = \frac{x^2 - 1}{x - 1}, \quad x \neq 1 \quad (4.28)$$

Untersuche Stetigkeit bei $x_0 = 1$:

$f(1)$ ist nicht definiert (0 durch 0 - Definitionslücke), also auch nicht stetig in 1.

Allerdings existiert $\lim_{x \rightarrow 1} f(x) = 2$ (s.o.); definiere nachträglich

$$f(1) := 2, \quad (4.29)$$

Übrigens, für die so *stetig fortgesetzte* Funktion gilt

$$\begin{aligned} f(x) &= \begin{cases} \frac{x^2-1}{x-1} & , \quad x \neq 1 \\ 2 & , \quad x = 1 \end{cases} \\ &= x + 1 \end{aligned} \quad (4.30)$$

7.

$$f(x) = \begin{cases} 1 & , \quad x < 1 \\ x & , \quad x \geq 1 \end{cases} \quad (4.31)$$

Untersuche Stetigkeit bei $x_0 = 1$:

$f(1) = 1$, $\lim_{x \rightarrow 1^+} = 1$ und $\lim_{x \rightarrow 1^-} = 1$,

also stetig in 1

Analog zu den Rechenregeln für Grenzwerte gilt:

1. Sind f und g stetig auf $I \subseteq \mathbb{R}$, so sind auch $f + g$, αf ($\alpha \in \mathbb{R}$) und fg stetig.

Außerdem ist $\frac{f}{g}$ stetig $\forall x \in I$ mit $g(x) \neq 0$.

2. Sind $f : I \rightarrow \mathbb{R}$ und $g : J \rightarrow \mathbb{R}$ stetig auf I bzw. J , und ist $g(J) \subseteq I$, dann ist auch $f \circ g$ stetig auf J .

4.4 Asymptoten

Nähert sich eine Funktion einer Geraden beliebig nahe an, so nennt man die Gerade eine Asymptote der Funktion (bzw. des Funktionsgraphs).

Waagerechte Asymptoten

Gilt

$$\lim_{x \rightarrow \infty} f(x) = a \quad \text{oder} \quad \lim_{x \rightarrow -\infty} f(x) = b, \quad (4.32)$$

so hat f eine waagerechte Asymptote $y = a$ bzw. $y = b$.

Senkrechte Asymptoten

Hat f eine Polstelle bei $x = a$, d.h.

$$\lim_{x \rightarrow a+} |f(x)| = \infty \quad \text{oder} \quad \lim_{x \rightarrow a-} |f(x)| = \infty, \quad (4.33)$$

so hat f eine senkrechte Asymptote $x = a$.

Schiefe Asymptoten

Gilt ($a \neq 0$)

$$\lim_{x \rightarrow \infty} [f(x) - (ax + b)] = 0 \quad \text{oder} \quad \lim_{x \rightarrow -\infty} [f(x) - (ax + b)] = 0, \quad (4.34)$$

so hat f die schiefe Asymptote $g(x) = ax + b$.⁵

Beispiele:

1. $f(x) = \frac{1}{x}$
waagerechte Asymptote $y = 0$ (x -Achse)
senkrechte Asymptote $x = 0$ (y -Achse).

2. $f(x) = \frac{3x^2 + 4}{2x^2 - x}$
waagerechte Asymptote $y = \frac{3}{2}$
senkrechte Asymptoten $x = 0$ und $x = \frac{1}{2}$

⁵Wir schreiben dafür auch $f(x) \sim ax + b$ und meinen dann nicht nur $\lim_{x \rightarrow \pm\infty} \frac{f(x)}{ax+b} = 1$ (vgl. Abschnitt 3, Stirlingsche Formel) sondern auch $f(x) - ax \sim b$, d.h. $\lim_{x \rightarrow \pm\infty} \frac{f(x) - ax}{b} = 1$.

3. $f(x) = \frac{3x^2 + 4}{x - 1}$
 senkrechte Asymptote $x = 1$
 schiefe Asymptote... Polynomdivision

$$\begin{array}{r} (3x^2 + 4) : (x - 1) = 3x + 3 + \frac{7}{x - 1} \\ \underline{-3x^2 + 3x} \\ 3x + 4 \\ \underline{-3x + 3} \\ 7 \end{array} \quad (4.35)$$

schiefe Asymptote $g(x) = 3x + 3$

4.5 Differentiation

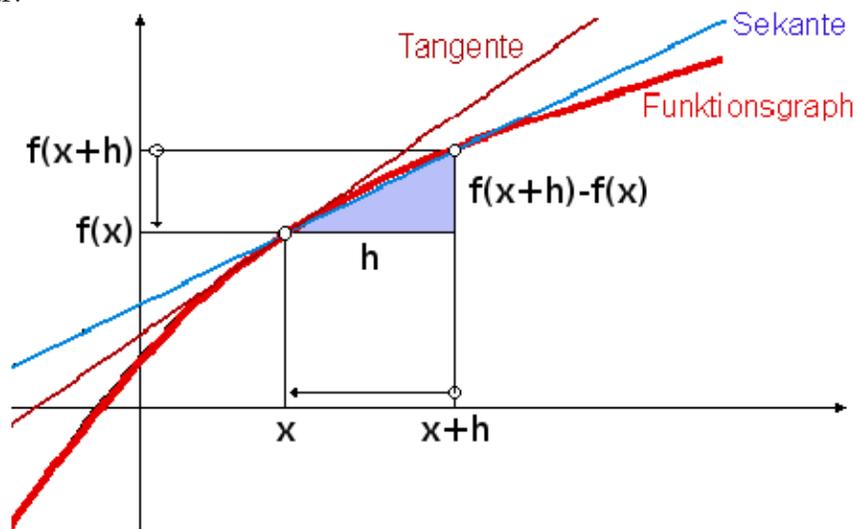
Definition: Sei $I \subset \mathbb{R}$ offen, $x \in I$. Die Funktion $f : I \rightarrow \mathbb{R}$ heißt differenzierbar in x , falls

$$\lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \quad (4.36)$$

existiert. Wir nennen dann

$$f'(x) := \frac{d}{dx} f(x) := \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \quad (4.37)$$

die Ableitung von f an der Stelle x . Ist $f \forall x \in I$ differenzierbar, so heißt f auf I differenzierbar.



Anschaulich: $f'(x)$ ist die Steigung der Tangente an f im Punkt $(x, f(x))$.

Analog für höhere Ableitungen:

$$\begin{aligned}
 f''(x) &= f^{(2)}(x) = \left(\frac{d}{dx}\right)^2 f(x) \\
 f''' &= f^{(3)} = \frac{d^3 f}{dx^3} \quad \text{etc.}
 \end{aligned}
 \tag{4.38}$$

Bemerkung: Damit die Ableitung existieren kann, muss (notwendigerweise)

$$\lim_{h \rightarrow 0} f(x+h) = f(x)$$

gelten. (Stetigkeit in x)

Beispiele:

1. $f : x \mapsto x^n, n \in \mathbb{N}_0$

$$n = 0: f(x) \equiv 1$$

$$\Rightarrow f(x+h) - f(x) = 0 \quad \forall h \text{ und } \forall x$$

$$\Rightarrow f'(x) \equiv 0$$

beliebige $n \geq 1$:

$$\begin{aligned}
 \lim_{h \rightarrow 0} \frac{(x+h)^n - x^n}{h} &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\sum_{\nu=0}^n \binom{n}{\nu} x^{n-\nu} h^\nu - x^n \right) \\
 &= \lim_{h \rightarrow 0} \sum_{\nu=1}^n \binom{n}{\nu} x^{n-\nu} h^{\nu-1} = \lim_{h \rightarrow 0} \sum_{\nu=0}^{n-1} \binom{n}{\nu+1} x^{n-\nu-1} h^\nu \\
 &= \lim_{h \rightarrow 0} \left[\binom{n}{1} x^{n-1} + \underbrace{\sum_{\nu=1}^{n-1} \binom{n}{\nu+1} x^{n-\nu-1} h^\nu}_{\xrightarrow{h \rightarrow 0} 0} \right] \\
 &= n x^{n-1}
 \end{aligned}
 \tag{4.39}$$

Gilt übrigens sogar für $n \in \mathbb{Z} \setminus \{0\}$, z.B. $f(x) = \frac{1}{x}$:

$$\begin{aligned}
 f'(x) &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{1}{x+h} - \frac{1}{x} \right) \\
 &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{x - (x+h)}{x(x+h)} \right) \\
 &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{-h}{x(x+h)} \right) = -\frac{1}{x^2}
 \end{aligned}
 \tag{4.40}$$

2.

$$f(x) = |x| = \begin{cases} x & , \quad x \geq 0 \\ -x & , \quad x < 0 \end{cases} \quad (4.41)$$

$$f'(x) = \begin{cases} 1 & , \quad x > 0 \\ -1 & , \quad x < 0 \end{cases} \quad (4.42)$$

$f'(0)$ existiert nicht.

3. $1 \leq k < n$, $f(x) = x^n$

$$\begin{aligned} f^{(k)}(x) &= n(n-1) \cdots (n-k+1) x^{n-k} \\ &= \binom{n}{k} k! x^{n-k} \end{aligned} \quad (4.43)$$

Satz 3. (Differentiationsregeln)

Seien f und g in x differenzierbar. Dann gilt

(i) *Linearität:* $af + bg$ ist differenzierbar in $x \forall a, b \in \mathbb{R}$, und es gilt

$$(af + bg)'(x) = af'(x) + bg'(x) \quad (4.44)$$

(ii) *Produktregel:* $f \cdot g$ ist differenzierbar in x , und es gilt

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x) \quad (4.45)$$

(iii) *Kettenregel*

Sei f differenzierbar in $g(x) \Rightarrow f \circ g$ ist differenzierbar in x mit

$$\frac{d}{dx} f(g(x)) = f'(g(x)) g'(x) \quad (4.46)$$

Beweise folgen direkt aus Rechenregeln für Grenzwerte (i); oder gleich mit Klein-o (i)–(iii).

Bemerkung: Aus (iii) und der Ableitung von $\frac{1}{x}$ folgt

$$\left(\frac{1}{g(x)} \right)' = -\frac{g'(x)}{g^2(x)} \quad (\text{falls } g(x) \neq 0) \quad (4.47)$$

und daraus mit (ii)

$$\left(\frac{f}{g} \right)' = \left(f \frac{1}{g} \right)' = f' \frac{1}{g} + f \left(-\frac{g'}{g^2} \right) = \frac{f'g - fg'}{g^2} \quad (\text{Quotientenregel}) \quad (4.48)$$

Beispiele:

$$1. ((x^2 + 1)^5 + 7x)' \stackrel{(i)}{=} ((x^2 + 1)^5)' + 7 \stackrel{(iii)}{=} 5(x^2 + 1)^4 (2x) + 7$$

$$2. f(x) = \sqrt[n]{x} = x^{\frac{1}{n}}, \quad x > 0 \quad (n \in \mathbb{Z} \setminus \{0\})$$

definiere $g(x) = x^n \Rightarrow g \circ f$ existiert

$$g(f(x)) = \left(x^{\frac{1}{n}}\right)^n = x, \quad x > 0 \quad (4.49)$$

$$\Rightarrow (g \circ f)'(x) = 1$$

$$\stackrel{(iii)}{=} g'(f(x)) f'(x) = n[f(x)]^{n-1} f'(x) = nx^{\frac{n-1}{n}} f'(x) \quad (4.50)$$

$$\Rightarrow f'(x) = \frac{1}{n} x^{\frac{1-n}{n}} = \frac{1}{n} x^{\frac{1}{n}-1}$$

3.

$$f(x) = \frac{1}{x^2 + 1}, \quad f'(x) \stackrel{(iii)}{=} -\frac{2x}{(x^2 + 1)^2} \quad (4.51)$$

4.5.1 Alternative Definition mit Klein-o

Differenzierbarkeit von f in x_0 bedeutet Approximierbarkeit in der Nähe von x_0 durch eine Gerade, nämlich die Tangente t im Punkt $(x_0, f(x_0))$,

$$t(x) := f(x_0) + f'(x_0)(x - x_0). \quad (4.52)$$

Lemma 4. f diffbar in $x_0 \Leftrightarrow \exists a \in \mathbb{R}$, so dass

$$f(x) = f(x_0) + a(x - x_0) + o(x - x_0), \quad x \rightarrow x_0. \quad (4.53)$$

Wir nennen dann $a = f'(x_0)$ die Ableitung von f an der Stelle x_0 .

Dabei sagen wir $f(x) = o(g(x))$, $x \rightarrow x_0$, falls gilt

$$\lim_{x \rightarrow x_0} \frac{|f(x)|}{|g(x)|} = 0. \quad (4.54)$$

Beispiele:

1. $x^2 = o(x)$, $x \rightarrow 0$
2. $x^3 = o(x)$, $x \rightarrow 0$
3. $x = o(x^2)$, $x \rightarrow \infty$
4. $e^{-x} = o(x^{-n})$, $x \rightarrow \infty$, $\forall n \in \mathbb{N}$ (Beweis später)
5. $x = o(1)$, $x \rightarrow 0$
6. $x o(x) = o(x^2)$, $x \rightarrow 0$

Was bedeutet diese Gleichung eigentlich?

Jedes Klein-o, das irgendwo auftaucht, ist ein Platzhalter für eine Funktion, die wir nicht genau kennen, oder die wir nicht genauer angeben wollen, für die wir aber etwas über ihr Verhalten in einem bestimmten Limes wissen.

In Worten bedeutet $x o(x) = o(x^2)$, $x \rightarrow 0$, also:

Wenn $g(x)$ für $x \rightarrow 0$ schneller verschwindet als x ,
dann verschwindet $x g(x)$ für $x \rightarrow 0$ sogar schneller als x^2 .

Und wie zeigt man das? Zum Beispiel so (hier alles für $x \rightarrow 0$):

$$g(x) = o(x) \Leftrightarrow \lim_{x \rightarrow 0} \frac{g(x)}{x} = 0 \Leftrightarrow \lim_{x \rightarrow 0} \frac{x g(x)}{x^2} = 0 \Leftrightarrow x g(x) = o(x^2)$$

Und nachdem wir so etwas einmal gezeigt haben, merken wir uns, dass wir mit diesem Argument beliebig Faktoren in das Argument des Klein-o hinein (oder aus ihm heraus) ziehen dürfen.

7. Analog zeigt man auch $o(x) + o(x^2) = o(x)$, $x \rightarrow 0$ oder $o(x) o(x) = o(x^2)$ (siehe Übungen). Zusammen können wir das dann z.B. wie folgt verwenden:

$(1 + x + o(x))(5 - x + o(x)) = 5 + 4x + o(x)$, $x \rightarrow 0$, denn

$$[1 + x + o(x)][5 - x + o(x)] = 5 + (5-1)x \underbrace{- x^2}_{=o(x)} + \underbrace{(1+x)o(x)}_{=o(x)+o(x^2)} + \underbrace{(5-x)o(x)}_{=o(x)+o(x^2)} + \underbrace{o(x)o(x)}_{=o(x^2)}$$

$\underbrace{\hspace{15em}}_{=o(x)}$

Beweis von Lemma 4

$$\begin{aligned} & f(x) = f(x_0) + a(x - x_0) + o(x - x_0) && | - f(x_0) \\ \Leftrightarrow & f(x) - f(x_0) = a(x - x_0) + o(x - x_0) && | : (x - x_0) \\ \Leftrightarrow & \frac{f(x) - f(x_0)}{x - x_0} = a + o(1) \\ \Leftrightarrow & \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = a, \end{aligned}$$

d.h. “ \Rightarrow ” wenn f diffbar in x_0 , dann \exists so ein a , und $a = f'(x_0)$,

“ \Leftarrow ” wenn so ein a existiert, dann ist f diffbar in x_0 , und $f'(x_0) = a$.

Beispiele:

1. Beweis der Kettenregel. (Satz 3 (iii))

Sei $f(x) = g(h(x))$, h diffbar in x_0 und g diffbar in $y_0 := h(x_0)$.

$$\begin{aligned} f(x) &= g(\underbrace{h(x_0) + h'(x_0)(x - x_0) + o(x - x_0)}_{=:y}) && \text{(da } h \text{ diffbar in } x_0) \\ &= g(y_0) + g'(y_0)(y - y_0) + o(y - y_0) && \text{(da } g \text{ diffbar in } y_0) \\ &= g(h(x_0)) + g'(h(x_0))(h'(x_0)(x - x_0) + o(x - x_0)) + \underbrace{o(h'(x_0)(x - x_0) + o(x - x_0))}_{=o(x-x_0)} \\ &= \underbrace{g(h(x_0))}_{f(x_0)} + \underbrace{g'(h(x_0))h'(x_0)}_{f'(x_0)}(x - x_0) + \underbrace{g'(h(x_0))o(x - x_0) + o(x - x_0)}_{=o(x-x_0)} && \square \end{aligned}$$

(4.55)

2. l'Hospitalsche Regel.⁶ Wir suchen

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \text{“ } \frac{0}{0} \text{”}, \quad (4.56)$$

also $f(x_0) = 0 = g(x_0)$. Falls f und g diffbar in x_0 und $g'(x_0) \neq 0$, so gilt

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} &= \lim_{x \rightarrow x_0} \frac{\overset{0}{f(x_0)} + f'(x_0)(x - x_0) + o(x - x_0)}{\overset{0}{g(x_0)} + g'(x_0)(x - x_0) + o(x - x_0)} \\ &= \lim_{x \rightarrow x_0} \frac{f'(x_0) + o(1)}{g'(x_0) + o(1)} \\ &= \frac{f'(x_0)}{g'(x_0)}. \end{aligned} \quad (4.57)$$

Gilt analog für⁷

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \text{“ } \frac{\infty}{\infty} \text{”}. \quad (4.58)$$

⁶Die eigentlich auf Johann Bernoulli zurückgeht. . .

⁷Denn $(\lim_{x \rightarrow x_0} f(x) = \pm\infty, \lim_{x \rightarrow x_0} g(x) = \pm\infty)$

$$\begin{aligned} \alpha &:= \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{\frac{1}{\frac{1}{g(x)}}}{\frac{1}{\frac{1}{f(x)}}} \quad \left(\text{vom Typ “ } \frac{0}{0} \text{”} \right) \\ &= \lim_{x \rightarrow x_0} \frac{\frac{g'(x)}{(g(x))^2}}{\frac{f'(x)}{(f(x))^2}} = \lim_{x \rightarrow x_0} \frac{(f(x))^2 g'(x)}{(g(x))^2 f'(x)} = \alpha^2 \lim_{x \rightarrow x_0} \frac{g'(x)}{f'(x)} \\ \Rightarrow \alpha &= \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} \quad (\text{falls existent!}) \end{aligned}$$

4.6 Die Exponentialfunktion

Eulersche Zahl:

$$a_n := \left(1 + \frac{1}{n}\right)^n \quad (4.59)$$

Wir zeigen zunächst, dass a_n nach oben und unten beschränkt ist:

$$\left(1 + \frac{1}{n}\right)^n \geq 1 + 1 = 2 \quad (\text{Bernoullische Ungleichung, Satz 2}) \quad (4.60)$$

$$\begin{aligned} \left(1 + \frac{1}{n}\right)^n &= \sum_{\nu=0}^n \binom{n}{\nu} \left(\frac{1}{n}\right)^\nu \quad (\text{Binomi, Satz 1}) \\ &= \sum_{\nu=0}^n \underbrace{\frac{n(n-1)\cdots(n-\nu+1)}{n^\nu}}_{\leq 1} \frac{1}{\nu!} \\ &\leq \sum_{\nu=0}^n \frac{1}{\nu!} = 1 + \sum_{\nu=1}^n \frac{1}{\nu!} \quad (4.61) \\ &\stackrel{(*)}{\leq} 1 + \sum_{\nu=1}^n \frac{1}{2^{\nu-1}} = 1 + \sum_{\nu=0}^{n-1} \left(\frac{1}{2}\right)^\nu = 1 + \frac{1 - \left(\frac{1}{2}\right)^n}{\frac{1}{2}} \quad (\text{geom. Summe}) \\ &\leq 1 + 2 = 3 \end{aligned}$$

In (*) haben wir verwendet, dass $\nu! \geq 2^{\nu-1}$ für $\nu \geq 1$.

Beweis dafür: (vollständige Induktion)

$$\nu = 1: 1! = 1 = 2^0$$

$$\nu \rightarrow \nu + 1: (\nu + 1)! = (\nu + 1) \cdot \nu! \underset{\text{i.V.}}{\geq} (\nu + 1) \cdot 2^{\nu-1} \underset{\nu \geq 1}{\geq} 2^\nu. \quad \square$$

Bis jetzt wissen wir also:

$$2 \leq a_n \leq 3 \quad (4.62)$$

Jetzt zeigen wir noch, dass a_n monoton wachsend ist, d.h. z.z. ist $a_{n+1} \geq a_n$ bzw. $\frac{a_{n+1}}{a_n} \geq 1$:

$$\begin{aligned} \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{\left(1 + \frac{1}{n}\right)^n} &= \frac{(n+1+1)^{n+1} n^{n+1}}{(n+1)^{n+1} (n+1)^{n+1}} \frac{n+1}{n} \\ &= \frac{n+1}{n} \left(\frac{(n+1+1)(n+1-1)}{(n+1)^2}\right)^{n+1} \quad (4.63) \\ &= \frac{n+1}{n} \left(\frac{(n+1)^2 - 1}{(n+1)^2}\right)^{n+1} = \frac{n+1}{n} \left(1 - \frac{1}{(n+1)^2}\right)^{n+1} \\ &\geq \frac{n+1}{n} \left(1 - \frac{1}{n+1}\right) = 1 \quad (\text{Bernoullische Ungleichung, Satz 2}) \end{aligned}$$

Die Folge a_n wächst monoton und ist nach oben beschränkt – es bleibt ihr nichts anderes übrig, als zu konvergieren. Den Grenzwert nennen wir e (Eulersche Zahl), und wir wissen bereits

$$2 \leq e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \leq 3 \quad (4.64)$$

Numerisch findet man $e = 2,718281828459\dots$

Exponentialfunktion: Wir definieren

$$\exp(x) := \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n, \quad x \in \mathbb{R}, \quad (4.65)$$

und schreiben später auch $\exp(x) = e^x$.

(Dass e^x eine berechtigte Schreibweise ist, ist erst noch zu zeigen – vgl. Satz 5, (ii).)

Bemerkungen:

1. Analog zu oben zeigt man auch hier, dass die Folge für alle $x \in \mathbb{R}$ monoton wächst und beschränkt ist \Rightarrow Konvergenz. (Details nicht schwierig, nur unhübsch)⁸

2. $\exp(0) = 1$ (offensichtlich)

3.

$$\exp(x) \geq 1 + x, \quad (4.66)$$

8

$$\begin{aligned} \left| \left(1 + \frac{x}{n}\right)^n \right| &\leq \sum_{\nu=0}^n \binom{n}{\nu} \left|\frac{x}{n}\right|^\nu = \sum_{\nu=0}^n \underbrace{\frac{n(n-1)\cdots(n-\nu+1)}{n^\nu}}_{\leq 1} \frac{|x|^\nu}{\nu!} \\ &\leq \sum_{\nu=0}^n \frac{|x|^\nu}{\nu!} = \sum_{\nu=0}^{N-2} \frac{|x|^\nu}{\nu!} + \sum_{\nu=N-1}^n \frac{|x|^\nu}{\nu!} \\ &\stackrel{(+)}{\leq} \sum_{\nu=0}^{N-2} \frac{|x|^\nu}{\nu!} + \sum_{\nu=N-1}^n \frac{|x|^\nu}{N^{\nu-N+1}} = \sum_{\nu=0}^{N-2} \frac{|x|^\nu}{\nu!} + \sum_{\nu=0}^{n-N+1} \frac{|x|^{\nu+N-1}}{N^\nu} = \sum_{\nu=0}^{N-2} \frac{|x|^\nu}{\nu!} + |x|^{N-1} \frac{1 - \left(\frac{|x|}{N}\right)^{n-N+2}}{1 - \frac{|x|}{N}} \\ &\leq \sum_{\nu=0}^{N-2} \frac{|x|^\nu}{\nu!} + |x|^{N-1} \frac{N}{N - |x|}, \quad (+): \quad \nu! \geq N^{\nu-N+1} \quad \text{für } \nu \geq N-1 \quad (\text{vollst. Ind.}), \end{aligned}$$

also nach oben beschränkt, da Ergebnis nicht von n abhängt

(N wird allein durch x festgelegt, $|x| < N \leq n$)

$$\begin{aligned} \frac{\left(1 + \frac{x}{n+1}\right)^{n+1}}{\left(1 + \frac{x}{n}\right)^n} &= \frac{(n+1+x)^{n+1} n^n}{(n+x)^n (n+1)^{n+1}} = \frac{n+x}{n} \left(\frac{n^2+n+nx}{n^2+nx+n+x}\right)^{n+1} = \frac{n+x}{n} \left(1 - \frac{x}{n^2+nx+n+x}\right)^{n+1} \\ &= \frac{n+x}{n} \left(1 - \frac{x}{(n+1)(n+x)}\right)^{n+1} \geq \frac{n+x}{n} \left(1 - \frac{x}{(n+x)}\right) = 1 \end{aligned}$$

(Bernoulli klappt wieder für $|x| < n$, also für hinreichend große n .)

denn

$$\left(1 + \frac{x}{n}\right)^n \geq 1 + x, \quad \text{falls } x > -n, \quad (\text{Bernoulli}) \quad (4.67)$$

d.h. falls n groß genug; $\Rightarrow \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \geq 1 + x$.

4.

$$\exp(x) \leq x \exp(x) + 1 \quad (4.68)$$

bzw. $\exp(x) \cdot (1 - x) \leq 1$. Beweis:

$$\left(1 + \frac{x}{n}\right)^n (1 - x) \stackrel{(*)}{\leq} \left(1 + \frac{x}{n}\right)^n \left(1 - \frac{x}{n}\right)^n = \left(1 - \frac{x^2}{n^2}\right)^n, \quad (4.69)$$

wobei (*): Bernoulli, falls n groß genug.

Für hinreichend große n gilt weiter

$$0 \leq \left(1 - \frac{x^2}{n^2}\right) \leq 1 \quad (4.70)$$

und damit auch die Potenz, und für $n \rightarrow \infty$ folgt die Behauptung.

5.

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n^2}\right)^n = 1 \quad (4.71)$$

Beweis:

$$\begin{aligned} \left(1 + \frac{x}{n^2}\right)^n &= \sum_{\nu=0}^n \binom{n}{\nu} \left(\frac{x}{n^2}\right)^\nu = 1 + \sum_{\nu=1}^n \frac{1}{n^\nu} \binom{n}{\nu} \left(\frac{x}{n}\right)^\nu \\ &\leq 1 + \frac{1}{n} \underbrace{\sum_{\nu=1}^n \binom{n}{\nu} \left(\frac{x}{n}\right)^\nu}_{\text{beschränkt durch } \exp(x)} \xrightarrow{n \rightarrow \infty} 1 \end{aligned} \quad (4.72)$$

Analog gilt $\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n^{1+\epsilon}}\right)^n = 1 \quad \forall \epsilon > 0$.

Satz 5. (Eigenschaften der Exponentialfunktion)

- (i) $\exp(0) = 1, \exp(1) = e$
- (ii) *Funktionalgleichung* $\exp(x) \exp(y) = \exp(x + y)$
- (iii) $\exp(x) > 0 \quad \forall x \in \mathbb{R}$
- (iv) \exp ist streng monoton wachsend, d.h.

$$x > y \quad \Rightarrow \quad \exp(x) > \exp(y) \quad (4.73)$$

- (v) \exp ist stetig $\forall x \in \mathbb{R}$

- (vi) \exp ist diffbar $\forall x \in \mathbb{R}$ mit $\exp' = \exp$
- (vii) $\lim_{x \rightarrow \infty} \exp(x) = \infty$ und $\lim_{x \rightarrow -\infty} \exp(x) = 0$
- (viii) $\forall n \in \mathbb{N}$ gilt $\lim_{x \rightarrow \infty} \frac{\exp(x)}{x^n} = \infty$ und $\lim_{x \rightarrow -\infty} |x|^n \exp(x) = 0$

Beweise:

(i) klar lt. Def.

(ii)

$$\begin{aligned}
 \left(1 + \frac{x}{n}\right)^n \left(1 + \frac{y}{n}\right)^n &\xrightarrow{n \rightarrow \infty} \exp(x) \exp(y) \\
 &= \left(1 + \frac{x+y}{n} + \frac{xy}{n^2}\right)^n = \left(1 + \frac{x+y}{n}\right)^n \underbrace{\left(1 + \frac{xy}{n^2 \left(1 + \frac{x+y}{n}\right)}\right)^n}_{\xrightarrow{n \rightarrow \infty} 1 \text{ (Bem. 5)}} \\
 &\xrightarrow{n \rightarrow \infty} \exp(x+y)
 \end{aligned}
 \tag{4.74}$$

(iii) $x \geq 0$: $e^x > 0$ (lt. Def.) $\Rightarrow e^{-x} = \frac{1}{e^x} > 0$, d.h. $e^x > 0$ auch für $x < 0$

Bemerkung: Aus der Definition von \exp (oder aus Bem. 3) folgt sogar $e^x > 1 \forall x > 0$.

(iv) $x < y$: $e^y = e^{y+x-x} = e^x \underbrace{e^{y-x}}_{>1} > e^x$

(v) z.z.: $\lim_{h \rightarrow 0} e^{x+h} = e^x$

$$\lim_{h \rightarrow 0} e^{x+h} = e^x \lim_{h \rightarrow 0} e^h \tag{4.75}$$

z.z. bleibt Stetigkeit bei Null.

Für $|h| < 1$ gilt lt. Bem. 3 und 4:

$$\underbrace{1+h}_{\xrightarrow{h \rightarrow 0} 1} \leq e^h \leq \underbrace{\frac{1}{1-h}}_{\xrightarrow{h \rightarrow 0} 1}, \tag{4.76}$$

also auch $\lim_{h \rightarrow 0} e^h = 1$.

(vi)

$$\frac{e^{x+h} - e^x}{h} = e^x \frac{e^h - 1}{h} \tag{4.77}$$

d.h. e^x diffbar in $x \Leftrightarrow e^x$ diffbar in Null.

Wieder lt. Bem. 3 und 4 gilt für $h > 0$:

$$\frac{1+h-1}{h} \leq \frac{e^h-1}{h} \leq \frac{he^h+1-1}{h} \quad (4.78)$$

d.h. $1 \leq \frac{e^h-1}{h} \leq e^h \xrightarrow{h \rightarrow 0} 1,$

also $\lim_{h \rightarrow 0^+} \frac{e^h-1}{h} = 1$ (und analog für $h < 0$).

(vii) lt. Bem 3 gilt: $e^x \geq (1+x) \Rightarrow \lim_{x \rightarrow \infty} e^x = \infty$

und auch $\lim_{x \rightarrow -\infty} e^x = \lim_{x \rightarrow \infty} e^{-x} = \lim_{x \rightarrow \infty} \frac{1}{e^x} = 0$

(viii) $\lim_{x \rightarrow \infty} \frac{e^x}{x^n} \stackrel{\text{l'H.}}{=} \lim_{x \rightarrow \infty} \frac{e^x}{nx^{n-1}} \stackrel{\text{l'H.}}{=} \dots = \lim_{x \rightarrow \infty} \frac{e^x}{n!} = \infty$

$\lim_{x \rightarrow -\infty} |x|^n e^x = \lim_{x \rightarrow \infty} x^n e^{-x} = \lim_{x \rightarrow \infty} \frac{x^n}{e^x} = 0$

4.7 Umkehrfunktionen

Definition: $f : A \rightarrow B$ (z.B. $A, B \subseteq \mathbb{R}$) heißt

- surjektiv, falls jedes $b \in B$ als Bild auftritt ($b \in B \Rightarrow \exists a \in A$ mit $f(a) = b$),
- injektiv, falls: $a \neq \tilde{a} \Rightarrow f(a) \neq f(\tilde{a}) \forall a, \tilde{a} \in A$,
- bijektiv, falls f surjektiv und injektiv ist.

Bemerkungen:

1. $f : A \rightarrow B$ nicht surjektiv – "heilbar", denn

$$\tilde{B} := f(A) = \{b \in B \mid \exists a \in A, \text{ so dass } f(a) = b\} \quad (4.79)$$

$\Rightarrow f : A \rightarrow \tilde{B}$ ist surjektiv. Z.B.

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto x^2 \end{aligned} \quad (4.80)$$

ist nicht surjektiv, denn $-1 \in B = \mathbb{R}$ tritt nicht als Funktionswert auf.
 $f(A) = \mathbb{R}_0^+$, und damit ist

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R}_0^+ \\ x &\mapsto x^2 \end{aligned} \quad (4.81)$$

surjektiv.

2. fehlende Injektivität ist kritischer, z.B.

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R}_0^+ \\ x &\mapsto x^2 \end{aligned} \quad (4.82)$$

ist nicht injektiv, denn $(-1)^2 = 1^2$ aber $-1 \neq 1$.

$$f_1 : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+ , \quad f_2 : \mathbb{R}_0^- \rightarrow \mathbb{R}_0^+ \quad (4.83)$$

$$x \mapsto x^2 \quad x \mapsto x^2 \quad (4.84)$$

sind injektiv (und bijektiv).

3. Ist $f : A \rightarrow B$ bijektiv, so existiert in natürlicher Weise eine Funktion

$$\begin{aligned} f^{-1} : B &\rightarrow A \\ b = f(a) &\mapsto a \end{aligned} \quad (4.85)$$

genannt Umkehrfunktion von f . $\left(f^{-1} \neq \frac{1}{f}\right)$

4. $f : A \rightarrow B$ bijektiv

\Rightarrow es existiert $f^{-1} : B \rightarrow A$

d.h. $f^{-1} \circ f : A \rightarrow A$ sowie $f \circ f^{-1} : B \rightarrow B$

mit $(f^{-1} \circ f)(x) = f^{-1}(f(x)) = x \quad \forall x \in A$

und $(f \circ f^{-1})(x) = f(f^{-1}(x)) = x \quad \forall x \in B$

5. Der Graph von f^{-1} ist der Graph von f gespiegelt an der ersten Winkelhalbierenden.

6. Streng monotone Funktionen sind injektiv.

7. Sei $f : A \rightarrow B$ diffbar, dann gilt

$$\begin{aligned} f'(x) > 0 \quad \forall x \in I \subseteq A &\Rightarrow f \text{ ist streng monoton wachsend auf } I \\ f'(x) < 0 \quad \forall x \in I \subseteq A &\Rightarrow f \text{ ist streng monoton fallend auf } I \end{aligned} \quad (4.86)$$

Beispiele:

1.

$$\begin{aligned} f_1 : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+ , & \quad f_1^{-1} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+ \\ x \mapsto x^2 & \quad x \mapsto \sqrt{x} \end{aligned} \quad (4.87)$$

$$\begin{aligned} f_2 : \mathbb{R}_0^- \rightarrow \mathbb{R}_0^+ , & \quad f_2^{-1} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^- \\ x \mapsto x^2 & \quad x \mapsto -\sqrt{x} \end{aligned} \quad (4.88)$$

$$\begin{aligned} f_1^{-1} \circ f_2 : \mathbb{R}_0^- \rightarrow \mathbb{R}_0^- , & \quad f_2 \circ f_1^{-1} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+ \end{aligned} \quad (4.89)$$

2. $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = x \quad \Rightarrow \quad f^{-1} = f$

3. $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2 + x$

$$\begin{aligned} f'(x) = 2x + 1 : & \quad f'(x) > 0 \quad \forall \quad x > -\frac{1}{2} \\ & \quad f'(x) < 0 \quad \forall \quad x < -\frac{1}{2} \end{aligned} \quad (4.90)$$

$$f\left(-\frac{1}{2}\right) = -\frac{1}{4}, \quad \lim_{x \rightarrow \pm\infty} = \infty, \quad f \text{ stetig} \quad (4.91)$$

\Rightarrow jedes $y \in [-\frac{1}{4}, \infty)$ tritt als Funktionswert auf; d.h.

$$f_1 : [-\frac{1}{2}, \infty) \rightarrow [-\frac{1}{4}, \infty) \quad \text{und} \quad f_2 : (-\infty, -\frac{1}{2}] \rightarrow [-\frac{1}{4}, \infty), \quad (4.92)$$

beide mit derselben Abbildungsvorschrift wie f , sind surjektiv und (da monoton) auch injektiv, d.h. bijektiv.

$$\Rightarrow \quad f_1^{-1} : [-\frac{1}{4}, \infty) \rightarrow [-\frac{1}{2}, \infty) \quad \text{und} \quad f_2^{-1} : [-\frac{1}{4}, \infty) \rightarrow (-\infty, -\frac{1}{2}] \quad (4.93)$$

Wie sehen die f_j^{-1} aus?

Nenne $y = f(x)$, und da $(f^{-1} \circ f)(x) = x$ gilt $f^{-1}(y) = x$.

Löse also $y = f(x)$ nach x auf!

$$y = x^2 + x \quad \Leftrightarrow \quad x^2 + x - y = 0 \quad (4.94)$$

$$\text{quadratische Gleichung: } x = \frac{-1 \pm \sqrt{1+4y}}{2} = -\frac{1}{2} \pm \sqrt{y + \frac{1}{4}}$$

$$\begin{aligned} \Rightarrow \quad f_1^{-1}(x) &= -\frac{1}{2} + \sqrt{x + \frac{1}{4}} \\ f_2^{-1}(x) &= -\frac{1}{2} - \sqrt{x + \frac{1}{4}} \end{aligned} \quad (4.95)$$

Satz 6. (Ableitung der Umkehrfunktion)

Ist $f : I \rightarrow J$ bijektiv und diffbar, dann folgt

$f^{-1} : J \rightarrow I$ ist diffbar $\forall x \in J$ mit $f'(f^{-1}(x)) \neq 0$ und es gilt

$$f^{-1}'(x) = \frac{1}{f'(f^{-1}(x))}. \quad (4.96)$$

Beweis:

Man zeigt zunächst, dass f^{-1} diffbar ist (ohne Beweis).

Dann wenden wir auf $f(f^{-1}(x)) = x$ die Kettenregel an,

$$f'(f^{-1}(x)) \cdot f^{-1}'(x) = 1. \quad (4.97)$$

Mit Division durch $f'(f^{-1}(x))$ (das war $\neq 0$) folgt die Behauptung. \square

4.8 Der Logarithmus

Definition: Die Umkehrfunktion von $\exp : \mathbb{R} \rightarrow \mathbb{R}^+$ heißt (natürlicher) Logarithmus,

$$\log : \mathbb{R}^+ \rightarrow \mathbb{R} \quad (4.98)$$

Manchmal schreibt man statt \log auch \ln (*logarithmus naturalis*).

also: $e^{\log x} = x$ für $x > 0$ und $\log(e^x) = x$ für $x \in \mathbb{R}$.

Bemerkung:

e^x ist injektiv, da streng monoton wachsend

$$\lim_{x \rightarrow -\infty} e^x = 0, \quad \lim_{x \rightarrow \infty} e^x = \infty$$

$$\Rightarrow \exp(\mathbb{R}) = (0, \infty)$$

Satz 7. (Eigenschaften von \log)

- (i) $\log 1 = 0, \quad \log e = 1$
- (ii) $\log(xy) = \log x + \log y$
- (iii) $\lim_{x \rightarrow 0^+} \log x = -\infty, \quad \lim_{x \rightarrow \infty} \log x = \infty$
- (iv) $(\log x)' = \frac{1}{x}, \quad x > 0$
 $\Rightarrow \log x$ ist streng monoton wachsend

Beweis:

- (i) $e^0 = 1 \Rightarrow \log 1 = 0$
 $e^1 = e \Rightarrow \log e = 1$
- (ii) $e^{\log x + \log y} = e^{\log x} e^{\log y} = xy = e^{\log(xy)}$
Da e^x injektiv folgt Behauptung.
- (iii) $\lim_{x \rightarrow -\infty} e^x = 0 \Leftrightarrow \lim_{x \rightarrow 0^+} \log x = -\infty$
 $\lim_{x \rightarrow \infty} e^x = \infty \Leftrightarrow \lim_{x \rightarrow \infty} \log x = \infty$
- (iv) e^x diffbar und $(e^x)' > 0$, also insbesondere $\neq 0$
 $\xRightarrow{\text{Satz 6}}$ $\log x$ ist diffbar, mit

$$(\log x)' = \frac{1}{\exp'(\log(x))} = \frac{1}{\exp(\log(x))} = \frac{1}{x}. \quad (4.99)$$

4.9 Weitere elementare Funktionen

Allgemeine Exponentialfunktion

Für $a > 0$, definieren wir

$$\begin{aligned}\mathbb{R} &\rightarrow \mathbb{R}^+ \\ x &\mapsto a^x := e^{x \log a}.\end{aligned}\tag{4.100}$$

Es gilt dann

$$\log(a^x) = \log(e^{x \log a}) = x \log a.\tag{4.101}$$

Umkehrfunktion: **Logarithmus zur Basis a**

Für $a > 0$, $a \neq 1$ definieren wir

$$\begin{aligned}\log_a : \mathbb{R}^+ &\rightarrow \mathbb{R} \\ x &\mapsto \log_a x := \frac{\log x}{\log a}.\end{aligned}\tag{4.102}$$

Denn, z.B.,

$$\log_a(a^x) = \frac{\log(a^x)}{\log a} = \frac{x \log a}{\log a} = x.\tag{4.103}$$

Also: $\log_e = \log = \ln$.

Auch gebräuchlich: $\text{lb} := \log_2$, $\text{ld} := \log_2$, $\text{lg} := \log_{10}$.

Allgemeine Potenz Für $\alpha \in \mathbb{R}$ definieren wir

$$\begin{aligned}\mathbb{R}^+ &\rightarrow \mathbb{R}^+ \\ x &\mapsto x^\alpha := e^{\alpha \log x}.\end{aligned}\tag{4.104}$$

Umkehrfunktion (falls $\alpha \neq 0$):

$$x \mapsto x^{\frac{1}{\alpha}}\tag{4.105}$$

Hyperbolische Funktionen. Siehe Übungen.

Eigenschaften und Rechenregeln. $x \in \mathbb{R}^+$, $\alpha, \beta \in \mathbb{R}$

1. $(x^\alpha)^\beta = x^{\alpha\beta}$
2. $x^{\alpha+\beta} = x^\alpha x^\beta$
3. $\log(x^\alpha) = \alpha \log x$

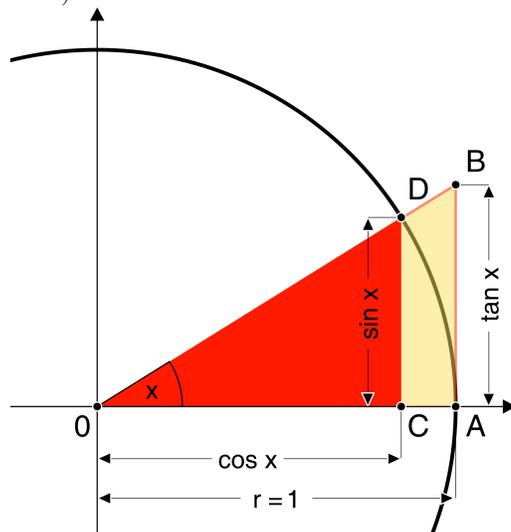
Bemerkungen zur Notation:

$$x^{\alpha\beta} = x^{(\alpha\beta)}, \quad \log x^\alpha = \log(x^\alpha)\tag{4.106}$$

4.10 Trigonometrische Funktionen

Einheitskreis (Kreis mit Radius 1)

Umfang 2π (Definition von π)



definiert $\sin x$ und $\cos x$ für $0 \leq x < 2\pi$
 2π -periodisch fortgesetzt, d.h.

$$\sin(x + 2\pi n) = \sin(x) \quad \forall n \in \mathbb{Z}. \quad (4.107)$$

(ebenso für \cos)

$$\tan x := \frac{\sin x}{\cos x} \quad \text{und} \quad \cot x = \frac{\cos x}{\sin x}, \quad (4.108)$$

wo definiert.

Spezielle Werte:

$$\begin{aligned} \sin 0 &= 0, & \cos 0 &= 1 \\ \sin \frac{\pi}{2} &= 1, & \cos \frac{\pi}{2} &= 0 \end{aligned} \quad (4.109)$$

Translationen:

$$\sin(x + \pi) = -\sin x \quad \cos(x + \pi) = -\cos x \quad (4.110)$$

$$\sin(x + \frac{\pi}{2}) = \cos x \quad \cos(x + \frac{\pi}{2}) = -\sin x \quad (4.111)$$

Weitere wichtige Beziehungen:

1. $\sin^2 x + \cos^2 x = 1 \quad \forall x$ (Pythagoras)
 und damit auch $|\sin x| \leq 1, |\cos x| \leq 1 \quad \forall x \in \mathbb{R}$
2. $\sin(-x) = -\sin x, \quad \cos(-x) = \cos x$

3. Für $0 < x < \frac{\pi}{2}$ gilt (siehe Abbildung):

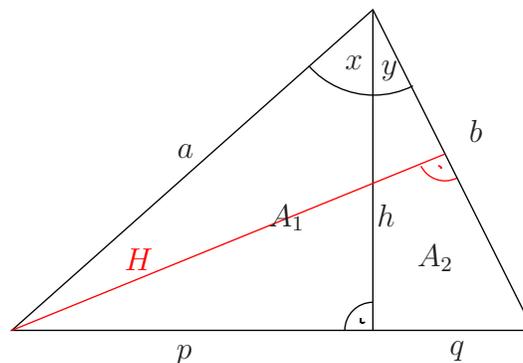
$$\sin x < x < \tan x .$$

Additionstheoreme:

$$\begin{aligned} \sin(x + y) &= \sin x \cos y + \cos x \sin y \\ \cos(x + y) &= \cos x \cos y - \sin x \sin y \end{aligned} \quad (4.112)$$

Beweis:

Laut Translationseigenschaften genügt es, die Additionstheoreme für $x, y > 0, x + y < \frac{\pi}{2}$ zu beweisen.



$$\begin{aligned} h &= a \cos x, & h &= b \cos y \\ p &= a \sin x, & q &= b \sin y \\ A_1 &= \frac{ph}{2} = \frac{ab \sin x \cos y}{2}, & A_2 &= \frac{qh}{2} = \frac{ab \cos x \sin y}{2} \end{aligned} \quad (*)$$

$$A_1 + A_2 = \frac{Hb}{2} = \frac{ab \sin(x + y)}{2} \quad (+)$$

(*) und (+) \Rightarrow Behauptung

$$\begin{aligned} \cos(x + y) &= \sin(x + y + \frac{\pi}{2}) \\ &= \sin x \cos(y + \frac{\pi}{2}) + \cos x \sin(y + \frac{\pi}{2}) \\ &= \sin x (-\sin y) + \cos x \cos y \end{aligned} \quad (4.113)$$

□

Ableitungen:

$$(\sin x)' = \cos x \quad \text{und} \quad (\cos x)' = -\sin x \quad (4.114)$$

Beweis: (für $\sin x$; $\cos x$ folgt aus $\sin(x + \frac{\pi}{2}) = \cos x$)

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} &= \lim_{h \rightarrow 0} \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \sin x \lim_{h \rightarrow 0} \frac{\cos h - 1}{h} + \cos x \lim_{h \rightarrow 0} \frac{\sin h}{h} \end{aligned} \quad (4.115)$$

reduziert auf Ableitungen von \sin und \cos bei Null.

Aus Beziehung 3 folgt für $0 < x < \frac{\pi}{2}$ (Kehrwerte):

$$\begin{aligned} \frac{1}{\sin x} &> \frac{1}{x} > \frac{1}{\tan x} \\ \Rightarrow 1 &> \frac{\sin x}{x} > \cos x \xrightarrow{x \rightarrow 0} 1 \end{aligned} \quad (4.116)$$

und damit

$$\lim_{h \rightarrow 0^+} \frac{\sin h}{h} = 1 \quad (\text{analog für } h \rightarrow 0^-) \quad (4.117)$$

Der andere Limes wird Null:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\cos h - 1}{h} &= \lim_{h \rightarrow 0} \frac{\sqrt{1 - \sin^2 h} - 1}{h} \cdot \frac{\sqrt{1 - \sin^2 h} + 1}{\sqrt{1 - \sin^2 h} + 1} \\ &= \lim_{h \rightarrow 0} \frac{(1 - \sin^2 h) - 1}{h(\sqrt{1 - \sin^2 h} + 1)} \\ &= - \left(\lim_{h \rightarrow 0} \frac{\sin h}{h} \right) \cdot \left(\lim_{h \rightarrow 0} \frac{\sin h}{\sqrt{1 - \sin^2 h} + 1} \right) \\ &= -1 \cdot \frac{0}{2} = 0 \end{aligned} \quad (4.118)$$

□

Die Umkehrfunktion des *Sinus* heißt *Arcus Sinus*, z.B.

$$\arcsin : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad (\text{Hauptzweig}) \quad (4.119)$$

... anderer Zweig: $[-1, 1] \rightarrow \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]$

... viele Zweige: $[-1, 1] \rightarrow \left[\frac{2n-1}{2}\pi, \frac{2n+1}{2}\pi\right], n \in \mathbb{Z}$

Hauptzweige der anderen trigonometrischen Funktionen:

$$\begin{aligned} \arccos &: [-1, 1] \rightarrow [0, \pi] \\ \arctan &: \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \\ \operatorname{arccot} &: \mathbb{R} \rightarrow (0, \pi) \end{aligned} \quad (4.120)$$

Ableitungen (Hauptzweige): (mit Satz 6)

$$\begin{aligned}
 (\arcsin x)' &= \frac{1}{\cos(\arcsin(x))} = \frac{1}{\sqrt{1-x^2}} \\
 (\arccos x)' &= \frac{1}{-\sin(\arccos(x))} = -\frac{1}{\sqrt{1-x^2}} \\
 (\tan x)' &= \left(\frac{\sin x}{\cos x}\right)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x} \\
 &= 1 + \tan^2 x \\
 (\arctan x)' &= \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}
 \end{aligned} \tag{4.121}$$

4.11 Potenzreihen

Beispiel: Die endliche geometrische Reihe war ($x \neq 1$)

$$\sum_{\nu=0}^n x^\nu = \frac{1-x^{n+1}}{1-x}. \tag{4.122}$$

Für $|x| < 1$ gilt $\lim_{n \rightarrow \infty} x^n = 0$, und damit

$$\sum_{\nu=0}^{\infty} x^\nu = \frac{1}{1-x}, \tag{4.123}$$

wobei $\sum_{\nu=0}^{\infty} a_\nu := \lim_{n \rightarrow \infty} \sum_{\nu=0}^n a_\nu$, d.h. wir untersuchen das Verhalten

der Folge der Partialsummen, $s_n := \sum_{\nu=0}^n a_\nu$ für $n \rightarrow \infty$.

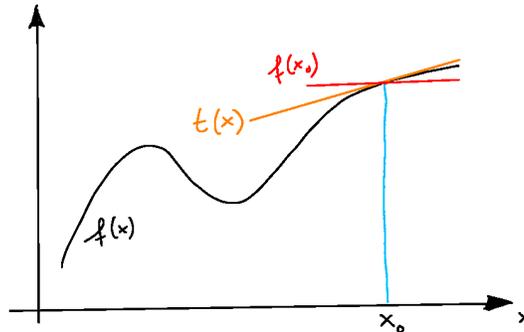
$$\text{Reihe: } \sum_{n=m}^{\infty} a_n \quad (\text{Summe mit Grenze } \infty)$$

$$\text{Potenzreihe in } x: \sum_{n=m}^{\infty} a_n x^n \tag{4.124}$$

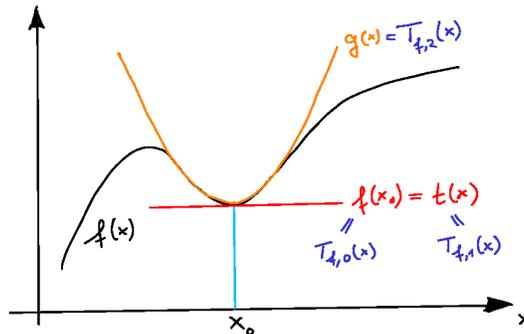
$$\text{oder auch } \sum_{n=m}^{\infty} a_n (x-x_0)^n \text{ mit festem } x_0 \in \mathbb{R}$$

Allgemeine Idee: Suche Approximation für Funktion f in der Nähe von x_0 (so oft diffbar wie nötig):

- f stetig in $x_0 \Rightarrow$ kann dort nicht springen – einfachste Näherung: $f(x) \approx f(x_0)$ für x in der Nähe von x_0



- f diffbar \Rightarrow Tangente an der Stelle x_0 existiert – etwas bessere Näherung: $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$ für x in der Nähe von x_0



- Näherung, $g(x)$, soll auch Änderung der Steigung (f'') berücksichtigen – Ansatz:

$$g(x) = f(x_0) + f'(x_0)(x - x_0) + c(x - x_0)^2 \quad (4.125)$$

Offensichtlich gilt bereits $f(x_0) = g(x_0)$ und $f'(x_0) = g'(x_0)$. Bestimme c aus Forderung

$$f''(x_0) \stackrel{!}{=} g''(x_0) \quad (4.126)$$

$$g''(x) = 2c \quad \Rightarrow \quad c = \frac{f''(x_0)}{2}. \quad (4.127)$$

Höhere Näherungen: Ansatz,

$$T_{f,n}(x) = \sum_{\nu=0}^n c_{\nu}(x - x_0)^{\nu} \quad (4.128)$$

und bestimme die c_{ν} aus Forderungen

$$f^{(\mu)}(x_0) \stackrel{!}{=} T_{f,n}^{(\mu)}(x_0), \quad \mu = 0, \dots, n \quad (4.129)$$

$$T_{f,n}^{(\mu)}(x) = \sum_{\nu=\mu}^n \nu(\nu-1) \cdots (\nu-\mu+1) c_{\nu} (x - x_0)^{\nu-\mu} \quad (4.130)$$

$$\Rightarrow T_{f,n}^{(\mu)}(x_0) = \mu! c_{\mu}$$

d.h. wir erhalten

$$c_\nu = \frac{f^{(\nu)}(x_0)}{\nu!} \quad (4.131)$$

Wir nennen

$$T_{f,n}(x) = \sum_{\nu=0}^n \frac{f^{(\nu)}(x_0)}{\nu!} (x - x_0)^\nu \quad (4.132)$$

das n -te Taylor-Polynom von f und

$$T_f(x) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(x_0)}{\nu!} (x - x_0)^\nu \quad (4.133)$$

die (formale) Taylor-Reihe von f .

Fragen:

- Ist $T_{f,n}(x)$ eine gute Näherung für $f(x)$ in der Nähe von x_0 ?
- Konvergiert $T_f(x)$? (Und für welche x , und gegen welchen Wert?)

Vermutung: Nächster Term der Entwicklung als Maß für den Fehler ...?

Satz 8. (Satz von Taylor)

Sei $f : I \rightarrow \mathbb{R}$, $I = (a, b) \subseteq \mathbb{R}$, $n + 1$ mal diffbar. Dann gilt für $x, x_0 \in I$

$$f(x) = \sum_{\nu=0}^n \frac{f^{(\nu)}(x_0)}{\nu!} (x - x_0)^\nu + R_n(x), \quad (4.134)$$

und für $x > x_0 \exists \xi \in (x_0, x)$ (für $x < x_0 \exists \xi \in (x, x_0)$) mit

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}. \quad (4.135)$$

R_n heißt (Lagrangesches) Restglied, x_0 heißt Entwicklungspunkt.

(momentan ohne Beweis)

Bemerkungen:

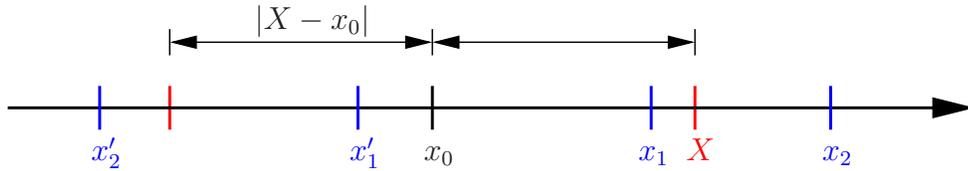
1. Insbesondere gilt $f(x) = T_{f,n}(x) + o((x - x_0)^n)$.
2. Es gibt auch andere Restgliedformeln (z.B. später Integralrestglied, dann mit Beweis).
3. $T_f(x)$ konvergiert also genau dann gegen $f(x)$, wenn

$$\lim_{n \rightarrow \infty} R_n(x) = 0 \quad \text{bzw. wenn} \quad \lim_{n \rightarrow \infty} |R_n(x)| = 0. \quad (4.136)$$

4. Falls $\lim_{n \rightarrow \infty} |R_n(x)| > 0$ (oder gar nicht existent), dann divergiert $T_f(x)$ (meistens) oder konvergiert gegen eine andere Funktion (selten).

5. Sei $|x_1 - x_0| < |X - x_0| < |x_2 - x_0|$, dann gilt:

$$\begin{aligned} \text{aus } \lim_{n \rightarrow \infty} |R_n(X)| = 0 &\Rightarrow \lim_{n \rightarrow \infty} |R_n(x_1)| = 0 \\ \text{aus } \lim_{n \rightarrow \infty} |R_n(X)| > 0 &\Rightarrow \lim_{n \rightarrow \infty} |R_n(x_2)| > 0 \end{aligned} \quad (4.137)$$



d.h. wir finden einen Radius $R \geq 0$ mit

$$\begin{aligned} T_f(x) &= f(x) \quad \forall |x - x_0| < R \\ T_{f,n}(x) &\not\rightarrow f(x) \quad \forall |x - x_0| > R \end{aligned} \quad (4.138)$$

6. Sei f in x_1 singulär (Pol, ...), dann gilt $R \leq |x_0 - x_1|$.

Beispiele:

1. $|x| < 1$:

$$\frac{1}{1-x} = \sum_{\nu=0}^{\infty} x^\nu \quad (\text{s.o.}) \quad (4.139)$$

Für $|x| > 1$ divergiert die Reihe (Summanden wachsen an).

2. $f(x) = e^x \Rightarrow f^{(\nu)}(x) = e^x$

$$T_f(x) = \sum_{\nu=0}^{\infty} \frac{f^{(\nu)}(0)}{\nu!} x^\nu = \sum_{\nu=0}^{\infty} \frac{x^\nu}{\nu!} \quad (4.140)$$

$|x| < K \in \mathbb{R}^+$:

$$\begin{aligned} |R_n(x)| &= \left| \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1} \right| = \left| e^\xi \frac{x \cdot x \cdots x}{1 \cdot 2 \cdots (n+1)} \right| \\ &\stackrel{K < N \leq n}{\leq} e^K \left| \frac{x \cdot x \cdots x}{1 \cdot 2 \cdots (N-1)} \frac{x \cdot x \cdots x}{N \cdot (N+1) \cdots (n+1)} \right| \\ &\leq \underbrace{e^K \left| \frac{x^{N-1}}{(N-1)!} \right|}_{\text{hängt nicht von } n \text{ ab}} \underbrace{\left(\frac{K}{N} \right)^{n-N+2}}_{\xrightarrow{n \rightarrow \infty} 0, \text{ da } K/N < 1} \xrightarrow{n \rightarrow \infty} 0 \end{aligned} \quad (4.141)$$

Also

$$e^x = \sum_{\nu=0}^{\infty} \frac{x^\nu}{\nu!} \quad \forall x \in \mathbb{R}. \quad (4.142)$$

3. $f(x) = \log(1+x)$, $|x| < 1$

$$f'(x) = \frac{1}{1+x} = \sum_{\nu=0}^{\infty} (-x)^{\nu} = \sum_{\nu=0}^{\infty} (-1)^{\nu} x^{\nu} \quad (4.143)$$

Da $\sum_{\nu=0}^{\infty} (-1)^{\nu} \frac{x^{\nu+1}}{\nu+1} + c$ gliedweise differenziert $f'(x)$ ergibt, vermuten wir

$$\log(1+x) = c - \sum_{\nu=1}^{\infty} (-1)^{\nu} \frac{x^{\nu}}{\nu}, \quad (4.144)$$

und folgern aus $f(0) = \log 1 = 0$, dass $c = 0$.

Bemerkung: Falls Reihe hübsch, dürfen wir gliedweise differenzieren (und integrieren); eigentlich brauchen wir dafür gleichmäßige Konvergenz (machen wir aber nicht). Für Taylorreihen und x innerhalb des Konvergenzradius' geht alles gut.

4. **Binomische Reihe:** $f(x) = (1+x)^{\alpha}$, $\alpha \in \mathbb{R} \setminus \{0\}$

$$f^{(\nu)}(x) = \alpha(\alpha-1)\cdots(\alpha-\nu+1)(1+x)^{\alpha-\nu} \quad \Rightarrow \quad \frac{f^{(\nu)}(0)}{\nu!} = \binom{\alpha}{\nu} \quad (4.145)$$

also

$$(1+x)^{\alpha} = \sum_{\nu=0}^{\infty} \binom{\alpha}{\nu} x^{\nu} \quad \forall |x| < 1 \quad (4.146)$$

Konvergenzradius ohne Beweis

Spezialfälle:

- $\alpha = n \in \mathbb{N}$: binomischer Satz (sogar für alle $x \in \mathbb{R}$)
- $\alpha = -1$: geometrische Reihe,

$$\binom{-1}{\nu} = \frac{(-1)(-2)\cdots(-1-\nu+1)}{\nu!} = (-1)^{\nu} \quad (4.147)$$

- $\alpha = \frac{1}{2}$:

$$\sqrt{1+x} = \sum_{\nu=0}^{\infty} \binom{1/2}{\nu} x^{\nu}, \quad |x| < 1 \quad (4.148)$$

5. $\sin^{(2\nu)}(x) = (-1)^{\nu} \sin x$, $\sin^{(2\nu+1)}(x) = (-1)^{\nu} \cos x$, also

$$\sin x = \sum_{\nu=0}^{\infty} \frac{(-1)^{\nu}}{(2\nu+1)!} x^{2\nu+1} \quad \forall x \in \mathbb{R} \quad (4.149)$$

analog:

$\cos^{(2\nu)}(x) = (-1)^{\nu} \cos x$, $\cos^{(2\nu+1)}(x) = (-1)^{\nu+1} \sin x$, also

$$\cos x = \sum_{\nu=0}^{\infty} \frac{(-1)^{\nu}}{(2\nu)!} x^{2\nu} \quad \forall x \in \mathbb{R} \quad (4.150)$$

6.

$$(\arctan x)' = \frac{1}{1+x^2} = \sum_{\nu=0}^{\infty} (-1)^\nu x^{2\nu}, \quad |x| < 1 \quad (4.151)$$

folgere

$$\arctan x = \sum_{\nu=0}^{\infty} (-1)^\nu \frac{x^{2\nu+1}}{2\nu+1} + c \quad (4.152)$$

mit $c = 0$ da $\arctan 0 = 0$.

7.

$$\frac{\sin x}{x} = \frac{1}{x} \sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{(2\nu+1)!} x^{2\nu+1} = \sum_{\nu=0}^{\infty} \frac{(-1)^\nu}{(2\nu+1)!} x^{2\nu} = 1 - \frac{x^2}{6} + o(x^2) \quad (4.153)$$

8. Anwendung: Grenzwerte berechnen

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{x^2 \sin^4 x}{(\cos x - 1)^3} &= \lim_{x \rightarrow 0} \frac{x^2 (x + o(x))^4}{(1 - \frac{x^2}{2} + o(x^2) - 1)^3} = \lim_{x \rightarrow 0} \frac{x^2 (x^4 + o(x^4))}{-\frac{x^6}{8} + o(x^6)} \\ &= \lim_{x \rightarrow 0} \frac{x^6 + o(x^6)}{-\frac{x^6}{8} + o(x^6)} = -8 \end{aligned} \quad (4.154)$$

... besser als sechsmal l'Hospital!

9. $\frac{1}{1-x}$ um $x_0 = 3$

$$\begin{aligned} \frac{1}{1-x} &= \frac{1}{1 - ((x-3) + 3)} = \frac{1}{-2 - (x-3)} = -\frac{1}{2} \cdot \frac{1}{1 + \frac{x-3}{2}} \\ &= -\frac{1}{2} \sum_{n=0}^{\infty} \left(-\frac{x-3}{2}\right)^n = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{2^{n+1}} (x-3)^n \end{aligned} \quad (4.155)$$

für $|\frac{x-3}{2}| < 1 \Leftrightarrow |x-3| < 2$

10.

$$\begin{aligned} \frac{e^x}{1-x} &= \left(\sum_{\nu=0}^{\infty} \frac{x^\nu}{\nu!} \right) \left(\sum_{\mu=0}^{\infty} x^\mu \right), \quad \text{für } |x| < 1 \text{ (wg. geom. Reihe)} \\ &= \left(1 + x + \frac{x^2}{2} + o(x^2) \right) (1 + x + x^2 + o(x^2)) \\ &= 1 + 2x + \frac{5}{2}x^2 + o(x^2), \quad x \rightarrow 0 \end{aligned} \quad (4.156)$$

allgemein (**Cauchy-Produkt**):

$$\begin{aligned} \left(\sum_{\nu=0}^{\infty} a_{\nu} x^{\nu} \right) \left(\sum_{\mu=0}^{\infty} b_{\mu} x^{\mu} \right) &= \sum_{\nu=0}^{\infty} \sum_{\mu=0}^{\infty} a_{\nu} b_{\mu} x^{\nu+\mu} & \left. \begin{array}{l} n := \nu + \mu \\ n: 0 \dots \infty \\ \nu: 0 \dots n \\ \mu = n - \nu \end{array} \right\} & (4.157) \\ &= \sum_{n=0}^{\infty} \sum_{\nu=0}^n a_{\nu} b_{n-\nu} x^n \end{aligned}$$

Erlaubt, falls alles hübsch (beide Reihen absolut konvergent, d.h. z.B. $\sum_{\nu=0}^{\infty} |a_{\nu} x^{\nu}| < \infty$). Damit nochmal

$$\begin{aligned} \frac{e^x}{1-x} &= \left(\sum_{\nu=0}^{\infty} \frac{x^{\nu}}{\nu!} \right) \left(\sum_{\mu=0}^{\infty} x^{\mu} \right), \quad |x| < 1 \\ &= \sum_{n=0}^{\infty} \sum_{\nu=0}^n \frac{1}{\nu!} x^n \\ &= \frac{1}{0!} + \left(\frac{1}{0!} + \frac{1}{1!} \right) x + \left(\frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} \right) x^2 + o(x^2) \end{aligned} \quad (4.158)$$

11. Weitere Anwendung: Lokale Extrema

Definition: x_0 heißt lokale Maximal- bzw. Minimalstelle von f , wenn gilt:

$$f(x) < f(x_0) \text{ bzw. } f(x) > f(x_0) \quad \forall x \neq x_0 \text{ in einer Umgebung von } x_0.$$

x_0 heißt Extremalstelle von f , falls x_0 lokale Maximal- oder Minimalstelle ist.

Sei $f : (a, b) \rightarrow \mathbb{R}$ zweimal stetig diffbar. Es gilt

- (i) x_0 ist Extremalstelle $\Rightarrow f'(x_0) = 0$. (notwendige Bedingung)
- (ii) $f'(x_0) = 0$ und $f''(x_0) < 0 \Rightarrow x_0$ ist lokale Maximalstelle
- $f'(x_0) = 0$ und $f''(x_0) > 0 \Rightarrow x_0$ ist lokale Minimalstelle
- (hinreichenden Bedingung)

Analog: Sei f n mal stetig diffbar, n gerade, mit

$$f^{(\nu)}(x_0) = 0 \quad \forall \nu = 1, \dots, n-1 \text{ und } f^{(n)}(x_0) > 0 \text{ (bzw. } f^{(n)}(x_0) < 0),$$

so ist x_0 ist lokale Minimalstelle (bzw. Maximalstelle).

Warum? Taylor!

$$f(x) = f(x_0) + \underbrace{f'(x_0)(x-x_0) + \dots + \frac{f^{(n-1)}(x_0)}{(n-1)!}(x-x_0)^{n-1}}_{=0} + \frac{f^{(n)}(\xi)}{n!} \underbrace{(x-x_0)^n}_{>0} \quad (4.159)$$

ξ liegt zwischen x_0 und x , und

wenn $f^{(n)}(x_0) < 0$ dann ist auch $f^{(n)}(x) < 0 \quad \forall x$ in einer Umgebung von x_0 ,

d.h. das VZ von $f^{(n)}(\xi)$ wird durch das VZ von $f^{(n)}(x_0)$ festgelegt. \square

Oft ist es allerdings besser, gleich die Taylor-Reihe der fraglichen Funktion zu betrachten, vgl. erstes Beispiel im nächsten Abschnitt...

4.12 Kurvendiskussion

- Definitionsbereich
- Verhalten am Rand des Definitionsbereichs
(Asymptoten, Definitionslücken, stetige Fortsetzbarkeit, ...)
- Nullstellen
- Extrema
- Skizze

Beispiele:

1. $f(x) = \frac{1 - x^4}{1 + x^4}$ definiert $\forall x \in \mathbb{R}$

Spiegelsymmetrie (Spiegelung an y-Achse), denn $f(-x) = f(x)$

Verhalten am "Rand" des Definitionsbereichs: $\lim_{x \rightarrow \pm\infty} f(x) = -1$

Nullstellen: $f(x) = 0 \Leftrightarrow x^4 = 1 \Leftrightarrow x = \pm 1$

Extrempunkte:

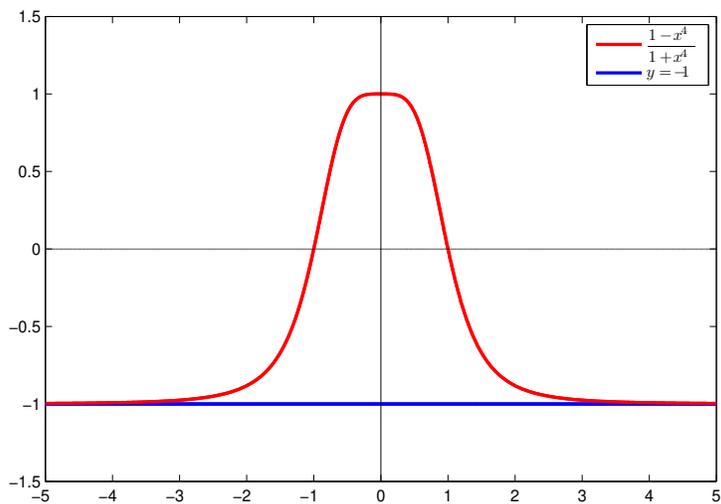
$$\begin{aligned} f'(x) &= \frac{-4x^3(1+x^4) - (1-x^4)4x^3}{(1+x^4)^2} \\ &= \frac{-8x^3}{(1+x^4)^2} \stackrel{!}{=} 0 \Leftrightarrow x = 0 \end{aligned} \tag{4.160}$$

z.B. Taylorentwicklung:

$$f(x) \underset{\text{geom.R.}}{=} (1-x^4)\left(1-x^4 + \underbrace{x^8 - x^{12} + \dots}_{=o(x^4)}\right) = 1 - 2x^4 + o(x^4) \tag{4.161}$$

also Hochpunkt $(0, 1)$.

Skizze:



$$2. f(x) = \frac{x^3 + 1}{x^2 - 1}$$

Definitionsbereich: $\mathbb{R} \setminus \{-1, 1\}$

Zähler bei $x = -1$ auch Null:

$$\left(\begin{array}{r} x^3 \\ -x^3 - x^2 \\ \hline -x^2 \\ x^2 + x \\ \hline x + 1 \\ -x - 1 \\ \hline 0 \end{array} + 1 \right) : (x + 1) = x^2 - x + 1 \quad (4.162)$$

stetig fortsetzbar durch $f(-1) := -\frac{3}{2}$ und damit

$$f(x) = \frac{x^2 - x + 1}{x - 1} \quad (4.163)$$

Pol (erster Ordnung, also mit Vorzeichenwechsel) bei $x = 1$,

d.h. $f(x) \sim \frac{1}{x-1} \rightarrow \pm\infty$ für $x \rightarrow 1\pm$

Verhalten für große $|x|$:

entweder durch Polynomdivision...

$$\left(\begin{array}{r} x^2 - x + 1 \\ -x^2 + x \\ \hline \end{array} \right) : (x - 1) = x + \frac{1}{x - 1} \quad (4.164)$$

...oder durch Taylorentwicklung

$$\begin{aligned} f(x) &= \frac{x - 1 + \frac{1}{x}}{1 - \frac{1}{x}} \\ &\stackrel{|x|>1}{=} \left(x - 1 + \frac{1}{x}\right) \left(1 + \frac{1}{x} + \frac{1}{x^2} + \dots\right) \\ &= x + (1 - 1) + \frac{1 - 1 + 1}{x} + \dots \\ &= x + \frac{1}{x} + \dots \end{aligned} \tag{4.165}$$

also $f(x) \sim x \rightarrow \pm\infty$ für $x \rightarrow \pm\infty$

Nullstellen:

$$f(x) = 0 \Leftrightarrow x^2 - x + 1 = 0 \Leftrightarrow x = \frac{1 \pm \sqrt{1-4}}{2} \notin \mathbb{R} \tag{4.166}$$

keine (in \mathbb{R})

Ableitung:

$$f'(x) = \frac{(2x-1)(x-1) - (x^2-x+1)}{(x-1)^2} = \frac{2x^2 - 3x + 1 - x^2 + x - 1}{(x-1)^2} = \frac{x^2 - 2x}{(x-1)^2} \tag{4.167}$$

Extrema: $f' = 0 \Leftrightarrow x = 0$ oder $x = 2$

Lokales Maximum $f(0) = -1$, lokales Minimum $f(2) = 3$

Warum (Min/Max)? 4 Möglichkeiten:

(i) wegen Stetigkeit, Pol und Asymptotik

(ii) Untersuche VZ von $f'(x) = \frac{x(x-2)}{(x-1)^2}$:

bei 0: $+$ \rightarrow $-$ also Hochpunkt

bei 2: $-$ \rightarrow $+$ also Tiefpunkte

(iii) Untersuche f'' :

$$\begin{aligned} f''(0) < 0 &\Rightarrow \text{Hochpunkt} \\ f''(2) > 0 &\Rightarrow \text{Tiefpunkt} \end{aligned} \tag{4.168}$$

Vorsicht: Hinreichend, nicht notwendig – falls $f'' = 0$ folgt gar nichts!

(iv) Taylorentwicklung

um Null:

$$\begin{aligned} f(x) &= \frac{x^2 - x + 1}{x - 1} = \frac{-1 + x - x^2}{1 - x} \\ &= (-1 + x - x^2) (1 + x + x^2 + x^3 + \dots) \\ &= -1 + (-x + x) + (-x^2 + x^2 - x^2) + \dots \\ &= -1 - x^2 + \dots \end{aligned} \tag{4.169}$$

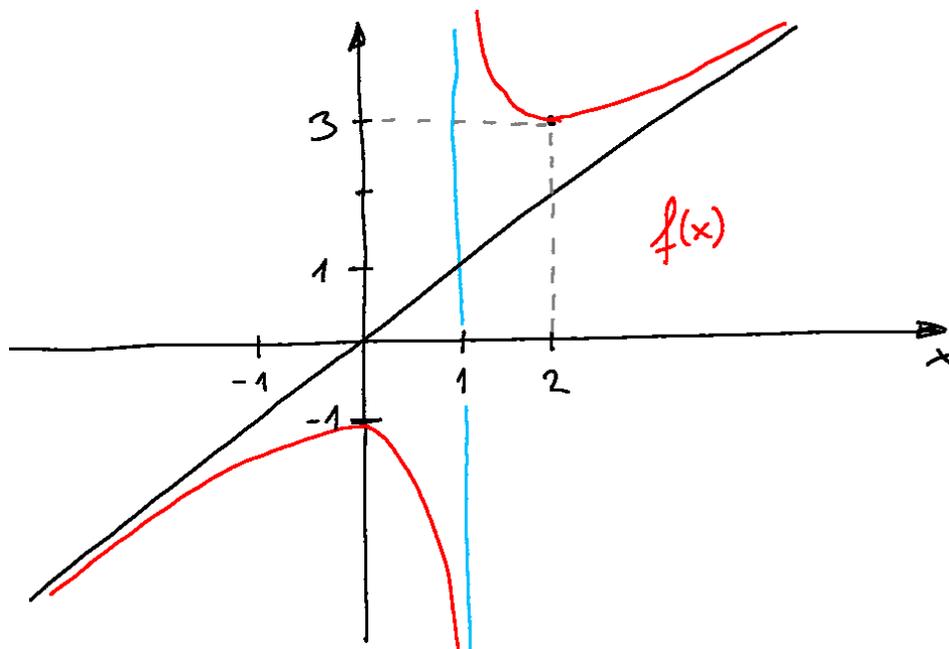
also Hochpunkt $(0, -1)$

um 2:

$$\begin{aligned} f(x) &= \frac{x^2 - x + 1}{x - 1} \\ &= \frac{(x - 2)^2 + 4x - 4 + x + 1}{x - 2 + 1} \\ &= \frac{(x - 2)^2 + 3(x - 2) + 3}{1 + (x - 2)} \\ &= (3 + 3(x - 2) + (x - 2)^2) (1 - (x - 2) + (x - 2)^2 - (x - 2)^3 + \dots) \\ &= 3 + (3 - 3)(x - 2) + (3 - 3 + 1)(x - 2)^2 + \dots \\ &= 3 + (x - 2)^2 + \dots \end{aligned} \tag{4.170}$$

also Tiefpunkt $(2, 3)$.

Skizze:



$$3. f : x \mapsto \sqrt{x - \frac{1}{x}}$$

Definitionsbereich: $x \neq 0$ und $x - \frac{1}{x} \geq 0 \Leftrightarrow x \geq \frac{1}{x}$

(i) für $x > 0$: $x^2 \geq 1 \Rightarrow x \geq 1$

(ii) für $x < 0$: $x^2 \leq 1 \Rightarrow x \geq -1$

also: Definitionsbereich $[-1, 0) \cup [1, \infty)$

Verhalten am Rand des Definitionsbereichs / Nullstellen:

$$f(-1) = 0$$

$$f(x) \sim \sqrt{-\frac{1}{x}} \rightarrow \infty \text{ für } x \rightarrow 0^-$$

$$f(1) = 0$$

$$f(x) \sim \sqrt{x} \rightarrow \infty \text{ für } x \rightarrow \infty$$

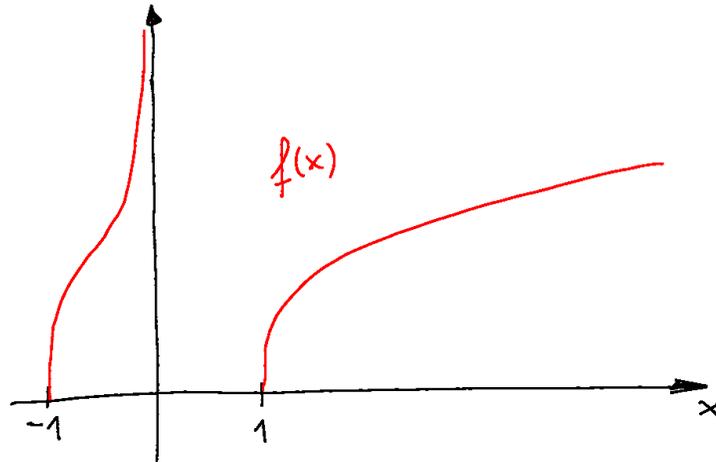
Ableitung:

$$f'(x) = \frac{1 + \frac{1}{x^2}}{2\sqrt{x - \frac{1}{x}}} > 0 \quad \forall x \in \mathbb{D} \quad (4.171)$$

...am Rand des Definitionsbereichs:

$$\lim_{x \rightarrow -1^+} f'(x) = \infty, \quad \lim_{x \rightarrow 1^+} f'(x) = \infty. \quad (4.172)$$

Skizze:



5 Vektorrechnung

5.1 Vektorräume

im WS 22/23 ohne Gruppen und Körper – Vektorräume immer über \mathbb{R} , später auch über \mathbb{C}

Definition: (Gruppe)

Sei $G \neq \emptyset$ eine Menge und \circ eine Verknüpfung $\circ : G \times G \rightarrow G$.

(G, \circ) heißt Gruppe falls gilt:

(G1) aus $a, b \in G \Rightarrow a \circ b \in G$ (Abgeschlossenheit)

(folgt eigentlich bereits aus $\circ : G \times G \rightarrow G$, ist aber gut, das explizit festzuhalten)

(G2) $(a \circ b) \circ c = a \circ (b \circ c) \forall a, b, c \in G$ (Assoziativität)

(G3) $\exists e \in G$ mit $a \circ e = a = e \circ a \forall a \in G$ (neutrales Element)

(G4) für jedes $a \in G \exists a^{-1} \in G$ mit $a \circ a^{-1} = e = a^{-1} \circ a$ (inverses Element)

(wobei e das neutrale Element aus (G3) ist)

Definition: (abelsche Gruppe)

Eine Gruppe (G, \circ) heißt kommutativ oder abelsch falls zusätzlich gilt:

(G5) $a \circ b = b \circ a \forall a, b \in G$ (Kommutativität)

Eigenschaften: (von Gruppen)

1. Das neutrale Element e ist eindeutig bestimmt.

Beweis: Annahme $\exists \tilde{e} \neq e$ mit $\forall a \in G$:

$$\begin{aligned}
 & \tilde{e} \circ a = a && | \circ a^{-1} \text{ von rechts} \\
 \Leftrightarrow & (\tilde{e} \circ a) \circ a^{-1} = a \circ a^{-1} \\
 \stackrel{(G2)}{\Leftrightarrow} & \tilde{e} \circ (a \circ a^{-1}) = a \circ a^{-1} && (5.1) \\
 \stackrel{(G4)}{\Leftrightarrow} & \tilde{e} \circ e = e \\
 \stackrel{(G3)}{\Leftrightarrow} & \tilde{e} = e && \text{Widerspruch}
 \end{aligned}$$

2. Zu jedem $a \in G$ ist das inverse Element eindeutig bestimmt.

Beweis: Annahme $\exists b \neq a^{-1}$ mit

$$\begin{aligned}
 & b \circ a = e && | \circ a^{-1} \text{ von rechts} \\
 \Leftrightarrow & (b \circ a) \circ a^{-1} = e \circ a^{-1} \\
 \stackrel{(G2)}{\Leftrightarrow} & b \circ (a \circ a^{-1}) = e \circ a^{-1} && (5.2) \\
 \stackrel{(G4)}{\Leftrightarrow} & b \circ e = e \circ a^{-1} \\
 \stackrel{(G3)}{\Leftrightarrow} & b = a^{-1} && \text{Widerspruch}
 \end{aligned}$$

3. Gleichungen lassen sich eindeutig lösen:

$\forall a, b \in G \exists x, y \in G$, so dass $a \circ x = b$ und $y \circ a = b$.

Beweis: $x = a^{-1} \circ b$ und $y = b \circ a^{-1}$ tun's – Eindeutigkeit?

Annahme: $\exists z \neq y$ mit

$$\begin{aligned}
& z \circ a = b && | \circ a^{-1} \text{ von rechts} \\
\Leftrightarrow & z \circ a \circ a^{-1} = b \circ a^{-1} && (5.3) \\
\Leftrightarrow & z = b \circ a^{-1} = y && \text{Widerspruch}
\end{aligned}$$

Bemerkung: für nicht-abelsche Gruppen gilt i.A. $x \neq y$.

Beispiele

1. $(\mathbb{Z}, +)$: $e = 0$, $a^{-1} = -a$ für $a \in \mathbb{Z}$
natürlich ebenso $(\mathbb{R}, +)$ aber nicht $(\mathbb{N}, +)$!
2. $(\mathbb{Q} \setminus \{0\}, \cdot)$: $e = 1$, für $\mathbb{Q} \ni x = \frac{p}{q}$ mit $p, q \in \mathbb{Z}$ (vollständig gekürzt) ist $x^{-1} = \frac{q}{p}$
ebenso $(\mathbb{R} \setminus \{0\}, \cdot)$ aber nicht $(\mathbb{Z} \setminus \{0\}, \cdot)$!
3. (Menge aller Polynome vom Grad $\leq n$, $+$)
4. G : Menge aller Symmetrieoperationen (Drehungen, Spiegelungen etc.), die ein bestimmtes Objekt (Atom, Molekül, ...) invariant lassen.
 \circ : Nacheinanderausführen der Operationen.
Beispiel: Rechteck \rightsquigarrow Gruppe mit 4 Elementen
5. Gruppe mit zwei Elementen, z.B. $(\mathbb{Z}_2, +)$
6. triviale Gruppe

Definition: (Körper)

Sei $K \neq \emptyset$ eine Menge mit zwei Verknüpfungen, $+$: $K \times K \rightarrow K$ und \cdot : $K \times K \rightarrow K$.

$(K, +, \cdot)$ heißt Körper, falls gilt:

(K1) $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.

(K2) $(K \setminus \{0\}, \cdot)$ ist ebenfalls eine abelsche Gruppe.

(K3) $(a + b) \cdot c = a \cdot c + b \cdot c \quad \forall a, b, c \in K$ (Distributivität)

Bemerkungen:

1. Nenne das zu $a \in K$ additiv inverse Element $(-a)$.
2. Jeder Körper hat mindestens zwei Elemente, nämlich $0 \in K$ und, da $(K \setminus \{0\}, \cdot)$ Gruppe ist, $\exists e \in K$, $e \neq 0$ (multiplikativ neutrales Element).
3. Multiplikation mit 0: $a \in K \setminus \{0\}$

$$0 = e + (-e) \quad | \cdot a \quad (5.4)$$

$$0 \cdot a = (e + (-e)) \cdot a \stackrel{(K3)}{=} a + (-e) \cdot a =: \tilde{a} \in K \quad (5.5)$$

Annahme: $\tilde{a} \in K \setminus \{0\} \Rightarrow \tilde{a} \cdot a^{-1} = e + (-e) = 0 \notin K \setminus \{0\}$ – Widerspruch!

$\Rightarrow \tilde{a} = 0$ also $0 \cdot a = 0$.

außerdem folgt wegen $a + (-e) \cdot a = 0$, dass $(-a) = (-e) \cdot a$.

4. Üblicherweise nennen wir $e = 1$ (multiplikativ neutrales Element).
5. aus $a \cdot b = 0 \Rightarrow a = 0$ oder $b = 0$, denn
anderfalls $a, b \in K \setminus \{0\} \Rightarrow a \cdot b \in K \setminus \{0\}$ da Gruppe – Widerspruch!

Beispiele

1. $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, oder auch: $(\mathbb{C}, +, \cdot)$
2. $\mathbb{Z}_2 = \{0, 1\}$ mit $(+, \cdot)$ modulo 2, d.h.

$$0 + 0 = 0, \quad 1 + 0 = 1, \quad \mathbf{1 + 1 = 0} \quad (5.6)$$

$$0 \cdot 0 = 0, \quad 0 \cdot 1 = 0, \quad 1 \cdot 1 = 1 \quad (5.7)$$

$$\Rightarrow (-1) = 1 \text{ und } 1^{-1} = 1$$

3. $\mathbb{Z}_3 = \{0, 1, 2\}$ mit $(+, \cdot)$ modulo 3, d.h.

$$\begin{array}{c|ccc} + & 0 & 1 & 2 \\ \hline 0 & 0 & 1 & 2 \\ 1 & 1 & 2 & 0 \\ 2 & 2 & 0 & 1 \end{array} \quad \begin{array}{c|ccc} \cdot & 0 & 1 & 2 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 2 \\ 2 & 0 & 2 & 1 \end{array} \quad (5.8)$$

$$\Rightarrow (-1) = 2, (-2) = 1 \text{ (additiv Inverse)}$$

$$\text{und } 1^{-1} = 1, 2^{-1} = 2 \text{ (multiplikativ Inverse)}$$

Bemerkungen:

- abelsche Gruppen \Rightarrow Tabellen symmetrisch bzgl. Spiegelung an der Diagonalen \
- Gruppen \Rightarrow Gleichungen eindeutig lösbar
 \Rightarrow jedes Element tritt in jeder Zeile und Spalte genau einmal auf!
(In \cdot -Tabelle natürlich der Block ohne Nullen, da nur $(K \setminus \{0\}, \cdot)$ Gruppe ist!)

4. $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ mit $(+, \cdot)$ modulo 4 ist **kein** Körper, denn
 $2 \cdot 2 = 0 \notin \mathbb{Z}_4 \setminus \{0\}$ – Widerspruch zu $(\mathbb{Z}_4 \setminus \{0\}, \cdot)$ Gruppe

im WS 22/23 ohne Gruppen und Körper – Vektorräume immer über \mathbb{R} , später auch über \mathbb{C}

Definition: (Vektorraum)

Sei $(V, +)$ eine abelsche Gruppe mit neutralem Element $\vec{0}$,

K ein Körper und $\cdot : K \times V \rightarrow V$ (skalare Multiplikation).

V heißt Vektorraum über K falls zusätzlich gilt:

$$(V1) \lambda \cdot \vec{a} \in V \quad \forall \lambda \in K \text{ und } \forall \vec{a} \in V \text{ (Abgeschlossenheit)}$$

$$(V2) \lambda \cdot (\vec{a} + \vec{b}) = \lambda \cdot \vec{a} + \lambda \cdot \vec{b}$$

$$(\lambda + \mu) \cdot \vec{a} = \lambda \cdot \vec{a} + \mu \cdot \vec{a}$$

$$\forall \lambda, \mu \in K \text{ und } \forall \vec{a}, \vec{b} \in V \text{ (Distributivgesetze)}$$

$$(V3) 1 \cdot \vec{a} = \vec{a} \quad \forall \vec{a} \in V$$

$$(V4) \lambda \cdot (\mu \cdot \vec{a}) = (\lambda \cdot \mu) \cdot \vec{a} \quad \forall \lambda, \mu \in K \text{ und } \forall \vec{a} \in V \text{ (Assoziativgesetz)}$$

Bemerkung

1. Multiplikationspunkte dürfen wie üblich weggelassen werden.
2. Aus (G1) für $(V, +)$ und (V1) $\Rightarrow \lambda \vec{a} + \mu \vec{b} \in V \quad \forall \lambda, \mu \in K \text{ und } \forall \vec{a}, \vec{b} \in V$

Beispiele

1. \mathbb{R} über \mathbb{R}

2. Ebene \mathbb{R}^2 , Raum \mathbb{R}^3 über \mathbb{R}
 allgemein \mathbb{R}^n mit komponentenweiser Vektor-Addition und skalarer Multiplikation

$$\vec{a}, \vec{b} \in \mathbb{R}^n : \quad \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad (5.9)$$

$$\vec{a} + \vec{b} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}, \quad \lambda \vec{a} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix} \quad (5.10)$$

3. Polynome (beliebigen Grads) über \mathbb{R}

$$P(x) = \sum_{\nu=0}^n a_\nu x^\nu, \quad Q(x) = \sum_{\nu=0}^m b_\nu x^\nu, \quad a_\nu, b_\nu, x \in \mathbb{R} \quad (5.11)$$

$$(\lambda P)(x) = \lambda P(x) = \sum_{\nu=0}^n \lambda a_\nu x^\nu, \quad (P+Q)(x) = P(x)+Q(x) = \sum_{\nu=0}^{\max\{n,m\}} (a_\nu + b_\nu) x^\nu, \quad (5.12)$$

(wobei $a_\nu = 0 \forall \nu > n$ und $b_\nu = 0 \forall \nu > m$)
 übrigens ∞ -dimensional (später)

5.2 Lineare Unabhängigkeit

Definitionen:

1. $\vec{0}$ ist linear abhängig (l.a.)
 $\vec{a} \neq \vec{0}$ ist linear unabhängig (l.u.)
2. Zwei Vektoren \vec{a}, \vec{b} heißen l.u. \Leftrightarrow aus $\lambda \vec{a} + \mu \vec{b} = \vec{0}$ folgt $\lambda = \mu = 0$
 d.h. der Nullvektor $\vec{0}$ lässt sich nur als triviale
 Linearkombination von \vec{a} & \vec{b} schreiben
 anderfalls heißen sie l.a.
 \vec{a}, \vec{b} l.a. $\Leftrightarrow \vec{a} \parallel \vec{b}$ (oder mindestens einer ist Nullvektor)
 \vec{a}, \vec{b} l.u. $\Leftrightarrow \vec{a}, \vec{b}$ spannen die Ebene $\left\{ \lambda \vec{a} + \mu \vec{b} \mid \lambda, \mu \in K \right\}$ auf.
3. $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ heißen l.u. \Leftrightarrow aus $\sum_{j=1}^n \lambda_j \vec{a}_j = \vec{0}$ folgt $\lambda_j = 0 \forall j = 1, \dots, n$

Bemerkungen:

1. Sind $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ l.u., so ist Koeffizientenvergleich möglich, d.h.

$$\sum_{j=1}^n \lambda_j \vec{a}_j = \sum_{j=1}^n \mu_j \vec{a}_j \quad \Leftrightarrow \quad \lambda_j = \mu_j \quad \forall j = 1, \dots, n \quad (5.13)$$

2. Sind $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n \in \mathbb{R}^n$ l.u. so spannen sie den gesamten \mathbb{R}^n auf, d.h. für jeden Vektor $\vec{b} \in \mathbb{R}^n$ gibt es eindeutig bestimmte Koeffizienten λ_j , so dass

$$\vec{b} = \sum_{j=1}^n \lambda_j \vec{a}_j \quad (5.14)$$

Bestimmung der λ_j durch Lösen eines linearen Gleichungssystems.

Beispiel:

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{a}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad (5.15)$$

Fragen:

1. Sind die Vektoren l.u.?
2. Können wir z.B. $\vec{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ als Linearkombination (LK) der \vec{a}_j darstellen?

Zu 2:

Suche λ_j mit

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad (5.16)$$

d.h. löse das lineare Gleichungssystem (LGS)

$$\begin{array}{rcl} \lambda_1 & + & \lambda_2 & & = & 1 \\ & & \lambda_2 & + & \lambda_3 & = & 2 \\ \lambda_1 & & & + & \lambda_3 & = & 3 \end{array} \quad (5.17)$$

Umformen: erlaubte Operationen

- Zeilen vertauschen
- Zeile mit Faktor $\neq 0$ multiplizieren
- Vielfache einer Zeile zu einer anderen addieren
- **Vorsicht:** Zu einer Zeile, die wir zu einer anderen Zeile addieren, im gleichen Schritt nicht auch etwas hinzuaddieren!

Gauß-Algorithmus:

- Ziel 1: Zeilenstufenform (durch Vorwärtselimination)
- Rückwärtselimination

kompakte Schreibweise

$$\begin{array}{c}
 \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 3 \end{array} \right) \begin{array}{l} \rceil^{-1} \\ \leftarrow_+ \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & -1 & 1 & 2 \end{array} \right) \begin{array}{l} \rceil_+ \\ \leftarrow_+ \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 2 & 4 \end{array} \right) \begin{array}{l} | \cdot 1/2 \\ \end{array} \quad (*) \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{array} \right) \begin{array}{l} \leftarrow_+ \\ \rceil_{-1} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array} \right) \begin{array}{l} \leftarrow_+ \\ \rceil_{-1} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array} \right)
 \end{array} \tag{5.18}$$

Die Lösung steht in der letzten Spalte:

$$\lambda_1 = 1, \quad \lambda_2 = 0, \quad \lambda_3 = 2. \tag{5.19}$$

Alternativ können wir auch in (*) – Zeilenstufenform – ablesen

$$\begin{array}{rcl}
 2\lambda_3 = 4 & \Rightarrow & \lambda_3 = 2 \\
 \lambda_2 + \lambda_3 = 2 & \Rightarrow & \lambda_2 = 2 - \lambda_3 = 0 \\
 \lambda_1 + \lambda_2 = 1 & \Rightarrow & \lambda_1 = 1 - \lambda_2 = 1
 \end{array} \tag{5.20}$$

Zu 1: Die drei Vektoren sind l.u.

Dazu lösen wir das LGS $(\cdots | \vec{0})$ und prüfen, ob es nur die Lösung $\lambda_1 = \lambda_2 = \lambda_3 = 0$ hat. (Sehen wir auch bereits in (*). – Wäre unterwegs eine Zeile mit ausschließlich Nullen links von | aufgetreten, so wären die Vektoren l.a.)

Neues Beispiel dazu: \vec{a}_1, \vec{a}_2 wie oben,

$$\vec{a}_3 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \tag{5.21}$$

also

$$\begin{array}{c}
 \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{array} \right) \begin{array}{l} \leftarrow \boxed{-1} \\ \leftarrow \boxed{+} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 \end{array} \right) \begin{array}{l} \leftarrow \boxed{+} \\ \leftarrow \boxed{+} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \begin{array}{l} \leftarrow \boxed{+} \\ \leftarrow \boxed{-1} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)
 \end{array} \tag{5.22}$$

d.h.

$$\lambda_3 = t, \quad \lambda_1 = \lambda_2 = -t \tag{5.23}$$

löst $\forall t \in \mathbb{R}$, d.h. die Vektoren $\vec{a}_1, \vec{a}_2, \vec{a}_3$ sind l.a.

Lässt sich $\vec{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ als LK darstellen? Nein, denn...

$$\begin{array}{c}
 \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 1 & 0 & 1 & 3 \end{array} \right) \begin{array}{l} \leftarrow \boxed{-1} \\ \leftarrow \boxed{+} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & -1 & -1 & 2 \end{array} \right) \begin{array}{l} \leftarrow \boxed{+} \\ \leftarrow \boxed{+} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 1 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 4 \end{array} \right)
 \end{array} \tag{5.24}$$

...letzte Zeile bedeutet

$$0 \cdot \lambda_1 + 0 \cdot \lambda_2 + 0 \cdot \lambda_3 = 4 \quad \text{Widerspruch!} \tag{5.25}$$

Vorsicht: Keine "Ringoperationen" machen!

$$\begin{array}{c}
 \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array} \right) \begin{array}{l} \leftarrow \boxed{+} \\ \leftarrow \boxed{+} \end{array} \begin{array}{l} \leftarrow \boxed{+} \\ \leftarrow \boxed{-1} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 1 & 1 & 0 & 3 \\ 0 & 1 & 1 & 5 \end{array} \right) \begin{array}{l} \leftarrow \boxed{-1} \\ \leftarrow \boxed{+} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 5 \\ 0 & 1 & 1 & 5 \end{array} \right) \begin{array}{l} \leftarrow \boxed{-1} \\ \leftarrow \boxed{+} \end{array} \\
 \hline
 \left(\begin{array}{ccc|c} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{array} \right) \begin{array}{l} \leftarrow \boxed{-1} \\ \leftarrow \boxed{+} \end{array}
 \end{array} \tag{5.26}$$

oben: eindeutige Lösung $x_1 = 1, x_2 = 2, x_3 = 3$.

unten: wähle $x_3 = t \in \mathbb{R}$ beliebig, $x_2 = 5 - t, x_1 = t - 2$, d.h. die allgemeine Lösung wäre

$$\vec{x} = \begin{pmatrix} -2 \\ 5 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} t \quad - \text{ zu viel!!!} \quad (5.27)$$

Erster Schritt war verboten!

inkl. Beispiele in $V = C(\mathbb{R})$: (i) $e^x, e^x, \sinh(x)$ l.a., (ii) $1, e^x, e^x$ l.u.

5.3 Lineare Gleichungssysteme und allgemeiner Gauß-Algorithmus

Allgemein: $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n, \vec{b} \in \mathbb{R}^m$

$$\sum_{j=1}^n x_j \vec{a}_j = \vec{b} \quad (5.28)$$

heißt lineares Gleichungssystem (LGS) – m Gleichungen für n Unbekannte.

$\vec{b} = 0$: homogenes LGS

$\vec{b} \neq 0$: inhomogenes LGS

Andere Schreibweise für LGS:

$$L(\vec{x}) = \vec{b}, \quad (5.29)$$

wobei

$$L : \quad \mathbb{R}^n \quad \rightarrow \quad \mathbb{R}^m$$

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto L(\vec{x}) = \sum_{j=1}^n x_j \vec{a}_j \quad (5.30)$$

eine **lineare** Abbildung ist, denn offensichtlich gilt

$$L(\lambda \vec{x} + \mu \vec{y}) = \lambda L(\vec{x}) + \mu L(\vec{y}). \quad (5.31)$$

Die Mengen (für $\vec{b} \neq \vec{0}$)

$$\mathcal{L}_{\vec{b}} := \left\{ \vec{x} \in \mathbb{R}^n \mid L(\vec{x}) = \vec{b} \right\} \quad \text{und} \quad \mathcal{L}_0 := \left\{ \vec{x} \in \mathbb{R}^n \mid L(\vec{x}) = \vec{0} \right\} \quad (5.32)$$

heißen Lösungsmenge des inhomogenen bzw. des homogenen LGS.

Übrigens: Wir nennen \mathcal{L}_0 auch $\ker L$, den *Kern von L*.

Satz 9. Sei $\vec{b} \neq \vec{0}$, seien $\vec{x}_1^h, \vec{x}_2^h \in \mathcal{L}_0$ und $\vec{y}^p \in \mathcal{L}_{\vec{b}}$, dann gilt:

(i) $\lambda \vec{x}_1^h + \mu \vec{x}_2^h \in \mathcal{L}_0 \quad \forall \lambda, \mu \in \mathbb{R}$,

(ii) $\vec{y} \in \mathcal{L}_{\vec{b}} \Leftrightarrow \exists \vec{x} \in \mathcal{L}_0$ mit $\vec{y} = \vec{y}^p + \vec{x}$.

Beweis:

(i) $L(\lambda \vec{x}_1^h + \mu \vec{x}_2^h) = \lambda L(\vec{x}_1^h) + \mu L(\vec{x}_2^h) = \vec{0} + \vec{0}$

(ii) “ \Leftarrow ”: $L(\vec{y}) = L(\vec{y}^p + \vec{x}) = L(\vec{y}^p) + L(\vec{x}) = \vec{b} + \vec{0} = \vec{b}$

“ \Rightarrow ”: $\vec{x} := \vec{y} - \vec{y}^p, L(\vec{x}) = L(\vec{y}) - L(\vec{y}^p) = \vec{b} - \vec{b} = \vec{0}$ □

Bemerkung: Gilt auch alles für $L : V \rightarrow W$ linear, mit beliebigen Vektorräumen V und W (über demselben Körper), denn $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ wurde nicht benötigt.

Zur Lösung: Bringe

$$\left(\vec{a}_1 \quad \vec{a}_2 \quad \cdots \quad \vec{a}_n \mid \vec{b} \right) \tag{5.33}$$

auf Zeilenstufenform (ZSF)

$$\left(\begin{array}{cccccccc|c} \blacksquare & * & * & * & \cdots & * & \cdots & * & \tilde{b}_1 \\ 0 & 0 & \blacksquare & * & \cdots & * & \cdots & * & \tilde{b}_2 \\ 0 & 0 & 0 & \blacksquare & * & \cdots & * & \cdots & * & \tilde{b}_3 \\ \vdots & & & \ddots & \ddots & \ddots & \vdots & & \vdots & \vdots \\ 0 & & & & 0 & \blacksquare & * & \cdots & * & \tilde{b}_r \\ 0 & & & & & & 0 & \cdots & 0 & \tilde{b}_{r+1} \\ 0 & & & & & & & 0 & \cdots & 0 \\ 0 & & \cdots & & & & & & 0 & \vdots \\ 0 & & & & & & & & 0 & \tilde{b}_m \end{array} \right) \left. \begin{array}{l} \vphantom{\left(\right)} \\ \vphantom{\left(\right)} \end{array} \right\} \begin{array}{l} r \text{ Zeilen} \\ \\ \\ \\ \\ m - r \text{ Zeilen} \end{array} \tag{5.34}$$

Kennzeichen der ZSF:

1. \blacksquare sind Zahlen $\neq 0$.
2. $*$ sind irgendwelche Zahlen.
3. Links (und unterhalb) von \blacksquare stehen nur Nullen.
4. Stufen von \blacksquare zu \blacksquare : eine Zeile nach unten, mindestens eine Spalte nach rechts.

Lösung des LGS:

1. Zeile der Form $(0 \cdots 0 \mid \tilde{b}_j)$ mit $\tilde{b}_j \neq 0$ bedeutet: LGS hat keine Lösung.
2. Spalten ohne \blacksquare entsprechen frei wählbaren Variablen (parametrisiere so die Lösung).
3. Variablen, die Spalten mit \blacksquare entsprechen, sind durch die Zeile, in der \blacksquare steht, festgelegt (von unten nach oben arbeiten).

Insbesondere hat das LGS mit ZSF

$$\left(\begin{array}{ccc|c} \blacksquare & & * & \tilde{b}_1 \\ & \ddots & & \vdots \\ 0 & & \blacksquare & \tilde{b}_m \end{array} \right) \tag{5.35}$$

(mit $n = m$) eine eindeutige Lösung.

Beispiele:

1.

$$\left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 2 \end{array} \right) \quad (5.36)$$

hat keine Lösung.

2.

$$\left(\begin{array}{ccc|c} \boxed{1} & 2 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \\ 0 & 0 & 0 & 0 \end{array} \right), \quad (5.37)$$

$x_3 = 2, x_2 = t \in \mathbb{R}$ beliebig, $x_1 = 1 - x_3 - 2t = -1 - 2t$, bzw.

$$\vec{x} = \begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix} + \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} t, \quad t \in \mathbb{R}. \quad (5.38)$$

3.

$$\left(\begin{array}{ccc|c} \boxed{1} & 2 & 1 & 1 \\ 0 & \boxed{1} & 2 & 4 \\ 0 & 0 & \boxed{1} & 0 \end{array} \right), \quad (5.39)$$

$x_3 = 0, x_2 = 4 - 2x_3 = 4, x_1 = 1 - 2x_2 - x_3 = -7$ (eindeutig).

5.4 Unterräume, Dimension und Basis

Definition: (Dimension)

Sei V ein Vektorraum. Die maximale Anzahl linear unabhängiger Vektoren in V heißt Dimension von V , $\dim V$.

Bemerkung: $\dim \mathbb{R}^n = n$

Beweis: Die kanonischen Einheitsvektoren

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (5.40)$$

sind linear unabhängig $\Rightarrow \dim \mathbb{R}^n \geq n$.

Seinen $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_{n+1} \in \mathbb{R}^n$. Bringe das LGS

$$\left(\vec{a}_1 \quad \vec{a}_2 \quad \dots \quad \vec{a}_{n+1} \mid \vec{0} \right) \quad (5.41)$$

auf Zeilenstufenform. Diese enthält höchstens n ■, da das LGS n Zeilen hat.

\Rightarrow Mindestens eine Variable bleibt frei wählbar.

\Rightarrow Es gibt nichttriviale Lösungen, d.h. die Vektoren sind l.a.
 Also ist $\dim \mathbb{R}^n < n + 1$ und damit $= n$. □

Definition: (Unterraum)

Sei V Vektorraum über K . U heißt Unterraum von V , falls gilt:

- (i) $U \subseteq V$
- (ii) U ist ein Vektorraum (über K).

Bemerkungen:

- 1. Zu gegebenem $U \subseteq V$ ist lediglich zu prüfen, ob aus $\vec{a}, \vec{b} \in U, \lambda \in K$, folgt, dass auch $\vec{a} + \vec{b}$ und $\lambda \vec{a} \in U$ sind. (Rechenregeln erbt U von V .)
- 2. $\dim U \leq \dim V$

Beispiele:

- 1. $U = \{\vec{0}\}$ ist Unterraum, $\dim U = 0$.
- 2. Die Lösungsmenge \mathcal{L}_0 eines homogenen LGS, $L(\vec{x}) = \vec{0}, L : \mathbb{R}^n \rightarrow \mathbb{R}^m$, ist Unterraum des \mathbb{R}^n , denn offensichtlich

$$\mathcal{L}_0 = \ker L = \left\{ \vec{x} \in \mathbb{R}^n \mid L(\vec{x}) = \vec{0} \right\} \subseteq \mathbb{R}^n, \tag{5.42}$$

und Abgeschlossenheit folgt aus Satz 9 (i).
 $\dim \mathcal{L}_0 =$ Anzahl der Spalten ohne ■ in ZSF.

Definitionen: (Lineare Hülle, Erzeugendensystem, Basis)

Sei V ein Vektorraum über K und $\vec{a}_1, \dots, \vec{a}_n \in V$.

- 1. Die Menge aller Linearkombinationen,

$$\text{span}(\vec{a}_1, \dots, \vec{a}_n) := \left\{ \vec{x} \in V \mid \vec{x} = \sum_{j=1}^n \lambda_j \vec{a}_j, \lambda_j \in K \right\}, \tag{5.43}$$

heißt lineare Hülle (oder Aufspann) von $\vec{a}_1, \dots, \vec{a}_n$.

- 2. Die Vektoren $\vec{a}_1, \dots, \vec{a}_n$ heißen Erzeugendensystem von V , falls gilt

$$\text{span}(\vec{a}_1, \dots, \vec{a}_n) = V. \tag{5.44}$$

Wir sagen auch, sie spannen V auf.

- 3. Sind die Vektoren $\vec{a}_1, \dots, \vec{a}_n$ l.u. und bilden ein Erzeugendensystem von V , so heißen sie Basis von V .

Bemerkungen:

- 1. $\text{span}(\vec{a}_1, \dots, \vec{a}_n)$ ist ein Unterraum von V .
- 2. Die Anzahl der Vektoren einer Basis von V ist gleich $\dim V$.

Beispiele:

1. Die Vektoren

$$\vec{a} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \vec{c} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad (5.45)$$

spannen den \mathbb{R}^2 auf, sind aber keine Basis, da $\vec{a} + \vec{b} = \vec{c}$

2. Die kanonischen Einheitsvektoren \vec{e}_j (s.o.), $j = 1, \dots, n$, mit Komponenten $(\vec{e}_j)_k = \delta_{jk}$, wobei

$$\delta_{jk} = \begin{cases} 1 & \text{für } j = k \\ 0 & \text{sonst} \end{cases} \quad (5.46)$$

das Kronecker-Symbol ist, bilden eine Basis des \mathbb{R}^n (die kanonische Basis).

3. Die Vektoren

$$\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{x}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{x}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad (5.47)$$

bilden eine Basis des \mathbb{R}^3 .

4. $\{1\}$ ist eine Basis von \mathbb{C} als VR über \mathbb{C} , d.h. $\dim \mathbb{C} = 1$.

5. $\{1, i\}$ ist eine Basis von \mathbb{C} als VR über \mathbb{R} , d.h. $\dim \mathbb{C} = 2$.

Satz 10.

Seien $\vec{a}_1, \dots, \vec{a}_n \in \mathbb{R}^m$. Die Dimension von $\text{span}(\vec{a}_1, \dots, \vec{a}_n)$ ist gleich der Anzahl der \blacksquare in der Zeilenstufenform von

$$\left(\vec{a}_1 \quad \vec{a}_2 \quad \dots \quad \vec{a}_n \mid \vec{0} \right). \quad (5.48)$$

Bemerkung: Damit gilt $\dim \mathcal{L}_0 + \dim \text{span}(\vec{a}_1, \dots, \vec{a}_n) = n$.

Es gilt außerdem $\text{span}(\vec{a}_1, \dots, \vec{a}_n) = \text{im } L$, wobei

im $L := \{\vec{y} \in \mathbb{R}^m \mid \exists \vec{x} \in \mathbb{R}^n \text{ mit } L(\vec{x}) = \vec{y}\}$ das Bild der linearen Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $L(\vec{x}) = \sum_{j=1}^n x_j \vec{a}_j$, ist. Oben hatten wir bereits bemerkt, dass $\mathcal{L}_0 = \ker L$.

Analog gilt für eine beliebige lineare Abbildung $L : V \rightarrow W$

$$\dim \ker L + \dim \text{im } L = \dim V.$$

Beweisskizze: maximale Anzahl \blacksquare : n

1. $n \blacksquare \Rightarrow$ Vektoren l.u. $\Rightarrow \dim \text{span}(\vec{a}_1, \dots, \vec{a}_n) = n$

2. eine Spalte ohne $\blacksquare \Rightarrow$ Vektoren l.a. $\Rightarrow \dim \text{span}(\vec{a}_1, \dots, \vec{a}_n) < n$.

Lasse diese Spalte weg: restliche $n - 1$ Vektoren sind l.u.

$$\Rightarrow \dim \text{span}(\vec{a}_1, \dots, \vec{a}_n) = n - 1.$$

3. etc.

5.5 Skalarprodukt und Norm

Wdh: Länge/Betrag eines Vektors aus \mathbb{R}^n und Skalarprodukt zweier Vektoren. Das motiviert die folgenden Definitionen...

Definition: (Norm)

Sei V ein Vektorraum über \mathbb{R} . Eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}_0^+$ heißt Norm,

wenn $\forall \vec{a}, \vec{b} \in V$ und $\forall \lambda \in \mathbb{R}$ gilt:

$$(N1) \|\vec{a}\| = 0 \Leftrightarrow \vec{a} = \vec{0}$$

$$(N2) \|\lambda\vec{a}\| = |\lambda|\|\vec{a}\|$$

$$(N3) \|\vec{a} + \vec{b}\| \leq \|\vec{a}\| + \|\vec{b}\| \text{ (Dreiecksungleichung)}$$

Definition: (Skalarprodukt)

Sei V ein Vektorraum über \mathbb{R} . Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ heißt Skalarprodukt,

wenn $\forall \vec{a}, \vec{b}, \vec{c} \in V$ und $\forall \lambda \in \mathbb{R}$ gilt:

$$(S1) \langle \vec{a}, \vec{b} \rangle = \langle \vec{b}, \vec{a} \rangle \text{ (symmetrisch)}^9$$

$$(S2) \langle \vec{a}, \lambda\vec{b} \rangle = \lambda\langle \vec{a}, \vec{b} \rangle$$

$$\langle \vec{a}, \vec{b} + \vec{c} \rangle = \langle \vec{a}, \vec{b} \rangle + \langle \vec{a}, \vec{c} \rangle \text{ (linear)}$$

$$(S3) \langle \vec{a}, \vec{a} \rangle \geq 0, \text{ wobei } \langle \vec{a}, \vec{a} \rangle = 0 \Leftrightarrow \vec{a} = \vec{0} \text{ (positiv definit)}$$

Satz 11. (Norm)

Sei V ein Vektorraum über \mathbb{R} . Jedes Skalarprodukt, $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ induziert eine Norm, $\|\cdot\| : V \rightarrow \mathbb{R}_0^+$, gegeben durch

$$\|\vec{a}\| := \sqrt{\langle \vec{a}, \vec{a} \rangle}. \quad (5.49)$$

Beweis:

(N1) folgt aus (S3). Zu (N2):

$$\|\lambda\vec{a}\| = \sqrt{\langle \lambda\vec{a}, \lambda\vec{a} \rangle} \stackrel{(S2)}{=} \sqrt{\lambda\langle \lambda\vec{a}, \vec{a} \rangle} \stackrel{(S1)}{=} \sqrt{\lambda\langle \vec{a}, \lambda\vec{a} \rangle} \stackrel{(S2)}{=} \sqrt{\lambda^2\langle \vec{a}, \vec{a} \rangle} = |\lambda|\sqrt{\langle \vec{a}, \vec{a} \rangle} = |\lambda|\|\vec{a}\| \quad (5.50)$$

Für (N3) beweisen wir zunächst:

Lemma 12. (Cauchy-Schwarzsche Ungleichung)

Skalarprodukt und Norm erfüllen

$$|\langle \vec{a}, \vec{b} \rangle| \leq \|\vec{a}\|\|\vec{b}\| \quad \forall \vec{a}, \vec{b} \in V. \quad (CS)$$

Beweis: Fall $\vec{b} = \vec{0}$ klar; also $\vec{b} \neq \vec{0}$:

$$0 \leq \langle \vec{a} - \lambda\vec{b}, \vec{a} - \lambda\vec{b} \rangle = \|\vec{a}\|^2 + \lambda^2\|\vec{b}\|^2 - 2\lambda\langle \vec{a}, \vec{b} \rangle, \quad (5.51)$$

⁹Später werden wir auch Normen und Skalarprodukte für Vektorräume über den komplexen Zahlen \mathbb{C} betrachten; dann muss lediglich (S1) modifiziert werden.

wähle $\lambda = \frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{b}\|^2}$,

$$0 \leq \|\vec{a}\|^2 + \frac{(\langle \vec{a}, \vec{b} \rangle)^2}{\|\vec{b}\|^2} - 2 \frac{(\langle \vec{a}, \vec{b} \rangle)^2}{\|\vec{b}\|^2} = \|\vec{a}\|^2 - \frac{(\langle \vec{a}, \vec{b} \rangle)^2}{\|\vec{b}\|^2} \quad (5.52)$$

$$\Leftrightarrow (\langle \vec{a}, \vec{b} \rangle)^2 \leq \|\vec{a}\|^2 \|\vec{b}\|^2$$

□

Damit (N3):

$$\begin{aligned} \|\vec{a} + \vec{b}\|^2 &= \langle \vec{a} + \vec{b}, \vec{a} + \vec{b} \rangle = \|\vec{a}\|^2 + \|\vec{b}\|^2 + 2\langle \vec{a}, \vec{b} \rangle \\ &\stackrel{\text{(CS)}}{\leq} \|\vec{a}\|^2 + \|\vec{b}\|^2 + 2\|\vec{a}\|\|\vec{b}\| = (\|\vec{a}\| + \|\vec{b}\|)^2 \end{aligned} \quad (5.53)$$

□

Beispiel: (kanonisches Skalarprodukt im \mathbb{R}^n)

Für $\vec{a}, \vec{b} \in \mathbb{R}^n$ definiert

$$\langle \vec{a}, \vec{b} \rangle = \sum_{j=1}^n a_j b_j \quad (5.54)$$

ein Skalarprodukt. Wir schreiben

$$\vec{a} \cdot \vec{b} := \langle \vec{a}, \vec{b} \rangle \quad (5.55)$$

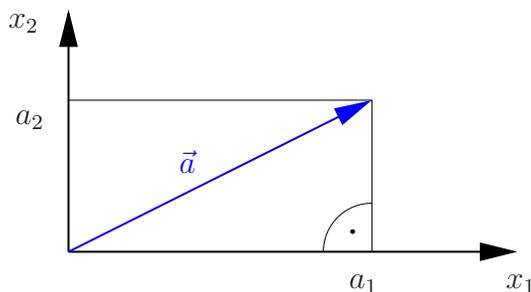
(bzw. auch $\vec{a}\vec{b} = \vec{a} \cdot \vec{b}$) und für die zugehörige Norm

$$|\vec{a}| := \|\vec{a}\| = \sqrt{\sum_{j=1}^n a_j^2} \quad (5.56)$$

Interpretation:

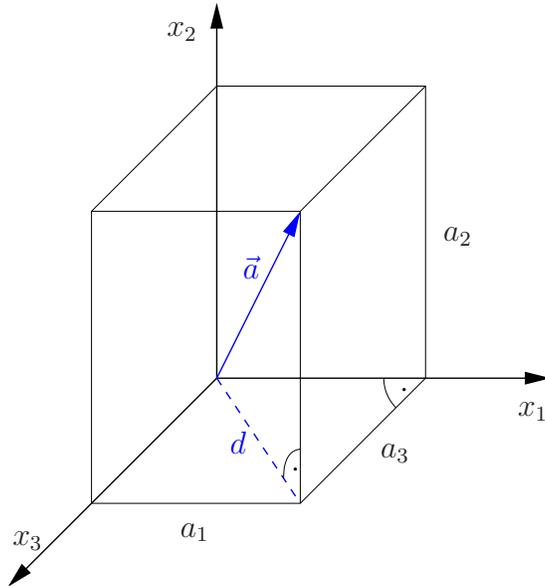
1. Norm/Betrag: Länge des Vektors

- \mathbb{R}^2 Pythagoras



$$|\vec{a}|^2 = a_1^2 + a_2^2$$

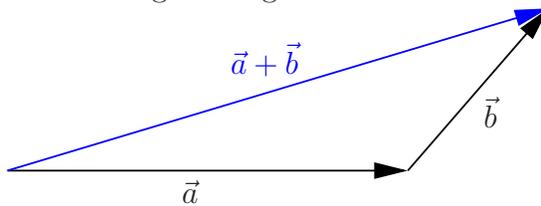
- \mathbb{R}^3 zweimal Pythagoras (Raumdiagonale)



$$\begin{aligned} d^2 &= a_1^2 + a_3^2 \\ |\vec{a}|^2 &= d^2 + a_2^2 \\ &= a_1^2 + a_2^2 + a_3^2 \end{aligned}$$

- \mathbb{R}^n analog $(n - 1)$ -mal Pythagoras

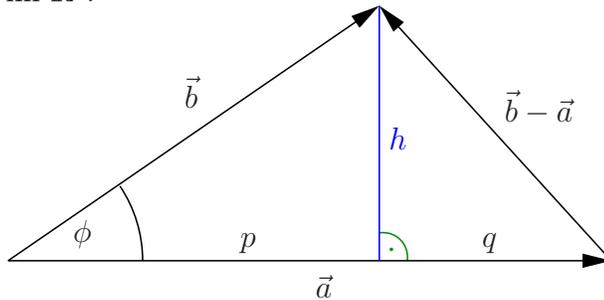
2. Dreiecksungleichung



$$|\vec{a} + \vec{b}| \leq |\vec{a}| + |\vec{b}|$$

3. Winkel zwischen \vec{a} und \vec{b}

- im \mathbb{R}^2 :



$$|\vec{a}| = p + q$$

$$h^2 + q^2 = |\vec{b} - \vec{a}|^2 \quad (5.57)$$

$$\Leftrightarrow h^2 + (|\vec{a}| - p)^2 = |\vec{b} - \vec{a}|^2 \quad (5.58)$$

$$\Leftrightarrow \cancel{h^2} + \cancel{|\vec{a}|^2} + p^2 - 2p|\vec{a}| = \cancel{|\vec{b}|^2} + \cancel{|\vec{a}|^2} - 2\vec{a}\vec{b} \quad | \left(-\frac{1}{2}\right) \quad (5.59)$$

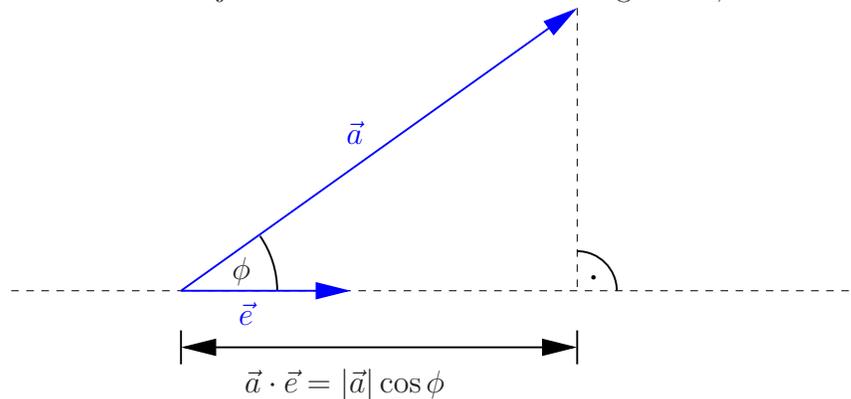
$$\Leftrightarrow \underset{\cos \phi = p/|\vec{b}|}{|\vec{a}| |\vec{b}| \cos \phi} = \vec{a}\vec{b} \quad (5.60)$$

- im \mathbb{R}^n : verlege Skizze in die Ebene, in der \vec{a} und \vec{b} liegen
- allgemein: (CS) stellt sicher, dass $\frac{|\langle \vec{a}, \vec{b} \rangle|}{\|\vec{a}\| \|\vec{b}\|} \leq 1$, d.h. wir können immer $\arccos\left(\frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \|\vec{b}\|}\right)$ als Winkel zwischen den Vektoren \vec{a} , \vec{b} betrachten.

4. \vec{e} heißt Einheitsvektor, falls $\|\vec{e}\| = 1$.

$\vec{a}, \vec{e} \in \mathbb{R}^n$, $\|\vec{e}\| = 1$:

$\vec{a} \cdot \vec{e}$ ist die Projektion von \vec{a} auf die Richtung von \vec{e} , denn



Allgemein für $\vec{a}, \vec{e} \in V$, $\|\vec{e}\| = 1$:

Nenne $\langle \vec{a}, \vec{e} \rangle = \|\vec{a}\| \cos \phi$ die Projektion von \vec{a} auf die Richtung von \vec{e} .

Definition: (Orthogonalität)

Zwei Vektoren $\vec{a}, \vec{b} \in V$ heißen orthogonal zueinander, falls $\langle \vec{a}, \vec{b} \rangle = 0$.

Bemerkungen:

1. Für den Winkel zwischen den Vektoren gilt dann

$$\phi = \arccos\left(\frac{\langle \vec{a}, \vec{b} \rangle}{\|\vec{a}\| \|\vec{b}\|}\right) = \arccos(0) = \frac{\pi}{2} \hat{=} 90^\circ.$$

2. Aus Orthogonalität folgt Lineare Unabhängigkeit (für Vektoren $\neq \vec{0}$)

Seien $\vec{a}_1, \dots, \vec{a}_n \in V$, $\vec{a}_j \neq \vec{0}$, paarweise orthogonal, d.h. $\langle \vec{a}_j, \vec{a}_k \rangle = 0 \forall k \neq j$:

$$\vec{0} = \sum_{j=1}^n \lambda_j \vec{a}_j \tag{5.61}$$

\Rightarrow

$$0 = \left\langle \vec{a}_k, \sum_{j=1}^n \lambda_j \vec{a}_j \right\rangle = \sum_{j=1}^n \lambda_j \underbrace{\langle \vec{a}_k, \vec{a}_j \rangle}_{=\delta_{jk} \|\vec{a}_k\|^2} = \lambda_k \underbrace{\|\vec{a}_k\|^2}_{>0} \tag{5.62}$$

$\Rightarrow \lambda_k = 0 \forall k = 1, \dots, n$, d.h. die Vektoren sind l.u. □

Definition: (ON-Basis)

Eine Basis, deren Elemente paarweise orthogonal sind, heißt Orthogonal-Basis. Sind die Vektoren zusätzlich normiert, d.h. haben sie alle Norm 1, dann bilden sie eine Orthonormal-Basis (ONB).

Kurz: $\{\vec{a}_j\}_{j=1, \dots, \dim V}$ ONB $\Leftrightarrow \langle \vec{a}_j, \vec{a}_k \rangle = \delta_{jk} \forall j, k = 1, \dots, \dim V$

Beispiele: Betrachte $V = \mathbb{R}^2$ mit dem kanonischen Skalarprodukt:

$$1. \vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ bilden eine ONB,} \quad 2. \vec{a}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \vec{a}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ auch.}$$

Basis-Entwicklung eines Vektors:

gegeben: $\vec{a} \in V$ und ONB $\{\vec{c}_j\}$ von V

gesucht: \vec{a} als LK der \vec{c}_j , d.h.

$$\vec{a} = \sum_{j=1}^n \lambda_j \vec{c}_j$$

Lösung: Bilde Skalarprodukt mit \vec{c}_k ,

$$\langle \vec{c}_k, \vec{a} \rangle = \left\langle \vec{c}_k, \sum_{j=1}^n \lambda_j \vec{c}_j \right\rangle = \sum_{j=1}^n \lambda_j \underbrace{\langle \vec{c}_k, \vec{c}_j \rangle}_{=\delta_{jk}} = \lambda_k. \quad (5.63)$$

Beispiel: $\vec{b} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$ mit Basen aus vorherigem Beispiel:

$$\begin{aligned} \vec{b} &= 3\vec{e}_1 - \vec{e}_2 \\ \vec{b} &= \sqrt{2}\vec{a}_1 + 2\sqrt{2}\vec{a}_2 \end{aligned} \quad (5.64)$$

Schmidtsches Orthogonalisierungsverfahren:

gegeben: Basis $\{\vec{a}_1, \dots, \vec{a}_n\}$ von V

gesucht: ON-Basis $\{\vec{c}_1, \dots, \vec{c}_n\}$ von V

$$\vec{c}_1 = \frac{\vec{a}_1}{\|\vec{a}_1\|} \quad (5.65)$$

$$\vec{b}_2 = \vec{a}_2 - \langle \vec{c}_1, \vec{a}_2 \rangle \vec{c}_1 \quad \vec{c}_2 = \frac{\vec{b}_2}{\|\vec{b}_2\|} \quad (5.66)$$

$$\vec{b}_3 = \vec{a}_3 - \langle \vec{c}_1, \vec{a}_3 \rangle \vec{c}_1 - \langle \vec{c}_2, \vec{a}_3 \rangle \vec{c}_2 \quad \vec{c}_3 = \frac{\vec{b}_3}{\|\vec{b}_3\|} \quad (5.67)$$

$$\vdots \quad (5.68)$$

$$\vec{b}_n = \vec{a}_n - \sum_{j=1}^{n-1} \langle \vec{c}_j, \vec{a}_n \rangle \vec{c}_j \quad \vec{c}_n = \frac{\vec{b}_n}{\|\vec{b}_n\|} \quad (5.69)$$

Bemerkung: Sind die Start-Vektoren $\{\vec{a}_1, \dots, \vec{a}_n\}$ l.a. so ergibt sich irgendwann der Nullvektor $\vec{0}$ (evt. mehrfach) – weglassen! \Rightarrow Die $\vec{c}_j \neq \vec{0}$ bilden eine ONB von $\text{span}(\vec{a}_1, \dots, \vec{a}_n)$.

Beispiel:

$$\vec{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{a}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad V = \text{lineare Hülle der beiden} \quad (5.70)$$

ONB bezüglich kanonischem Skalarprodukt:

$$\begin{aligned}\vec{c}_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \\ \vec{b}_2 &= \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} - \frac{3}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \\ \vec{c}_2 &= \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}.\end{aligned}\tag{5.71}$$

5.6 Kreuzprodukt und Spatprodukt im \mathbb{R}^3

Definition: (Kreuzprodukt)

Für $\vec{a}, \vec{b} \in \mathbb{R}^3$ definieren wir das Kreuzprodukt (oder Vektorprodukt) durch

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} := \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.\tag{5.72}$$

Satz 13. (Eigenschaften von \times)

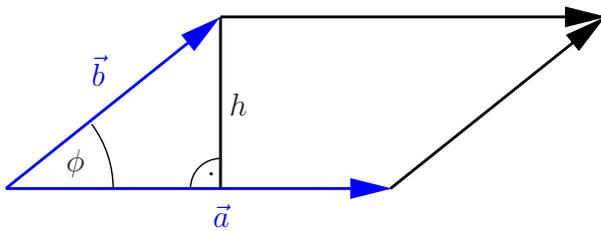
$\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3, \lambda \in \mathbb{R}$:

- (i) $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$ *(antikommutativ)*
- (ii) $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$ *(Distributivgesetz)*
- (iii) $\lambda(\vec{a} \times \vec{b}) = (\lambda\vec{a}) \times \vec{b} = \vec{a} \times (\lambda\vec{b})$
- (iv) $|\vec{a} \times \vec{b}| = \text{Fläche des von } \vec{a} \text{ und } \vec{b} \text{ aufgespannten Parallelogramms}$
- (v) $\vec{a} \times \vec{b}$ ist orthogonal zu \vec{a} und zu \vec{b}
(bzgl. des kanonischen Skalarprodukts)
- (vi) falls $\vec{a} \times \vec{b} \neq \vec{0}$, bilden \vec{a}, \vec{b} und $\vec{a} \times \vec{b}$ ein Rechtssystem *(rechte Hand-Regel)*

Bemerkung: (ii) & (iii): Linearität

Beweis:

- (i) klar lt. Def. (vertausche b_j und a_j)
übrigens: $\Rightarrow \vec{a} \times \vec{a} = \vec{0}$
- (ii) & (iii) explizit nachrechnen
- (iv) Fläche:



$$\begin{aligned}
 F &= |\vec{a}|h \\
 &= |\vec{a}||\vec{b}|\sin\phi \\
 &= |\vec{a}||\vec{b}|\sqrt{1-\cos^2\phi} \\
 &= \sqrt{|\vec{a}|^2|\vec{b}|^2 - |\vec{a}\vec{b}|^2}
 \end{aligned} \tag{5.73}$$

einerseits:

$$\begin{aligned}
 (\vec{a} \times \vec{b})^2 &= (a_2b_3)^2 + (a_3b_2)^2 - 2a_2b_2a_3b_3 \\
 &\quad + (a_3b_1)^2 + (a_1b_3)^2 - 2a_3b_3a_1b_1 \\
 &\quad + (a_1b_2)^2 + (a_2b_1)^2 - 2a_1b_1a_2b_2
 \end{aligned} \tag{5.74}$$

andererseits:

$$\begin{aligned}
 \vec{a}^2\vec{b}^2 - (\vec{a}\vec{b})^2 &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \\
 &= \cancel{a_1^2b_1^2} + a_1^2b_2^2 + a_1^2b_3^2 \\
 &\quad + a_2^2b_1^2 + \cancel{a_2^2b_2^2} + a_2^2b_3^2 \\
 &\quad + a_3^2b_1^2 + a_3^2b_2^2 + \cancel{a_3^2b_3^2} \\
 &\quad - \cancel{a_1^2b_1^2} - \cancel{a_2^2b_2^2} - \cancel{a_3^2b_3^2} \\
 &\quad - 2a_2b_2a_3b_3 - 2a_3b_3a_1b_1 - 2a_1b_1a_2b_2
 \end{aligned} \tag{5.75}$$

also $F = |\vec{a} \times \vec{b}|$

(v)

$$\begin{aligned}
 \vec{a}(\vec{a} \times \vec{b}) &= a_1a_2b_3 - a_1a_3b_2 + a_2a_3b_1 - a_2a_1b_3 + a_3a_1b_2 - a_3a_2b_1 = 0 \\
 \vec{b}(\vec{a} \times \vec{b}) &= -\vec{b}(\vec{b} \times \vec{a}) = 0
 \end{aligned} \tag{5.76}$$

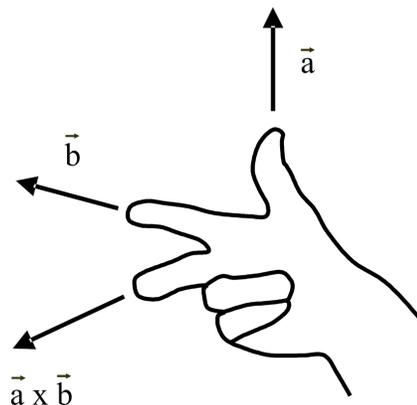
(vi) Zunächst überprüfen wir: $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$
(sowie $\vec{e}_2 \times \vec{e}_3 = \vec{e}_1$ und $\vec{e}_3 \times \vec{e}_1 = \vec{e}_2$)

Allgemein:

Nenne die Richtung von \vec{a} die x_1 -Richtung,

$$\text{bilde } \vec{a} \times \vec{b} = \vec{a} \times \left(\underbrace{\vec{b} - \frac{(\vec{a} \cdot \vec{b})}{|\vec{a}|} \frac{\vec{a}}{|\vec{a}|}}_{\text{nenne } x_2\text{-Richtung}} \right),$$

dann zeigt $\vec{a} \times \vec{b}$ wieder in x_3 -Richtung.



Definition: (Spatprodukt)

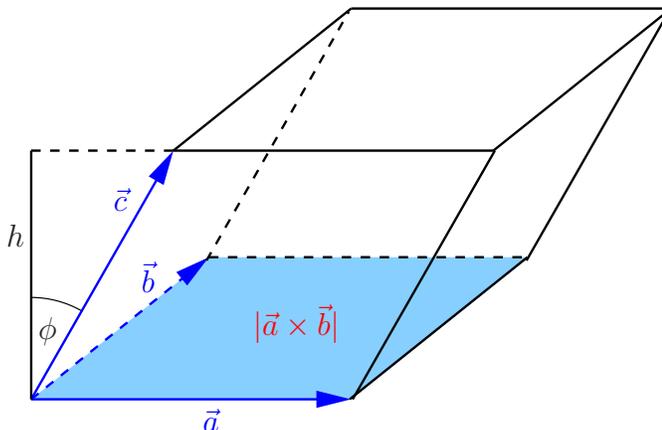
Die Abbildung $|\cdot, \cdot, \cdot| : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}$,

$$|\vec{a}, \vec{b}, \vec{c}| = (\vec{a} \times \vec{b}) \cdot \vec{c} \quad (5.77)$$

heißt Spatprodukt.

Eigenschaften:

1. $|\vec{a}, \vec{b}, \vec{c}| = |\vec{b}, \vec{c}, \vec{a}| = |\vec{c}, \vec{a}, \vec{b}|$
2. $|\vec{b}, \vec{a}, \vec{c}| = -|\vec{a}, \vec{b}, \vec{c}|$
3. Der Betrag von $|\vec{a}, \vec{b}, \vec{c}|$ ist gleich dem Volumen des, von den Vektoren aufgespannten, Parallelepipeds bzw. Spats.



Beweis:

1. explizit nachrechnen
2. folgt aus $\vec{b} \times \vec{a} = -\vec{a} \times \vec{b}$
3. $||\vec{a}, \vec{b}, \vec{c}|| = \underbrace{|\vec{a} \times \vec{b}|}_{\text{Grundfläche}} \underbrace{|\vec{c}| \cos \phi}_{\text{Höhe}} = V$

5.7 Geraden und Ebenen

Geraden:

- (i) gegeben: Punkt $\vec{p}_1 \in \mathbb{R}^n$ und Richtung $\vec{v} \in \mathbb{R}^n$,

$$g = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} = \vec{p}_1 + \vec{v}t, t \in \mathbb{R}\} \quad (5.78)$$

- (ii) gegeben: zwei Punkte $\vec{p}_1, \vec{p}_2 \in \mathbb{R}^n$,

$$g = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} = \vec{p}_1 + (\vec{p}_2 - \vec{p}_1)t, t \in \mathbb{R}\} \quad (5.79)$$

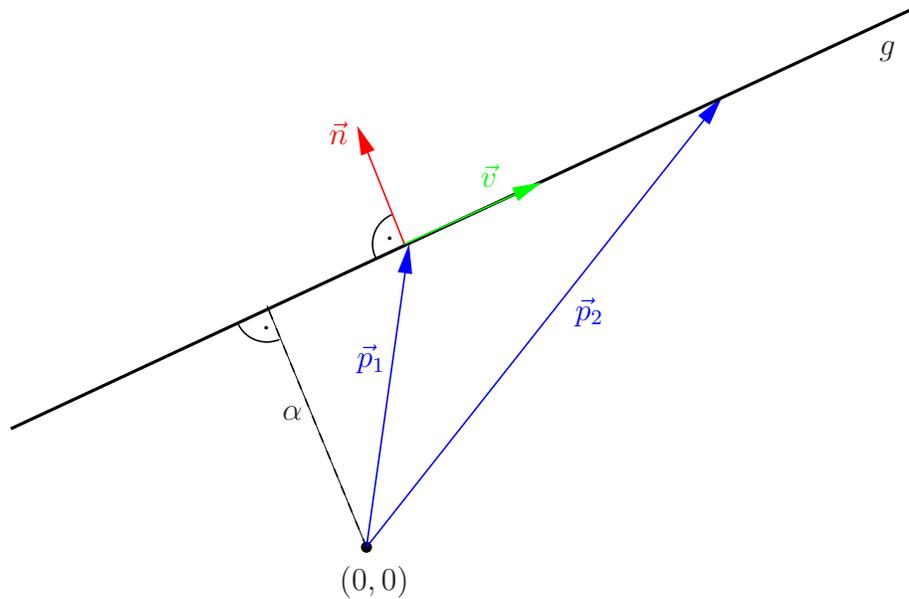
- (iii) **speziell im \mathbb{R}^2**

gegeben: Punkt \vec{p}_1 und Normalenrichtung \vec{n} (kurz: $\vec{n} \perp g$)

$$\begin{aligned} g &= \{\vec{x} \in \mathbb{R}^2 \mid (\vec{x} - \vec{p}_1)\vec{n} = 0\} \\ &= \{\vec{x} \in \mathbb{R}^2 \mid \vec{x}\vec{n} = \alpha, \alpha = \vec{p}_1\vec{n}\} . \end{aligned} \quad (5.80)$$

- parameterfreie Darstellung

- Falls $|\vec{n}| = 1$, dann ist $|\alpha|$ der Abstand der Gerade vom Ursprung.



Ebenen:

- (i) gegeben: Punkt $\vec{p}_1 \in \mathbb{R}^n$ und 2 Richtungen $\vec{v}, \vec{w} \in \mathbb{R}^n$ l.u.,

$$E = \{ \vec{x} \in \mathbb{R}^n \mid \vec{x} = \vec{p}_1 + \vec{v}s + \vec{w}t, s, t \in \mathbb{R} \} \quad (5.81)$$

- (ii) gegeben: zwei Punkte $\vec{p}_1, \vec{p}_2 \in \mathbb{R}^n$ und eine Richtung $\vec{v} \in \mathbb{R}^n$ mit $\vec{p}_2 - \vec{p}_1, \vec{v}$ l.u.,

$$E = \{ \vec{x} \in \mathbb{R}^n \mid \vec{x} = \vec{p}_1 + (\vec{p}_2 - \vec{p}_1)s + \vec{v}t, s, t \in \mathbb{R} \} \quad (5.82)$$

- (iii) gegeben: drei Punkte $\vec{p}_1, \vec{p}_2, \vec{p}_3 \in \mathbb{R}^n$ mit $\vec{p}_2 - \vec{p}_1, \vec{p}_3 - \vec{p}_1$ l.u.,

$$E = \{ \vec{x} \in \mathbb{R}^n \mid \vec{x} = \vec{p}_1 + (\vec{p}_2 - \vec{p}_1)s + (\vec{p}_3 - \vec{p}_1)t, s, t \in \mathbb{R} \} \quad (5.83)$$

- (iv) **speziell im \mathbb{R}^3**

gegeben: Punkt \vec{p}_1 und $\vec{n} \perp E$,

$$\begin{aligned} E &= \{ \vec{x} \in \mathbb{R}^3 \mid (\vec{x} - \vec{p}_1)\vec{n} = 0 \} \\ &= \{ \vec{x} \in \mathbb{R}^3 \mid \vec{x}\vec{n} = \alpha, \alpha = \vec{p}_1\vec{n} \} . \end{aligned} \quad (5.84)$$

- parameterfreie Darstellung
- Falls $|\vec{n}| = 1$, dann ist $|\alpha|$ der Abstand der Ebene vom Ursprung.
- Darstellung heißt Hessesche Normalform, falls $|\vec{n}| = 1$ und \vec{n} so, dass $\alpha \geq 0$.

Geraden im \mathbb{R}^3 : Schnitte zweier nicht paralleler Ebenen, z.B. parameterfreie Darstellung,

$$g = \{ \vec{x} \in \mathbb{R}^3 \mid \vec{x}\vec{n}_1 = \alpha_1, \vec{x}\vec{n}_2 = \alpha_2 \} , \quad (5.85)$$

mit \vec{n}_1, \vec{n}_2 l.u.

5.8 Kurven und Spezielle Koordinatensysteme

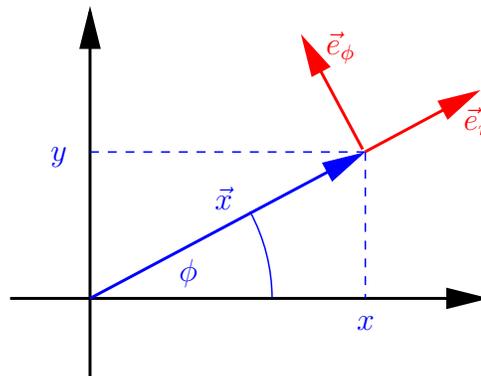
Polarkoordinaten im \mathbb{R}^2 :

$\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ lässt sich auch durch Radius $r \in [0, \infty)$ (Abstand zum Ursprung) und Winkel $\phi \in [0, 2\pi)$ charakterisieren,

$$\begin{aligned} r &= \sqrt{x^2 + y^2}, \\ \phi &= \arctan \frac{y}{x} \quad (\text{Zweig beachten}). \end{aligned} \quad (5.86)$$

Umgekehrt:

$$x = r \cos \phi, \quad y = r \sin \phi \quad (5.87)$$



Bemerkung: Für $(0, 0)$ ist ϕ nicht definiert.

Einheitsvektoren (an jedem Punkt – mitgeführtes Koordinatensystem):

$$\vec{e}_r = \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix}, \quad \vec{e}_\phi = \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix} \quad (5.88)$$

mit

$$\begin{aligned} |\vec{e}_r|^2 &= \cos^2 \phi + \sin^2 \phi = 1, & |\vec{e}_\phi|^2 &= 1, \\ \vec{e}_r \cdot \vec{e}_\phi &= \cos \phi (-\sin \phi) + \sin \phi \cos \phi = 0, \end{aligned} \quad (5.89)$$

also ONB, und

$$\vec{x} = \begin{pmatrix} r \cos \phi \\ r \sin \phi \end{pmatrix} = r \vec{e}_r \quad (5.90)$$

Definition: (Kurven im \mathbb{R}^n)

Eine Abbildung

$$\begin{aligned} \vec{x}: \quad \mathbb{R} \supseteq I &\rightarrow \mathbb{R}^n \\ t &\mapsto \vec{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} \end{aligned} \quad (5.91)$$

heißt Kurve im \mathbb{R}^n .

Die (Momentan-)Geschwindigkeit (entlang der Kurve) ist die (komponentenweise) Ableitung nach t ,

$$\dot{\vec{x}}(t) = \frac{d}{dt} \vec{x}(t) = \begin{pmatrix} \dot{x}_1 \\ \vdots \\ \dot{x}_n \end{pmatrix}. \quad (5.92)$$

Analog: Beschleunigung $\ddot{\vec{x}}(t) = \frac{d^2}{dt^2} \vec{x}(t)$.

Beispiele:

1.

$$\begin{aligned} \vec{x}(t) &= \vec{x}_0 + \vec{v}t, & \vec{x}_0, \vec{v} \in \mathbb{R}^3 \text{ fest} \\ \Rightarrow \dot{\vec{x}}(t) &= \vec{v}, & \ddot{\vec{x}}(t) = \vec{0} \end{aligned} \quad (5.93)$$

(Gerade, gleichförmige Bewegung)

2.

$$\begin{aligned} \vec{x}(t) &= \frac{1}{2}\vec{a}t^2, & \vec{a} \in \mathbb{R}^3 \text{ fest} \\ \Rightarrow \dot{\vec{x}}(t) &= \vec{a}t, & \ddot{\vec{x}}(t) = \vec{a} \end{aligned} \quad (5.94)$$

(Gerade, gleichförmig beschleunigte Bewegung)

3. Kurve im \mathbb{R}^2 gegeben durch $r(t)$ und $\phi(t)$:

$$\vec{x}(t) = \begin{pmatrix} r(t) \cos(\phi(t)) \\ r(t) \sin(\phi(t)) \end{pmatrix} \quad (5.95)$$

Geschwindigkeit:

$$\begin{aligned} \dot{\vec{x}}(t) &= \begin{pmatrix} \dot{r}(t) \cos(\phi(t)) - r(t) \sin(\phi(t)) \dot{\phi}(t) \\ \dot{r}(t) \sin(\phi(t)) + r(t) \cos(\phi(t)) \dot{\phi}(t) \end{pmatrix} \\ &= \dot{r}(t) \begin{pmatrix} \cos(\phi(t)) \\ \sin(\phi(t)) \end{pmatrix} + r(t) \dot{\phi}(t) \begin{pmatrix} -\sin(\phi(t)) \\ \cos(\phi(t)) \end{pmatrix} \\ &= \dot{r}(t) \vec{e}_r + r(t) \dot{\phi}(t) \vec{e}_\phi \end{aligned} \quad (5.96)$$

Beispiel zum Beispiel: $\phi(t) = \omega t$, $r(t) = R$, ω, R fest

$$\dot{\vec{x}}(t) = 0 \vec{e}_r + R\omega \vec{e}_\phi \quad (\text{bzw. } \vec{e}_{\phi(t)=\omega t}) \quad (5.97)$$

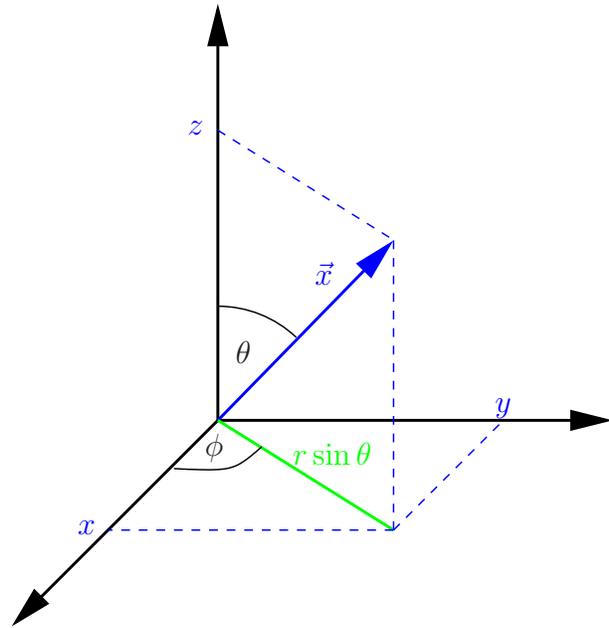
(Kreis, mit Winkelgeschwindigkeit ω durchlaufen)

Kugelkoordinaten im \mathbb{R}^3 :

$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ lässt sich durch Radius $r = |\vec{x}|$ und zwei Winkel, θ und ϕ , charakterisieren,

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \sin \theta \cos \phi \\ r \sin \theta \sin \phi \\ r \cos \theta \end{pmatrix}, \quad (5.98)$$

$r \in [0, \infty)$, $\theta \in [0, \pi]$, $\phi \in [0, 2\pi)$.



Umgekehrt:

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2}, \\ \phi &= \arctan \frac{y}{x} \quad (\text{Zweig beachten!}), \\ \theta &= \arccos \frac{z}{\sqrt{x^2 + y^2 + z^2}} \quad (\text{da } \frac{z}{r} = \cos \theta). \end{aligned} \quad (5.99)$$

6 Matrizen und Determinanten

6.1 Matrizen

Definition: Eine $m \times n$ -Matrix A ist ein rechteckiges Zahlenschema,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad (6.1)$$

mit m Zeilen und n Spalten. Die $a_{ij} \in \mathbb{R}$ (oder $\in \mathbb{C}$) heißen Elemente (oder Komponenten) der Matrix. Wir schreiben

$$A = (a_{ij}) \quad (6.2)$$

und sagen $A \in \mathbb{R}^{m \times n}$ (bzw. $A \in \mathbb{C}^{m \times n}$).

Bemerkungen:

1. Vektoren $\in \mathbb{R}^n$ lassen sich als Matrizen auffassen, entweder als Spaltenvektor,

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^{n \times 1}, \quad (6.3)$$

oder als Zeilenvektor,

$$(x_1, x_2, \dots, x_n) \in \mathbb{R}^{1 \times n}. \quad (6.4)$$

2. Die Addition zweier Matrizen gleichen Typs definieren wir komponentenweise,

$$\begin{aligned} A &= (a_{ij}), & B &= (b_{ij}) \\ C &= A + B & \text{mit } C &= (c_{ij}), \quad c_{ij} = a_{ij} + b_{ij}, \end{aligned} \quad (6.5)$$

z.B.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 6 & 8 \\ 10 & 12 \end{pmatrix}. \quad (6.6)$$

3. Ebenso die Multiplikation mit Skalaren,

$$\begin{aligned} A &= (a_{ij}), \quad \lambda \in \mathbb{R} \\ D &= \lambda A \quad \text{mit } D = (d_{ij}), \quad d_{ij} = \lambda a_{ij}, \end{aligned} \quad (6.7)$$

z.B.

$$2 \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \end{pmatrix} \quad (6.8)$$

4. Damit wird $\mathbb{R}^{m \times n}$ Vektorraum über \mathbb{R} mit den üblichen Rechenregeln und Dimension $\dim \mathbb{R}^{m \times n} = mn$.

Neutrales Element der Addition bzw. "Nullvektor" ist die Nullmatrix:

$$0 = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}. \quad (6.9)$$

Definition: Für zwei Matrizen $A = (a_{ij}) \in \mathbb{R}^{m \times \ell}$ und $B = (b_{ij}) \in \mathbb{R}^{\ell \times n}$ definieren wir das Matrixprodukt durch

$$C = AB \quad \text{mit} \quad C = (c_{ij}), \quad c_{ij} = \sum_{k=1}^{\ell} a_{ik} b_{kj}, \quad (6.10)$$

kurz: *Zeile mal Spalte*.

Bemerkung: Ebenso wird für $A \in \mathbb{C}^{m \times \ell}$ und $B \in \mathbb{C}^{\ell \times n}$ das Produkt $AB \in \mathbb{C}^{m \times n}$ definiert.

Beispiel:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 3}, \quad B = \begin{pmatrix} 1 & 4 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 6 & 0 & 1 \end{pmatrix} \in \mathbb{R}^{3 \times 4} \quad (6.11)$$

$$AB = \begin{pmatrix} 1 & 32 & 2 & 3 \\ 0 & 5 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 4}$$

zum Rechnen:

$$\begin{array}{c|c} & \begin{pmatrix} 1 & 4 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 6 & 0 & 1 \end{pmatrix} \\ \hline \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 32 & 2 & 3 \\ 0 & 5 & 1 & 0 \end{pmatrix} \end{array} \quad (6.12)$$

$$32 = 1 \cdot 4 + 2 \cdot 5 + 3 \cdot 6$$

BA ist hier gar nicht definiert!

Bemerkungen:

1. Ist z.B. $A \in \mathbb{R}^{2 \times 3}$ und $B \in \mathbb{R}^{3 \times 2}$ so existieren zwar sowohl AB als auch BA , haben aber unterschiedliche Form: $AB \in \mathbb{R}^{2 \times 2}$, $BA \in \mathbb{R}^{3 \times 3}$.
2. Sind A und B quadratisch, so gilt i.A. trotzdem $AB \neq BA$ (nichtkommutativ), z.B.

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 0 \end{pmatrix}, \quad (6.13)$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 1 & 0 \end{pmatrix}.$$

3.

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n \times n} \quad (6.14)$$

heißt Einheitsmatrix, mit Elementen

$$I = (\delta_{ij}), \quad \delta_{jk} = \begin{cases} 1 & \text{für } j = k \\ 0 & \text{sonst} \end{cases}, \quad (\text{Kronecker-}\delta) \quad (6.15)$$

Für jede quadratische Matrix $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ (oder $\in \mathbb{C}^{n \times n}$) gilt

$$AI = IA = A, \quad (6.16)$$

denn, z.B. mit $C = (c_{ij}) = AI$,

$$c_{ij} = \sum_k a_{ik} \delta_{kj} = a_{ij}. \quad (6.17)$$

Rechenregeln:

1. $(A_1 + A_2)B = A_1B + A_2B$
 $A(B_1 + B_2) = AB_1 + AB_2$
2. $\lambda(AB) = (\lambda A)B = A(\lambda B)$
3. $A(BC) = (AB)C$

wobei $\lambda \in \mathbb{R}$, alles andere Matrizen, Typen so, dass alle Produkte definiert.

Beweis: Nachrechnen.

Definition: Sei $A = (a_{ij}) \in \mathbb{R}^{m \times n}$. Die Matrix $A^T = (a_{ij}^T) \in \mathbb{R}^{n \times m}$ mit Elementen

$$a_{ij}^T = a_{ji} \quad (6.18)$$

heißt Transponierte von A .

Beispiele:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = (1 \ 2 \ 3)^T = (1, 2, 3)^T \quad (6.19)$$

Rechenregeln: ($\lambda \in \mathbb{R}$, A, B Matrizen – so, dass AB definiert)

1. $(A + B)^T = A^T + B^T$
2. $(\lambda A)^T = \lambda A^T$
3. $(A^T)^T = A$

4. $(AB)^T = B^T A^T$, denn, mit $C = AB$

$$c_{ij}^T = c_{ji} = \sum_k a_{jk} b_{ki} = \sum_k a_{kj}^T b_{ik}^T = \sum_k b_{ik}^T a_{kj}^T, \quad (6.20)$$

also $C^T = B^T A^T$. □

Anwendung (Matrixprodukt): Lineare Gleichungssysteme.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned} \quad (6.21)$$

ist äquivalent zu

$$A\vec{x} = \vec{b} \quad (6.22)$$

mit $A = (a_{ij}) \in \mathbb{R}^{m \times n}$, $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ bzw. $\in \mathbb{R}^{n \times 1}$ und $\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m$
(vgl. Kurzschreibweise für LGS).

Spezialfall: $A \in \mathbb{R}^{n \times n}$, $\vec{x}, \vec{b} \in \mathbb{R}^n$: quadratisches LGS (n Gleichungen für n Unbekannte),

$$A\vec{x} = \vec{b} \quad (6.23)$$

Jetzt wäre es schön, wenn wir durch A "teilen" könnten, um \vec{x} zu bekommen.

Definition: Sei $A \in \mathbb{R}^{n \times n}$. Wir nennen A^{-1} die zu A inverse Matrix, falls gilt

$$A^{-1}A = I \quad \text{und} \quad AA^{-1} = I. \quad (6.24)$$

Bemerkungen:

1. Nicht alle Matrizen sind invertierbar, z.B.

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \Rightarrow \quad A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (6.25)$$

Annahme: $\exists B \in \mathbb{R}^{2 \times 2}$ mit

$$\begin{array}{l} BA = I \quad | \cdot A \text{ von rechts} \\ B \underbrace{A^2}_{=0} = A \quad \text{Widerspruch!} \end{array} \quad (6.26)$$

2. Falls existent, so ist Inverse eindeutig bestimmt, denn:

Annahme: $BA = I$ und $AC = I$. (z.z.: $B = C$)

$$B = BI = B(AC) = (BA)C = IC = C \quad \square \quad (6.27)$$

3. quadratisches LGS $A\vec{x} = \vec{b}$, A invertierbar $\Rightarrow \vec{x} = A^{-1}\vec{b}$, eindeutig!

Berechnung: $A \in \mathbb{R}^{n \times n}$, betrachte

$$AX = I \tag{6.28}$$

falls lösbar, dann $X = A^{-1}$. Spaltenweise:

$$X = (\vec{x}_1, \dots, \vec{x}_n), \quad I = (\vec{e}_1, \dots, \vec{e}_n), \tag{6.29}$$

d.h. $\vec{x}_i = i$ -te Spalte von X . Es gilt

$$AX = (A\vec{x}_1, \dots, A\vec{x}_n) \stackrel{!}{=} (\vec{e}_1, \dots, \vec{e}_n) \tag{6.30}$$

d.h. wir lösen simultan die n LGSe $A\vec{x}_i = \vec{e}_i$. Praktisch: Gauß-Algorithmus

$$\frac{\left(\begin{array}{c|c} A & I \end{array} \right)}{\vdots} \frac{\left(\begin{array}{c|c} I & A^{-1} \end{array} \right)}{\tag{6.31}}$$

Beispiel:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{6.32}$$

$$\begin{array}{c} \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array} \right) \begin{array}{l} \leftarrow -\frac{c}{a} \\ \rightarrow + \end{array} \cdot 1/a \\ \hline \left(\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & \frac{ad-bc}{a} & -\frac{c}{a} & 1 \end{array} \right) \cdot \frac{a}{ad-bc} \\ \hline \left(\begin{array}{cc|cc} 1 & \frac{b}{a} & \frac{1}{a} & 0 \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array} \right) \begin{array}{l} \leftarrow + \\ \rightarrow -\frac{b}{a} \end{array} \\ \hline \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{a} + \frac{bc}{a(ad-bc)} & -\frac{b}{ad-bc} \\ 0 & 1 & -\frac{c}{ad-bc} & \frac{a}{ad-bc} \end{array} \right) \end{array} \tag{6.33}$$

$$\frac{1}{a} + \frac{bc}{a(ad-bc)} = \frac{1}{a} \frac{ad-bc+bc}{ad-bc} = \frac{d}{ad-bc} \tag{6.34}$$

also

$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \quad \text{falls } ad-bc \neq 0. \tag{6.35}$$

Wir nennen die Zahl

$$\det A = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc \tag{6.36}$$

die Determinante von A .

Sie bestimmt, ob A invertierbar ist ($\det A \neq 0$) oder nicht ($\det A = 0$).

6.2 Determinanten

$$\det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = a_1 b_2 - a_2 b_1 = \left[\begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} \right]_3 = \left[\begin{pmatrix} 0 \\ 0 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \right]_3$$

ist also Fläche (mit Vorzeichen) des, von den Vektoren

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2$$

aufgespannten, Parallelogramms (vgl. Kreuzprodukt). Vorzeichen:

$$\det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1, \quad \det(\vec{a}, \vec{b}) = -\det(\vec{b}, \vec{a}).$$

Verallgemeinerung auf höhere Dimension? Suche Abbildung

$$\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}, \tag{6.37}$$

die das Volumen des, von den Spaltenvektoren der Matrix aufgespannten Parallelepipeds liefert (mit Orientierung, d.h. VZ) \Rightarrow Gewünschte Eigenschaften:

(i) Linearität in jeder Spalte, insbesondere

$$\det(\vec{a}_1, \dots, \lambda \vec{a}_j, \dots, \vec{a}_n) = \lambda \det(\vec{a}_1, \dots, \vec{a}_j, \dots, \vec{a}_n) \quad \forall j = 1, \dots, n$$

(ii) Normierung

$$\det I = \det(\vec{e}_1, \dots, \vec{e}_n) = 1$$

(Volumen des n -dimensionalen Einheitswürfels: $1 \cdot 1 \cdot \dots \cdot 1 = 1^n = 1$)

(iii) Alternierend, d.h. wechselt das Vorzeichen, wenn wir zwei Spalten vertauschen.

Folgerungen:

(i) \Rightarrow

$$\det A = \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \sum_{j_1, \dots, j_n=1}^n C_{j_1 \dots j_n} a_{j_1 1} a_{j_2 2} \cdots a_{j_n n}$$

(Jeder Summand enthält genau ein Element aus jeder Spalte.)

(ii) \Rightarrow

$$\det I = C_{123\dots n} = 1 \tag{6.38}$$

(iii) :

$$\det(\vec{e}_{j_1}, \dots, \vec{e}_{j_n}) = C_{j_1 \dots j_n} \tag{6.39}$$

vertausche zwei Spalten

$$\begin{aligned} C_{\dots j_k \dots j_l \dots} &= \det(\dots, \vec{e}_{j_k}, \dots, \vec{e}_{j_l}, \dots) \\ &= -\det(\dots, \vec{e}_{j_l}, \dots, \vec{e}_{j_k}, \dots) = -C_{\dots j_l \dots j_k \dots} \end{aligned} \quad (6.40)$$

Sind nun zwei der Indizes gleich (also $j_l = j_k, l \neq k$), dann ist das zugehörige C Null, d.h. jeder Summand enthält jetzt genau ein Element aus jeder Spalte **und** jeder Zeile. Weiter folgt mit $j_l \neq j_k$: Beträge aller $C_{j_1 \dots j_n}$ sind gleich (= 1).

Damit sind alle C festgelegt!

Beispiele:

- $n = 2$: $C_{11} = C_{22} = 0, \quad C_{12} = 1, \quad C_{21} = -1$ (passt, s.o.)
- $n = 3$

$$\begin{aligned} C_{123} &= 1 && \text{vertausche 1. und 2. Index} \\ C_{213} &= -1 && \text{vertausche 2. und 3. Index} \\ C_{231} &= 1 && \text{etc.} \\ C_{321} &= -1 && \\ C_{312} &= 1 && \\ C_{132} &= -1 && \end{aligned} \quad (6.41)$$

Damit erhalten wir tatsächlich das Spatprodukt, denn

$$\begin{aligned} \det(\vec{a}, \vec{b}, \vec{c}) &= a_1 b_2 c_3 - a_2 b_1 c_3 + a_2 b_3 c_1 - a_3 b_2 c_1 + a_3 b_1 c_2 - a_1 b_3 c_2 \\ &= \underbrace{\begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}}_{=\vec{a} \times \vec{b}} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = |\vec{a}, \vec{b}, \vec{c}|. \end{aligned} \quad (6.42)$$

Mit etwas Indexgymnastik ergeben sich folgende weitere

Eigenschaften der Determinante:

1. $\det A^T = \det A$
2. für (obere) Dreiecksmatrizen: Produkt der Diagonalelemente

$$\det \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} = a_{11} a_{22} \cdots a_{nn} \quad (6.43)$$

3. Spalten oder Zeilen vertauschen: Faktor (-1)
insbesondere: Sind zwei Spalten (Zeilen) gleich, so verschwindet die Determinante.

4. det ist linear in jeder Zeile und Spalte, vgl. Forderung (i), insbesondere

$$\det(\vec{a}_1, \dots, \vec{a}_j + \vec{b}, \dots, \vec{a}_n) = \det(\vec{a}_1, \dots, \vec{a}_j, \dots, \vec{a}_n) + \det(\vec{a}_1, \dots, \vec{b}, \dots, \vec{a}_n). \quad (6.44)$$

Ebenso für Zeilen.

5. Vielfaches einer Spalte (Zeile) zu einer anderen: det invariant, denn (4. mit $\vec{b} = \lambda \vec{a}_k$, $k \neq j$)

$$\det(\vec{a}_1, \dots, \vec{a}_j + \lambda \vec{a}_k, \dots, \vec{a}_n) = \det(\vec{a}_1, \dots, \vec{a}_j, \dots, \vec{a}_n) + \lambda \det(\vec{a}_1, \dots, \underbrace{\vec{a}_k}_{j\text{-te Spalte}}, \dots, \vec{a}_n). \quad (6.45)$$

=0, da \vec{a}_k doppelt vorkommt

also Gauß (ohne Durchmultiplizieren)

Beispiel:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ -1 & 2 & 3 \\ 1 & 1 & 5 \end{pmatrix} \begin{array}{l} \left[\begin{array}{l} \leftarrow_+ \\ \leftarrow_+ \end{array} \right]^{-1} \\ \leftarrow_+ \end{array} \quad (6.46)$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix}$$

d.h. $\det A = 1 \cdot 3 \cdot 4 = 12$.

6. $\det A \neq 0$

\Leftrightarrow Spalten von A sind l.u. (ebenso Zeilen)

$\Leftrightarrow A$ invertierbar

\Leftrightarrow LGS $A\vec{x} = \vec{b}$ hat für jedes \vec{b} eine eindeutige Lösung, nämlich $\vec{x} = A^{-1}\vec{b}$

Satz 14. (Determinantenmultiplikationssatz)

Seien $A, B \in \mathbb{R}^{n \times n}$ (oder $\in \mathbb{C}^{n \times n}$). Dann gilt

$$\det(AB) = \det A \cdot \det B. \quad (6.47)$$

Beweis:

$$\begin{aligned} \det(AB) &= \det(A\vec{b}_1, \dots, A\vec{b}_n), \quad \text{wobei } B = (\vec{b}_1, \dots, \vec{b}_n) = (b_{ij}), \text{ d.h. } \vec{b}_j = \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} \\ &= \det \left(\sum_{i_1=1}^n \vec{a}_{i_1} b_{i_1 1}, \dots, \sum_{i_n=1}^n \vec{a}_{i_n} b_{i_n n} \right), \quad \text{wobei } A = (\vec{a}_1, \dots, \vec{a}_n) \\ &= \sum_{i_1, \dots, i_n=1}^n b_{i_1 1} \cdots b_{i_n n} \underbrace{\det(\vec{a}_{i_1}, \dots, \vec{a}_{i_n})}_{=C_{i_1 \dots i_n} \det A} \\ &= \det A \underbrace{\sum_{i_1, \dots, i_n=1}^n b_{i_1 1} \cdots b_{i_n n} C_{i_1 \dots i_n}}_{=\det B} \end{aligned} \quad (6.48)$$

□

Bemerkung: Damit gilt auch (falls A invertierbar)

$$\det A^{-1} = \frac{1}{\det A}, \quad (6.49)$$

denn $\det A \cdot \det A^{-1} = \det(AA^{-1}) = \det I = 1$.

Satz 15. (Laplacescher Entwicklungssatz)

Sei $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. Die Matrix $A_{ij} \in \mathbb{R}^{(n-1) \times (n-1)}$ entstehe aus A durch Streichen der i -ten Zeile und der j -ten Spalte. Es gilt:

$$\begin{aligned} \det A &= \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij} \quad (\text{Entwicklung nach der } j\text{-ten Spalte}) \\ &= \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij} \quad (\text{Entwicklung nach der } i\text{-ten Zeile}) \end{aligned} \quad (6.50)$$

(folgt unmittelbar aus der Definition der Determinante)

Bemerkungen:

1. Vorzeichen sind schachbrettartig verteilt:

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{pmatrix} \quad (6.51)$$

2. Vor allem sinnvoll, für Zeilen (Spalten) mit vielen Nullen.

Beispiel:

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = 1 \det \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix} - 4 \det \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} + 7 \det \begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix}$$

(Entwicklung nach der 1. Spalte)

$$= -4 \det \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix} + 5 \det \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix} - 6 \det \begin{pmatrix} 1 & 2 \\ 7 & 8 \end{pmatrix}$$

(Entwicklung nach der 2. Zeile)

(6.52)

7 Komplexe Zahlen

Führe eine Zahl i ein, mit

$$i^2 = -1, \quad (7.1)$$

bzw. auch $i = \sqrt{-1}$. Komplexe Zahlen

$$\mathbb{C} = \{z = x + iy \mid x, y \in \mathbb{R}\} \quad (7.2)$$

Rechnen wie mit reellen Zahlen, d.h. $z_j = x_j + iy_j$, $x_j, y_j \in \mathbb{R}$,

$$\begin{aligned} z_1 + z_2 &= x_1 + iy_1 + x_2 + iy_2 \\ &= (x_1 + x_2) + i(y_1 + y_2) \\ z_1 \cdot z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\ &= x_1x_2 + ix_1y_2 + iy_1x_2 + i^2y_1y_2 \\ &= (x_1x_2 - y_1y_2) + i(x_1y_2 + y_1x_2) \end{aligned} \quad (7.3)$$

$\bar{z} = x - iy$ heißt **komplex konjugiert** zu $z = x + iy$. Es folgt

$$\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2, \quad (7.4)$$

und insbesondere $\overline{z^n} = \bar{z}^n$ ($n \in \mathbb{N}_0$).

$(\mathbb{C}, +, \cdot)$ ist ein Körper mit

- neutralem Element der Addition: $0 = 0 + i0$
- neutralem Element der Multiplikation: $1 = 1 + i0$
- additiv Inversem (zu $z = x + iy$, $x, y \in \mathbb{R}$): $-z = -x - iy$
- multiplikativ Inversem (zu $z \neq 0$, $z = x + iy$, $x, y \in \mathbb{R}$):

$$z^{-1} = \frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{x - iy}{(x + iy)(x - iy)} = \frac{x - iy}{x^2 + y^2} \quad (\text{insbesondere: } i^{-1} = \frac{1}{i} = -i). \quad (7.5)$$

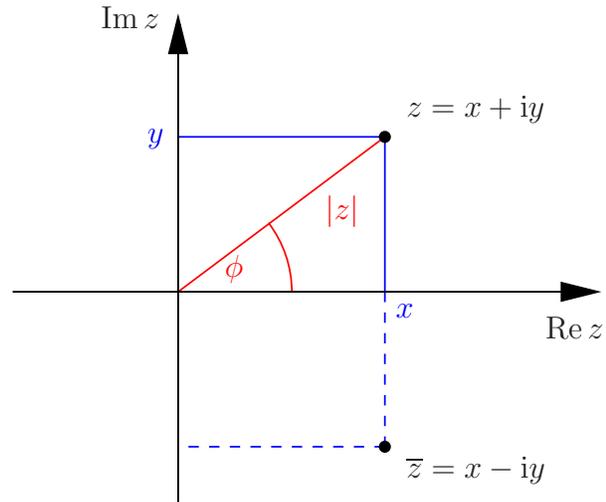
Wir nennen $x \in \mathbb{R}$ **Realteil** und $y \in \mathbb{R}$ **Imaginärteil** von $z = x + iy$ und schreiben

$$\operatorname{Re} z = x, \quad \operatorname{Im} z = y \quad (7.6)$$

Offensichtlich gilt

$$\operatorname{Re} z = \frac{z + \bar{z}}{2}, \quad \operatorname{Im} z = \frac{z - \bar{z}}{2i} \quad (7.7)$$

Anschaulich:
Gaußsche Zahlenebene



Polardarstellung: Charakterisiere z durch Betrag und Argument (Phase)

$$\begin{aligned} r = |z| &= \sqrt{z\bar{z}} = \sqrt{x^2 + y^2} \\ \phi &= \arg z \in [0, 2\pi) \end{aligned} \tag{7.8}$$

$\arg z = \arctan \frac{y}{x}$ (Zweig!)

Damit $z^{-1} = \frac{\bar{z}}{|z|^2}$ und $z = r \cos \phi + ir \sin \phi$.

Einheitskreis: $\{z = \cos \phi + i \sin \phi \mid \phi \in [0, 2\pi)\}$,
denn $|\cos \phi + i \sin \phi|^2 = \cos^2 \phi + \sin^2 \phi = 1$.

7.1 Komplexe e-Funktion

Behauptung: $e^{i\phi} = \cos \phi + i \sin \phi$, $\forall \phi \in \mathbb{R}$, (vgl. Prolog)
wobei die komplexe e-Funktion über die bekannte Reihe definiert wird,

$$e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}, \quad z \in \mathbb{C} \tag{7.9}$$

$$\begin{aligned} e^{i\phi} &= \sum_{n=0}^{\infty} \frac{i^n \phi^n}{n!} \\ &= \sum_{n=0}^{\infty} \frac{i^{2n}}{(2n)!} \phi^{2n} + \sum_{n=0}^{\infty} \frac{i^{2n+1}}{(2n+1)!} \phi^{2n+1} \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \phi^{2n} + \sum_{n=0}^{\infty} \frac{(-1)^n i}{(2n+1)!} \phi^{2n+1} \\ &= \cos \phi + i \sin \phi. \quad \square \end{aligned} \tag{7.10}$$

Damit nochmal Polardarstellung: $z = re^{i\phi} = |z|e^{i\arg z}$.

Definieren wir auch $\sin z$ und $\cos z$ für $z \in \mathbb{C}$ durch die Reihen, so gilt auch

$$e^{iz} = \cos z + i \sin z \quad \forall z \in \mathbb{C}. \quad (7.11)$$

Rechnung genau gleich wie (7.10)

Eigenschaften: (für $\phi \in \mathbb{R}$ und $z, w \in \mathbb{C}$)

1. $|e^{i\phi}| = 1$
2. $e^z \cdot e^w = e^{z+w}$
3. $\overline{e^z} = e^{\bar{z}}$, insbesondere $\overline{e^{i\phi}} = e^{-i\phi}$
4. $\arg e^{i\phi} = \phi$ (bis auf Vielfache von 2π)
5. $\cos z = \frac{e^{iz} + e^{-iz}}{2}$, $\sin z = \frac{e^{iz} - e^{-iz}}{2i}$

Beispiele:

1. Moivre-Formel

$$(\cos \phi + i \sin \phi)^n = \cos(n\phi) + i \sin(n\phi), \quad \forall \phi \in \mathbb{R}, \forall n \in \mathbb{N} \quad (7.12)$$

denn $e^{in\phi} = (e^{i\phi})^n$

2. Additionstheoreme (früher geometrisch bewiesen) lassen sich nun direkt nachrechnen, z.B. ($z, w \in \mathbb{C}$)

$$\begin{aligned} \sin z \cos w + \cos z \sin w &= \frac{e^{iz} - e^{-iz}}{2i} \frac{e^{iw} + e^{-iw}}{2} + \frac{e^{iz} + e^{-iz}}{2} \frac{e^{iw} - e^{-iw}}{2i} \\ &= \frac{1}{4i} \left[e^{i(z+w)} + e^{i(z-w)} - e^{i(-z+w)} - e^{i(-z-w)} \right. \\ &\quad \left. + e^{i(z+w)} - e^{i(z-w)} + e^{i(-z+w)} - e^{i(-z-w)} \right] \\ &= \frac{e^{i(z+w)} - e^{-i(z+w)}}{2i} \\ &= \sin(w + z). \end{aligned} \quad (7.13)$$

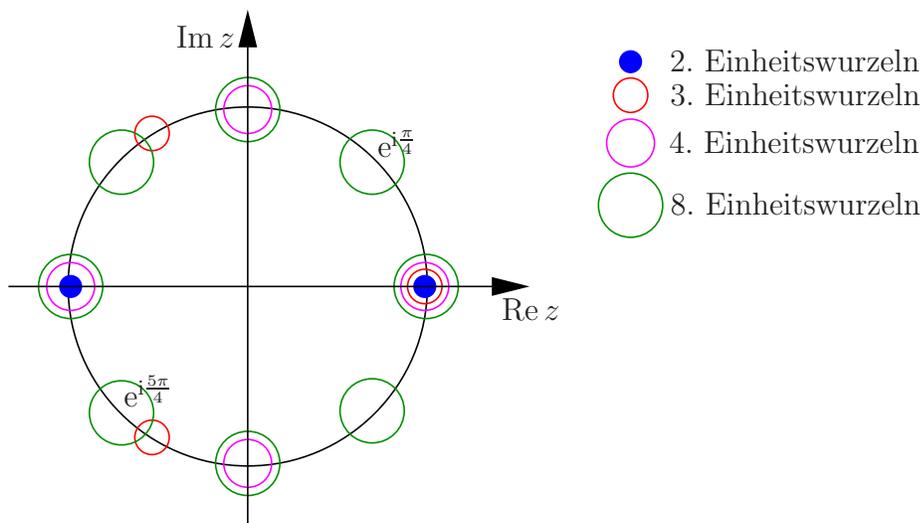
3. $x \in \mathbb{R}$:

$$\begin{aligned}
 \sum_{\nu=0}^n \sin(\nu x) &= \operatorname{Im} \sum_{\nu=0}^n e^{i\nu x} = \operatorname{Im} \frac{e^{i(n+1)x} - 1}{e^{ix} - 1} && (x \neq 2\pi k, k \in \mathbb{Z}) \\
 &= \operatorname{Im} \frac{e^{i\frac{n+1}{2}x} (e^{i\frac{n+1}{2}x} - e^{-i\frac{n+1}{2}x})}{e^{i\frac{x}{2}} (e^{i\frac{x}{2}} - e^{-i\frac{x}{2}})} \frac{2i}{2i} \\
 &= \operatorname{Im} e^{i\frac{n}{2}x} \underbrace{\frac{\sin(\frac{n+1}{2}x)}{\sin(\frac{x}{2})}}_{\text{reell}} \\
 &= \sin\left(\frac{n}{2}x\right) \frac{\sin\left(\frac{n+1}{2}x\right)}{\sin\left(\frac{x}{2}\right)}
 \end{aligned} \tag{7.14}$$

Einheitswurzeln. Gesucht sind alle Lösungen $z \in \mathbb{C}$ von $z^n = 1$.

Ansatz: $z = e^{i\phi} \Rightarrow z^n = e^{in\phi} \stackrel{!}{=} 1 \Leftrightarrow n\phi = 2\pi\nu, \nu \in \mathbb{Z}$

$$\Rightarrow \text{Lösungen: } \phi_\nu = \frac{2\pi}{n}\nu, \quad \nu = 0, 1, \dots, n-1 \tag{7.15}$$



n -te Wurzeln aus $w \in \mathbb{C}$:

$$z^n \stackrel{!}{=} w = |w|e^{i\psi}$$

$$z = \sqrt[n]{|w|} e^{i\frac{\psi}{n} + i\frac{2\pi\nu}{n}}, \quad \nu = 0, 1, \dots, n-1. \tag{7.16}$$

Beispiel \sqrt{i} : $i = e^{i\frac{\pi}{2}}$

$z_0 = e^{i\frac{\pi}{4}}$ und $z_1 = e^{i\frac{\pi}{4} + i\pi} = e^{i\frac{5\pi}{4}}$ erfüllen $z_j^2 = i$.

7.2 \mathbb{C}^n , $\mathbb{C}^{m \times n}$ und Verwandtes

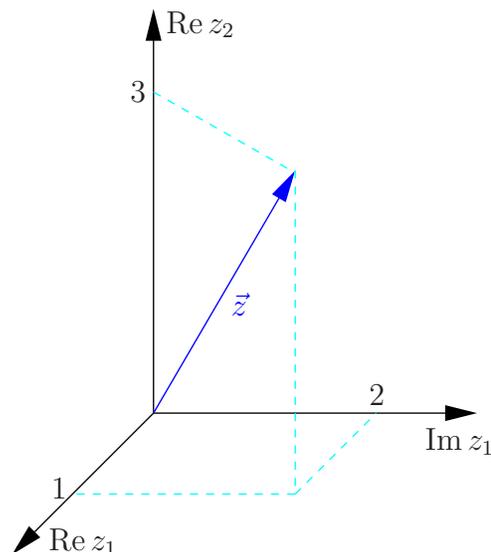
\mathbb{C} als Vektorraum über \mathbb{R} :

Dimension 2, Basis $\{1, i\}$, isomorph¹⁰ zu \mathbb{R}^2
(Gaußsche Zahlenebene);

analog ist z.B. \mathbb{C}^2 über \mathbb{R} isomorph¹⁰ zu \mathbb{R}^4 ,
z.B.

$$\vec{z} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} 1 + 2i \\ 3 \end{pmatrix} \in \mathbb{C}^2 \quad (7.17)$$

Mit dem Zeichnen der 4. Achse, $\text{Im } z_2$, tue ich
mich etwas schwer. . .



Weiter sind z.B.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} i \\ 0 \end{pmatrix} \quad (7.19)$$

l.u. in \mathbb{C}^2 als VR über \mathbb{R} , aber

l.a. in \mathbb{C}^2 als VR über \mathbb{C} .

Komplexe Vektorräume: Vektorräume über dem Körper \mathbb{C}

Alles wie bei reellen Vektorräumen, mit einer Ausnahme:

Skalarprodukte, $V \times V \rightarrow \mathbb{C}$, müssen

$$(S1)' \quad \langle \vec{z}, \vec{w} \rangle = \overline{\langle \vec{w}, \vec{z} \rangle}$$

erfüllen.

Wegen (S2), $\langle \vec{z}, \lambda \vec{w} \rangle = \lambda \langle \vec{z}, \vec{w} \rangle$, folgt

$$\langle \lambda \vec{z}, \vec{w} \rangle = \bar{\lambda} \langle \vec{z}, \vec{w} \rangle \quad (7.20)$$

Beispiel: $V = \mathbb{C}^n$ über \mathbb{C}

$$\vec{z} = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}, \quad z_j, w_j \in \mathbb{C} \quad (7.21)$$

¹⁰isomorph heißt hier: es gibt eine lineare bijektive Abbildung, z.B.

$$\mathbb{C}^2 \ni \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mapsto \begin{pmatrix} \text{Re } z_1 \\ \text{Im } z_1 \\ \text{Re } z_2 \\ \text{Im } z_2 \end{pmatrix} \in \mathbb{R}^4, \quad (7.18)$$

die die Vektorraumstruktur erhält.

Kanonisches Skalarprodukt:

$$\langle \vec{z}, \vec{w} \rangle = \vec{z}^T \vec{w} := \sum_{j=1}^n \bar{z}_j w_j \quad (7.22)$$

mittlerer Ausdruck in Matrixschreibweise

Zugehörige Norm:

$$\|\vec{z}\| = |\vec{z}| = \sqrt{\vec{z}^T \vec{z}} = \sqrt{\sum_{j=1}^n |z_j|^2} \quad (7.23)$$

Beispiele:

$$\begin{aligned} \left\langle \begin{pmatrix} 1 \\ i \\ 1+i \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ 1+i \end{pmatrix} \right\rangle &= (1, -i, 1-i) \begin{pmatrix} 3 \\ 5 \\ 1+i \end{pmatrix} \\ &= 3 - 5i + \underbrace{(1-i)(1+i)}_{=2} = 5 - 5i \end{aligned} \quad (7.24)$$

$$\left| \begin{pmatrix} 1+i \\ 2 \end{pmatrix} \right| = \sqrt{(1-i)(1+i) + 4} = \sqrt{6} \quad (7.25)$$

Matrizen mit komplexen Einträgen: $A = (a_{jk}) \in \mathbb{C}^{m \times n}$.

Komplex konjugierte Matrix:

$$\bar{A} = (\bar{a}_{jk}), \quad \overline{AB} = \bar{A} \bar{B}. \quad (7.26)$$

Alles wie gehabt, z.B. Determinanten,

$$\det \begin{pmatrix} 2+i & 1+3i \\ 4i & 2-i \end{pmatrix} = 4 + 1 - (4i - 12) = 17 - 4i \quad (7.27)$$

8 Integration

Definition: Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und $f : I \rightarrow \mathbb{R}$. Eine diffbare Funktion $F : I \rightarrow \mathbb{R}$ heißt Stammfunktion von f , falls gilt $F'(x) = f(x) \forall x \in I$.

Bemerkung: Ist F Stammfunktion von f , so auch \tilde{F} mit $\tilde{F}(x) = F(x) + c$ für jedes $c \in \mathbb{R}$, denn

$$\tilde{F}'(x) = (F(x) + c)' = F'(x) = f. \quad (8.1)$$

So erhalten wir alle Stammfunktionen von f , denn, sind F und G Stammfunktionen, so gilt

$$(F(x) - G(x))' = F'(x) - G'(x) = f(x) - f(x) = 0, \quad (8.2)$$

d.h. $F(x) - G(x)$ ist konstant.

(Nur konstante Funktionen haben Ableitung $\equiv 0$:

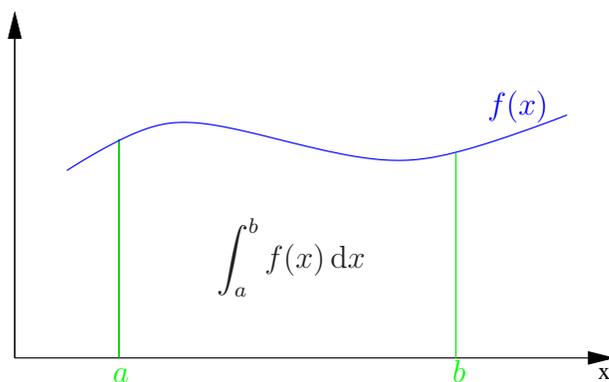
$$g'(x) = 0 \forall x \in I$$

$\Rightarrow g$ ist monoton wachsend und fallend $\forall x \in I$

$\Rightarrow g$ ist konstant.)

Fläche unter Funktionsgraph: Sei $I \subseteq \mathbb{R}$ ein Intervall, $a, b \in I$, $a < b$, $f : I \rightarrow \mathbb{R}$ stetig auf I und $f(x) \geq 0 \forall x \in I$. Wir bezeichnen den Flächeninhalt der Fläche, begrenzt durch den Graph von f , die x -Achse und die senkrechten Geraden $x = a$ und $x = b$, mit

$$\int_a^b f(x) dx. \quad (8.3)$$



Behauptung: (Hauptsatz der Differential- Integralrechnung)

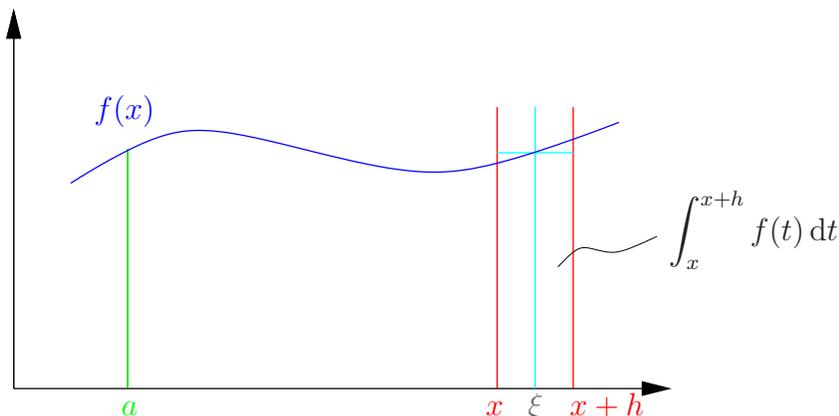
$$F(x) := \int_a^x f(t) dt \quad (8.4)$$

ist Stammfunktion von $f \forall x \in I$, d.h.

$$\frac{d}{dx} \int_a^x f(t) dt = f(x) \quad (8.5)$$

Beweisidee:

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} \frac{1}{h} \int_x^{x+h} f(t) dt. \quad (8.6)$$



Da f stetig, existiert sicher ein $\xi \in [x, x+h]$, so dass

$$\int_x^{x+h} f(t) dt = h f(\xi). \quad (8.7)$$

Damit gilt

$$F'(x) = \lim_{h \rightarrow 0} f(\xi) = f(\lim_{h \rightarrow 0} \xi) = f(x) \quad (8.8)$$

□

Berechnung von Integralen: Nun gilt¹¹

$$\int_a^b f(x) dx = F(b) - F(a), \quad (8.9)$$

wobei F eine beliebige Stammfunktion von f ist, denn:

Sei F eine beliebige Stammfunktion von f , so existiert ein $c \in \mathbb{R}$ mit

$$\begin{aligned} F(x) &= \int_a^x f(t) dt + c \\ \Rightarrow F(a) &= \int_a^a f(t) dt + c = c \\ \text{und } F(b) &= \int_a^b f(t) dt + c \\ \Rightarrow F(b) - F(a) &= \int_a^b f(t) dt. \quad \square \end{aligned} \quad (8.10)$$

¹¹Wenn wir vorher präzise definiert hätten, was ein Integral eigentlich ist, dann würden wir diese Formel *Hauptsatz der Differential- und Integralrechnung* nennen.

Beispiel: $a \in \mathbb{R}$

$$\int_0^{\frac{\pi}{a}} \sin(ax) \, dx = -\frac{\cos(ax)}{a} \Big|_0^{\frac{\pi}{a}} = \left(-\frac{\cos(\pi)}{a} \right) - \left(-\frac{\cos(0)}{a} \right) = \frac{2}{a} \quad (8.11)$$

Bemerkung: Wir schreiben auch (unbestimmtes Integral)

$$\int f(x) \, dx = F(x) \quad \text{falls} \quad F'(x) = f(x). \quad (8.12)$$

Beispiele:

$$1. \quad \int x^n \, dx = \frac{x^{n+1}}{n+1}, \quad n \neq -1 \quad (8.13)$$

$$2. \quad \int \frac{dx}{x} = \log x \quad (8.14)$$

Eigenschaften des Integrals: (f, g stückweise stetig und beschränkt, $b > a$)

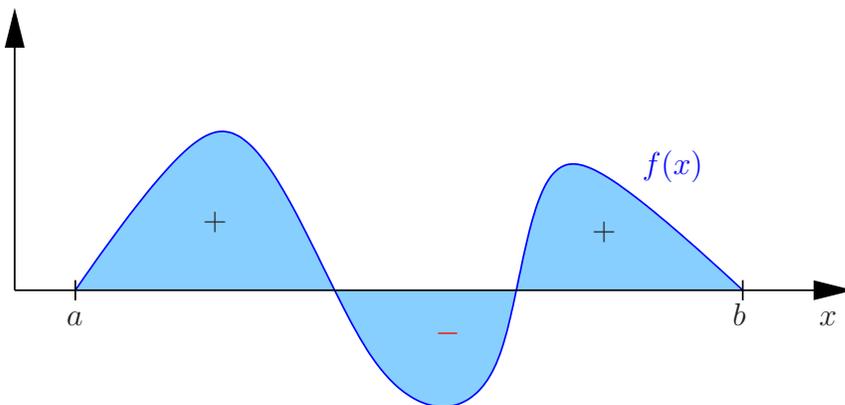
(i) Linearität:

$$\int_a^b [\lambda f(x) + \mu g(x)] \, dx = \lambda \int_a^b f(x) \, dx + \mu \int_a^b g(x) \, dx, \quad \lambda, \mu \in \mathbb{R} \quad (8.15)$$

Zunächst nur für nicht-negative Integranden, d.h. $f(x), g(x) \geq 0 \, \forall x \in [a, b]$ und $\lambda, \mu \geq 0$, aber wenn wir Flächen unter der x -Achse einen negativen Flächeninhalt zuordnen,

$$\int_a^b [-f(x)] \, dx = - \int_a^b f(x) \, dx, \quad (8.16)$$

dann dürfen f, g sowie λ, μ beliebiges Vorzeichen haben.



(ii) Stückweise integrieren: (dann ist uns auch egal, ob f in c unstetig ist)

$$\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \quad (8.17)$$

Zunächst nur für $c \in [a, b]$, aber wenn wir festlegen, dass

$$\int_b^a f(x) dx = - \int_a^b f(x) dx \quad (8.18)$$

dann können wir auch c außerhalb $[a, b]$ zulassen (falls beide Integrale auf der rechten Seite existieren).

(iii) Für Abschätzungen:

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx, \quad (8.19)$$

vgl. Skizze unter (i), sowie

$$f(x) \leq g(x) \quad \forall x \in [a, b] \quad \Rightarrow \quad \int_a^b f(x) dx \leq \int_a^b g(x) dx. \quad (8.20)$$

8.1 Uneigentliche Integrale

Wir definieren

$$\int_a^\infty f(x) dx := \lim_{b \rightarrow \infty} \int_a^b f(x) dx \quad (8.21)$$

falls $\int_a^b \dots$ für beliebig große b existiert.

Analog für $\lim_{x \rightarrow b} f(x) = \pm\infty$, $a < b$:

$$\int_a^b f(x) dx := \lim_{y \rightarrow b^-} \int_a^y f(x) dx \quad (8.22)$$

Beispiele

1.

$$\begin{aligned} \int_1^\infty \frac{1}{x^2} dx &= \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x^2} dx = \lim_{b \rightarrow \infty} \underbrace{\left[-\frac{1}{x} \right]_1^b}_{= -\frac{1}{x} \Big|_1^\infty} = \lim_{b \rightarrow \infty} \left(-\frac{1}{b} + 1 \right) = 1 \end{aligned} \quad (8.23)$$

2.

$$\int_0^1 \frac{1}{\sqrt{x}} dx \left(= \lim_{y \rightarrow 0^+} \int_y^1 \frac{1}{\sqrt{x}} dx \right) = 2\sqrt{x} \Big|_0^1 = 2 - 0 = 2 \quad (8.24)$$

Bemerkung: Falls uneigentlich an mehreren Stellen, müssen wir die Limes unabhängig bilden dürfen, z.B.

$$\int_{-1}^1 \frac{dx}{x} \neq \lim_{y \rightarrow 0^+} \left(\int_{-1}^{-y} \frac{dx}{x} + \int_y^1 \frac{dx}{x} \right) \quad (8.25)$$

Linke Seite existiert nicht,

$$\begin{aligned}
 \int_{-1}^1 \frac{dx}{x} &= \lim_{b \rightarrow 0^-} \int_{-1}^b \frac{dx}{x} + \lim_{a \rightarrow 0^+} \int_a^1 \frac{dx}{x} \\
 &= \lim_{b \rightarrow 0^-} \left[\log |x| \right]_{-1}^b + \lim_{a \rightarrow 0^+} \left[\log x \right]_a^1 \\
 &= \lim_{b \rightarrow 0^-} \log |b| - \lim_{a \rightarrow 0^+} \log a \\
 &= \text{“} -\infty + \infty \text{”}
 \end{aligned} \tag{8.26}$$

wohingegen

$$\begin{aligned}
 \lim_{y \rightarrow 0^+} \left(\int_{-1}^{-y} \frac{dx}{x} + \int_y^1 \frac{dx}{x} \right) &= \lim_{y \rightarrow 0^+} \left(\log |x| \Big|_{-1}^{-y} + \log x \Big|_y^1 \right) \\
 &= \lim_{y \rightarrow 0^+} (\log y - \log y) = 0
 \end{aligned} \tag{8.27}$$

Übrigens: $(\log |x|)' = \frac{1}{x}$, denn (Kettenregel)

$$\log |x| = \begin{cases} \log x, & x > 0 \\ \log(-x), & x < 0 \end{cases}, \quad (\log |x|)' = \begin{cases} \frac{1}{x}, & x > 0 \\ \frac{1}{(-x)}(-1), & x < 0 \end{cases} \tag{8.28}$$

8.2 Integrationstechniken

Satz 16. (Partielle Integration)

Seien f und g stetig auf $[a, b]$ und stetig diffbar auf (a, b) , dann gilt

$$\int_a^b f'(x) g(x) dx = f(x) g(x) \Big|_a^b - \int_a^b f(x) g'(x) dx \tag{8.29}$$

Beweis: Wegen $(fg)' = f'g + fg'$ gilt

$$\begin{aligned}
 \int_a^b [f'(x) g(x) + f(x) g'(x)] dx &= f(x) g(x) \Big|_a^b \\
 &= \int_a^b f'(x) g(x) dx + \int_a^b f(x) g'(x) dx \quad \square
 \end{aligned} \tag{8.30}$$

abziehen \Rightarrow Beh.

Beispiele:

1.

$$\begin{aligned}
 \int x \cos x dx &= \int \underset{g}{x} \underset{f'}{\cos x} dx = \int \underset{g}{x} \underset{f}{\sin x} dx - \int \underset{g'}{1} \underset{f}{\sin x} dx \\
 &= x \sin x + \cos x
 \end{aligned} \tag{8.31}$$

2.

$$\begin{aligned}\int \log x \, dx &= \int 1 \cdot \log x \, dx \\ &= x \log x - \int x \frac{1}{x} \, dx \\ &= x \log x - x\end{aligned}\tag{8.32}$$

Analog:

$$\begin{aligned}\int \arcsin x \, dx &= x \arcsin x - \int x \frac{1}{\sqrt{1-x^2}} \, dx \\ &= x \arcsin x + \sqrt{1-x^2}\end{aligned}\tag{8.33}$$

Uneigentliche Integrale zunächst übersprungen

3.

$$\begin{aligned}\int \sin^2 x \, dx &= -\sin x \cos x + \int \cos^2 x \, dx \\ &= -\sin x \cos x + \int (1 - \sin^2 x) \, dx \\ &= -\sin x \cos x + x - \int \sin^2 x \, dx \\ &= \frac{x}{2} - \frac{1}{2} \sin x \cos x\end{aligned}\tag{8.34}$$

Satz 17. (Integralrestglied für Taylor)

Sei f $n + 1$ -mal stetig diffbar auf (a, b) , dann gilt

$$f(x) - \sum_{\nu=0}^n \frac{f^{(\nu)}(x_0)}{\nu!} (x - x_0)^\nu = \frac{1}{n!} \int_{x_0}^x f^{(n+1)}(t) (x - t)^n \, dt,\tag{8.35}$$

wobei $x, x_0 \in (a, b)$.

Beweis: (vollständige Induktion)

$n = 0$:

$$f(x) - f(x_0) = \int_{x_0}^x f'(t) \, dt \quad \text{o.k.}\tag{8.36}$$

$n \rightarrow n + 1$:

$$\begin{aligned}
 f(x) &= \sum_{\nu=0}^{n+1} \frac{f^{(\nu)}(x_0)}{\nu!} (x - x_0)^\nu \\
 &= f(x) - \sum_{\nu=0}^n \frac{f^{(\nu)}(x_0)}{\nu!} (x - x_0)^\nu - \frac{f^{(n+1)}(x_0)}{(n+1)!} (x - x_0)^{n+1} \\
 &\stackrel{\text{I.V.}}{=} \frac{1}{n!} \int_{x_0}^x f^{(n+1)}(t) (x - t)^n dt - \frac{f^{(n+1)}(x_0)}{(n+1)!} (x - x_0)^{n+1} \\
 &\stackrel{\text{p.I.}}{=} \underbrace{\frac{1}{n!} f^{(n+1)}(t) \left(-\frac{(x-t)^{n+1}}{n+1} \right) \Big|_{t=x_0}^{t=x}}_{=0 + \frac{f^{(n+1)}(x_0)}{(n+1)!} (x-x_0)^{n+1}} - \frac{1}{n!} \int_{x_0}^x f^{(n+2)}(t) \left(-\frac{(x-t)^{n+1}}{n+1} \right) dt - \frac{f^{(n+1)}(x_0)}{(n+1)!} (x - x_0)^{n+1} \\
 &= \frac{1}{(n+1)!} \int_{x_0}^x f^{(n+2)}(t) (x - t)^{n+1} dt
 \end{aligned} \tag{8.37}$$

□

Satz 18. (Substitutionsregel)

Sei f stetig auf I und $g : [a, b] \rightarrow J \subset I$ stetig diffbar, dann gilt

$$\int_a^b f(g(t)) g'(t) dt = \int_{g(a)}^{g(b)} f(x) dx. \tag{8.38}$$

Beweis: Sei F Stammfunktion von f ,

- $\int_{g(a)}^{g(b)} f(x) dx = F(x) \Big|_{g(a)}^{g(b)} = F(g(b)) - F(g(a))$
- $\frac{d}{dt} F(g(t)) = F'(g(t)) g'(t) = f(g(t)) g'(t)$

$$\int_a^b f(g(t)) g'(t) dt = F(g(t)) \Big|_a^b = F(g(b)) - F(g(a)) \quad \square \tag{8.39}$$

Beispiele:

1.

$$\int \frac{g'(t)}{g(t)} dt = \log |g(t)|, \quad \text{mit } f(x) = \frac{1}{x}, F(x) = \log |x| \tag{8.40}$$

z.B.

$$\begin{aligned}
 \int \tan x dx &= \int \frac{\sin x}{\cos x} dx = -\log |\cos x|, \quad \text{mit } g(x) = \cos x, \\
 \int_e^{10} \frac{dx}{x \log x} &= \log \log x \Big|_e^{10} = \log \log 10, \quad \text{mit } g(x) = \log x,
 \end{aligned} \tag{8.41}$$

2. $n \neq -1$:

$$\begin{aligned} \int \sin^n x \cos x, dx & \quad \sin x = t, \quad \cos x dx = dt \quad \left(\frac{dt}{dx} = \cos x \right) \\ & = \int t^n dt = \frac{t^{n+1}}{n+1} = \frac{\sin^{n+1}(x)}{n+1} \end{aligned} \quad (8.42)$$

3.

$$\begin{aligned} \int_0^\pi \sin^2 x dx \\ \sin x = t, \quad \cos x dx = dt, \quad dx = \frac{dt}{\cos x} = \frac{dt}{\sqrt{1-\sin^2 x}} = \frac{dt}{\sqrt{1-t^2}} \quad (8.43) \\ \neq \int_{\sin 0=0}^{\sin \pi=0} \frac{t^2}{\sqrt{1-t^2}} dt = 0 ? \end{aligned}$$

Vorsicht: $\sin^2 x + \cos^2 x = 1 \quad \forall x$,

aber $\cos x = \sqrt{1-\sin^2 x}$ nur dort, wo $\cos x \geq 0$,

z.B. nicht für $\frac{\pi}{2} < x \leq \pi$.

Merke: Vorsicht beim Wurzelziehen!

Übrigens: Das Integral lösen wir z.B. mit partieller Integration, siehe Beispiel dort.

Einschub: Ableiten nach oberer und/oder unterer Grenze

$$\frac{d}{dx} \int_{\phi(x)}^{\psi(x)} f(t) dt \quad (8.44)$$

Sei F Stammfunktion von f , d.h. $F' = f$:

$$\begin{aligned} \frac{d}{dx} \int_{\phi(x)}^{\psi(x)} f(t) dt & = \frac{d}{dx} \left(F(\psi(x)) - F(\phi(x)) \right) \\ & = F'(\psi(x)) \psi'(x) - F'(\phi(x)) \phi'(x) \\ & = f(\psi(x)) \psi'(x) - f(\phi(x)) \phi'(x) \end{aligned} \quad (8.45)$$

F taucht nicht mehr auf! Geht also auch, wenn wir F nicht explizit kennen!

Beispiel:

$$\frac{d}{dx} \int_0^{\sin x} e^{-t^2} dt = e^{-\sin^2 x} \cos x. \quad (8.46)$$

Uneigentliche Integrale zunächst übersprungen

Partialbruchzerlegung

Idee:

$$\begin{aligned}\int \frac{1}{(x-1)(x-2)} dx &= \int \left(\frac{1}{x-2} - \frac{1}{x-1} \right) dx \\ &= \log|x-2| - \log|x-1| \\ &= \log \left| \frac{x-2}{x-1} \right|\end{aligned}\tag{8.47}$$

Wie finden wir die Zerlegung? (inkl. weiterer Beispiele)

Ziel: Berechne

$$\int \frac{P(x)}{Q(x)} dx\tag{8.48}$$

mit Polynomen, P, Q , mit reellen Koeffizienten – o.B.d.A Zählergrad < Nennergrad (sonst Polynomdivision). Sei Q von Grad n , d.h.

$$Q(x) = c_n \prod_{j=1}^r (x - x_j)^{\nu_j}, \quad \sum_{j=1}^r \nu_j = n, \quad c_n \neq 0,\tag{8.49}$$

Nullstellen x_j mit Vielfachheiten ν_j :

- Entweder $x_j \in \mathbb{R}$
- oder $x_j \notin \mathbb{R} \Rightarrow \overline{x_j}$ ist auch Nullstelle,

denn mit

$$Q(x) = \sum_{k=0}^n c_k x^k, \quad c_j \in \mathbb{R},\tag{8.50}$$

gilt: Aus $Q(x_j) = \sum_{k=0}^n c_k x_j^k = 0$ folgt

$$0 = \overline{Q(x_j)} = \sum_{k=0}^n \overline{c_k x_j^k} = \sum_{k=0}^n c_k \overline{x_j}^k = Q(\overline{x_j}). \quad \square\tag{8.51}$$

Beispiel: $Q(x) = 1 + x^2$, Nullstellen $\pm i$.

Nun gilt:

$$\begin{aligned}\frac{P(x)}{Q(x)} &= \frac{a_1^{(1)}}{x - x_1} + \frac{a_1^{(2)}}{(x - x_1)^2} + \dots + \frac{a_1^{(\nu_1)}}{(x - x_1)^{\nu_1}} \\ &\quad + \frac{a_2^{(1)}}{x - x_2} + \dots + \frac{a_2^{(\nu_2)}}{(x - x_2)^{\nu_2}} \\ &\quad \vdots \\ &\quad + \frac{a_r^{(1)}}{x - x_r} + \dots + \frac{a_r^{(\nu_r)}}{(x - x_r)^{\nu_r}}\end{aligned}\tag{8.52}$$

mit Koeffizienten $a_j^{(k)} \in \mathbb{C}$. (Beweis: Auf Hauptnenner bringen.)

Falls Nullstelle komplex, so sind Koeffizienten der komplex konjugierten Nullstelle auch das Komplexkonjugierte der Koeffizienten (nur dann wird Summe der Partialbrüche wieder reell).

Partialbrüche lassen sich integrieren

Beispiele

1.

$$\frac{1}{(x-1)(x-2)} = \frac{a}{x-1} + \frac{b}{x-2} \quad (8.53)$$

1. Methode: Hauptnenner & Koeffizientenvergleich

$$\begin{aligned} a(x-2) + b(x-1) &= (a+b)x - (2a+b) \stackrel{!}{=} 1 \\ \Rightarrow a+b &= 0, \quad 2a+b = -1 \quad (\text{LGS}) \\ \Rightarrow b &= -a, \quad a = -1 \quad \Rightarrow (b = 1) \end{aligned} \quad (8.54)$$

2. Methode: Grenzübergänge ("Zuhaltmethode")

$$\frac{1}{(x-1)(x-2)} = \frac{a}{x-1} + \frac{b}{x-2} \quad \left| (x-1) \text{ danach } x=1 \right. \quad (8.55)$$

$$\frac{1}{-1} = a$$

...

$$\left| (x-2) \text{ danach } x=2 \right. \quad (8.56)$$

$$\frac{1}{1} = b$$

2.

$$\frac{x+1}{(x-1)^2} = \frac{a}{x-1} + \frac{b}{(x-1)^2} \quad \left| (x-1)^2, \quad x=1 \right. \quad (8.57)$$

$$1+1=2=b$$

...

$$\left| (x-1), \quad x \rightarrow \infty \right. \quad (8.58)$$

$$\lim_{x \rightarrow \infty} \frac{x+1}{x-1} = 1 = a$$

also

$$\begin{aligned} \int \frac{x+1}{(x-1)^2} dx &= \int \frac{1}{x-1} dx + \int \frac{2}{(x-1)^2} dx \\ &= \log|x-1| - \frac{2}{x-1} \end{aligned} \quad (8.59)$$

3.

$$\frac{x^2 - 1}{(x^2 + 1)^2} = \frac{a}{x - i} + \frac{b}{(x - i)^2} + \frac{\bar{a}}{x + i} + \frac{\bar{b}}{(x + i)^2} \quad \left| (x - i)^2, \quad x = i \quad (8.60) \right.$$

$$\frac{i^2 - 1}{(i + i)^2} = \frac{1}{2} = b$$

$$\dots \quad \left| (x - i), \quad x \rightarrow \infty \quad (8.61) \right.$$

$$0 = a + \bar{a} \quad \Leftrightarrow \quad \operatorname{Re} a = 0$$

$$\dots \quad \left| x = 0 \quad (8.62) \right.$$

$$-1 = ia - \frac{1}{2} - i\bar{a} - \frac{1}{2} \quad \Leftrightarrow \quad \operatorname{Im} a = 0 \quad (8.63)$$

also

$$\begin{aligned} \int \frac{x^2 - 1}{(x^2 + 1)^2} dx &= \frac{1}{2} \int \frac{dx}{(x - i)^2} + \frac{1}{2} \int \frac{dx}{(x + i)^2} \\ &\stackrel{(*)}{=} -\frac{1}{2} \left(\frac{1}{x - i} + \frac{1}{x + i} \right) = -\frac{x}{x^2 + 1} \end{aligned} \quad (8.64)$$

Da wir uns nicht sicher sind, ob (*) o.k. war, machen wir zumindest den Test:

$$\left(-\frac{x}{x^2 + 1} \right)' = -\frac{(x^2 + 1) - x \cdot 2x}{(x^2 + 1)^2} = \frac{x^2 - 1}{(x^2 + 1)^2} \quad \text{o.k.} \quad (8.65)$$

Übrigens, auch gut für Reihenentwicklung:

$$\begin{aligned} \frac{1}{(x - 1)(x - 2)} &= \frac{1}{1 - x} - \frac{1}{2} \frac{1}{1 - \frac{x}{2}} \\ &= \sum_{\nu=0}^{\infty} x^{\nu} - \frac{1}{2} \sum_{\nu=0}^{\infty} x^{\nu} 2^{-\nu} \\ &= \sum_{\nu=0}^{\infty} (1 - 2^{-(\nu+1)}) x^{\nu} \end{aligned} \quad (8.66)$$

Bisher hätten wir das mit dem Cauchy-Produkt gerechnet.

weiteres Beispiel

$$\int_3^{\infty} \frac{7x + 1}{(x^2 - 1)(x - 2)} dx \quad (8.67)$$

Nenner-Nullstellen: $-1, 1, 2$, d.h.

$$\frac{7x + 1}{(x^2 - 1)(x - 2)} = \frac{A}{x + 1} + \frac{B}{x - 1} + \frac{C}{x - 2}, \quad (8.68)$$

$|\cdot(x+1), x \rightarrow -1: \frac{-7+1}{(-1-1)(-1-2)} = A$, d.h. $A = -1$.

Analog: $B = -4, C = 5$. Damit

$$\begin{aligned} \int_3^\infty \frac{7x+1}{(x^2-1)(x-2)} dx &= \int_3^\infty \left(-\frac{1}{x+1} - \frac{4}{x-1} + \frac{5}{x-2} \right) dx \\ &= \left[-\log(x+1) - 4\log(x-1) + 5\log(x-2) \right]_3^\infty \\ &= \left[\log \frac{(x-2)^5}{(x+1)(x-1)^4} \right]_3^\infty = \log 1 - \log \frac{1}{4 \cdot 2^4} \\ &= 6 \log 2 \end{aligned} \tag{8.69}$$

8.3 Riemannsche Zwischensummen

Bisher haben wir das Integral etwas unscharf als Fläche unter dem Funktionsgraph definiert. In diesem Abschnitt präzisieren wir, was wir damit genau meinen. Dabei erkennen wir, dass hinter jedem Integral ein *komplizierter Grenzwert* steckt, und motivieren neben auch die Schreibweise “ dx ”

Definition: (Riemannsche Zwischensummen)

Sei $f : [a, b] \rightarrow \mathbb{R}_0^+$ beschränkt und stückweise stetig (d.h. höchstens an endlich vielen Stellen unstetig). Zerlege $[a, b]$ in n Teilintervalle $[x_{j-1}, x_j]$, $j = 1, \dots, n$, mit

$$a = x_0 < x_1 < \dots < x_{n-1} < x_n = b, \quad (\text{Partition}) \tag{8.70}$$

und wähle aus jedem eine Zwischenstelle

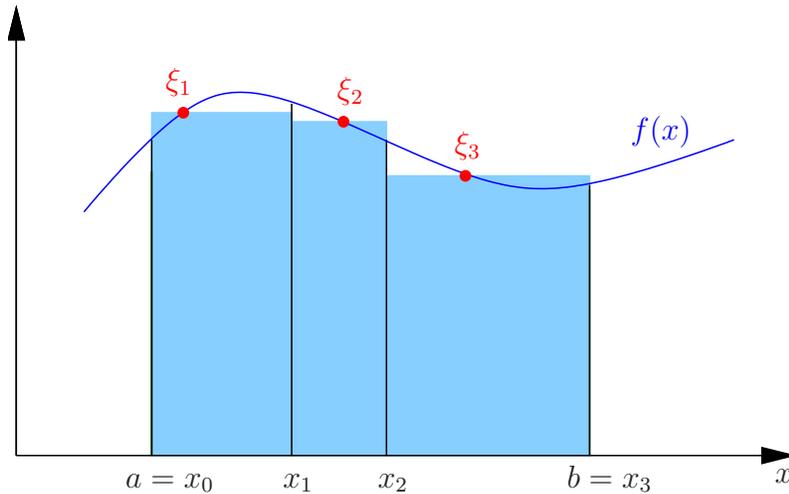
$$\xi_j \in [x_{j-1}, x_j]. \quad (\text{Belegung}) \tag{8.71}$$

Wir nennen

$$Z_n := \sum_{j=1}^n f(\xi_j)(x_j - x_{j-1}) \tag{8.72}$$

eine (Riemannsche) Zwischensumme. Die Länge des größten Teilintervalls heißt Feinheit μ_n der Partition,

$$\mu_n = \max_j (x_j - x_{j-1}). \tag{8.73}$$

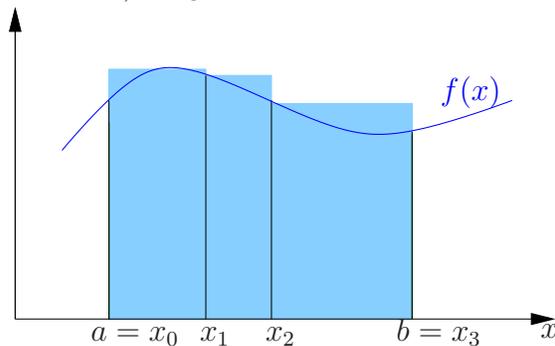
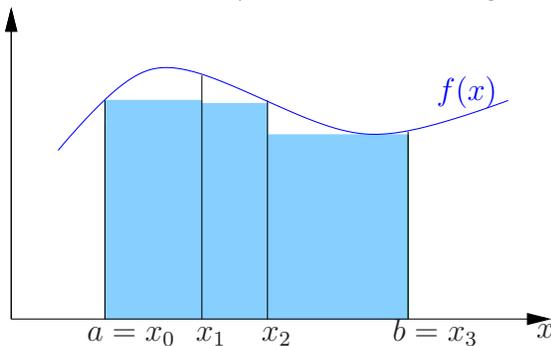


Satz 19. Sei alles wie in obiger Definition, dann gilt: Für jede Folge von Partitionen und Belegungen mit $\lim_{n \rightarrow \infty} \mu_n = 0$ existiert $\lim_{n \rightarrow \infty} Z_n$, der Grenzwert ist für alle solchen Folgen gleich, und wir schreiben

$$\lim_{n \rightarrow \infty} Z_n = \lim_{n \rightarrow \infty} \sum_{j=1}^n f(\xi_j)(x_j - x_{j-1}) = \int_a^b f(x) dx. \quad (8.74)$$

Bemerkungen:

1. Aus $\Delta x = x_{j+1} - x_j$ wird dx : infinitesimale x -Änderung.
2. "Stützstellen" ξ_j müssen nicht äquidistant gewählt werden (z.B. bei numerischer Approximation).
3. Funktionen f , für die Satz 19 gilt, heißt (Riemann-)integrierbar.



Beweisidee:

Bilde Ober- und Untersummen,

$$S_n^O := \sum_{j=1}^n \max_{x \in [x_{j-1}, x_j]} f(x) \cdot (x_j - x_{j-1}) \quad \text{und} \quad S_n^U := \sum_{j=1}^n \min_{x \in [x_{j-1}, x_j]} f(x) \cdot (x_j - x_{j-1}), \quad (8.75)$$

dann gilt

$$\min_{x \in [a, b]} f(x) \cdot (b - a) \leq S_n^U \leq Z_n \leq S_n^O \leq \max_{x \in [a, b]} f(x) \cdot (b - a). \quad (8.76)$$

S_n^U nach oben beschränkt. Bilden wir das $n + 1$ -te Folgenglied durch Teilen eines Teilintervalls, so ist $S_{n+1}^U \geq S_n^U \Rightarrow$ Konvergenz. I.A. ist Konvergenz etwas aufwendiger. S_n^O analog. Die beiden Limites sind gleich (und damit auch der von Z_n), da

$$f \text{ stückweise stetig} \Rightarrow \max_{x \in [x_{j-1}, x_j]} f(x) - \min_{x \in [x_{j-1}, x_j]} f(x) \xrightarrow[n \rightarrow \infty]{} 0 \text{ für fast alle } j. \quad (8.77)$$

9 Differentialgleichungen (DGLn)

9.1 DGLn erster Ordnung mit getrennten Veränderlichen

Beispiel:

$$y'(x) = x^2(1 + y^2(x)) \quad (9.1)$$

Idee:

$$y' = \frac{dy}{dx}, \quad (9.2)$$

wir probieren mal...

$$\begin{aligned} \frac{dy}{dx} &= x^2(1 + y^2) \\ \frac{dy}{1 + y^2} &= x^2 dx \quad (\text{“getrennte Veränderliche”}) \\ \int \frac{dy}{1 + y^2} &= \int x^2 dx \end{aligned} \quad (9.3)$$

$$\arctan y = \frac{x^3}{3} + c, \quad c \in \mathbb{R}$$

$$y = \tan\left(\frac{x^3}{3} + c\right)$$

Ist das gut gegangen? Test:¹²

$$y'(x) = \underbrace{\left[1 + \tan^2\left(\frac{x^3}{3} + c\right)\right]}_{=1+y^2} x^2 \quad (9.4)$$

Allgemein:

$$y'(x) = f(x) g(y) \quad (9.5)$$

1. Gibt es ein y_0 mit $g(y_0) = 0 \Rightarrow y(x) = y_0$ (konstant) löst die DGL
2. Daher nun $g(y) \neq 0$ (wo's interessiert)

$$\begin{aligned} \frac{dy}{dx} &= f(x) g(y) \\ \frac{dy}{g(y)} &= f(x) dx \end{aligned} \quad (9.6)$$

Definiere nun

$$F(x) := \int^x f(t) dt, \quad \Phi(y) := \int^y \frac{dt}{g(t)}. \quad (9.7)$$

¹²Wdh: $(\tan x)' = \left(\frac{\sin x}{\cos x}\right)' = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \tan^2 x$

DGL wird äquivalent zu

$$\Phi(y) = F(x), \quad (9.8)$$

und

$$y = \Phi^{-1}(F(x)) \quad (9.9)$$

löst – falls Φ umkehrbar!

Wenn g stets das gleiche VZ hat, dann ist Φ streng monoton und damit umkehrbar.

Betrachte nun das Anfangswertproblem (AWP)

$$y'(x) = f(x)g(y), \quad y(x_0) = y_0. \quad (9.10)$$

Definiere neu

$$F(x) := \int_{x_0}^x f(t) dt, \quad \Phi(y) := \int_{y_0}^y \frac{dt}{g(t)}, \quad (9.11)$$

falls $g(t) > 0 \forall t \in [y_0, y]$ (oder $g(t) < 0 \forall t \in [y_0, y]$).

Behauptung: $y(x) = \Phi^{-1}(F(x))$ löst AWP.

Beweis: $y(x_0) = \Phi^{-1}(F(x_0)) = \Phi^{-1}(0) = y_0$, da $\Phi(y_0) = 0$.

$$\begin{aligned} y'(x) &= \Phi^{-1'}(F(x)) f(x) \\ &= \frac{1}{\underbrace{\Phi'(\Phi^{-1}(F(x)))}_{=y}} f(x), & \Phi'(y) &= \frac{1}{g(y)} \\ &= g(y(x)) f(x). \end{aligned} \quad (9.12)$$

□

Beispiele:

1.

$$\begin{aligned} & y' = y^2, & y(0) &= 1 & (9.13) \\ \Leftrightarrow & \frac{dy}{y^2} = dx, & y(0) &= 1 \\ \Leftrightarrow & \int_1^y \frac{d\tilde{y}}{\tilde{y}^2} = \int_0^x d\tilde{x} \\ \Leftrightarrow & \left[-\frac{1}{\tilde{y}} \right]_1^y = x \\ \Leftrightarrow & -\frac{1}{y} + 1 = x \\ \Leftrightarrow & y = \frac{1}{1-x} \end{aligned}$$

2. homogene lineare DGL (s.u.) ist auch von diesem Typ

$$\begin{aligned}
 & y' + f(x)y = 0, & y(0) = 3 & (9.14) \\
 \Leftrightarrow & \int_3^y \frac{d\tilde{y}}{\tilde{y}} = - \int_0^x f(\tilde{x}) d\tilde{x} \\
 \Leftrightarrow & \log y - \log 3 = - \int_0^x f(t) dt \\
 \Leftrightarrow & y = 3e^{-\int_0^x f(t) dt}
 \end{aligned}$$

9.2 Lineare DGLn erster Ordnung

$$y'(x) + f(x)y(x) = g(x) \quad (9.15)$$

heißt lineare inhomogene DGL 1. Ordnung.

$$y'(x) + f(x)y(x) = 0 \quad (9.16)$$

heißt zugehörige homogene lineare DGL 1. Ordnung.

Bemerkungen: (Linearität)

1. Sind y_1 und y_2 zwei Lösungen der homogenen Gleichung, so ist

$$\alpha y_1 + \beta y_2, \quad \alpha, \beta \in \mathbb{R} \text{ (oder } \mathbb{C}), \quad (9.17)$$

ebenfalls Lösung der homogenen Gleichung; also ist die Lösungsmenge

$$L_h = \{y \mid y' + f(x)y = 0\} \quad (9.18)$$

ein Vektorraum.

2. Sind y_1 und y_2 zwei Lösungen der inhomogenen Gleichung, so ist $y_1 - y_2$ Lösung der homogenen Gleichung; also ist, mit einer Lösung y_p der inhomogenen Gleichung, die Lösungsmenge

$$\begin{aligned}
 L_i &= \{y \mid y' + f(x)y = g(x)\} \\
 &= \{y \mid y = y_p + y_h, y_h \in L_h\}.
 \end{aligned} \quad (9.19)$$

3. Außerdem

$$L_h = \left\{ y \mid y(x) = ce^{-\int^x f(t) dt}, c \in \mathbb{R} \text{ (oder } \mathbb{C}) \right\} \quad (9.20)$$

da $\left(ce^{-\int^x f(t) dt} \right)' = ce^{-\int^x f(t) dt} (-f(x))$. Dies sind auch alle Lösungen (später).¹³

¹³Die Schreibweise $= \int^x f(t) dt =: F(x)$ bezeichnet eine Stammfunktion von f , d.h. $F'(x) = f(x)$.

Beispiele: ($c \in \mathbb{R}$ oder \mathbb{C} , nach Bedarf)

$$1. \quad y' - y = 0 \quad \Rightarrow \quad L_h = \left\{ y \mid y(x) = ce^{-\int^x (-1) dt} = ce^x \right\}$$

$$2. \quad y' - x^3 y = 0 \quad \Rightarrow \quad L_h = \left\{ y \mid y(x) = ce^{\int^x t^3 dt} = ce^{\frac{x^4}{4}} \right\}$$

$$3. \quad y' + \frac{y}{2x} = 0 \quad \Rightarrow \quad L_h = \left\{ y \mid y(x) = ce^{-\int^x \frac{dt}{2t}} = \frac{c}{\sqrt{|x|}} \right\}$$

$$4. \quad y' + \frac{\sin x}{x} y = 0 \quad \Rightarrow \quad L_h = \left\{ y \mid y(x) = ce^{-\int^x \frac{\sin t}{t} dt} \right\}$$

Bemerkungen:

1. einparametrische Lösungsmenge ($\dim L_h = 1$)
2. Oft hat man ein sogenanntes Anfangswertproblem (AWP): Gesucht ist Lösung von

$$y' + f(x)y = 0 \quad \text{mit} \quad y(x_0) = y_0 \quad (9.21)$$

Lösung:

$$y(x) = y_0 e^{-\int_{x_0}^x f(t) dt} \quad (9.22)$$

$$\text{denn } y(x_0) = y_0 e^{-\int_{x_0}^{x_0} f(t) dt} = y_0 e^0 = y_0$$

Spezielle Lösung der inhomogenen Gleichung: (Variation der Konstanten)

$$y' + f(x)y = g(x) \quad (9.23)$$

Sei $F(x)$ Stammfunktion von f , d.h. $F' = f$.

$y(x) = ce^{-F(x)}$ löst homogene Gleichung.

Ansatz für Lösung y_p der inhomogenen Gleichung:

$$y_p(x) = c(x) e^{-F(x)}. \quad (9.24)$$

$$y_p' = c' e^{-F} - ce^{-F} f \quad (9.25)$$

in DGL

$$(c' e^{-F} - ce^{-F} f) + fce^{-F} = g \quad \Rightarrow \quad c' = ge^F \quad (9.26)$$

also

$$c(x) = \int^x e^{F(t)} g(t) dt \quad (9.27)$$

und damit

$$y_p(x) = e^{-F(x)} \int^x e^{F(t)} g(t) dt \quad (9.28)$$

Lösungsmenge

$$L_i = y_p + L_h = \left\{ y \mid y(x) = e^{-F(x)} \int^x e^{F(t)} g(t) dt + ae^{-F(x)}, a \in \mathbb{R} \right\} \quad (\text{oder } a \in \mathbb{C}). \quad (9.29)$$

Will man das AWP mit $y(x_0) = y_0$ lösen, so ist Lösung eindeutig bestimmt,

$$y(x) = e^{-\int_{x_0}^x f(t) dt} \left(y_0 + \int_{x_0}^x e^{\int_{x_0}^t f(\tau) d\tau} g(t) dt \right) \quad (9.30)$$

Beispiele:

1. $y' - y = 3, \quad L_h = \{y \mid y(x) = ce^x\}$

$$y_p = e^x \int e^{-t} 3 dt = e^x (-e^{-x}) 3 = -3 \quad (9.31)$$

Oft einfacher: y_p raten.

Also $L_i = \{y \mid y(x) = -3 + ce^x \mid c \in \mathbb{R}\}$.

AWP: $y(1) = 5$:

$$y(1) = -3 + ce \stackrel{!}{=} 5 \quad \Rightarrow \quad c = \frac{8}{e}, \quad (9.32)$$

und $y(x) = -3 + 8e^{x-1}$ löst AWP.

2. $y' - x^3 y = \sin x, \quad L_h = \{y \mid y(x) = ce^{\frac{x^4}{4}}\}$

$$y_p = e^{\frac{x^4}{4}} \int e^{-\frac{t^4}{4}} \sin t dt. \quad (9.33)$$

9.3 Lineare DGLn zweiter Ordnung mit konstanten Koeffizienten

Beispiel: Federpendel

... auch Modell für Schwingung eines zweiatomigen Moleküls.

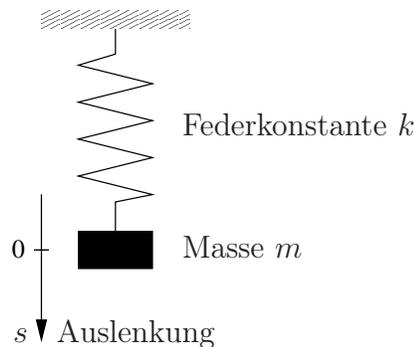
Rückstellkraft: $F = -ks$

Newton: $-ks = m\ddot{s}$

(Zeit $t, s = s(t), \ddot{s} = \frac{d^2s}{dt^2}$)

DGL:

$$\ddot{s}(t) + \frac{k}{m}s(t) = 0 \quad (9.34)$$



Lösungen: $s_1 = \sin(\omega t), s_2 = \cos(\omega t)$ mit $\omega = \sqrt{k/m}$ (und Linearkombinationen), denn

$$\ddot{s}_{1,2} = -\omega^2 s_{1,2} \quad (9.35)$$

Allgemein:

$$y'' + a_1 y' + a_0 y = f(x) \quad (9.36)$$

heißt inhomogene lineare DGL 2. Ordnung mit konstanten Koeffizienten, $a_{0,1} \in \mathbb{R}$ oder $\in \mathbb{C}$; zugehörige homogene Gleichung:

$$y'' + a_1 y' + a_0 y = 0. \quad (9.37)$$

Bemerkungen: (Linearität)

1. Sind y_1 und y_2 zwei Lösungen der homogenen Gleichung, so auch

$$\alpha y_1 + \beta y_2, \quad \alpha, \beta \in \mathbb{C}. \quad (9.38)$$

\Rightarrow Lösungsmenge L_h ist ein Vektorraum.

2. Sind y_1 und y_2 Lösungen der inhomogenen Gleichung, so ist $y_1 - y_2$ Lösung der homogenen Gleichung, d.h.

$$L_i = y_p + L_h \quad (9.39)$$

wobei y_p eine partikuläre Lösung der inhomogenen Gleichung ist.

Lösungsmenge der homogenen DGL: $y'' + a_1 y' + a_0 y = 0$.

$$\chi(x) = x^2 + a_1 x + a_0 \quad (9.40)$$

heißt zugehöriges charakteristisches Polynom.

Ist λ Nullstelle von χ , d.h. $\chi(\lambda) = 0$

$\Rightarrow y(x) = ce^{\lambda x}$, $c \in \mathbb{C}$ ist eine Lösung der DGL, denn

$$y'(x) = \lambda y(x), \quad y''(x) = \lambda^2 y(x), \quad (9.41)$$

in DGL:

$$y'' + a_1 y' + a_0 y = y \underbrace{(\lambda^2 + a_1 \lambda + a_0)}_{\chi(\lambda)=0}. \quad (9.42)$$

Funktioniert natürlich auch für lineare DGLn höherer Ordnung mit konstanten Koeffizienten.

Im obigen **Beispiel** (Federpendel):

$$\chi(x) = x^2 + \omega^2 \quad (9.43)$$

Nullstellen: $\lambda_{1,2} = \pm i\omega$.

Lösungsmenge:

$$L_h = \left\{ s \mid s(t) = c_1 e^{i\omega t} + c_2 e^{-i\omega t}, \quad c_{1,2} \in \mathbb{C} \right\} \quad (9.44)$$

Die speziellen (reellen) Lösungen $\sin(\omega t)$, $\cos(\omega t)$ sind LK von $e^{\pm i\omega t}$,

$$\begin{aligned} \sin(\omega t) &= \frac{e^{i\omega t} - e^{-i\omega t}}{2i}, & c_1 &= \frac{1}{2i} = -c_2, \\ \cos(\omega t) &= \frac{e^{i\omega t} + e^{-i\omega t}}{2}, & c_1 &= c_2 = \frac{1}{2}. \end{aligned} \quad (9.45)$$

Reelle Lösungsmenge:

$$L_h^{\text{reell}} = \left\{ s \mid s(t) = c_1 \sin(\omega t) + c_2 \cos(\omega t), \quad c_{1,2} \in \mathbb{R} \right\} \subset L_h. \quad (9.46)$$

Allgemein:

1. χ hat zwei verschiedene reelle Nullstellen, $\lambda_1 \neq \lambda_2$:

$$L_h^{\text{reell}} = \left\{ y \mid y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}, c_{1,2} \in \mathbb{C} \in \mathbb{R} \right\} \quad (9.47)$$

2. χ hat komplex konjugierte Nullstellen, $\lambda_2 = \bar{\lambda}_1$: L_h wie oben.
Reelle Lösungen?

$$\lambda_1 = u + iv \quad (\lambda_2 = u - iv) \quad (9.48)$$

$$y_1(x) = \frac{1}{2} (e^{\lambda_1 x} + e^{\lambda_2 x}) = e^{ux} \cos(vx) \quad (9.49)$$

$$y_2(x) = \frac{1}{2i} (e^{\lambda_1 x} - e^{\lambda_2 x}) = e^{ux} \sin(vx)$$

Federpendel: $u = 0, v = \omega$

$$L_h^{\text{reell}} = \left\{ y \mid y(x) = c_1 e^{ux} \sin(vx) + c_2 e^{ux} \cos(vx), c_{1,2} \in \mathbb{R} \right\} \quad (9.50)$$

3. χ hat eine doppelte Nullstelle, $\lambda \in \mathbb{R}$:

$$L_h^{\text{reell}} = \left\{ y \mid y(x) = c_1 e^{\lambda x} + c_2 x e^{\lambda x}, c_{1,2} \in \mathbb{C} \in \mathbb{R} \right\} \quad (9.51)$$

Bleibt z.z.: $y = x e^{\lambda x}$ löst DGL.

$$\begin{aligned} y' &= e^{\lambda x} + \lambda x e^{\lambda x} \\ y'' &= \lambda e^{\lambda x} + \lambda e^{\lambda x} + \lambda^2 x e^{\lambda x} \end{aligned} \quad (9.52)$$

in DGL

$$y'' + a_1 y' + a_0 y = \underbrace{\lambda^2 x e^{\lambda x} + a_1 \lambda x e^{\lambda x} + a_0 x e^{\lambda x}}_{=x e^{\lambda x} (\lambda^2 + \lambda a_1 + a_0) = x e^{\lambda x} \chi(\lambda) = 0} + \underbrace{2\lambda e^{\lambda x} + a_1 e^{\lambda x}}_{e^{\lambda x} (2\lambda + a_1) = e^{\lambda x} \chi'(\lambda) = 0} \quad (9.53)$$

denn

$$\begin{aligned} \chi(x) &= x^2 + a_1 x + a_0 &= (x - \lambda)^2 & \text{(doppelte Nullstelle)} \\ \chi'(x) &= 2x + a_1 &= 2(x - \lambda) \\ \chi'(\lambda) &= 2\lambda + a_1 &= 0. \end{aligned} \quad (9.54)$$

Bemerkungen:

1. Lösungsmenge ist stets zweidimensionaler Vektorraum.
2. Für ein AWP kann noch $y(x_0) = y_0$ und $y'(x_0) = v_0$ vorgegeben werden.

Beispiel: AWP für Federpendel

$$\ddot{s} + \omega^2 s = 0, \quad s(0) = 0, \quad \dot{s}(0) = 1 \quad (9.55)$$

allgemeine Lösung

$$s(t) = c_1 e^{i\omega t} + c_2 e^{-i\omega t} \quad (9.56)$$

$$s(0) = c_1 + c_2 \stackrel{!}{=} 0 \quad \Rightarrow \quad c_2 = -c_1 \quad (9.57)$$

$$\dot{s}(0) = i\omega c_1 - i\omega c_2 \stackrel{!}{=} 1 \quad \Rightarrow \quad 2i\omega c_1 = 1 \quad (9.58)$$

also

$$c_1 = \frac{1}{2i\omega} = -c_2 \quad \Rightarrow \quad s(t) = \frac{\sin(\omega t)}{\omega}. \quad (9.59)$$

Lösung der inhomogenen DGL: Variation der Konstanten

$$y'' + a_1 y' + a_0 y = f(x) \quad (9.60)$$

Seien y_1 und y_2 zwei l.u. Lösungen der homogenen DGL. Ansatz für partikuläre Lösung:

$$y(x) = c_1(x) y_1(x) + c_2(x) y_2(x) \quad (9.61)$$

Ableiten:

$$y'(x) = c_1 y_1' + c_2 y_2' + \underbrace{c_1' y_1 + c_2' y_2}_{=: R(x)} \quad (9.62)$$

$$y''(x) = c_1 y_1'' + c_1' y_1' + c_2 y_2'' + c_2' y_2' + R'$$

Einsetzen:

$$\begin{aligned} f(x) &= y'' + a_1 y' + a_0 y \\ f(x) &= \underbrace{c_1 y_1'' + a_1 c_1 y_1' + a_0 c_1 y_1}_{=0 \text{ da } y_1 \text{ homogene Lös.}} + \underbrace{c_2 y_2'' + a_1 c_2 y_2' + a_0 c_2 y_2}_{=0 \text{ da } y_2 \text{ homogene Lös.}} + c_1' y_1' + c_2' y_2' + R' + a_1 R \end{aligned} \quad (9.63)$$

Falls $R(x) \equiv 0$ (und damit $R' \equiv 0$): LGS für c_1' und c_2'

$$\begin{aligned} c_1' y_1 + c_2' y_2 &= 0 \quad (R(x) = 0) \\ c_1' y_1' + c_2' y_2' &= f(x) \end{aligned} \quad (9.64)$$

Lösung, z.B. durch Cramersche Regel:

$$\begin{aligned} c_1' &= \frac{1}{W(x)} \det \begin{pmatrix} 0 & y_2 \\ f & y_2' \end{pmatrix} = -\frac{y_2(x) f(x)}{W(x)} \\ c_2' &= \frac{1}{W(x)} \det \begin{pmatrix} y_1 & 0 \\ y_1' & f \end{pmatrix} = \frac{y_1(x) f(x)}{W(x)} \end{aligned} \quad (9.65)$$

mit der Wronski-Determinanten

$$W(x) = \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix} \quad (9.66)$$

Falls wir durch W teilen durften!

Lemma 20. (Wronski-Determinante)

Seien y_1 und y_2 Lösungen von

$$y'' + a_1y' + a_0y = 0$$

und $W(x) = \det \begin{pmatrix} y_1 & y_2 \\ y_1' & y_2' \end{pmatrix}$. Dann gilt:

1. $W'(x) = -a_1W(x)$ (und damit $W(x) = W(x_0)e^{-a_1(x-x_0)}$).
2. $W(x)$ besitzt keine Nullstelle $\Leftrightarrow y_1$ und y_2 sind l.u.

Beweis:

1.

$$\begin{aligned} W(x) &= y_1y_2' - y_1'y_2 \\ W'(x) &= y_1'y_2' + y_1y_2'' - y_1''y_2 - y_1'y_2' \\ &\stackrel{\text{DGL}}{=} y_1(-a_1y_2' - a_0y_2) - y_2(-a_1y_1' - a_0y_1) \\ &= -a_1(y_1y_2' - y_2y_1') \\ &= -a_1W(x) \end{aligned} \tag{9.67}$$

2.
 - Wegen Teil 1 gilt $W(x) = Ce^{-a_1x}$ mit einer Konstanten C .
 - Da die e-Funktion keine (reellen) Nullstellen hat, ist W entweder überall Null (falls $C = 0$) oder nirgends (falls $C \neq 0$).
 - $W(x) = 0$ überall $\Leftrightarrow \begin{pmatrix} y_2 \\ y_2' \end{pmatrix}$ ist Vielfaches von $\begin{pmatrix} y_1 \\ y_1' \end{pmatrix}$. (Eigenschaft von \det)

□

Also lautet eine partikuläre Lösung:

$$y_p(x) = -y_1(x) \int^x \frac{y_2(t)f(t)}{W(t)} dt + y_2(x) \int^x \frac{y_1(t)f(t)}{W(t)} dt \tag{9.68}$$

Beispiel:

$$y'' + 5y' + 6y = \sin x \tag{9.69}$$

charakteristisches Polynom: $\chi(x) = x^2 + 5x + 6$

$$\text{Nullstellen: } \lambda_{1,2} = \frac{-5 \pm \sqrt{25 - 24}}{2} = \frac{-5 \pm 1}{2} = \begin{cases} -2 \\ -3 \end{cases}$$

$$\Rightarrow L_h = \left\{ y \mid y(x) = c_1e^{-2x} + c_2e^{-3x}, c_{1,2} \in \mathbb{C} \right\} \tag{9.70}$$

$$W(x) = \det \begin{pmatrix} e^{-2x} & e^{-3x} \\ -2e^{-2x} & -3e^{-3x} \end{pmatrix} = -3e^{-5x} + 2e^{-5x} = -e^{-5x} \tag{9.71}$$

$$\begin{aligned}
y_p &= -e^{-2x} \int^x \frac{e^{-3t} \sin t}{(-e^{-5t})} dt + e^{-3x} \int^x \frac{e^{-2t} \sin t}{(-e^{-5t})} dt \\
&= e^{-2x} \int^x e^{2t} \sin t dt - e^{-3x} \int^x e^{3t} \sin t dt
\end{aligned} \tag{9.72}$$

NR:

$$\begin{aligned}
\int e^{\alpha x} \sin x dx &\stackrel{\text{p.I.}}{=} -e^{\alpha x} \cos x + \alpha \int e^{\alpha x} \cos x dx \\
&\stackrel{\text{p.I.}}{=} -e^{\alpha x} \cos x + \alpha e^{\alpha x} \sin x - \alpha^2 \int e^{\alpha x} \sin x dx \\
\Rightarrow \int e^{\alpha x} \sin x dx &= \frac{e^{\alpha x} (\alpha \sin x - \cos x)}{1 + \alpha^2}
\end{aligned} \tag{9.73}$$

... und damit:

$$y_p = \frac{2 \sin x - \cos x}{5} - \frac{3 \sin x - \cos x}{10} = \frac{1}{10} \sin x - \frac{1}{10} \cos x. \tag{9.74}$$

Bemerkung: Ansatz $y_p = A \sin x + B \cos x$ hätte auch zum Ziel geführt.

9.4 Existenz und Eindeutigkeit für DGLn 1. Ordnung

Hat das AWP $y' = f(x, y)$, $y(x_0) = y_0$, eine Lösung und ist diese eindeutig?

Beispiel: $y'(x) = 2\sqrt{|y(x)|}$, $y(0) = 0$.

Satz 21. (Picard-Lindelöf)

Die Funktion $f : D \rightarrow \mathbb{R}$ sei im Rechteck

$$D = \{(x, y) \mid |x - x_0| \leq a, |y - y_0| \leq b, a, b \in \mathbb{R}^+ \text{ fest}\} \subset \mathbb{R}^2 \tag{9.75}$$

stetig und dort stetig nach y differenzierbar.¹⁴ Weiter seien M und h durch

$$M = \max_{(x,y) \in D} |f(x, y)| \quad \text{und} \quad h = \min\left(a, \frac{b}{M}\right) \tag{9.76}$$

definiert. Dann gibt es in der Umgebung

$$U_h(x_0) = \{x \mid |x - x_0| < h\} \tag{9.77}$$

der Stelle x_0 genau eine Lösung des AWP

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0. \tag{9.78}$$

¹⁴Man sagt *partiell* nach y differenzierbar und meint: Man behandelt x wie eine Konstante und leitet nach y ab. Dafür schreibt man $\frac{\partial f}{\partial y}$ – siehe später.

Beweisskizze:

Ein gut lesbarer vollständiger Beweis findet sich z.B. in Burg, Haf und Wille: Höhere Mathematik für Ingenieure, Band III, B.G. Teubner, Stuttgart, 1993, S. 17ff.

- (i) Schreibe AWP als Integralgleichung.
- (ii) Definiere eine Folge von Näherungen $y_n(x)$ für die gesuchte Funktion $y(x)$.
- (iii) Zeige, dass die Folge wohldefiniert ist.
- (iv) Zeige, dass die Folge gegen das gesuchte $y(x)$ konvergiert.
- (v) Zeige die Eindeutigkeit der gefundenen Lösung.

ad (i):

$$y' = f(x, y), \quad y(x_0) = y_0 \quad \Leftrightarrow \quad y(x) = y_0 + \int_{x_0}^x f(t, y(t)) dt, \quad (9.79)$$

denn “ \Rightarrow ”

$$y(x) - y_0 \stackrel{\text{Anfangswert}}{=} y(x) - y(x_0) \stackrel{\text{Hauptsatz}}{=} \int_{x_0}^x y'(t) dt \stackrel{\text{DGL}}{=} \int_{x_0}^x f(t, y(t)) dt \quad (9.80)$$

und “ \Leftarrow ”

$$\begin{aligned} y'(x) &\stackrel{\text{Integralgn.}}{=} 0 + \frac{d}{dx} \int_{x_0}^x f(t, y(t)) dt \stackrel{\text{Diff. nach oberer Grenze}}{=} f(x, y(x)), \\ y(x_0) &\stackrel{\text{Integralgn.}}{=} y_0 + \int_{x_0}^{x_0} f(t, y(t)) dt = y_0. \end{aligned} \quad (9.81)$$

ad (ii): Picard-Iteration

$$\begin{aligned} y_0(x) &:= y_0 \quad (\text{also konstant}) \\ y_1(x) &:= y_0 + \int_{x_0}^x f(t, y_0) dt \\ &\vdots \\ y_n(x) &:= y_0 + \int_{x_0}^x f(t, y_{n-1}(t)) dt \end{aligned} \quad (9.82)$$

ad (iii):z.z.: $y_n(x)$ bleibt im Definitionsbereich von f falls $x \in U_h(x_0)$, d.h.

$$|y_n(x) - y_0| \leq b \quad \forall |x - x_0| < h.$$

Vollständige Induktion:

 $n = 0$:

$$|y_0(x) - y_0| = 0 \leq b$$

$n \rightarrow n + 1$:

$$|y_{n+1}(x) - y_0| = \left| \int_{x_0}^x f(t, y_n(t)) dt \right| \stackrel{\text{I.V.}}{\leq} M|x - x_0| < Mh \leq b \quad \square$$

ad (iv) & (v)

- Zeige, dass die Folge konvergiert – ähnliche Abschätzungen wie zu (iii) nur etwas aufwändiger, hier geht Stetigkeit von $\frac{\partial f}{\partial y}$ ein.
- So wie die Folge definiert war, muss sie dann gegen eine Lösung des AWP konvergieren!
- Zeige Eindeutigkeit – wieder ähnliche Abschätzungen, verwende dabei nochmals Stetigkeit von $\frac{\partial f}{\partial y}$.

Beispiel: AWP $y' + \pi y = 0$, $y(0) = 42$

- berechne Picard-Iterierte y_0, y_1, y_2, y_3
- rate Formel für y_n und beweise mit vollständiger Induktion
- bestimme $\lim_{n \rightarrow \infty} y_n$ und vergleiche mit Lösung nach Abschnitt 9.1 oder 9.2.

10 Hauptachsentransformation

Beispiel:

$$\begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (10.1)$$

$$\begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ -6 \end{pmatrix} = 6 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (10.2)$$

Allgemein: Gegeben eine quadratische Matrix A .

Gibt es Vektoren \vec{x} und Zahlen λ , so dass

$$A\vec{x} = \lambda\vec{x} ? \quad (10.3)$$

($\vec{x} \neq 0$, sonst langweilig). Wie findet man sie?

Inwiefern charakterisieren sie A ?

10.1 Eigenwerte und Eigenvektoren

Definition: (Eigenwert, Eigenvektor, charakteristisches Polynom)

Sei A eine quadratische Matrix, d.h. $A \in \mathbb{C}^{n \times n}$.

Eine Zahl $\lambda \in \mathbb{C}$ heißt Eigenwert (EW) von A wenn gilt:

$$\text{Es gibt ein } \vec{x} \neq \vec{0} \text{ mit } A\vec{x} = \lambda\vec{x}. \quad (10.4)$$

Jedes solche $\vec{x} \in \mathbb{C}^n$ heißt zugehöriger Eigenvektor (EV).

$\chi_A(\lambda) := \det(A - \lambda I)$ heißt charakteristisches Polynom von A .

Bemerkungen:

1. χ_A ist ein Polynom n -ten Grades, denn

$$\begin{aligned} \chi_A(\lambda) &= \det \begin{pmatrix} a_{11} - \lambda & & * \\ & \ddots & \\ * & & a_{nn} - \lambda \end{pmatrix}, & * : \text{Terme ohne } \lambda \\ &= (a_{11} - \lambda) \cdots (a_{nn} - \lambda) + \dots \\ &= (-1)^n \lambda^n + \dots \end{aligned} \tag{10.5}$$

2. Ist \vec{x} Eigenvektor zum Eigenwert λ , so auch $\mu \vec{x} \forall \mu \neq 0$.

Lemma 22. Sei $A \in \mathbb{C}^{n \times n}$: λ ist Eigenwert von $A \Leftrightarrow \chi_A(\lambda) = 0$

Beweis:

$$\begin{aligned} \lambda \text{ ist Eigenwert von } A &\Leftrightarrow A\vec{x} = \lambda\vec{x} \text{ für ein } \vec{x} \neq \vec{0} \\ &\Leftrightarrow A\vec{x} - \lambda\vec{x} = \vec{0} \text{ für ein } \vec{x} \neq \vec{0} \\ &\Leftrightarrow A\vec{x} - \lambda I\vec{x} = \vec{0} \text{ für ein } \vec{x} \neq \vec{0} \\ &\Leftrightarrow (A - \lambda I)\vec{x} = \vec{0} \text{ für ein } \vec{x} \neq \vec{0} \\ &\Leftrightarrow (A - \lambda I)\vec{x} = \vec{0} \text{ hat nichttriviale Lösung} \\ &\Leftrightarrow \det(A - \lambda I) = 0 \Leftrightarrow \chi_A(\lambda) = 0 \end{aligned} \tag{10.6}$$

□

Bemerkung:

Zu λ gehörende Eigenvektoren findet man durch Lösen des homogenen LGS $(A - \lambda I)\vec{x} = \vec{0}$.

Beispiel:

$$A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} \tag{10.7}$$

$$\chi_A(\lambda) = \det \begin{pmatrix} 5 - \lambda & -1 \\ -1 & 5 - \lambda \end{pmatrix} = (5 - \lambda)^2 - 1 = \lambda^2 - 10\lambda + 24 \stackrel{!}{=} 0 \tag{10.8}$$

$$\lambda_{1,2} = \frac{10 \pm \sqrt{100 - 96}}{2} = 5 \pm 1 \tag{10.9}$$

Eigenvektoren zu $\lambda_1 = 6$:

$$\frac{\left(\begin{array}{cc|c} 5-6 & -1 & 0 \\ -1 & 5-6 & 0 \end{array} \right)}{\left(\begin{array}{cc|c} -1 & -1 & 0 \\ -1 & -1 & 0 \end{array} \right)} \Rightarrow \vec{x}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ und Vielfache} \tag{10.10}$$

...zu $\lambda_2 = 4$:

$$\left(\begin{array}{cc|c} 1 & -1 & 0 \\ -1 & 1 & 0 \end{array} \right) \Rightarrow \vec{x}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ und Vielfache} \tag{10.11}$$

10.2 Einige Begriffe

Definition: Sei A eine quadratische Matrix, $A \in \mathbb{C}^{n \times n}$.

- (i) A heißt hermitesch, falls $\overline{A}^T = A$.
- (ii) A heißt symmetrisch, falls A reell, also $A \in \mathbb{R}^{n \times n}$, und $A^T = A$.
- (iii) A heißt unitär, falls $\overline{A}^T A = I$ (d.h. $A^{-1} = \overline{A}^T$).
- (iv) A heißt orthogonal, falls A reell, also $A \in \mathbb{R}^{n \times n}$, und $A^T A = I$ (d.h. $A^{-1} = A^T$).

Bemerkungen:

1. $A = (\vec{a}_1, \dots, \vec{a}_n) \in \mathbb{C}^{n \times n}$ ist unitär $\Leftrightarrow \vec{a}_1, \dots, \vec{a}_n$ bilden ON-Basis des \mathbb{C}^n ,
(bezüglich des kanonischen Skalarprodukts $\langle \vec{x}, \vec{y} \rangle = \overline{\vec{x}}^T \vec{y}$)
denn

$$\begin{aligned}
 A \text{ unitär} &\Leftrightarrow \overline{A}^T A = I \\
 &\Leftrightarrow \begin{pmatrix} \overline{\vec{a}_1}^T \\ \vdots \\ \overline{\vec{a}_n}^T \end{pmatrix} (\vec{a}_1, \dots, \vec{a}_n) = I \\
 &\Leftrightarrow (\overline{\vec{a}_j}^T \vec{a}_k) = I \\
 &\Leftrightarrow \vec{a}_1, \dots, \vec{a}_n \text{ bilden ONB}
 \end{aligned} \tag{10.12}$$

2. A unitär $\Rightarrow |\det A| = 1$, denn

$$\begin{aligned}
 \overline{A}^T A = I &\Rightarrow \det(\overline{A}^T A) = \det I = 1 \Leftrightarrow \det \overline{A}^T \cdot \det A = 1 \\
 &\Leftrightarrow \overline{\det A} \cdot \det A = 1 \\
 &\Leftrightarrow |\det A|^2 = 1 \Leftrightarrow |\det A| = 1
 \end{aligned} \tag{10.13}$$

3. Unitäre Matrizen bilden bezüglich \cdot eine Gruppe.

4. $A \in \mathbb{C}^{n \times n}$ unitär $\Leftrightarrow \|A\vec{x}\| = \|\vec{x}\| \quad \forall \vec{x} \in \mathbb{C}^n$, denn

$$\|A\vec{x}\|^2 = \langle A\vec{x}, A\vec{x} \rangle = \overline{A\vec{x}}^T A\vec{x} = \overline{\vec{x}}^T \overline{A}^T A \vec{x} \tag{10.14}$$

d.h.

$$\|A\vec{x}\|^2 = \|\vec{x}\|^2 \quad \forall \vec{x} \in \mathbb{C}^n \Leftrightarrow \overline{A}^T A = I \Leftrightarrow A \text{ unitär.} \tag{10.15}$$

5. $A \in \mathbb{C}^{n \times n}$ hermitesch $\Leftrightarrow \langle \vec{x}, A\vec{y} \rangle = \langle A\vec{x}, \vec{y} \rangle \quad \forall x, y \in \mathbb{C}^n$, denn

$$\text{“}\Rightarrow\text{”}: \langle \vec{x}, A\vec{y} \rangle = \overline{\vec{x}}^T A\vec{y} = \overline{\vec{x}}^T \overline{A}^T \vec{y} = \overline{A\vec{x}}^T \vec{y} = \langle A\vec{x}, \vec{y} \rangle$$

$$\text{“}\Leftarrow\text{”}: a_{jk} = \langle \vec{e}_j, A\vec{e}_k \rangle = \langle A\vec{e}_j, \vec{e}_k \rangle = \overline{\langle \vec{e}_k, A\vec{e}_j \rangle} = \overline{a_{kj}}$$

6. Ist $B \in \mathbb{C}^{n \times n}$ hermitesch, so auch $\overline{A}^T B A$ für beliebiges $A \in \mathbb{C}^{n \times n}$, denn

$$\overline{\overline{A}^T B A}^T = \overline{A}^T \overline{B}^T \overline{\overline{A}^T}^T = \overline{A}^T B A. \tag{10.16}$$

10.3 Diagonalisierbarkeit

Definition: $A \in \mathbb{C}^{n \times n}$ heißt diagonalisierbar $\Leftrightarrow \exists S$ regulär (d.h. invertierbar) mit

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}. \quad (10.17)$$

Bemerkungen:

1. $S = (\vec{s}_1, \dots, \vec{s}_n)$ ist regulär $\Leftrightarrow \det S \neq 0$
 $\Leftrightarrow \vec{s}_1, \dots, \vec{s}_n$ sind l.u.
2. Die λ_j sind Eigenwerte von A mit zugehörigen Eigenvektoren \vec{s}_j , denn

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \Leftrightarrow AS = S \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad (10.18)$$

und mit

$$AS = A(\vec{s}_1, \dots, \vec{s}_n) = (A\vec{s}_1, \dots, A\vec{s}_n) \quad \text{sowie} \quad (10.19)$$

$$S \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = (\vec{s}_1, \dots, \vec{s}_n) \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = (\vec{s}_1 \lambda_1, \dots, \vec{s}_n \lambda_n)$$

folgt $A\vec{s}_j = \lambda_j \vec{s}_j$.

3. Eigenvektoren zu verschiedenen Eigenwerten sind l.u.

Beweis:

$A\vec{x} = \lambda\vec{x}$, $A\vec{y} = \mu\vec{y}$ mit $\vec{x} \neq \vec{0}$, $\vec{y} \neq \vec{0}$, $\lambda \neq \mu$.

Sei $\alpha\vec{x} = \beta\vec{y}$ (z.z.: $\alpha = \beta = 0$)

$$\begin{aligned} A\alpha\vec{x} &= \alpha\lambda\vec{x} = A\beta\vec{y} = \beta\mu\vec{y} = \alpha\mu\vec{x} \\ \Rightarrow \alpha(\lambda - \mu)\vec{x} &= \vec{0} \\ \xRightarrow{\vec{x} \neq \vec{0}, \lambda \neq \mu} \alpha &= 0 \\ \Rightarrow \beta\vec{y} &= \vec{0} \\ \xRightarrow{\vec{y} \neq \vec{0}} \beta &= 0 \end{aligned}$$

□

Beispiel:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix} \quad (10.20)$$

Berechne: EWe, EVen, S , S^{-1} , $S^{-1}AS$.

Satz 23. Ist $A \in \mathbb{C}^{n \times n}$ hermitesch (oder symmetrisch), so folgt:

- (i) Alle Eigenwerte von A sind reell.
- (ii) Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal zueinander.

Beweis:

- (i) Sei $A\vec{x} = \lambda\vec{x}$, $\vec{x} \neq \vec{0}$

$$\begin{aligned} \langle \vec{x}, A\vec{x} \rangle &= \langle \vec{x}, \lambda\vec{x} \rangle = \lambda \|\vec{x}\|^2 \\ &= \langle A\vec{x}, \vec{x} \rangle = \langle \lambda\vec{x}, \vec{x} \rangle = \bar{\lambda} \|\vec{x}\|^2 \end{aligned} \quad (10.21)$$

$\Rightarrow \lambda = \bar{\lambda}$, also $\lambda \in \mathbb{R}$.

- (ii) Sei $A\vec{x} = \lambda\vec{x}$, $A\vec{y} = \mu\vec{y}$ mit $\vec{x} \neq \vec{0}$, $\vec{y} \neq \vec{0}$, $\lambda \neq \mu$.

$$(\lambda - \mu)\langle \vec{x}, \vec{y} \rangle \stackrel{\lambda \in \mathbb{R}}{=} \langle \lambda\vec{x}, \vec{y} \rangle - \langle \vec{x}, \mu\vec{y} \rangle = \langle A\vec{x}, \vec{y} \rangle - \langle \vec{x}, A\vec{y} \rangle \stackrel{A \text{ herm.}}{=} 0 \quad (10.22)$$

□

Satz 24. (Hauptachsentransformation)

Sei $A \in \mathbb{C}^{n \times n}$ hermitesch (symmetrisch), dann existiert eine unitäre (orthogonale) Matrix U mit

$$\bar{U}^T A U = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad (10.23)$$

wobei λ_j die Eigenwerte von A sind.

Bemerkung: Ist die Matrix nicht hermitesch, dann muss sie auch nicht unbedingt diagonalisierbar sein. Beispiel: $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. χ_A hat 1 als doppelte Nullstelle, aber der zugehörige Eigenraum ist nur eindimensional ($\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und Vielfache).

Beweis: Für den Fall, dass alle Eigenwerte verschieden sind, bereits gezeigt.

Sei daher nun λ eine ν -fache Nullstelle von χ_A , d.h.

$$\chi_A(t) = (\lambda - t)^\nu p(t), \quad p(\lambda) \neq 0, \quad (10.24)$$

mit einem Polynom p . Sei $\{\vec{x}_1, \dots, \vec{x}_\ell\}$ eine maximale Menge l.u. EVen zum EW λ , d.h.

$$A\vec{x}_j = \lambda\vec{x}_j \quad \forall j = 1, \dots, \ell. \quad (10.25)$$

Wir wissen: $1 \leq \ell \leq \nu$. Zu zeigen: $\ell = \nu$.

Orthonormiere: $\vec{u}_1, \dots, \vec{u}_\ell$ ONB von $\text{span}(\vec{x}_1, \dots, \vec{x}_\ell)$.

Ergänze zu ONB von \mathbb{C}^n durch $\vec{u}_{\ell+1}, \dots, \vec{u}_n$. $U = (\vec{u}_1, \dots, \vec{u}_n)$ ist unitär.

Betrachte $B = \overline{U}^T AU = (b_{jk})$.

Wir zeigen: B hat die Form

$$B = \underbrace{\left(\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & 0 \\ 0 & & \lambda & \\ \hline & & & \tilde{A} \end{array} \right)}_{\ell \text{ Spalten}} \left. \vphantom{\begin{array}{ccc|c} \lambda & & 0 & \\ & \ddots & & 0 \\ 0 & & \lambda & \\ \hline & & & \tilde{A} \end{array}} \right\} \ell \text{ Zeilen}, \quad (10.26)$$

denn mit

$$b_{jk} = \overline{u}_j^T A \vec{u}_k = \langle \overline{u}_j, A \vec{u}_k \rangle \quad (10.27)$$

folgt

$$b_{jk} = \langle \overline{u}_j, \lambda \vec{u}_k \rangle = \lambda \delta_{jk} \quad \forall k = 1, \dots, \ell, \quad (10.28)$$

was die Blöcke links oben und links unten liefert.

Der Block rechts oben folgt nur, weil A hermitesch ist:

$$b_{jk} = \langle \overline{u}_j, A \vec{u}_k \rangle = \langle A \overline{u}_j, \vec{u}_k \rangle = \langle \lambda \overline{u}_j, \vec{u}_k \rangle = \lambda \delta_{jk} \quad \forall j = 1, \dots, \ell. \quad (10.29)$$

Nun gilt einerseits

$$\begin{aligned} \chi_B(t) &= \det(B - tI) = \det(\overline{U}^T AU - t \overline{U}^T U) = \det(U^T (A - tI) U) \\ &= \det(U^T) \det(A - tI) \det(U) = \det(U^T U) \det(A - tI) = \det(A - tI) \\ &= \chi_A(t) \end{aligned} \quad (10.30)$$

und andererseits

$$\chi_B = (\lambda - t)^\ell \underbrace{\det(\tilde{A} - tI)}_{=: q(t)} \quad (10.31)$$

Noch zu zeigen: $q(\lambda) \neq 0$, denn daraus folgt $\ell = \nu$.

Annahme (führen wir gleich zum Widerspruch): $q(\lambda) = 0$

$\Rightarrow \exists \vec{x} \neq \vec{0}$, das EV von \tilde{A} zum EW λ ist, d.h.

$$\begin{aligned} \tilde{A} \vec{x} = \lambda \vec{x} &\Leftrightarrow B \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vec{x} \end{pmatrix} = \lambda \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vec{x} \end{pmatrix} &\Leftrightarrow \overline{U}^T AU \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vec{x} \end{pmatrix} = \lambda \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vec{x} \end{pmatrix} \\ &\Leftrightarrow AU \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vec{x} \end{pmatrix} = \lambda U \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vec{x} \end{pmatrix} \end{aligned} \quad (10.32)$$

d.h. $U \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vec{x} \end{pmatrix}$ ist EV von A zum EW λ . Da aber $U \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vec{x} \end{pmatrix}$ eine Linearkombination von $\vec{u}_{\ell+1}, \dots, \vec{u}_n$ ist, ist das ein *Widerspruch* dazu, dass $\text{span}(\vec{x}_1, \dots, \vec{x}_\ell)$ alle EVen zum EW λ enthält. \square

10.3.1 Praktische Durchführung der HAT

Gegeben A hermitesch (oder symmetrisch)

1. Berechne $\chi_A(x) = \det(A - xI)$.

2. Faktorisiere $\chi_A(x) = (-1)^n \prod_{\rho=1}^r (x - \lambda_\rho)^{n_\rho}$,

mit $\lambda_1, \dots, \lambda_r$ verschiedenen NST mit Vielfachheiten n_1, \dots, n_r .

3. Bestimme die Eigenräume $E_\rho = \{\vec{x} \mid (A - \lambda_\rho I)\vec{x} = \vec{0}\}$, d.h. löse LGSs.

Bemerkungen: (i) $\dim E_\rho = n_\rho$
(ii) $E_\rho \perp E_\sigma$ für $\rho \neq \sigma$

4. Bestimme für jedes E_ρ eine ON-Basis, $\{\vec{u}_1^\rho, \dots, \vec{u}_{n_\rho}^\rho\}$.

Alle zusammen, $\{\vec{u}_1^1, \dots, \vec{u}_{n_1}^1, \vec{u}_1^2, \dots, \vec{u}_{n_r}^r\}$, ergeben eine ON-Basis des \mathbb{C}^n (\mathbb{R}^n).

$U = (\vec{u}_1, \dots, \vec{u}_n) = (\vec{u}_1^1, \dots, \vec{u}_{n_1}^1, \vec{u}_1^2, \dots, \vec{u}_{n_r}^r)$ ist die gesuchte Transformationsmatrix.

Beispiel:

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{6}{5} & \frac{2}{5} \\ 0 & \frac{2}{5} & \frac{9}{5} \end{pmatrix} \quad (10.33)$$

$$\begin{aligned} \chi_A(x) &= \det(A - xI) \\ &= (1 - x) \left(x^2 - \frac{15}{5}x + \frac{54}{25} - \frac{4}{25} \right) \\ &= (1 - x)(x^2 - 3x + 2) \end{aligned} \quad (10.34)$$

$\lambda_1 = 1$, $\lambda_{3,2} = \frac{3 \pm \sqrt{9-8}}{2}$, d.h. $\lambda_2 = 1$, $\lambda_3 = 2$

Eigenvektor zu λ_3 :

$$\left(\begin{array}{ccc|c} -1 & 0 & 0 & 0 \\ 0 & -\frac{4}{5} & \frac{2}{5} & 0 \\ 0 & \frac{2}{5} & -\frac{1}{5} & 0 \end{array} \right) \Rightarrow \vec{u}_3 = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \quad (10.35)$$

Eigenvektoren zu $\lambda_1 = \lambda_2 = 1$:

$$\left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & 0 \\ 0 & \frac{2}{5} & \frac{4}{5} & 0 \end{array} \right) \Rightarrow \vec{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{u}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \quad (10.36)$$

Und damit

$$U = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 & \sqrt{5} & 0 \\ 1 & 0 & 2 \\ 2 & 0 & -1 \end{pmatrix} \Rightarrow \bar{U}^T A U = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (10.37)$$

Anwendungen:

A hermitesch, $\bar{U}^T A U = D$ mit Diagonalmatrix $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ der Eigenwerte.

$$A^n = \left(U D \bar{U}^T \right)^n = U D \underbrace{\bar{U}^T U}_{=I} D \bar{U}^T \dots U D \bar{U}^T = U D^n \bar{U}^T \quad (10.38)$$

$$\begin{aligned} e^{Ax} &= \sum_{n=0}^{\infty} \frac{x^n}{n!} A^n && (A^0 = I) \\ &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \left(U D \bar{U}^T \right)^n \\ &= \sum_{n=0}^{\infty} \frac{x^n}{n!} U D^n \bar{U}^T \\ &= U \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} D^n \right) \bar{U}^T \\ &= U e^{Dx} \bar{U}^T \\ &= U \begin{pmatrix} e^{\lambda_1 x} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n x} \end{pmatrix} \bar{U}^T \end{aligned} \quad (10.39)$$

... und falls alle $\lambda_j \neq 0$:

$$A^{-1} = U D^{-1} \bar{U}^T \quad \text{mit} \quad D^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\lambda_n} \end{pmatrix} \quad (10.40)$$

10.4 Quadratische Formen

Beispiel $q(x, y) = x^2 + 2xy + 3y^2 \dots$

Sei $B = (b_{jk}) \in \mathbb{R}^{n \times n}$, $\vec{x} \in \mathbb{R}^n$; betrachte

$$\vec{x}^T B \vec{x} = \sum_{j,k=1}^n x_j b_{jk} x_k = \sum_{j=1}^n b_{jj} x_j^2 + \sum_{j \neq k} b_{jk} x_j x_k \quad (10.41)$$

Mit

$$A = \frac{1}{2} (B + B^T) \quad \text{d.h.} \quad a_{jk} = \frac{b_{jk} + b_{kj}}{2}, \quad (10.42)$$

also A symmetrisch, gilt

$$\vec{x}^T B \vec{x} = \vec{x}^T A \vec{x}. \quad (10.43)$$

Im Folgenden interessieren uns daher nur symmetrische Matrizen (oder analog hermitesche Matrizen im Komplexen).

Bemerkung zur Notation: $q_A(\vec{x}) = \vec{x}^T A \vec{x} = \langle \vec{x}, A \vec{x} \rangle$

Definition: Sei $A \in \mathbb{R}^{n \times n}$ eine symmetrische Matrix. Dann heißt

$$q_A(\vec{x}) := \vec{x}^T A \vec{x} = \sum_{j,k=1}^n a_{jk} x_j x_k \quad (10.44)$$

die zugehörige quadratische Form.

Eine quadratische Form q_D heißt Diagonalform, falls D eine Diagonalmatrix ist, etwa

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \quad \text{und damit} \quad q_D(\vec{x}) = \sum_{j=1}^n \lambda_j x_j^2. \quad (10.45)$$

Bemerkung: $q_A(\vec{x}) = q_B(\vec{x}) \quad \forall \vec{x} \in \mathbb{R}^n \quad \Leftrightarrow \quad A = B$

Beweis:

“ \Leftarrow ”: klar

“ \Rightarrow ”:

(i) $\vec{x} = \vec{e}_j$ (kanonischer Basisvektor)

$$q_A(\vec{e}_j) = \vec{e}_j^T A \vec{e}_j = a_{jj} = q_B(\vec{e}_j) = b_{jj}, \quad (10.46)$$

d.h. Diagonalelemente sind gleich.

(ii) $\vec{x} = \vec{e}_j + \vec{e}_k, j \neq k$

$$\begin{aligned} q_A(\vec{e}_j + \vec{e}_k) &= (\vec{e}_j + \vec{e}_k)^T A (\vec{e}_j + \vec{e}_k) = a_{jj} + a_{jk} + a_{kj} + a_{kk} = a_{jj} + 2a_{kj} + a_{kk} \\ &= q_B(\vec{e}_j + \vec{e}_k) = b_{jj} + 2b_{kj} + b_{kk} \end{aligned} \quad (10.47)$$

\Rightarrow $2a_{jk} = 2b_{jk}$, also auch Nichtdiagonalelemente sind gleich. □

Geometrische Interpretation für $n = 2$: Kegelschnitte

Sei $A \in \mathbb{R}^{2 \times 2}$ symmetrisch und $\mathfrak{K}_A \subset \mathbb{R}^2$ die Kurve

$$\mathfrak{K}_A := \{ \vec{x} \in \mathbb{R}^2 \mid q_A(\vec{x}) = 1 \}. \quad (10.48)$$

Speziell:

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad (10.49)$$

- $\lambda_1, \lambda_2 > 0$: Ellipse
(Kreis für $\lambda_1 = \lambda_2 > 0$)
- $\lambda_1 \cdot \lambda_2 < 0$: Hyperbel

- $\lambda_1 \cdot \lambda_2 = 0$ ($\lambda_1 \neq \lambda_2$): Geraden

Aber was ist $5x_1^2 - 2x_1x_2 + 5x_2^2 = 1$?

Lemma 25. Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und $U = (\vec{u}_1, \dots, \vec{u}_n)$ orthogonal. Mit

$$\vec{x} = \sum_{j=1}^n y_j \vec{u}_j, \quad \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad (10.50)$$

gilt dann

$$q_A(\vec{x}) = q_B(\vec{y}), \quad \text{wobei } B = U^T A U. \quad (10.51)$$

Beweis:

$$\vec{x} = \sum_{j=1}^n y_j \vec{u}_j = U \vec{y} \quad \Leftrightarrow \quad \vec{y} = U^T \vec{x} \quad (10.52)$$

$$q_A(\vec{x}) = \vec{x}^T A \vec{x} \stackrel{UU^T=I}{=} \underbrace{\vec{x}^T U}_{=\vec{y}^T} \underbrace{U^T A U}_{=B} \underbrace{U^T \vec{x}}_{=\vec{y}} = q_B(\vec{y}) \quad (10.53)$$

□

Ist insbesondere B diagonal,

$$B = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad (10.54)$$

dann ist $q_B(\vec{y})$ in Diagonalform,

$$q_B(\vec{y}) = \sum_{j=1}^n \lambda_j y_j^2, \quad (10.55)$$

wobei die λ_j die Eigenwerte von A zu den Eigenvektoren \vec{u}_j sind,

$$A \vec{u}_j = \lambda_j \vec{u}_j. \quad (10.56)$$

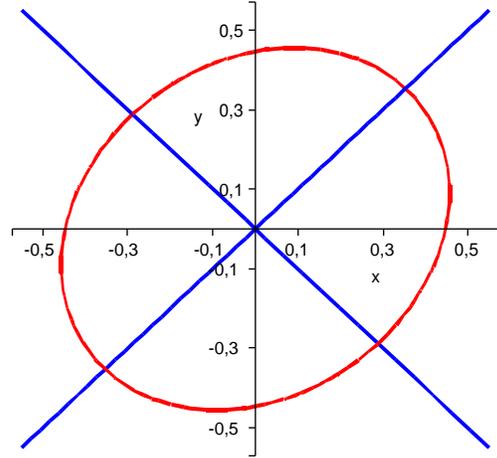
Im **Eingangsbeispiel** von Seite 118:

$$A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix} \quad (10.57)$$

Eigenwerte und (normierte) Eigenvektoren:

$$\lambda_1 = 4, \quad \lambda_2 = 6, \\ \vec{u}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{u}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (10.58)$$

Ellipse mit Halbachsen $\frac{1}{2}$ und $\frac{1}{\sqrt{6}} \approx 0,4$.



Definition: (Definitheit)

Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch, dann heißt A , beziehungsweise die quadratische Form q_A ,

- positiv definit $\Leftrightarrow q_A(\vec{x}) > 0 \quad \forall \vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$,
- positiv semidefinit $\Leftrightarrow q_A(\vec{x}) \geq 0 \quad \forall \vec{x} \in \mathbb{R}^n$,
- negativ definit $\Leftrightarrow q_A(\vec{x}) < 0 \quad \forall \vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\}$,
- negativ semidefinit $\Leftrightarrow q_A(\vec{x}) \leq 0 \quad \forall \vec{x} \in \mathbb{R}^n$,
- indefinit $\Leftrightarrow \exists \vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$ mit $q_A(\vec{x}_1) \cdot q_A(\vec{x}_2) < 0$

Satz 26. Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch, U orthogonal mit $U^T A U = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$,

dann sind äquivalent:

- (i) A ist positiv definit
- (ii) $\lambda_j > 0 \quad \forall j = 1, \dots, n$
- (iii) $\det \begin{pmatrix} a_{11} & \dots & a_{1\nu} \\ \vdots & & \vdots \\ a_{\nu 1} & \dots & a_{\nu\nu} \end{pmatrix} > 0 \quad \forall \nu = 1, \dots, n$ (Hauptunterdeterminanten > 0)

Bemerkung: A negativ definit $\Leftrightarrow -A$ positiv definit

Beweis:

(i) \Leftrightarrow (ii):

$$\begin{aligned} & q_A(\vec{x}) > 0 \quad \forall \vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\} \\ \Leftrightarrow & q_B(\vec{y}) > 0 \quad \forall \vec{x} = U\vec{y} \in \mathbb{R}^n \setminus \{\vec{0}\} \\ \Leftrightarrow & q_B(\vec{y}) = \sum_{j=1}^n \lambda_j y_j^2 > 0 \quad \forall \vec{y} = U^T \vec{x} \in \mathbb{R}^n \setminus \{\vec{0}\} \quad (10.59) \\ \Leftrightarrow & \lambda_j > 0 \quad \forall j = 1, \dots, n \end{aligned}$$

(i) \Leftrightarrow (iii):

Zunächst $n = 2$,

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \quad (10.60)$$

$$\begin{aligned} q_A(x, y) &= ax^2 + 2bxy + cy^2 \\ &= a \left(x^2 + 2\frac{b}{a}xy \right) + cy^2 \quad (\text{falls } a \neq 0) \\ &= a \left[\left(x + \frac{b}{a}y \right)^2 - \frac{b^2}{a^2}y^2 \right] + cy^2 \\ &= a \left(x + \frac{b}{a}y \right)^2 + \underbrace{\frac{ac - b^2}{a}}_{= \frac{1}{a} \det A} y^2 \end{aligned} \quad (10.61)$$

d.h. nur positiv definit, falls $a > 0$ und $\det A > 0$.

Bleibt noch $a = 0$:

$$q_A(x, y) = 2bxy + cy^2 \quad (10.62)$$

nimmt für gegebenes y durch geeignete Wahl von x jeden beliebigen Wert an (positiv wie negativ), also indefinit. Damit für $n = 2$ gezeigt.

Weiter mit Induktion, $n \rightarrow n + 1$:

Sei $A \in \mathbb{R}^{(n+1) \times (n+1)}$ symmetrisch und U die orthogonale Matrix, die den linken oberen $n \times n$ -Block diagonalisiert. Damit gilt

$$A \mapsto \begin{pmatrix} U & 0 \\ 0 \dots 0 & 1 \end{pmatrix}^T A \begin{pmatrix} U & 0 \\ 0 \dots 0 & 1 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 & b_1 \\ & \ddots & \vdots \\ 0 & & \lambda_n & b_n \\ b_1 & \dots & b_n & c \end{pmatrix} =: B \quad (10.63)$$

und

$$q_A \text{ positiv definit} \quad \Leftrightarrow \quad q_B \text{ positiv definit.} \quad (10.64)$$

$$\begin{aligned} q_B(\vec{x}) &= \sum_{j=1}^n \lambda_j x_j^2 + \sum_{j=1}^n 2b_j x_j x_{n+1} + cx_{n+1}^2 \\ &= \sum_{j=1}^n \lambda_j \left[\left(x_j + \frac{b_j}{\lambda_j} x_{n+1} \right)^2 - \frac{b_j^2}{\lambda_j^2} x_{n+1}^2 \right] + cx_{n+1}^2 \\ &= \sum_{j=1}^n \lambda_j \left(x_j + \frac{b_j}{\lambda_j} x_{n+1} \right)^2 + \underbrace{\frac{c \cdot \lambda_1 \cdots \lambda_n - \sum_{j=1}^n b_j^2 \frac{\lambda_1 \cdots \lambda_n}{\lambda_j}}{\lambda_1 \cdots \lambda_n}}_{= \frac{\det B}{\lambda_1 \cdots \lambda_n}} x_{n+1}^2 \end{aligned} \quad (10.65)$$

d.h. q_B positiv definit $\Leftrightarrow \lambda_1, \dots, \lambda_n > 0$ und $\det B > 0$.

Mit IV, (i) \Leftrightarrow (ii) und $\det A = \det B$ folgt die Behauptung. \square

Beispiele:

1.

$$A = \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 2 & 0 & 8 & 2 \\ 0 & 2 & 2 & 9 \end{pmatrix} \quad (10.66)$$

ist positiv definit, denn

$$\begin{aligned} a_{11} &= 1 > 0, \\ \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= 1 > 0, \\ \det \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 8 \end{pmatrix} &= 1(8 - 0) + 2(0 - 2) = 4 > 0, \end{aligned} \quad (10.67)$$

$$\begin{aligned} \det A &= \det \begin{pmatrix} 1 & 0 & 2 \\ 0 & 8 & 2 \\ 2 & 2 & 9 \end{pmatrix} + 2 \det \begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 2 \\ 2 & 2 & 9 \end{pmatrix} \\ &= 1(72 - 4) + 2(0 - 16) + 2(-2)(9 - 4) \\ &= 68 - 32 - 20 = 16 > 0. \end{aligned}$$

2.

$$B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -2 \end{pmatrix} \quad (10.68)$$

ist negativ definit, weil

$$-B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 2 \end{pmatrix} \quad (10.69)$$

positiv definit ist:

$$\begin{aligned} -b_{11} &= 1 > 0, \\ \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} &= 1 > 0, \\ \det(-B) &= 1 \det \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} = 1 > 0. \end{aligned} \quad (10.70)$$

Übrigens:

$$\begin{aligned} b_{11} &= -1 < 0 \\ \det \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} &= 1 > 0 \\ \det B &= (-1)(2 - 1) = -1 < 0 \end{aligned} \quad (10.71)$$

alternierende Hauptunterdeterminanten.

11 Differentiation in mehreren Variablen

11.1 Punktmengen und Folgen im \mathbb{R}^n

Wiederholung/Notation/Begriffe

$$\vec{x} \in \mathbb{R}^n, \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, x_i \in \mathbb{R}, \text{ d.h. } \vec{x} = \sum_{i=1}^n x_i \vec{e}_i$$

$$\|\vec{x}\| = |\vec{x}| = \sqrt{\sum_{i=1}^n x_i^2} \quad (\text{Norm})$$

$$\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$$

$$\text{Dreiecksungleichung: } \|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$$

$$\text{Cauchy-Schwarzsche Ungleichung: } |\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \cdot \|\vec{y}\|, \text{ wobei}$$

$$\vec{x} \cdot \vec{y} = \vec{x}^T \vec{y} = \sum_{i=1}^n x_i y_i$$

$$U_\varepsilon(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^n, \|\vec{x} - \vec{x}_0\| < \varepsilon\} \quad (11.1)$$

heißt offene ε -Umgebung von \vec{x}_0 .

$$\overline{U}_\varepsilon(\vec{x}_0) := \{\vec{x} \in \mathbb{R}^n, \|\vec{x} - \vec{x}_0\| \leq \varepsilon\} \quad (11.2)$$

heißt (ab-)geschlossene ε -Umgebung von \vec{x}_0 .

Definition:

Eine Menge $M \subseteq \mathbb{R}^n$ heißt

- offen, falls zu jedem $\vec{x} \in M$ ein $\varepsilon(\vec{x}) > 0$ existiert mit $U_\varepsilon(\vec{x}) \subset M$.
- abgeschlossen, falls $M^C := \{\vec{x} \in \mathbb{R}^n, \vec{x} \notin M\}$ offen ist.
- beschränkt, falls es ein $K > 0$ gibt, mit: $\vec{x} \in M \Rightarrow \|\vec{x}\| \leq K$.
- kompakt, falls M sowohl abgeschlossen als auch beschränkt ist.

Bemerkungen:

1. Die leere Menge \emptyset ist sowohl offen als auch abgeschlossen – ebenso \mathbb{R}^n .
2. Offene Mengen beinhalten ihre Randpunkte nicht.
3. Sind $M_\alpha \subset \mathbb{R}^n$ offen, so auch ist $\bigcup_{\alpha} M_\alpha$ offen.

Gilt aber i.A. nicht für Durchschnitte, z.B. $I_n := (-\frac{1}{n}, \frac{1}{n})$, $n \in \mathbb{N}$, offen, aber $\bigcap_{n=1}^{\infty} I_n = \{0\}$ (da $0 \in I_n \forall n \in \mathbb{N}$) abgeschlossen.

4. $U_\varepsilon(\vec{x})$ ist offen und beschränkt.

5. $\overline{U_\varepsilon(\vec{x})}$ ist abgeschlossen und beschränkt, also kompakt.

Definition:

Sei (\vec{x}_k) eine Folge, d.h.

$$\vec{x}_k = \begin{pmatrix} x_1^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix}, \quad x_i^{(k)} \in \mathbb{R} \quad \text{und (z.B.) } k = 1, 2, \dots \quad (11.3)$$

(\vec{x}_k) heißt konvergent gegen $\vec{a} \in \mathbb{R}^n$ falls gilt:

$$\text{Für jedes } \varepsilon > 0 \exists N(\varepsilon) \text{ mit } |\vec{x}_k - \vec{a}| < \varepsilon \forall k \geq N(\varepsilon). \quad (11.4)$$

Wir schreiben dann $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{a}$.

Bemerkung: $\lim_{k \rightarrow \infty} \vec{x}_k = \vec{a} \Leftrightarrow \lim_{k \rightarrow \infty} x_i^{(k)} = a_i \forall i = 1, \dots, n$, wobei $\vec{a} = (a_1, \dots, a_n)^T$.

11.2 Stetige Funktionen mehrerer Variabler

Definition: (Stetigkeit)

Sei $M \subseteq \mathbb{R}^n$, $f : M \rightarrow \mathbb{R}$. f heißt stetig in $\vec{x}_0 \in M$, falls gilt:

Zu jedem $\varepsilon > 0 \exists \delta(\varepsilon) > 0$ mit $|f(\vec{x}) - f(\vec{x}_0)| < \varepsilon \forall \vec{x} \in U_\delta(\vec{x}_0) \cap M$

(d.h. $|\vec{x} - \vec{x}_0| < \delta$, $\vec{x} \in M$).

Bemerkungen:

1. Polynome sind überall stetig, z.B. $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, $\vec{x} \mapsto x_1^2 - 2x_2 + 7x_3^3$; ebenso exp, sin, cos, $|\cdot|$; für positive Argumente auch log und $\sqrt{\quad}$.
2. Rechenregeln wie in 1D: Summen, Produkte und Verkettungen (Kompositionen) stetiger Funktionen sind stetig.
3. Ist f in $\vec{x}_0 \in M$ stetig, so auch als Funktion einer Variablen, z.B. $f(x, y) = xy$, $g(x) := f(x, y_0)$ stetig (als Funktion von x) und $h(y) := f(x_0, y)$ stetig.
4. Vorsicht: Umkehrung gilt i.A. nicht, z.B.

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & , \quad (x, y) \neq (0, 0) \\ 0 & , \quad x = y = 0 \end{cases} \quad (11.5)$$

stetig als Funktion einer der beiden Variablen, aber nicht stetig in Null als Funktion zweier Variabler, denn $\begin{pmatrix} x \\ x \end{pmatrix} \xrightarrow{x \rightarrow 0} \vec{0}$, aber

$$f(x, x) = \frac{x^2}{x^2 + x^2} = \frac{1}{2} \neq 0 \quad (11.6)$$

Satz 27. Sei $M \subset \mathbb{R}^n$ kompakt und $f : M \rightarrow \mathbb{R}$ stetig, dann gilt: f nimmt auf M ein Minimum und ein Maximum an, d.h. $\exists \vec{a}, \vec{b} \in M$ mit $f(\vec{a}) \leq f(\vec{x}) \leq f(\vec{b}) \forall \vec{x} \in M$.

(ohne Beweis) anschaulich klar ... Skizzen

11.3 Partielle, Richtungs- und totale Ableitung

Beispiel: (partielle Differentiation)

$$f(x, y, z) = x^2y + xy \sin z \quad (11.7)$$

$$\frac{\partial f}{\partial x} = 2xy + y \sin z, \quad \frac{\partial f}{\partial y} = x^2 + x \sin z, \quad \frac{\partial f}{\partial z} = xy \cos z \quad (11.8)$$

Partielle Ableitung nach...: Behandle die anderen Variablen als Konstanten.

Definition: Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}$, $\vec{y} \in M$. Existiert

$$\begin{aligned} \frac{\partial f}{\partial x_j}(\vec{y}) &:= \lim_{h \rightarrow 0} \frac{f(\vec{y} + h\vec{e}_j) - f(\vec{y})}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(y_1, \dots, y_{j-1}, y_j + h, y_{j+1}, \dots, y_n) - f(y_1, \dots, y_n)}{h} \end{aligned} \quad (11.9)$$

so heißt f partiell nach x_j differenzierbar und $\frac{\partial f}{\partial x_j}(\vec{y})$ heißt partielle Ableitung an der Stelle \vec{y} .

Andere Notation: $f_{x_k} := \frac{\partial f}{\partial x_k}$

Definition: Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}$, $\vec{x}_0 \in M$, $\vec{v} \in \mathbb{R}^n$ mit $|\vec{v}| = 1$, so heißt (falls existent)

$$\frac{\partial f}{\partial \vec{v}}(\vec{x}_0) := \lim_{h \rightarrow 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0)}{h} \quad (11.10)$$

Richtungsableitung von f in \vec{x}_0 in Richtung von \vec{v} .

Bemerkungen:

1. $\frac{\partial f}{\partial \vec{e}_j} = \frac{\partial f}{\partial x_j}$
2. **Vorsicht:** Aus der Existenz aller partiellen Ableitungen im Punkt \vec{x}_0 folgt i.A. noch nicht die Stetigkeit!

Beispiel:

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4} & , \quad x^2 + y^2 > 0 \\ 0 & , \quad x = y = 0 \end{cases} \quad (11.11)$$

$$\lim_{t \rightarrow 0} f(t^2, t) = \lim_{t \rightarrow 0} \frac{t^4}{t^4 + t^4} = \frac{1}{2} \neq 0, \quad (11.12)$$

also nicht stetig in $\vec{0}$. Richtungsableitungen: $\vec{v} = \begin{pmatrix} c \\ s \end{pmatrix}$, $c^2 + s^2 = 1$.

$$\frac{\partial f}{\partial \vec{v}}(\vec{0}) = \lim_{h \rightarrow 0} \frac{f(h\vec{v})}{h} = \lim_{h \rightarrow 0} \frac{h^3 c s^2}{(h^2 c^2 + h^4 s^4)h} = \begin{cases} \frac{s^2}{c} & , \quad c \neq 0 \\ 0 & , \quad c = 0 \end{cases}, \quad (11.13)$$

d.h. alle Richtungsableitungen (und insbesondere alle partiellen Ableitungen) existieren in $\vec{0}$.

Im \mathbb{R}^1 : f diffbar in $x_0 \Leftrightarrow$ Approximierbarkeit des Graphs in der Nähe von $(x_0, f(x_0))$ durch eine Gerade (Tangente),

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|), \quad x \rightarrow x_0. \quad (11.14)$$

Dabei war $f(x) = o(g(x))$, $x \rightarrow x_0$, falls gilt

$$\lim_{x \rightarrow x_0} \frac{|f(x)|}{|g(x)|} = 0. \quad (11.15)$$

Definition: (Totale Differenzierbarkeit)

Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}$, $\vec{x}_0 \in M$. f heißt (total) differenzierbar in \vec{x}_0 , falls es einen (Zeilen-)Vektor \vec{a} gibt, mit

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \vec{a}\vec{h} + o(|\vec{h}|), \quad |\vec{h}| \rightarrow 0. \quad (11.16)$$

Dann heißt $f'(\vec{x}_0) := \vec{a}$ Ableitung von f in \vec{x}_0 .

Die lineare Abbildung $df : \mathbb{R}^n \rightarrow \mathbb{R}$, definiert durch $\vec{h} \mapsto f'(\vec{x}_0)\vec{h}$, heißt totales Differential.

Bemerkungen:

1. " $f : \mathbb{R}^n \rightarrow \mathbb{R}$ in \vec{x}_0 diffbar" meint "total diffbar".
2. anschaulich: $f : \mathbb{R}^2 \supseteq M \rightarrow \mathbb{R}$,
 Fläche über M : $\{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in M, z = f(x, y)\}$ (Graph von f)
 Tangentialebene im Punkt $(x_0, y_0, f(x_0, y_0))$:
 $\{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in M, z = f(x_0, y_0) + f'(x_0, y_0) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}\}$

Satz 28. Sei $M \subseteq \mathbb{R}^n$ offen, $\vec{x}_0 \in M$, $f : M \rightarrow \mathbb{R}$ diffbar in \vec{x}_0 , dann gilt

- (i) f ist stetig in \vec{x}_0 .
- (ii) f ist nach allen Variablen partiell diffbar und es gilt

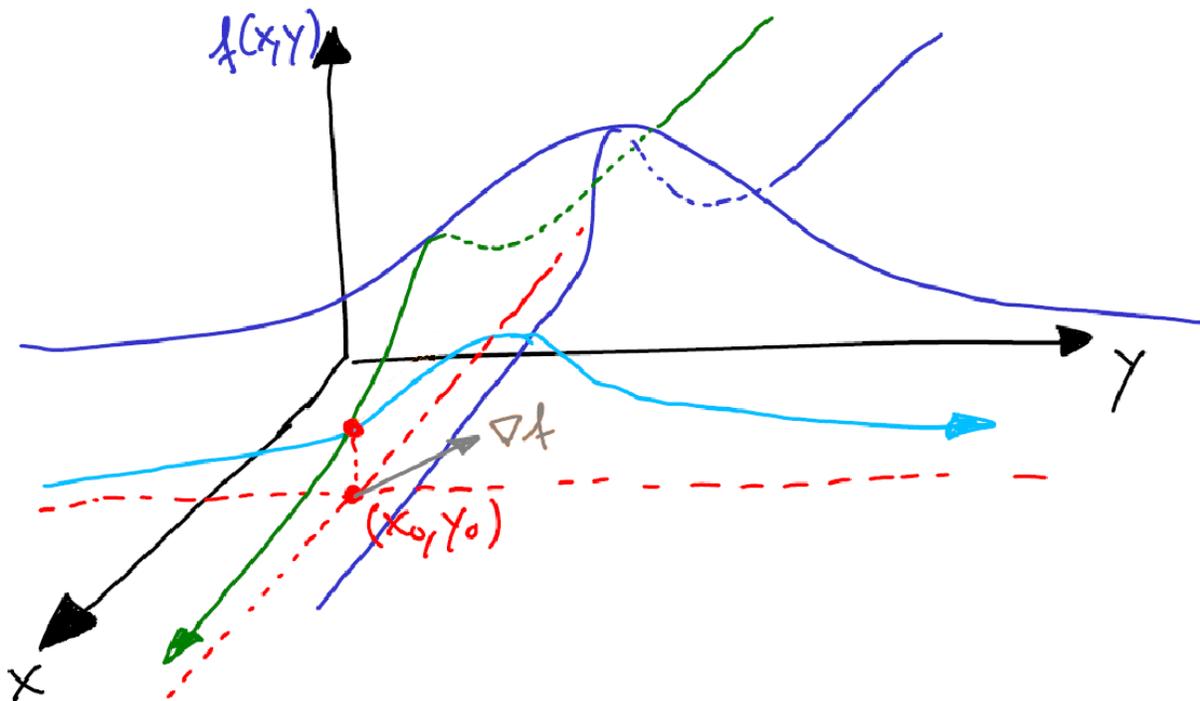
$$\begin{aligned} f'(\vec{x}_0) &= \left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \frac{\partial f}{\partial x_2}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0) \right) \\ &=: (\nabla f)(\vec{x}_0) && \text{(Nabla)} && (11.17) \\ &=: (\text{grad } f)(\vec{x}_0) && \text{(Gradient)}. \end{aligned}$$

(iii) Alle Richtungsableitungen existieren in \vec{x}_0 und es gilt ($|\vec{v}| = 1$)

$$\frac{\partial f}{\partial \vec{v}}(\vec{x}_0) = (\nabla f)(\vec{x}_0) \cdot \vec{v}. \quad (11.18)$$

Bemerkungen:

1. Nabla-Operator: $\nabla := \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right)$, bildet Skalarfeld f auf Vektorfeld ∇f ab
2. ∇f zeigt in die Richtung des steilsten Anstiegs der Funktion und steht senkrecht auf den Höhenlinien – folgt aus Gln. (11.16).



hier: $\frac{\partial f}{\partial x}(x_0, y_0) < 0$, $\frac{\partial f}{\partial y}(x_0, y_0) > 0$

Beweis:

(i)

$$\lim_{|\vec{h}| \rightarrow 0} f(\vec{x}_0 + \vec{h}) \stackrel{\text{diffbar}}{=} \lim_{|\vec{h}| \rightarrow 0} \left(f(\vec{x}_0) + f'(\vec{x}_0) \cdot \vec{h} + o(|\vec{h}|) \right) = f(\vec{x}_0) \quad (11.19)$$

(ii)

$$\begin{aligned} \frac{\partial f}{\partial x_j}(\vec{x}_0) &= \lim_{h \rightarrow 0} \frac{f(\vec{x}_0 + h\vec{e}_j) - f(\vec{x}_0)}{h} \\ &= \lim_{\text{diffbar } h \rightarrow 0} \frac{1}{h} (f'(\vec{x}_0) \cdot h\vec{e}_j + o(h)) \\ &= \lim_{h \rightarrow 0} (f'(\vec{x}_0) \cdot \vec{e}_j + o(1)) \\ &= (f'(\vec{x}_0))_j \end{aligned} \quad (11.20)$$

(iii)

$$\begin{aligned}\frac{\partial f}{\partial \vec{v}}(\vec{x}_0) &= \lim_{h \rightarrow 0} \frac{f(\vec{x}_0 + h\vec{v}) - f(\vec{x}_0)}{h} \\ &\stackrel{\text{diffbar}}{=} \lim_{h \rightarrow 0} \frac{1}{h} (f'(\vec{x}_0) \cdot h\vec{v} + o(h)) \\ &= f'(\vec{x}_0) \cdot \vec{v}\end{aligned}\tag{11.21}$$

□

Wann ist f total diffbar? Kriterium?

Satz 29. Sei $M \subseteq \mathbb{R}^n$ offen, $\vec{x}_0 \in M$, $f : M \rightarrow \mathbb{R}$. Weiter besitze f alle partiellen Ableitungen $\frac{\partial f}{\partial x_j}$ in einer Umgebung von \vec{x}_0 und diese seien stetig in \vec{x}_0 . Dann ist f diffbar in \vec{x}_0 .

Beweis:

$$\begin{aligned}f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0) &= f(\vec{x}_0 + \vec{h}) - f(\vec{x}_0 + \vec{h} - h_n \vec{e}_n) \\ &\quad + f(\vec{x}_0 + \vec{h} - h_n \vec{e}_n) - f(\vec{x}_0 + \vec{h} - h_n \vec{e}_n - h_{n-1} \vec{e}_{n-1}) \\ &\quad + \dots - \dots \\ &\quad + f(\vec{x}_0 + h_1 \vec{e}_1) - f(\vec{x}_0) \\ &\stackrel{\text{part. diffbar}}{=} \frac{\partial f}{\partial x_n}(\vec{x}_0 + \vec{h} - h_n \vec{e}_n) \cdot h_n + o(h_n) \\ &\quad + \frac{\partial f}{\partial x_{n-1}}(\vec{x}_0 + \vec{h} - h_n \vec{e}_n - h_{n-1} \vec{e}_{n-1}) \cdot h_{n-1} + o(h_{n-1}) \\ &\quad + \dots \\ &\quad + \frac{\partial f}{\partial x_1}(\vec{x}_0) \cdot h_1 + o(h_1) \\ &\stackrel{\text{part. Abl. stetig}}{=} \sum_{j=1}^n \left[\left(\frac{\partial f}{\partial x_j}(\vec{x}_0) + o(1) \right) h_j + o(h_j) \right] \\ &= \sum_{j=1}^n \left(\frac{\partial f}{\partial x_j}(\vec{x}_0) h_j + o(|\vec{h}|) \right) \\ &= (\nabla f)(\vec{x}_0) \cdot \vec{h} + o(|\vec{h}|)\end{aligned}\tag{11.22}$$

□

Im obigen **Beispiel**,

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4} & , \quad (x, y) \neq (0, 0) \\ 0 & , \quad x = y = 0 \end{cases}\tag{11.23}$$

existieren zwar die partiellen Ableitungen in $\vec{0}$ (s.o.), z.B. $\frac{\partial f}{\partial x}(0,0) = 0$, sind dort aber nicht stetig; z.B., für $(x,y) \neq (0,0)$,

$$\frac{\partial f}{\partial x}(x,y) = \frac{y^2(x^2 + y^4) - xy^2 2x}{(x^2 + y^4)^2} = \frac{y^6 - x^2 y^2}{x^4 + 2x^2 y^4 + y^8}, \quad (11.24)$$

$$\lim_{t \rightarrow 0} \frac{\partial f}{\partial x}(t,t) = -1 \neq 0 = \lim_{x \rightarrow 0} \left(\lim_{y \rightarrow 0} \frac{\partial f}{\partial x}(x,y) \right) \quad (11.25)$$

Also ist $\frac{\partial f}{\partial x}$ nicht stetig in $\vec{0}$.

Bemerkung: Hilfreich (und leicht zu zeigen) sind dann die folgenden Aussagen:

1. Polynome, exp, sin, cos sind überall diffbar;
für positive Argumente auch log und $\sqrt{\quad}$.
2. Summen, Produkte und Verkettungen diffbarer Funktionen sind diffbar.

Satz 30. (Kettenregel, Differentiation nach Kurve)

Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}$ diffbar in $\vec{x}_0 \in M$; weiter sei $\vec{x} : [a,b] \rightarrow M$ diffbar in $t_0 \in (a,b)$ mit $\vec{x}(t_0) = \vec{x}_0$. Dann ist $F := f \circ \vec{x} : [a,b] \rightarrow \mathbb{R}$ diffbar in t_0 und es gilt

$$F'(t_0) = (\nabla f)(\vec{x}_0) \cdot \dot{\vec{x}}(t_0) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\vec{x}_0) \dot{x}_j(t_0) \quad (11.26)$$

Beweis:

$$\begin{aligned} F(t_0 + h) - F(t_0) &= f(\vec{x}(t_0 + h)) - f(\vec{x}(t_0)) \\ &= f\left(\vec{x}_0 + \dot{\vec{x}}(t_0)h + o(h)\right) - f(\vec{x}_0) \\ &= f(\vec{x}_0) + (\nabla f)(\vec{x}_0) \left(\dot{\vec{x}}(t_0)h + o(h)\right) + o(h) - f(\vec{x}_0) \\ &= (\nabla f)(\vec{x}_0) \cdot \dot{\vec{x}}(t_0) h + o(h) \end{aligned} \quad (11.27)$$

□

Definition: (Kurvenintegral)

Sei $M \subseteq \mathbb{R}^n$ offen, $\vec{f} : M \rightarrow \mathbb{R}^n$ stetig (komponentenweise) und $\mathfrak{K} : [a,b] \ni t \mapsto \vec{x}(t) \in M$ diffbar auf (a,b) , dann heißt

$$\int_{\mathfrak{K}} \vec{f}(\vec{x}) \, d\vec{x} := \int_a^b \vec{f}(\vec{x}(t)) \dot{\vec{x}}(t) \, dt \quad (11.28)$$

Kurvenintegral längs der Kurve \mathfrak{K} .

Bemerkungen:

1. Ist \vec{f} Kraftfeld, so ist $\int_{\mathfrak{K}} \vec{f} d\vec{x}$ die längs \mathfrak{K} geleistete Arbeit.
2. Gibt es ein $F : M \rightarrow \mathbb{R}$, mit $\nabla F = \vec{f}$, dann gilt

$$\begin{aligned} \int_{\mathfrak{K}} \vec{f}(\vec{x}) d\vec{x} &= \int_a^b (\nabla F)(\vec{x}(t)) \cdot \dot{x}(t) dt \stackrel{\text{Satz 30}}{=} \int_a^b \left(\frac{d}{dt} F(\vec{x}(t)) \right) dt \\ &= F(\vec{x}(b)) - F(\vec{x}(a)), \end{aligned} \quad (11.29)$$

hängt also nur von Anfangs- und Endpunkt ab, nicht vom Weg. Solch ein \vec{f} heißt konservativ. Später: Kriterium für Existenz von F .

Bei Kraftfeldern (Physik) heißt $V = -F$ Potential.

3. $f : \mathbb{R}^n \supseteq M \rightarrow \mathbb{R}$ heißt Skalarfeld
 $\vec{f} : \mathbb{R}^n \supseteq M \rightarrow \mathbb{R}^n$ heißt Vektorfeld

Beispiel:

$$\vec{f}(x, y, z) = \begin{pmatrix} y \\ x \\ z^2 \end{pmatrix}, \quad \mathfrak{K} : \vec{x}(t) = \begin{pmatrix} \cos t \\ \sin t \\ t \end{pmatrix}, \quad 0 \leq t \leq 2\pi \quad (11.30)$$

$$\begin{aligned} \int_{\mathfrak{K}} \vec{f} d\vec{x} &= \int_0^{2\pi} \begin{pmatrix} \sin t \\ \cos t \\ t^2 \end{pmatrix} \cdot \begin{pmatrix} -\sin t \\ \cos t \\ 1 \end{pmatrix} dt = \int_0^{2\pi} (-\sin^2 t + \cos^2 t + t^2) dt \\ &= \int_0^{2\pi} (\cos(2t) + t^2) dt = \left[\frac{\sin(2t)}{2} + \frac{t^3}{3} \right]_0^{2\pi} = \frac{8\pi^3}{3} \end{aligned} \quad (11.31)$$

$$\begin{aligned} F &= \int y dx = xy + h_1(y, z) \\ &= \int x dy = xy + h_2(x, z) \\ &= \int z^2 dz = \frac{z^3}{3} + h_3(x, y) \\ \Rightarrow F &= xy + \frac{z^3}{3} \quad \text{tut's} \end{aligned} \quad (11.32)$$

$$\int_{\mathfrak{K}} \vec{f} d\vec{x} = F(\vec{x}(2\pi)) - F(\vec{x}(0)) = F(1, 0, 2\pi) - F(1, 0, 0) = \frac{8\pi^3}{3} \quad (11.33)$$

11.4 Höhere Ableitungen und Satz von Taylor

Definition Sei $M \subseteq \mathbb{R}^n$ offen, $\vec{x}_0 \in M$, $f : M \rightarrow \mathbb{R}$.

$$(f_{x_j})_{x_k}(\vec{x}_0) =: f_{x_j x_k}(\vec{x}_0) = \frac{\partial^2 f}{\partial x_k \partial x_j}(\vec{x}_0) =: \frac{\partial}{\partial x_k} \frac{\partial f}{\partial x_j}(\vec{x}_0) \quad (11.34)$$

heißt partielle Ableitung 2. Ordnung von f im Punkt \vec{x}_0 , falls f_{x_j} in $U_\varepsilon(\vec{x}_0)$ existiert und partiell nach x_k diffbar ist; analog für höhere Ableitungen.

Bemerkungen:

1. Beispiel im \mathbb{R}^3 : $f_{xxzy} = \frac{\partial}{\partial y} f_{xxz} = \frac{\partial}{\partial y} \frac{\partial}{\partial z} f_{xx} = \frac{\partial}{\partial y} \frac{\partial}{\partial z} \frac{\partial^2}{\partial x^2} f$
2. Es kommt auf die Reihenfolge an, d.h. $f_{xxzy} \underset{\text{i.A.}}{\neq} f_{zyyx}$.
3. Beispiel: $f(x, y) = e^{xy} + x \sin y$

$$\begin{aligned} f_x &= ye^{xy} + \sin y \\ f_{xx} &= y^2 e^{xy}, \quad f_{xxxx} = y^5 e^{xy} \\ f_{xy} &= (1 + xy)e^{xy} + \cos y \end{aligned} \tag{11.35}$$

Bezeichnung: f heißt k mal stetig (partiell) diffbar auf M , falls alle nur möglichen partiellen Ableitungen der Ordnung k auf M existieren und stetig sind.

Symbol: $f \in C^k(M)$, insbesondere

$f \in C^0(M)$: f stetig

$f \in C^1(M)$: alle part. Abl. existieren und sind stetig ($\Rightarrow f$ total diffbar auf M)

Satz 31. (Lemma von Schwarz)

Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}$, $f \in C^1(M)$; weiter existiere $f_{x_j x_k}$ auf M und sei stetig, dann folgt: $f_{x_k x_j}$ existiert ebenfalls, und es gilt $f_{x_k x_j} = f_{x_j x_k}$.

(ohne Beweis)

Beispiel: $f(x, y, z) = y^2 \sin((x^2 + z^4)e^{xz}) + 2ye^{x+z}$

Berechne: $\frac{\partial^{14} f}{\partial^5 x \partial^5 y \partial^4 z} = 0$, weil $f_{yyy} = 0$.

Bemerkung:

Falls ein Vektorfeld \vec{f} , komponentenweise C^1 , eine Stammfunktion¹⁵ hat, d.h. $\vec{f} = \nabla F$, dann gilt

$$\frac{\partial f_j}{\partial x_k} = \frac{\partial f_k}{\partial x_j}, \tag{11.36}$$

da $F \in C^2$ und deshalb $F_{x_j x_k} = F_{x_k x_j}$.

Umkehrung gilt auch (ohne Beweis), daher: Kriterium für konservatives Feld.

Bezeichnungen: (für Satz von Taylor)

- $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^k(M)$

- $\vec{h} = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbb{R}^n$

¹⁵bzw. ein Potential $\vec{f} = -\nabla V$, d.h. $V = -F$

- $(\nabla \vec{h})^k f(\vec{x}_0) = \sum_{j_1, \dots, j_k=1}^n f_{x_{j_1} \dots x_{j_k}}(\vec{x}_0) h_{j_1} \dots h_{j_k}$

- $n = 1$: $(\nabla \vec{h})^k f(x_0) = f^{(k)}(x_0) h^k$

- n beliebig, $k = 0$: $(\nabla \vec{h})^0 f(\vec{x}_0) = f(\vec{x}_0)$

$$k = 1: (\nabla \vec{h})^1 f(\vec{x}_0) = \sum_{j=1}^n f_{x_j}(\vec{x}_0) h_j = (\nabla f)(\vec{x}_0) \cdot \vec{h}$$

$$k = 2: (\nabla \vec{h})^2 f(\vec{x}_0) = \sum_{i,j=1}^n f_{x_i x_j}(\vec{x}_0) h_i h_j$$

$$\text{mit } f'' := (f_{x_i x_j}) = \begin{pmatrix} f_{x_1 x_1} & \dots & f_{x_1 x_n} \\ \vdots & & \vdots \\ f_{x_n x_1} & \dots & f_{x_n x_n} \end{pmatrix},$$

d.h. $f''(\vec{x}_0) \in \mathbb{R}^{n \times n}$ (Hesse-Matrix), gilt

$$(\nabla \vec{h})^2 f(\vec{x}_0) = \vec{h}^T f''(\vec{x}_0) \vec{h}.$$

Ist $f \in C^2(M) \Rightarrow f''$ ist symmetrisch, und damit

$$(\nabla \vec{h})^2 f(\vec{x}_0) = q_{f''(\vec{x}_0)}(\vec{h}).$$

Satz 32. (Satz von Taylor)

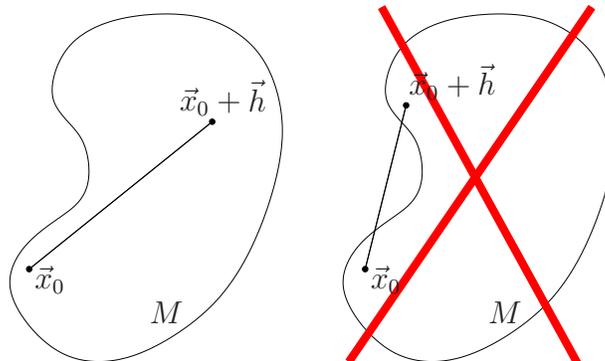
Sei $M \subseteq \mathbb{R}^n$ offen, $\vec{h} \in \mathbb{R}^n$, $f \in C^{k+1}(M)$, $\vec{x}_0 + t\vec{h} \in M \forall t \in [0, 1]$, dann gilt

$$f(\vec{x}_0 + \vec{h}) = \sum_{\kappa=0}^k \frac{((\nabla \vec{h})^\kappa f)(\vec{x}_0)}{\kappa!} + R_k \tag{11.37}$$

wobei

$$R_k = \frac{((\nabla \vec{h})^{k+1} f)(\vec{x}_0 + \vartheta \vec{h})}{(k+1)!} \text{ f\"ur ein } \vartheta \in (0, 1). \tag{11.38}$$

Bemerkung:



Beweis: $\phi(t) := f(\vec{x}_0 + t\vec{h})$, $t \in \mathbb{R}$ Taylorentwicklung (1D) von ϕ (Satz 8):

$$\phi(1) = \sum_{\kappa=0}^k \frac{\phi^{(\kappa)}(0)}{\kappa!} + \frac{\phi^{(k+1)}(\vartheta)}{(k+1)!} \quad \vartheta \in (0, 1). \quad (11.39)$$

Bleibt z.z.: $\phi^{(\kappa)}(t) = \left((\nabla \vec{h})^\kappa f \right) (\vec{x}_0 + t\vec{h})$

Induktion nach κ :

$\kappa = 0$: klar

$\kappa = 1$: Differentiation nach Kurve:

$$(f(\vec{x}(t)))' = (\nabla f)(\vec{x}(t)) \cdot \dot{\vec{x}}(t), \quad \text{hier } \dot{\vec{x}}(t) = \vec{h}, \quad (11.40)$$

also

$$\phi'(t) = (\nabla f)(\vec{x}_0 + t\vec{h}) \cdot \vec{h}. \quad (11.41)$$

$\kappa \rightarrow \kappa + 1$:

$$\begin{aligned} \phi^{(\kappa+1)}(t) &= \frac{d}{dt} (\phi^{(\kappa)}(t)) \\ &\stackrel{\text{i.V.}}{=} \frac{d}{dt} \left((\nabla \vec{h})^\kappa f \right) (\vec{x}_0 + t\vec{h}) \\ &= \frac{d}{dt} \sum_{j_1, \dots, j_\kappa=1}^n f_{x_{j_1} \dots x_{j_\kappa}} (\vec{x}_0 + t\vec{h}) h_{j_1} \cdots h_{j_\kappa} \\ &= \sum_{j_1, \dots, j_\kappa=1}^n h_{j_1} \cdots h_{j_\kappa} \sum_{j_{\kappa+1}=1}^n f_{x_{j_1} \dots x_{j_{\kappa+1}}} (\vec{x}_0 + t\vec{h}) h_{j_{\kappa+1}} \end{aligned} \quad (11.42)$$

□

Spezialfälle:

1. $n = 1$ bekannt

$$2. f \in C^1(M): \quad f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \underbrace{(\nabla f)(\vec{x}_0 + \vartheta \vec{h}) \cdot \vec{h}}_{=o(1), |\vec{h}| \rightarrow 0}$$

$$3. f \in C^2(M): \quad f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + (\nabla f)(\vec{x}_0) \cdot \vec{h} + \underbrace{\frac{1}{2} \vec{h}^T f''(\vec{x}_0 + \vartheta \vec{h}) \vec{h}}_{=o(|\vec{h}|), |\vec{h}| \rightarrow 0}$$

$$4. f \in C^3(M): \quad f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + (\nabla f)(\vec{x}_0) \cdot \vec{h} + \frac{1}{2} \vec{h}^T f''(\vec{x}_0) \vec{h} + \underbrace{R_2}_{=o(|\vec{h}|^2), |\vec{h}| \rightarrow 0}$$

Beispiele:

$$1. f(x, y) = e^{xy} + x \sin(\pi y)$$

$$(\nabla f)(x, y) = (ye^{xy} + \sin(\pi y), xe^{xy} + \pi x \cos(\pi y)) \quad (11.43)$$

$$f''(x, y) = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} y^2 e^{xy} & (1+xy)e^{xy} + \pi \cos(\pi y) \\ (1+xy)e^{xy} + \pi \cos(\pi y) & x^2 e^{xy} - \pi^2 x \sin(\pi y) \end{pmatrix} \quad (11.44)$$

Entwickle um $\vec{x}_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$$\begin{aligned} f(x, y) &= e^2 + (2e^2 - e^2 + \pi) \begin{pmatrix} x-1 \\ y-2 \end{pmatrix} \\ &+ \frac{1}{2} \begin{pmatrix} x-1 & y-2 \end{pmatrix} \begin{pmatrix} 4e^2 & 3e^2 + \pi \\ 3e^2 + \pi & e^2 \end{pmatrix} \begin{pmatrix} x-1 \\ y-2 \end{pmatrix} + o(|x-1|^2 + |y-2|^2) \end{aligned} \quad (11.45)$$

2. Einfacher oft mit bekannten Reihen, z.B. $f(x, y) = e^{x+y^2}$ um $\vec{0}$:

$$\begin{aligned} f(x, y) &= \sum_{n=0}^{\infty} \frac{(x+y^2)^n}{n!} \\ &= 1 + (x+y^2) + \frac{1}{2}(x+y^2)^2 + o((x+y^2)^2) \\ &= 1 + x + \left(\frac{1}{2}x^2 + y^2\right) + o(x^2 + y^2) \end{aligned} \quad (11.46)$$

Daraus lesen wir ab:

$$(\nabla f)(0, 0) = (1, 0) \quad \text{und} \quad f''(0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad (11.47)$$

11.5 Lokale Minima und Maxima

Definition: Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}$. $\vec{x}_0 \in M$ heißt lokale Maximalstelle von f , falls ein $\varepsilon > 0$ existiert, mit

$$f(\vec{x}) < f(\vec{x}_0) \quad \forall \vec{x} \in U_\varepsilon(\vec{x}_0) \setminus \{\vec{x}_0\} \quad (11.48)$$

(analog: lokale Minimalstelle... $f(\vec{x}) > f(\vec{x}_0)$...)

Satz 33. (Lokale Extrema)

Sei $M \subseteq \mathbb{R}^n$ offen, $f \in C^1(M)$.

(i) notwendige Bedingung:

$$\text{Ist } \vec{x}_0 \in M \text{ lokale Extremstelle von } f \quad \Rightarrow \quad f'(\vec{x}_0) = \vec{0}^T.$$

$$\left(\text{Ist sogar } f \in C^2(M) \quad \Rightarrow \quad f''(\vec{x}_0) \text{ ist } \begin{array}{l} \text{positiv semidefinit (Minimum).} \\ \text{negativ semidefinit (Maximum).} \end{array} \right)$$

(ii) *hinreichende Bedingung:*

Ist $f \in C^2(M)$ und gilt $f'(\vec{x}_0) = \vec{0}^T$, $f''(\vec{x}_0)$ positiv definit (negativ definit)
 $\Rightarrow \vec{x}_0$ ist lokale Minimalstelle (Maximalstelle).

Bemerkung: Ist $(\nabla f)(\vec{x}_0) = \vec{0}$ so heißt \vec{x}_0 *kritischer Punkt* von f .

Beweis:

Idee: Taylor

$$f(\vec{x}) = f(\vec{x}_0) + \underbrace{f'(\vec{x}_0)}_{=\vec{0}^T} \cdot (\vec{x} - \vec{x}_0) + \underbrace{\frac{1}{2}(\vec{x} - \vec{x}_0)^T f''(\vec{x}_0)(\vec{x} - \vec{x}_0)}_{\text{pos./neg. definit}} + \dots \quad (11.49)$$

(i) Sei \vec{x}_0 lokale Extremstelle von f

\Rightarrow die eindimensionale Funktion $\phi(t) := f(\vec{x}_0 + t\vec{e}_j)$ hat lokales Extremum bei $t = 0$

$\Rightarrow 0 = \frac{d\phi}{dt}(0) = \frac{\partial f}{\partial x_j}(\vec{x}_0)$ (für jedes $j = 1, \dots, n$)

(ii) Taylor:

$$\begin{aligned} f(\vec{x}) &= f(\vec{x}_0) + f'(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \frac{1}{2}(\vec{x} - \vec{x}_0)^T f''(\vec{x}_0 + \vartheta(\vec{x} - \vec{x}_0))(\vec{x} - \vec{x}_0) \\ &= f(\vec{x}_0) + \frac{1}{2}(\vec{x} - \vec{x}_0)^T f''(\vec{x}_0 + \vartheta(\vec{x} - \vec{x}_0))(\vec{x} - \vec{x}_0) \end{aligned} \quad (11.50)$$

(z.B. für Minimum:) $f''(\vec{x}_0)$ pos. definit, außerdem f'' stetig, da $f \in C^2(M)$, d.h.

$\exists U_\varepsilon(\vec{x}_0)$ mit $f''(\vec{x}_0 + \vartheta(\vec{x} - \vec{x}_0))$ pos. definit $\forall \vec{x} \in U_\varepsilon(\vec{x}_0)$

$\Rightarrow f(\vec{x}) > f(\vec{x}_0) \forall \vec{x} \in U_\varepsilon(\vec{x}_0) \setminus \{\vec{x}_0\}$

□

Beispiel: $f(x, y) = 2x^2 + 4y^2 + 4xy - 4x + 4y + 10$

$$(\nabla f)(x, y) = (4x + 4y - 4, 8y + 4x + 4) \stackrel{!}{=} (0, 0) \quad (11.51)$$

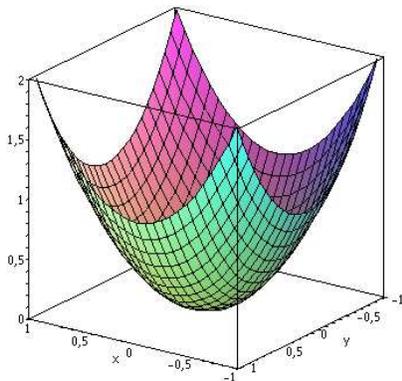
$$\begin{aligned} \Rightarrow 4x + 4y - 4 = 0 &\Rightarrow x + y = 1 \quad \text{in} \quad 2y + x + 1 = 0 \\ \text{also } 2y + 1 - y + 1 = 0 &\Leftrightarrow y = -2 \quad \text{und damit } x = 3 \end{aligned} \quad (11.52)$$

$\begin{pmatrix} 3 \\ -2 \end{pmatrix}$ ist einzige potentielle lokale Extremstelle.

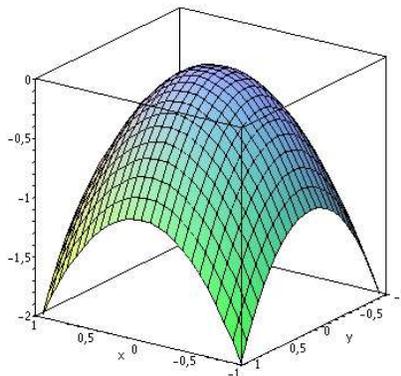
$$f''(x, y) = \begin{pmatrix} 4 & 4 \\ 4 & 8 \end{pmatrix} = f''(3, -2) \quad (11.53)$$

Hauptunterdeterminanten: $4 > 0$, $32 - 16 = 16 > 0$, also ist f'' pos. definit und f hat ein lokales Minimum bei $\begin{pmatrix} 3 \\ -2 \end{pmatrix}$ (sogar global).

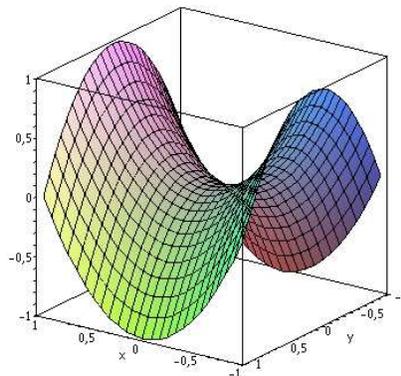
Skizzen für typische Fälle, $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f'(0,0) = 0, \dots$



$f''(0,0)$ positiv definit
lokales Minimum



$f''(0,0)$ negativ definit
lokales Maximum



$f''(0,0)$ indefinit
Sattel

MAPLE-Code:

```
> plot3d(x^2+y^2,x=-1..1,y=-1..1,view=0..2,axes=boxed,orientation=[125,70]);
> plot3d(-x^2-y^2,x=-1..1,y=-1..1,view=-2..0,axes=boxed,orientation=[125,70]);
> plot3d(x^2-y^2,x=-1..1,y=-1..1,view=-1..1,axes=boxed,orientation=[125,70]);
```

11.6 Vektorwertige Funktionen

Betrachte

$$\vec{f} : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\vec{x} \mapsto \vec{f}(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{pmatrix} \quad (11.54)$$

d.h. m Funktionen $f_j : \mathbb{R}^n \rightarrow \mathbb{R}$.

Stetigkeit, Diffbarkeit etc. komponentenweise.

Also: Sei $M \subseteq \mathbb{R}^n$ offen, $\vec{x}_0 \in M$. $\vec{f} : M \rightarrow \mathbb{R}^m$ heißt (total) diffbar in \vec{x}_0 , falls es eine Matrix $A \in \mathbb{R}^{m \times n}$ gibt, mit

$$\vec{f}(\vec{x}) = \vec{f}(\vec{x}_0) + A(\vec{x} - \vec{x}_0) + o(|\vec{x} - \vec{x}_0|), \quad \vec{x} \rightarrow \vec{x}_0. \quad (11.55)$$

Es gilt dann $A = (a_{ij})$, $a_{ij} = \frac{\partial f_i}{\partial x_j}(\vec{x}_0)$, $i = 1, \dots, m$, $j = 1, \dots, n$.

$A = \left(\frac{\partial f_i}{\partial x_j}(\vec{x}_0) \right) =: \vec{f}'(\vec{x}_0) =: \frac{\partial \vec{f}}{\partial \vec{x}}(\vec{x}_0)$ heißt Fundamental- oder Jacobi-Matrix.

Spezialfälle:

1. $m = 1$, n beliebig: Skalarfeld, $f' = \nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)$

2. $n = 1$, m beliebig: Kurve, $f' = \begin{pmatrix} f'_1 \\ \vdots \\ f'_m \end{pmatrix}$

3. $n = m \geq 2$: Vektorfeld, $\vec{f}' \in \mathbb{R}^{n \times n}$, also quadratisch

Guter Zeitpunkt, um mal ein paar Vektorpfeile wegzulassen. . .

Satz 34. (Kettenregel)

Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}^m$, diffbar in $\vec{y}_0 \in M$.

Weiter sei $K \subseteq \mathbb{R}^k$ offen, $g : K \rightarrow M$, diffbar in $\vec{x}_0 \in K$ mit $\vec{y}_0 = g(\vec{x}_0)$.

Dann ist die Funktion $h : \mathbb{R}^k \supseteq K \rightarrow \mathbb{R}^m$, definiert durch $h = f \circ g$, d.h. $h(\vec{x}) = f(g(\vec{x}))$, diffbar in \vec{x}_0 mit

$$h'(\vec{x}_0) = f'(g(\vec{x}_0)) g'(\vec{x}_0) \tag{11.56}$$

Bemerkung: Matrizen. . .

- $h'(\vec{x}_0) \in \mathbb{R}^{m \times k}$
- $f'(\vec{y}_0) \in \mathbb{R}^{m \times n}$
- $g'(\vec{x}_0) \in \mathbb{R}^{n \times k}$

Dimensionen passen also!

Beweis: ($\vec{x} \rightarrow \vec{x}_0$)

$$\begin{aligned} h(\vec{x}) - h(\vec{x}_0) &= f(g(\vec{x})) - f(g(\vec{x}_0)) \\ &\stackrel{g \text{ diffbar}}{=} f(g(\vec{x}_0) + g'(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + o(|\vec{x} - \vec{x}_0|)) - f(g(\vec{x}_0)) \\ &\stackrel{f \text{ diffbar}}{=} \underbrace{f(g(\vec{x}_0)) + f'(g(\vec{x}_0)) \cdot (g'(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + o(|\vec{x} - \vec{x}_0|))}_{=o(|\vec{x}-\vec{x}_0|)} \\ &\quad + \underbrace{o(|g'(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)| + o(|\vec{x} - \vec{x}_0|))}_{=o(|\vec{x}-\vec{x}_0|)} - \underbrace{f(g(\vec{x}_0))}_{=o(|\vec{x}-\vec{x}_0|)} \\ &= f'(g(\vec{x}_0)) \cdot g'(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \underbrace{f'(g(\vec{x}_0)) o(|\vec{x} - \vec{x}_0|) + o(|\vec{x} - \vec{x}_0|)}_{=o(|\vec{x}-\vec{x}_0|)} \end{aligned} \tag{11.57}$$

□

Beispiele:

1. $n = m = k = 1$: klar eindimensionale Kettenregel
2. $g : \mathbb{R}^2 \rightarrow \mathbb{R}^4$, $f : \mathbb{R}^4 \rightarrow \mathbb{R}^3$

$$f(x_1, x_2, x_3, x_4) = \begin{pmatrix} x_1^2 - x_2 \\ x_2 \\ x_3 x_4 \end{pmatrix}, \quad g(x, y) = \begin{pmatrix} \sin x \\ \cos x \\ e^y \\ xy \end{pmatrix} \tag{11.58}$$

$$h := f \circ g, \quad h'(\pi, 0) = ?$$

$$f'(x_1, x_2, x_3, x_4) = \begin{pmatrix} 2x_1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & x_4 & x_3 \end{pmatrix}, \quad g'(x, y) = \begin{pmatrix} \cos x & 0 \\ -\sin x & 0 \\ 0 & e^y \\ y & x \end{pmatrix} \quad (11.59)$$

$$\begin{aligned} h'(\pi, 0) &= f'(g(\pi, 0)) \cdot g'(\pi, 0) \\ &= f'(0, -1, 1, 0) \cdot g'(\pi, 0) \\ &= \begin{pmatrix} 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & \pi \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \pi \end{pmatrix} \end{aligned} \quad (11.60)$$

Hier auch explizit möglich:

$$h(x, y) = \begin{pmatrix} \sin^2 x - \cos x \\ \cos x \\ xy e^y \end{pmatrix} \quad (11.61)$$

$$h'(x, y) = \begin{pmatrix} 2 \sin x \cos x + \sin x & 0 \\ -\sin x & 0 \\ y e^y & x(1+y)e^y \end{pmatrix} \quad (11.62)$$

$$h'(\pi, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & \pi \end{pmatrix} \quad (11.63)$$

11.7 Implizit definierte Funktionen

Idee: $F(\vec{x}, \vec{y}) = 0$ wobei $\vec{x} \in \mathbb{R}^n$, $\vec{y} \in \mathbb{R}^m$ und $F(\vec{x}, \vec{y}) \in \mathbb{R}^m$. Löse nach \vec{y} auf.

Eindeutig möglich? Was können wir über $\vec{y} = f(\vec{x})$ lernen, ohne $f(\vec{x})$ explizit zu bestimmen? (da oft schwierig bis unmöglich)

Beispiele:

1. ($n = m = 1$)

$$\sin^2 x - y^3 + e^x = 0 \quad (11.64)$$

global nach y auflösbar: $y = f(x) = (e^x + \sin^2 x)^{1/3}$

Dann erhalten wir natürlich auch leicht $f'(x) = \frac{1}{3}(e^x + \sin^2 x)^{2/3}(e^x + 2 \sin x \cos x)$.

2. ($n = m = 1$)

$$\sin^2 x - y^2 + e^x = 0 \quad (11.65)$$

Löse in Umgebung von $(x_0, y_0) = (0, -1)$, also lokal, nach y auf:

$$y = f(x) = -\sqrt{e^x + \sin^2 x} \quad (11.66)$$

Dann erhalten wir direkt

$$f'(x) = -\frac{e^x + 2 \sin x \cos x}{2\sqrt{e^x + \sin^2 x}}, \quad (11.67)$$

insbesondere an der Stelle $x_0 = 0$: $f'(0) = -\frac{1}{2}$.

Geht aber auch ohne explizites Auflösen!

Leite $F(x, f(x)) = 0$ nach x ab:

$$2 \sin x \cos x - 2f(x)f'(x) + e^x = 0 \quad \Rightarrow \quad f'(0) = \frac{1}{2f(0)} = -\frac{1}{2} \quad (11.68)$$

3. ($n = m = 1$)

$$\sin^2 x - y \cos(\pi y) + e^x = 0 \quad (11.69)$$

Löse in Umgebung von $(x_0, y_0) = (0, -1)$ nach y auf – sicher nicht explizit, aber vielleicht $f'(0) = ?$ Setze $y = f(x)$ und leite nach x ab:

$$2 \sin x \cos x - f'(x) \cos(\pi f(x)) + f(x) \sin(\pi f(x)) \pi f'(x) + e^x = 0 \quad (11.70)$$

An der Stelle $(0, -1)$:

$$-f'(0) \cos(-\pi) - \sin(-\pi) \pi f'(0) + 1 = 0 \quad \Rightarrow \quad f'(0) = -1 \quad (11.71)$$

Falls es so ein $f(x)$ gibt...

4. ($n = 1, m = 2$)

$$\begin{aligned} x^3 + xe^y + \sin z &= 0 \\ z^2 + y \cos x &= 0 \end{aligned} \quad (11.72)$$

In Umgebung von $(0, 0, 0)$ nach y und z auflösbar, dh. definiert das dort Funktionen $y(x)$ und $z(x)$? Vielleicht, aber wohl nicht explizit. Falls ja: Einsetzen und ableiten (nach x):

$$\begin{aligned} 3x^2 + e^{y(x)} + xy'(x)e^{y(x)} + z'(x) \cos(z(x)) &= 0 \\ 2z(x)z'(x) + y'(x) \cos(x) - y(x) \sin(x) &= 0 \end{aligned} \quad (11.73)$$

An der Stelle $x = 0$: $y(0) = z(0) = 0$, d.h.

$$\left. \begin{aligned} 0 + 1 + 0 + z'(0) &= 0 \\ 0 + y'(0) - 0 &= 0 \end{aligned} \right\} \Rightarrow z'(0) = -1, y'(0) = 0 \quad (11.74)$$

Falls solche Funktionen existieren!

Satz 35. (Implizite Funktionen)

Sei $M \subseteq \mathbb{R}^{n+m}$ offen, $F : M \rightarrow \mathbb{R}^m$.

Seien weiter $(\vec{x}_0, \vec{y}_0) \in M$ mit $\vec{x}_0 \in \mathbb{R}^n$ und $\vec{y}_0 \in \mathbb{R}^m$ gegeben mit $F(\vec{x}_0, \vec{y}_0) = 0$.

Ferner sei $F \in C^1(M)$ und $\det \frac{\partial F}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \neq 0$.

Dann existiert ein $\varepsilon > 0$, so dass es genau eine Funktion $f : \mathbb{R}^n \supset U_\varepsilon(\vec{x}_0) \rightarrow \mathbb{R}^m$ gibt mit $f(\vec{x}_0) = \vec{y}_0$ und

$$F(\vec{x}, f(\vec{x})) = 0 \quad \forall \vec{x} \in U_\varepsilon(\vec{x}_0). \quad (11.75)$$

Weiter gilt $f \in C^1(U_\varepsilon(\vec{x}_0))$ mit

$$f'(\vec{x}) = - \left[\frac{\partial F}{\partial \vec{y}}(\vec{x}, f(\vec{x})) \right]^{-1} \frac{\partial F}{\partial \vec{x}}(\vec{x}, f(\vec{x})). \quad (11.76)$$

Kurz: $F(\vec{x}, \vec{y}) = 0$ ist (lokal) nach \vec{y} auflösbar bzw.

definiert (lokal) eine eindeutige Funktion $\vec{y} = f(\vec{x})$.

Bemerkungen:

- In Beispiel 3 oben war $F(x, y) = \sin^2 x - y \cos(\pi y) + e^x$ und damit

$$\begin{aligned} \frac{\partial F}{\partial y}(x, y) &= -\cos(\pi y) + y \sin(\pi y) \pi \\ \Rightarrow \frac{\partial F}{\partial y}(0, -1) &= -\sin(-\pi) - \sin(-\pi) \pi = 1 \neq 0, \end{aligned} \quad (11.77)$$

d.h. $y = f(x)$ existierte tatsächlich (lokal).

- In Beispiel 4 oben war $F(x, y, z) = \begin{pmatrix} x^3 + xe^y + \sin z \\ z^2 + y \cos x \end{pmatrix}$, $\vec{y} := \begin{pmatrix} y \\ z \end{pmatrix}$:

$$\frac{\partial F}{\partial (y, z)}(x, y, z) = \begin{pmatrix} xe^y & \cos z \\ \cos x & 2z \end{pmatrix}, \quad \frac{\partial F}{\partial (y, z)}(0, 0, 0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad (11.78)$$

$\Rightarrow \det \left(\frac{\partial F}{\partial (y, z)}(\vec{0}) \right) = -1 \neq 0$, d.h. solche Funktionen $y(x)$ und $z(x)$ existieren tatsächlich lokal. Weiter:

$$\frac{\partial F}{\partial (x)}(x, y, z) = \begin{pmatrix} 3x^2 + e^y \\ -y \sin x \end{pmatrix}, \quad \frac{\partial F}{\partial (x)}(\vec{0}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (11.79)$$

$$\Rightarrow \begin{pmatrix} y'(0) \\ z'(0) \end{pmatrix} = - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \quad (\text{wie oben}). \quad (11.80)$$

Beweisidee:

(Existenz und Diffbarkeit von f ohne Bew., **aber:** Grund für Bedingung und Herleitung

der Formel für f')

$$0 \stackrel{!}{=} F(\vec{x}, \vec{y}) \underset{F \text{ diffbar}}{=} F(\vec{x}_0, \vec{y}_0) + \frac{\partial F}{\partial \vec{x}}(\vec{x}_0, \vec{y}_0) \cdot (\vec{x} - \vec{x}_0) + \frac{\partial F}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \cdot (\vec{y} - \vec{y}_0) + o\left(\sqrt{|\vec{x} - \vec{x}_0|^2 + |\vec{y} - \vec{y}_0|^2}\right) \quad (11.81)$$

Auflösen nach \vec{y} erfordert Multiplikation von links mit $\left[\frac{\partial F}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0)\right]^{-1}$,
daher Bedingung $\det \frac{\partial F}{\partial \vec{y}}(\vec{x}_0, \vec{y}_0) \neq 0$

Leite dann $F(\vec{x}, f(\vec{x})) = 0$ nach \vec{x} ab (Kettenregel):

$$\begin{aligned} \frac{\partial F}{\partial \vec{x}}(\vec{x}, f(\vec{x})) + \frac{\partial F}{\partial \vec{y}}(\vec{x}, f(\vec{x})) \underbrace{\frac{\partial f}{\partial \vec{x}}(\vec{x})}_{f'(\vec{x})} &= 0 \\ \Rightarrow f'(\vec{x}) &= - \left[\frac{\partial F}{\partial \vec{y}}(\vec{x}, f(\vec{x})) \right]^{-1} \frac{\partial F}{\partial \vec{x}}(\vec{x}, f(\vec{x})) \quad \forall \vec{x} \in U_\varepsilon(\vec{x}_0), \end{aligned} \quad (11.82)$$

insbesondere für $\vec{x} = \vec{x}_0$ □

Korollar zu Satz 35. (Umkehrfunktion)

Sei $M \subseteq \mathbb{R}^n$ offen, $f : M \rightarrow \mathbb{R}^n$, $f \in C^1(M)$.

Weiter sei $\vec{x}_0 \in M$ mit $\det f'(\vec{x}_0) \neq 0$ und $\vec{y}_0 := f(\vec{x}_0)$.

Dann existiert lokal eine stetig diffbare Umkehrfunktion f^{-1} von f , d.h.

- es gibt offene Mengen $U, V \subseteq \mathbb{R}^n$, mit $\vec{x}_0 \in U$ und $\vec{y}_0 \in V$,
- $f : U \rightarrow V$ bijektiv, d.h. insbesondere $f^{-1} : V \rightarrow U$ existiert und
- $f^{-1} \in C^1(V)$ mit $f^{-1}(\vec{y}) = (f'(f^{-1}(\vec{y})))^{-1}$.

Beweis:

Löse $\vec{y} = f(\vec{x})$ nach $\vec{x} = g(\vec{y})$ auf, d.h.

löse $F(\vec{x}, \vec{y}) = 0$ mit

$$F(\vec{x}, \vec{y}) := f(\vec{x}) - \vec{y} \quad (11.83)$$

nach $\vec{x} = g(\vec{y})$ auf. Möglich lt. Satz 35, falls $0 \neq \det \frac{\partial F}{\partial \vec{x}}(\vec{x}_0, \vec{y}_0) = \det f'(\vec{x}_0)$. Weiter gilt

$$g'(\vec{y}) = - \underbrace{\left[\frac{\partial F}{\partial \vec{x}}(g(\vec{y}), \vec{y}) \right]^{-1}}_{=f'(\vec{x})} \underbrace{\frac{\partial F}{\partial \vec{y}}(g(\vec{y}), \vec{y})}_{=-I} = [f'(\vec{x})]^{-1} \quad (11.84)$$

□

Beispiel:

$$f(x_1, x_2) = \begin{pmatrix} x_1 - x_2^2 \\ 2x_2 \end{pmatrix} \quad (11.85)$$

$$\det(f'(x_1, x_2)) = \det \begin{pmatrix} 1 & -2x_2 \\ 0 & 2 \end{pmatrix} = 2 \quad (11.86)$$

also überall lokal umkehrbar, z.B. $\vec{x}_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$$f(\vec{x}_0) = f(1, 2) = \begin{pmatrix} -3 \\ 4 \end{pmatrix} = \vec{y}_0 \quad (11.87)$$

$$f^{-1'}(-3, 4) = [f'(1, 2)]^{-1} = \begin{pmatrix} 1 & -4 \\ 0 & 2 \end{pmatrix}^{-1} = \frac{1}{2} \begin{pmatrix} 2 & 4 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & \frac{1}{2} \end{pmatrix} \quad (11.88)$$

geht hier auch explizit

$$f(x_1, x_2) = \begin{pmatrix} x_1 - x_2^2 \\ 2x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad (11.89)$$

(2. Komp.) $x_2 = \frac{1}{2}y_2$, (in 1. Komp.) $x_1 - \frac{1}{4}y_2^2 = y_1 \Rightarrow x_1 = y_1 + \frac{1}{4}y_2^2$, also

$$f^{-1}(y_1, y_2) = \begin{pmatrix} y_1 + \frac{1}{4}y_2^2 \\ \frac{1}{2}y_2 \end{pmatrix} \quad (11.90)$$

und damit

$$f^{-1'}(y_1, y_2) = \begin{pmatrix} 1 & \frac{1}{2}y_2 \\ 0 & \frac{1}{2} \end{pmatrix} \Rightarrow f^{-1'}(-3, 4) = \begin{pmatrix} 1 & 2 \\ 0 & \frac{1}{2} \end{pmatrix} \quad (11.91)$$

11.8 Extremwerte unter Nebenbedingungen

Definition: Sei $M \subseteq \mathbb{R}^n$ offen und f, g_1, \dots, g_m jeweils $M \rightarrow \mathbb{R}$. $\vec{x}_0 \in M$ heißt lokale Maximalstelle (Minimalstelle) von f unter den Nebenbedingungen $g_j = 0$, $j = 1, \dots, m$, falls

- (i) $g_j(\vec{x}_0) = 0 \forall j = 1, \dots, m$ und
- (ii) es existiert ein $\varepsilon > 0$, so dass
 - $f(\vec{x}) < f(\vec{x}_0) \forall \vec{x} \in U_\varepsilon(\vec{x}_0) \setminus \{\vec{x}_0\}$ mit $g_j(\vec{x}) = 0 \forall j = 1, \dots, m$.
 - ($f(\vec{x}) > f(\vec{x}_0) \dots$)

Beispiel: (n=2, m=1) Suche Extremstellen von

$$f(x, y) = e^{-(x-5)^2 - (y-4)^2} \quad (11.92)$$

unter der Nebenbedingung $x^2 + y^2 = 1$ (d.h. $g(x, y) = x^2 + y^2 - 1$).

1. Idee: Löse NB nach y auf, setze in f ein, und löse "normales" Extremwertproblem für $f(x, y(x))$.

Aber: Auflösen geht nicht immer, und wenn dann oft hässlich, z.B.

- Wurzeln
- Fallunterscheidungen
- Randpunkte

2. Idee: Geometrisch

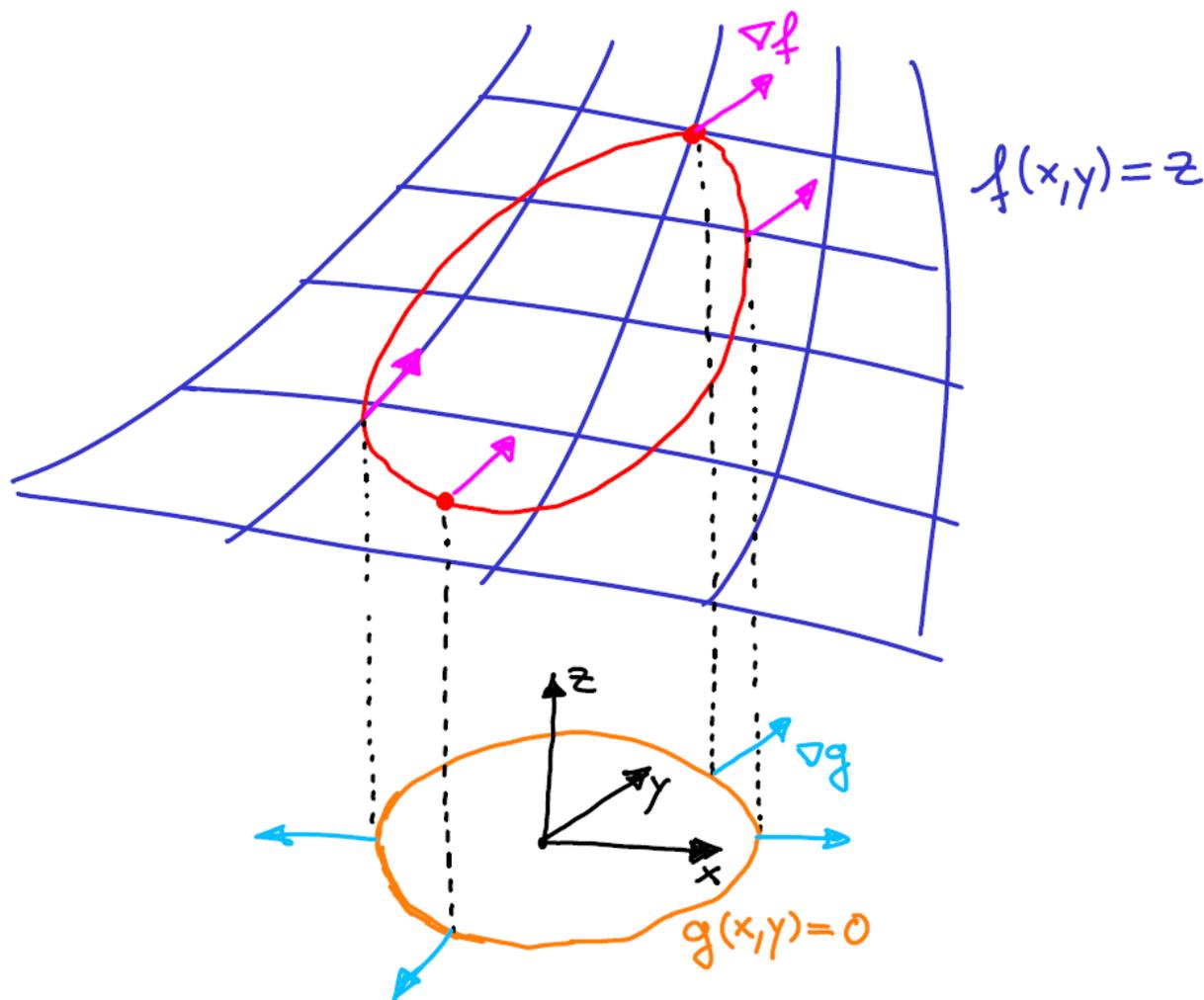
Angenommen wir haben NB parametrisiert, hier als Kurve $\mathfrak{K} : \vec{x}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$, d.h.

$$\{\vec{x} \in \mathbb{R}^2 \mid g(\vec{x}) = 0\} = \{\vec{x}(t) \mid t = \dots\} \quad (11.93)$$

brauchen wir nicht explizit – hier geht's aber z.B. mit $x = \cos t$, $y = \sin t$, $t \in [0, 2\pi]$

Suche Extremstelle von $f(\vec{x}(t))$, notwendige Bedingung:

$$\frac{d}{dt} f(\vec{x}(t)) = (\nabla f)(\vec{x}(t)) \cdot \dot{\vec{x}}(t) \stackrel{!}{=} 0 \quad (11.94)$$



andererseits

$$g(\vec{x}(t)) = 0 \quad \Rightarrow \quad (\nabla g)(\vec{x}(t)) \cdot \dot{\vec{x}}(t) = 0 \quad (11.95)$$

d.h. $\nabla g \perp \dot{\vec{x}}(t)$ (immer) und an Extremstelle auch $\nabla f \perp \dot{\vec{x}}(t)$
 also gilt an Extremstelle ($n = 2$): $\nabla f \parallel \nabla g$, d.h. $\exists \lambda \in \mathbb{R}$, so dass

$$(\nabla f) + \lambda(\nabla g) = 0 \quad (\text{an einer Extremstelle}) \quad (11.96)$$

Suche nun Extrema von

$$L(\vec{x}, \lambda) := f(\vec{x}) + \lambda g(\vec{x}), \quad (11.97)$$

denn $\nabla L = 0$ liefert obige Bedingung und $\frac{\partial L}{\partial \lambda} = 0$ reproduziert NB.

Allgemein...

Satz 36. (Lagrange-Multiplikatoren)

Sei $M \subseteq \mathbb{R}^n$ offen, $f, g_1, \dots, g_m \in C^1(M)$ skalare Funktionen mit $m < n$. Ist $\vec{x}_0 \in M$ ein lokales Extremum von f unter den Nebenbedingungen $g_j = 0$, $j = 1, \dots, m$, und sind $(\nabla g_1)(\vec{x}_0), \dots, (\nabla g_m)(\vec{x}_0)$ linear unabhängig, so gibt es sogenannte Lagrange-Multiplikatoren $\lambda_1, \dots, \lambda_m$ für die gilt:

Definieren wir (die Lagrange-Funktion)

$$L(\vec{x}, \vec{\lambda}) := f(\vec{x}) + \sum_{j=1}^m \lambda_j g_j(\vec{x}) \quad (11.98)$$

so gibt es ein $\vec{\lambda}_0 \in \mathbb{R}^m$, so dass gilt

$$\frac{\partial L}{\partial \vec{x}}(\vec{x}_0, \vec{\lambda}_0) = 0 \quad \text{und} \quad \frac{\partial L}{\partial \lambda_j}(\vec{x}_0, \vec{\lambda}_0) = 0 \quad \forall j = 1, \dots, m. \quad (11.99)$$

(ohne Beweis; Beweisidee oben)

Kurz: Statt Extremwertaufgabe mit Nebenbedingungen für f untersuche gewöhnliches Extremalproblem für L .

Beispiel: (von oben)

$$\begin{aligned} f(x, y) &= e^{-(x-5)^2 - (y-4)^2} \\ g(x, y) &= x^2 + y^2 - 1 \end{aligned} \quad (11.100)$$

$$L(x, y, \lambda) = e^{-(x-5)^2 - (y-4)^2} + \lambda(x^2 + y^2 - 1) \quad (11.101)$$

$$\frac{\partial L}{\partial x} = -2(x-5) e^{-(x-5)^2 - (y-4)^2} + 2\lambda x \stackrel{!}{=} 0 \quad (\text{i})$$

$$\frac{\partial L}{\partial y} = -2(y-4) e^{-(x-5)^2 - (y-4)^2} + 2\lambda y \stackrel{!}{=} 0 \quad (\text{ii})$$

$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - 1 \stackrel{!}{=} 0 \quad (\text{iii})$$

Für $x \neq 0 \neq y$ bilde

$$(i) \cdot y - (ii) \cdot x : -2y(x-5) e^{-(x-5)^2 - (y-4)^2} + 2x(y-4) e^{-(x-5)^2 - (y-4)^2} = 0$$

$$\Leftrightarrow 5y = 4x \quad \Leftrightarrow y = \frac{4}{5}x$$

in (iii)

$$x^2 + \frac{16}{25}x^2 = 1 \quad \Leftrightarrow \quad x = \pm \frac{5}{\sqrt{41}} \quad (11.102)$$

Also: Potentielle Extremstellen von f unter NB $g = 0$ sind

$$\vec{x}_+ = \left(\frac{5}{\sqrt{41}}, \frac{4}{\sqrt{41}} \right) \quad \text{und} \quad \vec{x}_- = \left(-\frac{5}{\sqrt{41}}, -\frac{4}{\sqrt{41}} \right). \quad (11.103)$$

Bei $x = 0$ oder $y = 0$ finden wir keine weiteren Extrema.

Skizze zeigt, dass Maximum bei \vec{x}_+ und Minimum bei \vec{x}_- , ... oder besser:

Wir sehen, dass

$$f(\vec{x}_+) > f(\vec{x}_-). \quad (11.104)$$

Da f stetig ist und der Einheitskreis kompakt, nimmt die Funktion auf dem Einheitskreis nach Satz 27 Minimum und Maximum an, d.h. das Minimum wird bei \vec{x}_- und das Maximum bei \vec{x}_+ angenommen.

Weitere Erläuterungen zu den Fällen $(n = 3, m = 2)$, $(n = 3, m = 1)$ und der allgemeinen Situation.

9 Nachtrag: Differentialgleichungen (DGLn)

9.5 Existenz und Eindeutigkeit für Systeme und DGLn beliebiger Ordnung

In Abschnitt 9.4 haben wir gezeigt, dass das AWP

$$y' = f(x, y), \quad y(x_0) = y_0 \quad (9.83)$$

unter gewissen Bedingungen an die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ eine eindeutige Lösung $y : U_\varepsilon(x_0) \rightarrow \mathbb{R}$ hat.

Ersetzen wir nun y durch $\vec{y} : \mathbb{R} \rightarrow \mathbb{R}^n$ und f durch $\vec{f} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$, so erhalten wir ein System aus n DGLn erster Ordnung und das AWP

$$\vec{y}' = \vec{f}(x, \vec{y}), \quad \vec{y}(x_0) = \vec{y}_0. \quad (9.84)$$

Satz 21 und seinen Beweis kann man dann fast wörtlich übernehmen:

Satz 37. (Picard-Lindelöf für Systeme)

Die Funktion $\vec{f} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ sei im $(n+1)$ -dimensionalen Rechteck

$$D := \{(x, \vec{y}) \mid |x - x_0| \leq a, |y_j - y_{0j}| \leq b, j = 1, \dots, n\} \quad (9.85)$$

stetig und dort stetig partiell nach y_1, y_2, \dots, y_n differenzierbar. Weiter seien M und h durch

$$M := \max_{j=1, \dots, n} \max_{(x, \vec{y}) \in D} |f_j(x, \vec{y})| \quad \text{und} \quad h := \min\left(a, \frac{b}{M}\right) \quad (9.86)$$

definiert. Dann gibt es in der Umgebung

$$U_h(x_0) := \{x \mid |x - x_0| < h\} \quad (9.87)$$

der Stelle x_0 genau eine Lösung des AWP

$$\vec{y}' = \vec{f}(x, \vec{y}), \quad \vec{y}(x_0) = \vec{y}_0. \quad (9.88)$$

Bemerkung: (DGLn höherer Ordnung)

Wie wir in Übungsaufgabe 29 gesehen haben, können wir eine DGL n -ter Ordnung

$$y^{(n)}(x) = f(x, y(x), y'(x), \dots, y^{(n-1)}(x)) \quad (9.89)$$

umschreiben als System von n DGLn erster Ordnung, durch

$$\vec{y}(x) = \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix} := \begin{pmatrix} y(x) \\ y'(x) \\ \vdots \\ y^{(n-1)}(x) \end{pmatrix}, \quad (9.90)$$

auf

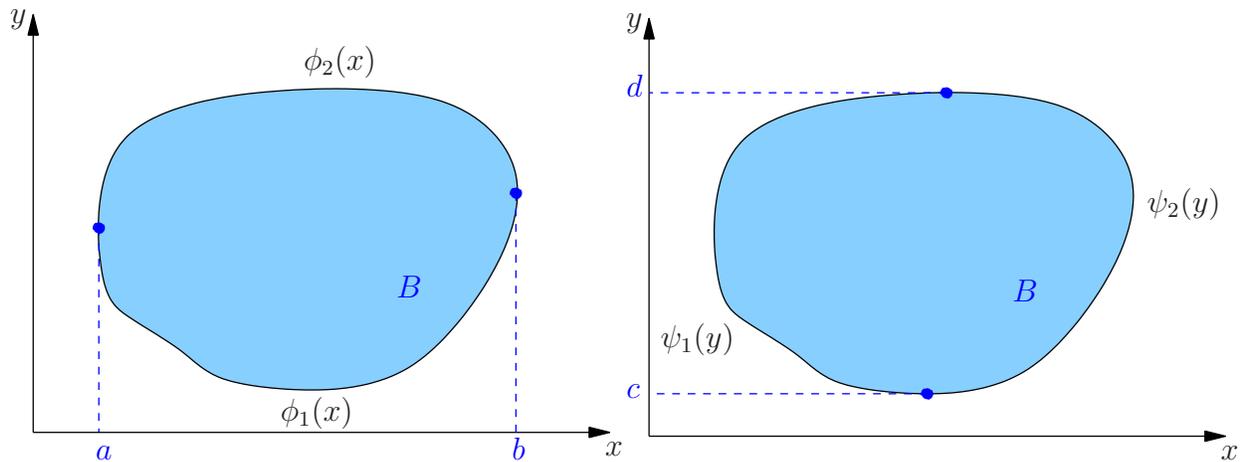
$$\vec{y}' = \begin{pmatrix} y_1'(x) \\ y_2'(x) \\ \vdots \\ y_{n-1}'(x) \\ y_n'(x) \end{pmatrix} = \begin{pmatrix} y_2(x) \\ y_3(x) \\ \vdots \\ y_n(x) \\ f(x, y_1(x), \dots, y_{n-1}(x)) \end{pmatrix}. \quad (9.91)$$

Um dann (unter gewissen Voraussetzungen) ein eindeutiges AWP zu erhalten, müssen wir einen Anfangswert(-vektor) $\vec{y}(x_0) = \vec{y}_0$ vorgeben, d.h. **eine eindeutige Lösung der DGL n -ter Ordnung erhalten wir, wenn wir n Anfangswerte, $y(x_0), y'(x_0), \dots, y^{(n)}(x_0)$, vorgeben** (vgl. Abschnitt 9.3)

12 Integration im \mathbb{R}^n

12.1 Bereichs- und iterierte Integrale

Beispiel:



Flächeninhalt:

$$\begin{aligned}
 |B| &= \int_a^b (\phi_2(x) - \phi_1(x)) \, dx \\
 &= \int_a^b \left(\int_{\phi_1(x)}^{\phi_2(x)} dy \right) dx,
 \end{aligned}
 \tag{12.1}$$

andererseits

$$\begin{aligned}
 |B| &= \int_c^d (\psi_2(y) - \psi_1(y)) \, dy \\
 &= \int_c^d \left(\int_{\psi_1(y)}^{\psi_2(y)} dx \right) dy,
 \end{aligned}
 \tag{12.2}$$

d.h. Fläche gegeben durch iteriertes Integral.

Mit Integrand... (zunächst über Rechtecke)

Satz 38. (Fubini)

Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\vec{x} \mapsto f(\vec{x})$ stetig auf dem Rechteck $a_j \leq x_j \leq b_j$, $j = 1, 2$, dann gilt

$$\int_{a_2}^{b_2} \left(\int_{a_1}^{b_1} f(x_1, x_2) \, dx_1 \right) dx_2 = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} f(x_1, x_2) \, dx_2 \right) dx_1
 \tag{12.3}$$

(ohne Beweis)

Verallgemeinerungen: Gilt auch für

- $f : \mathbb{R}^n \rightarrow \mathbb{R}$ und den n -dimensionalen Quader $a_j \leq x_j \leq b_j$, $j = 1, \dots, n$ und beliebige Vertauschung der Integrationsreihenfolge in

$$\int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \dots dx_1 \quad (12.4)$$

- Integration über sogenannte **Normalbereiche** $B \subset \mathbb{R}^n$ der Form (hier $n = 2$)

$$\begin{aligned} B &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x) \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y) \right\}, \quad \phi_j, \psi_j \text{ stetig} \end{aligned} \quad (12.5)$$

(vgl. obige Skizze). d.h. verschieden parametrisierbar

Beispiel:

Inneres des Einheitskreises

$$\begin{aligned} B &= \{ \vec{x} \in \mathbb{R}^2 \mid |\vec{x}|^2 \leq 1 \} \\ &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid -1 \leq x \leq 1, -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2} \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid -1 \leq y \leq 1, -\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2} \right\} \end{aligned} \quad (12.6)$$

ist Normalbereich.

Definition: (Bereichsintegral)

Seien $\phi_{1,2}(x)$ stetig auf $[a, b]$ und $\psi_{1,2}(x, y)$ stetig auf

$$B = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid a \leq x \leq b, \phi_1(x) \leq y \leq \phi_2(x) \right\}, \quad (12.7)$$

und sei

$$K := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} x \\ y \end{pmatrix} \in B, \psi_1(x, y) \leq z \leq \psi_2(x, y) \right\}, \quad (12.8)$$

sowie f stetig auf B und g stetig auf K . Dann heißt

$$\int_B f dV := \iint_B f dx dy := \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} f(x, y) dy dx \quad (12.9)$$

bzw.

$$\int_K g dV := \iiint_K g dx dy dz := \int_a^b \int_{\phi_1(x)}^{\phi_2(x)} \int_{\psi_1(x,y)}^{\psi_2(x,y)} g(x, y, z) dz dy dx \quad (12.10)$$

Bereichsintegral von f über B (von g über K – analog für höhere Dimensionen). Speziell heißen $|B| = \int_B dV$ Fläche von B und $\int_K dV$ Volumen von K .

Bemerkung:

Falls B oder K kein Normalbereich: Versuche, in Normalbereiche zu zerschneiden.

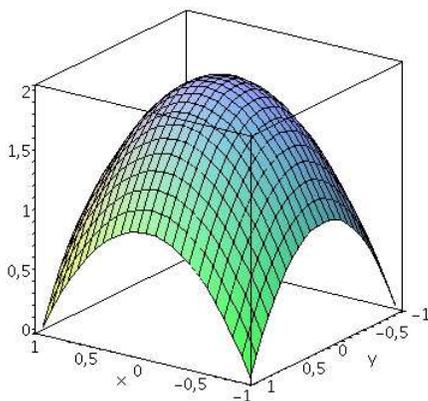
Beispiele:

1. Fläche des Einheitskreises

$$B = \{\vec{x} \in \mathbb{R}^2 \mid |\vec{x}|^2 \leq 1\} \tag{12.11}$$

$$\begin{aligned} |B| &= \int_B dV = \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} dy dx \\ &= 2 \int_{-1}^1 \sqrt{1-x^2} dx, & x = \sin t \\ &= 2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sqrt{1-\sin^2 t} \cos t dt = 2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t dt \\ &= 2 \left[\frac{t}{2} + \frac{1}{4} \sin(2t) \right]_{-\frac{\pi}{2}}^{+\frac{\pi}{2}} = \pi \end{aligned} \tag{12.12}$$

2. $B = \{(x, y)^T \in \mathbb{R}^2 \mid |x| \leq 1, |y| \leq 1\}$, $f(x, y) = 2 - x^2 - y^2$



anschaulich: Volumen zwischen $z = f(x, y)$ und xy -Ebene

$$\begin{aligned} \int_B f dV &= \int_{-1}^1 \int_{-1}^1 (2 - x^2 - y^2) dx dy \\ &= \int_{-1}^1 \left[2x - \frac{x^3}{3} - y^2 x \right]_{x=-1}^{x=+1} dy \\ &= \int_{-1}^1 \left(4 - \frac{2}{3} - 2y^2 \right) dy \\ &= \left[\frac{10}{3} y - \frac{2}{3} y^3 \right]_{-1}^{+1} \\ &= \frac{20}{3} - \frac{4}{3} = \frac{16}{3} \end{aligned} \tag{12.13}$$

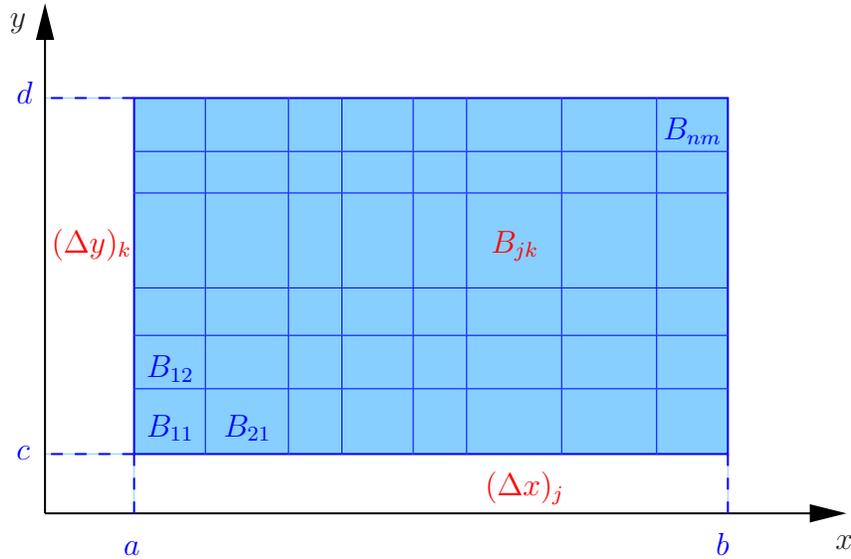
Bemerkung:

Damit haben wir das Bereichsintegral als iteriertes Integral definiert.

Für weniger hübsche Ränder komplizierter...

Die Konstruktion war angepasst an kartesische Koordinaten. Wie sollten wir für krummlinige vorgehen? Dazu ahmen wir zunächst für $\int_c^d \int_a^b f(x, y) dx dy$ die Konstruktion aus einer Dimension nach:

1. Zerlege Rechteck in kleine Rechtecke



2. Bilde Ober- und Untersummen

$$S_{nm}^O := \sum_{j=1}^n \sum_{k=1}^m \max_{\vec{x} \in B_{jk}} f(\vec{x}) \cdot |B_{jk}|, \quad S_{nm}^U := \sum_{j=1}^n \sum_{k=1}^m \min_{\vec{x} \in B_{jk}} f(\vec{x}) \cdot |B_{jk}| \quad (12.14)$$

3. Verfeinere: $n, m \rightarrow \infty$, alle Kantenlängen $\rightarrow 0$

4. Hoffe, dass $\lim_{n, m \rightarrow \infty} S_{nm}^O = \lim_{n, m \rightarrow \infty} S_{nm}^U =: \int_B f dV$.

Dabei wird Fläche $|B_{jk}| = (\Delta x)_j (\Delta y)_k$ zum

Flächenelement in kartesischen Koordinaten: $dx dy$, "infinitesimales Quadrat/Rechteck".

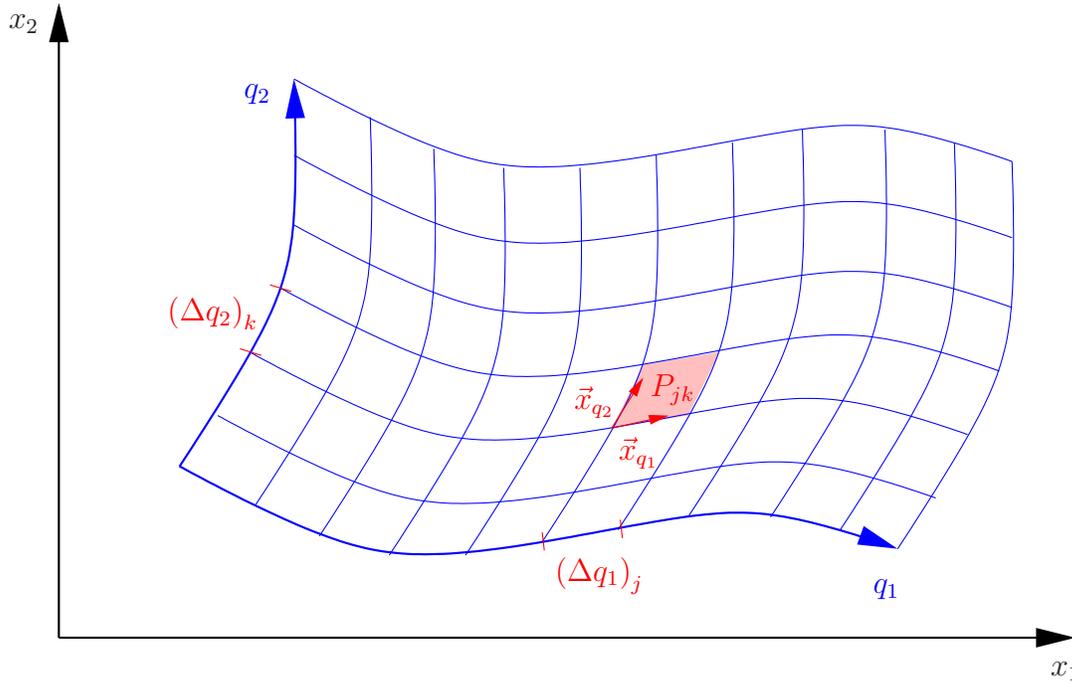
Analog:

Volumenelement in kartesischen Koordinaten: $dx dy dz$ "infinitesimaler Würfel/Quader",
bzw. im \mathbb{R}^n : $dx_1 dx_2 \dots dx_n$.

Frage: Integration in anderen (krümmigen) Koordinaten (q_1, \dots, q_n) ?

D.h. bijektive Abbildung

$$\begin{aligned} \vec{x} : \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ \vec{q} &\mapsto \vec{x}(\vec{q}) \end{aligned} \quad (12.15)$$



1. Zerlege in kleine Parallelogramme P_{jk} , aufgespannt durch Tangentialvektoren an die Koordinatenlinien.
2. Fläche des Parallelograms (an der Stelle \vec{q})

$$|P_{jk}| = |\det(\vec{x}_{q_1}(\vec{q}), \vec{x}_{q_2}(\vec{q}))| (\Delta q_1)_j (\Delta q_2)_k \quad (12.16)$$

Damit z.B. Obersumme

$$S_{nm}^O := \sum_{j=1}^n \sum_{k=1}^m \max_{\vec{q} \in P_{jk}} f(\vec{x}(\vec{q})) \cdot |P_{jk}|. \quad (12.17)$$

3. Verfeinere: Das Flächenelement wird

$$dV = dx dy = |\det(\vec{x}_{q_1}(\vec{q}), \vec{x}_{q_2}(\vec{q}))| dq_1 dq_2. \quad (12.18)$$

Gilt tatsächlich...

Satz 39. (Transformationssatz bzw. Substitutionsregel)

Sei $\tilde{K} \subset \mathbb{R}^n$, $\vec{x} : \tilde{K} \rightarrow K \subset \mathbb{R}^n$, $\vec{q} \mapsto \vec{x}(\vec{q})$ bijektiv und $\vec{x} \in C^1(\tilde{K})$. Weiter sei $f : K \rightarrow \mathbb{R}$ stetig, dann gilt

$$\begin{aligned} \int_K f dV &= \int \cdots \int_K f(\vec{x}) dx_1 \dots dx_n \\ &= \int \cdots \int_{\tilde{K}} f(\vec{x}(\vec{q})) \left| \det \frac{\partial \vec{x}}{\partial \vec{q}}(\vec{q}) \right| dq_1 \dots dq_n. \end{aligned} \quad (12.19)$$

Beispiele: (Siehe auch Abschnitt 5.8, insbesondere Skizzen dort)

1. Ebene Polarkoordinaten (r, φ)

$$\vec{x} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix}, \quad \begin{array}{l} 0 \leq r < \infty \\ 0 \leq \varphi < 2\pi \end{array} \quad (12.20)$$

$$\begin{aligned} dV = dx dy &= \left| \det \frac{\partial(x, y)}{\partial(r, \varphi)} \right| dr d\varphi \\ &= \left| \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} \right| dr d\varphi = r dr d\varphi \end{aligned} \quad (12.21)$$

(a) nochmal Fläche des Kreises, diesmal mit Radius R : $K = \{\vec{x} \in \mathbb{R}^2 \mid |\vec{x}| \leq R\}$

$$\int_K dV = \int_0^{2\pi} \int_0^R r dr d\varphi = 2\pi \left[\frac{r^2}{2} \right]_0^R = \pi R^2 \quad (12.22)$$

viel einfacher als in kartesischen Koordinaten...

(b)

$$J := \int_{-\infty}^{+\infty} e^{-x^2} dx = ? \quad (12.23)$$

keine elementare Stammfunktion

$$J^2 = \left(\int_{-\infty}^{+\infty} e^{-x^2} dx \right)^2 = \left(\int_{-\infty}^{+\infty} e^{-x^2} dx \right) \left(\int_{-\infty}^{+\infty} e^{-y^2} dy \right) = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy \quad (12.24)$$

Polarkoordinaten...

$$J^2 = \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\varphi = 2\pi \left[\frac{e^{-r^2}}{(-2)} \right]_0^{\infty} = \pi, \quad (12.25)$$

und damit $J = \sqrt{\pi}$, also

$$\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi} \quad (12.26)$$

2. Flächeninhalt einer Ellipse mit Halbachsen a und b

3. Kugelkoordinaten

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \sin \vartheta \cos \varphi \\ r \sin \vartheta \sin \varphi \\ r \cos \vartheta \end{pmatrix}, \quad \begin{array}{l} 0 \leq r < \infty \\ 0 \leq \vartheta \leq \pi \\ 0 \leq \varphi < 2\pi \end{array} \quad (12.27)$$

$$\begin{aligned}
dV &= dx \, dy \, dz = \left| \det \left(\frac{\partial(x, y, z)}{\partial(r, \vartheta, \varphi)} \right) \right| dr \, d\vartheta \, d\varphi \\
&= \left| \det \begin{pmatrix} \sin \vartheta \cos \varphi & r \cos \vartheta \cos \varphi & -r \sin \vartheta \sin \varphi \\ \sin \vartheta \sin \varphi & r \cos \vartheta \sin \varphi & r \sin \vartheta \cos \varphi \\ \cos \vartheta & -r \sin \vartheta & 0 \end{pmatrix} \right| dr \, d\vartheta \, d\varphi \\
&= r^2 \sin \vartheta \left| \cos \vartheta (\cos \vartheta \cos^2 \varphi + \cos \vartheta \sin^2 \varphi) \right. \\
&\quad \left. + \sin \vartheta (\sin \vartheta \cos^2 \varphi + \sin \vartheta \sin^2 \varphi) \right| dr \, d\vartheta \, d\varphi \\
&= \left| r^2 \sin \vartheta (\cos^2 \vartheta + \sin^2 \vartheta) \right| dr \, d\vartheta \, d\varphi \\
&= r^2 \sin \vartheta \, dr \, d\vartheta \, d\varphi \quad (\text{vgl. ÜA 45})
\end{aligned} \tag{12.28}$$

Mit $0 \leq \vartheta \leq \pi$ sind Betragsstriche nicht notwendig, da $\sin \vartheta$ positiv.

(a) Volumen einer Kugel mit Radius R :

$$\int_0^{2\pi} \int_0^\pi \int_0^R r^2 \sin \vartheta \, dr \, d\vartheta \, d\varphi = 2\pi \underbrace{\int_0^R r^2 \, dr}_{=R^3/3} \underbrace{\int_0^\pi \sin \vartheta \, d\vartheta}_{=2} = \frac{4\pi}{3} R^3. \tag{12.29}$$

4. Volumen des Vivianischen Fensters

12.2 Oberflächenintegrale

Gegeben: 2-dimensionale (Ober-)Fläche \mathcal{F} im \mathbb{R}^3 , z.B.

- implizit, $F(x, y, z) = \text{const.}$
- explizit, $z = f(x, y)$
- oder parametrisiert...

Definition: Eine parametrisierte Fläche $\mathcal{F} \subset \mathbb{R}^3$ ist gegeben durch $\vec{x}: \mathbb{R}^2 \supset B \rightarrow \mathbb{R}^3$, d.h.

$$\vec{x}(u, v) = \begin{pmatrix} x_1(u, v) \\ x_2(u, v) \\ x_3(u, v) \end{pmatrix}, \quad x_j \text{ kartesisch}, \quad (u, v) \in B. \tag{12.30}$$

- (i) $\vec{x}_u = \frac{\partial \vec{x}}{\partial u}$ und $\vec{x}_v = \frac{\partial \vec{x}}{\partial v}$ heißen Tangentialvektoren.
 \mathcal{F} bzw. Parametrisierung heißt regulär, falls \vec{x}_u, \vec{x}_v l.u.,
d.h. spannen Tangentialebene auf.
- (ii) Ist \mathcal{F} regulär, so heißt $\vec{n} := \frac{\vec{x}_u \times \vec{x}_v}{|\vec{x}_u \times \vec{x}_v|}$ Normale(-neinheitsvektor)
- (iii) Ist \mathcal{F} Oberfläche eines Körpers K im \mathbb{R}^3 , so heißt \vec{n}_i innere Normale, falls \vec{n}_i Normaleinheitsvektor ist und ins Innere von K zeigt, d.h. $\vec{n}_i = \pm \vec{n}$.

Beispiele:

1. \mathcal{F} explizit gegeben, $z = f(x, y)$, $(x, y) \in B \subset \mathbb{R}^2$

$$\vec{x}(u, v) = \begin{pmatrix} u \\ v \\ f(u, v) \end{pmatrix} \quad (12.31)$$

$$\vec{x}_u = \begin{pmatrix} 1 \\ 0 \\ f_u \end{pmatrix}, \quad \vec{x}_v = \begin{pmatrix} 0 \\ 1 \\ f_v \end{pmatrix} \quad \text{regulär!} \quad (12.32)$$

$$\vec{x}_u \times \vec{x}_v = \begin{pmatrix} -f_u \\ -f_v \\ 1 \end{pmatrix}, \quad \vec{n} = \frac{1}{\sqrt{1 + f_u^2 + f_v^2}} \begin{pmatrix} -f_u \\ -f_v \\ 1 \end{pmatrix} \quad (12.33)$$

2. Kugeloberfläche (Radius R), $\mathcal{F} := \{\vec{x} \in \mathbb{R}^3 \mid |\vec{x}| = R\}$,
speziell für $R = 1$: $S^2 := \{\vec{x} \in \mathbb{R}^3 \mid |\vec{x}| = 1\}$

$$\vec{x}(u, v) = R \begin{pmatrix} \sin v \cos u \\ \sin v \sin u \\ \cos v \end{pmatrix}, \quad 0 \leq u < 2\pi, \quad 0 \leq v \leq \pi \quad (12.34)$$

$$\vec{x}_u = R \sin v \begin{pmatrix} -\sin u \\ \cos u \\ 0 \end{pmatrix}, \quad \vec{x}_v = R \begin{pmatrix} \cos v \cos u \\ \cos v \sin u \\ -\sin v \end{pmatrix} \quad (12.35)$$

$$\vec{x}_u \times \vec{x}_v = -R^2 \sin v \underbrace{\begin{pmatrix} \sin v \cos u \\ \sin v \sin u \\ \cos v \end{pmatrix}}_{=\vec{n}} = -R \sin v \vec{x}(u, v), \quad \text{regulär für } v \neq 0, \pi. \quad (12.36)$$

3. Kegelmantel ($h > 0$)

$$\vec{x}(u, v) = \begin{pmatrix} v \cos u \\ v \sin u \\ v \end{pmatrix}, \quad 0 \leq u < 2\pi, \quad 0 \leq v \leq h \quad (12.37)$$

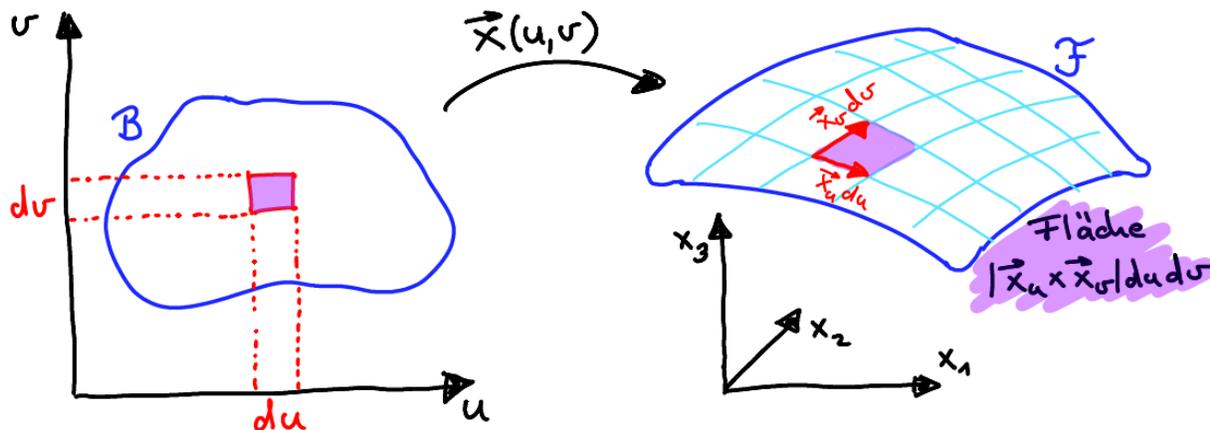
$$\vec{x}_u = v \begin{pmatrix} -\sin u \\ \cos u \\ 0 \end{pmatrix}, \quad \vec{x}_v = \begin{pmatrix} \cos u \\ \sin u \\ 1 \end{pmatrix} \quad (12.38)$$

$$\vec{x}_u \times \vec{x}_v = \begin{pmatrix} v \cos u \\ v \sin u \\ -v \end{pmatrix}, \quad \vec{n} = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos u \\ \sin u \\ -1 \end{pmatrix} \quad (12.39)$$

regulär für $v \neq 0$

Integration über \mathcal{F} : Konstruktion wie bei Substitutionsregel

1. zerlege \mathcal{F} in Parallelogramme, aufgespannt durch \vec{x}_u, \vec{x}_v
2. Flächenelement $dO = |\vec{x}_u \times \vec{x}_v| du dv$



Definition: (Oberfläche / Flächenintegral)

Sei $\mathcal{F} \subset \mathbb{R}^3$ eine parametrisierte Fläche, regulär, d.h. $\vec{x} : \mathbb{R}^2 \supset B \rightarrow \mathbb{R}^3$, und sei f stetig auf \mathcal{F} ($\stackrel{\text{Def.}}{\Leftrightarrow} f \circ \vec{x}$ stetig auf B), dann heißt

$$\int_{\mathcal{F}} f dO := \iint_B f(\vec{x}(u, v)) |\vec{x}_u \times \vec{x}_v| du dv \quad (12.40)$$

Flächenintegral von f über \mathcal{F} ; speziell heißt

$$O(\mathcal{F}) := \int_{\mathcal{F}} dO \quad (12.41)$$

Oberfläche von \mathcal{F} .

Beispiele:

1. Oberfläche einer Kugel mit Radius R (parametrisiert wie in Bsp. 2 oben):
 $|\vec{x}_u \times \vec{x}_v| = R^2 \sin v$,

$$O(\mathcal{F}) = \int_0^{2\pi} \int_0^\pi R^2 \sin v dv du = 4\pi R^2 \quad (12.42)$$

2. $f = \arctan(\frac{y}{x})$ auf (Einheits-)Kugelhälfte – parametrisiert wie oben, mit $0 \leq u \leq \pi$:

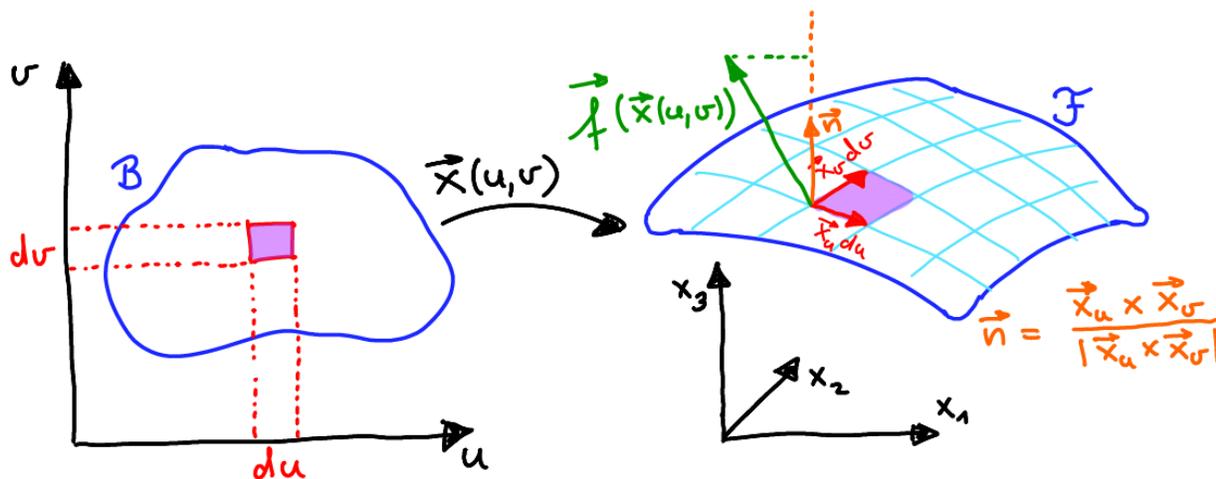
$$\begin{aligned} \iint_{\mathcal{F}} f dO &= \int_0^\pi \int_0^\pi \arctan\left(\frac{\cos v \sin u}{\cos v \cos u}\right) \sin v dv du \\ &= 2 \int_0^\pi u du = \pi^2 \end{aligned} \quad (12.43)$$

Definition: (Oberflächenintegral 2. Art / Fluss)

Sei $\mathcal{F} \subset \mathbb{R}^3$ eine parametrisierte Fläche, regulär, d.h. $\vec{x}: \mathbb{R}^2 \supset B \rightarrow \mathbb{R}^3$, und sei $\vec{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ein stetiges Vektorfeld auf \mathcal{F} ($\stackrel{\text{Def.}}{\Leftrightarrow} \vec{f} \circ \vec{x}$ stetig auf B), dann heißt

$$\begin{aligned} \int_{\mathcal{F}} \vec{f} d\vec{O} &:= \iint_B \left(\vec{f}(\vec{x}(u, v)) \cdot \vec{n}(u, v) \right) \underbrace{|\vec{x}_u \times \vec{x}_v|}_{=dO} du dv \\ &= \iint_B \vec{f}(\vec{x}(u, v)) \cdot \underbrace{(\vec{x}_u \times \vec{x}_v)}_{=d\vec{O}} du dv \end{aligned} \quad (12.44)$$

Fluss von \vec{f} durch \mathcal{F} .



Bemerkung: Vorzeichen von $\int_{\mathcal{F}} \vec{f} d\vec{O}$ hängt von Parametrisierung ab.

Beispiel: $\vec{f}(\vec{x}) = a \frac{\vec{x}}{|\vec{x}|^3}$, $a \in \mathbb{R}$,

\mathcal{F} : Kugeloberfläche einer Kugel mit Radius R (s.o.): $\vec{x}_u \times \vec{x}_v = -R \sin v \vec{x}(u, v)$,

$$\begin{aligned} \int_{\mathcal{F}} \vec{f} d\vec{O} &= \int_0^{2\pi} \int_0^\pi a \frac{\vec{x}(u, v)}{|\vec{x}(u, v)|^3} (-R \sin v) \vec{x}(u, v) dv du \\ &\stackrel{|\vec{x}(u, v)|=R}{=} -a \int_0^{2\pi} \int_0^\pi \sin v dv du = -4\pi a. \end{aligned} \quad (12.45)$$

12.3 Integralsätze

Motivation: 1D

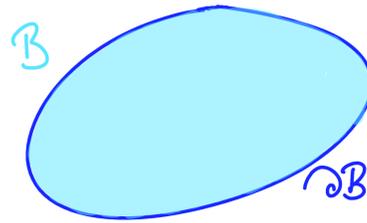
$$\int_a^b f'(x) dx = f(b) - f(a) \quad (12.46)$$

(Hauptsatz der Differential- und Integralrechnung)
 a, b : Endpunkte / "Rand" des Integrationsbereichs

Mehrdimensional?

z.B. endlicher 2D Bereich, $B \subset \mathbb{R}^2$,
 mit Randkurve $\partial B \dots$

... oder 3D Bereich, $K \subset \mathbb{R}^3$,
 mit Oberfläche ∂K .



Satz 40. (Gaußscher Integralsatz)

Sei $K \subset \mathbb{R}^3$ ein Körper mit parametrisierter Oberfläche ∂K (stückweise glatt) und mit innerer Normale \vec{n}_i , und sei $\vec{f}: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ein stetig diffbares Vektorfeld. Dann gilt

$$\iiint_K \operatorname{div} \vec{f} dV = - \oiint_{\partial K} \underbrace{\vec{f} \cdot \vec{n}_i}_{=d\vec{o}} \quad (12.47)$$

Notation \oiint betont, dass es sich um eine *geschlossene* Oberfläche handelt.

Beweisskizze:

Sei K Normalbereich (falls nicht: zerschneiden), d.h. insbesondere

$$\begin{aligned} K &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} x \\ y \end{pmatrix} \in B \subset \mathbb{R}^2, \varphi_1(x, y) \leq z \leq \varphi_2(x, y) \right\} \\ &= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} x \\ z \end{pmatrix} \in \tilde{B} \subset \mathbb{R}^2, \psi_1(x, z) \leq y \leq \psi_2(x, z) \right\}, \end{aligned} \quad (12.48)$$

und damit

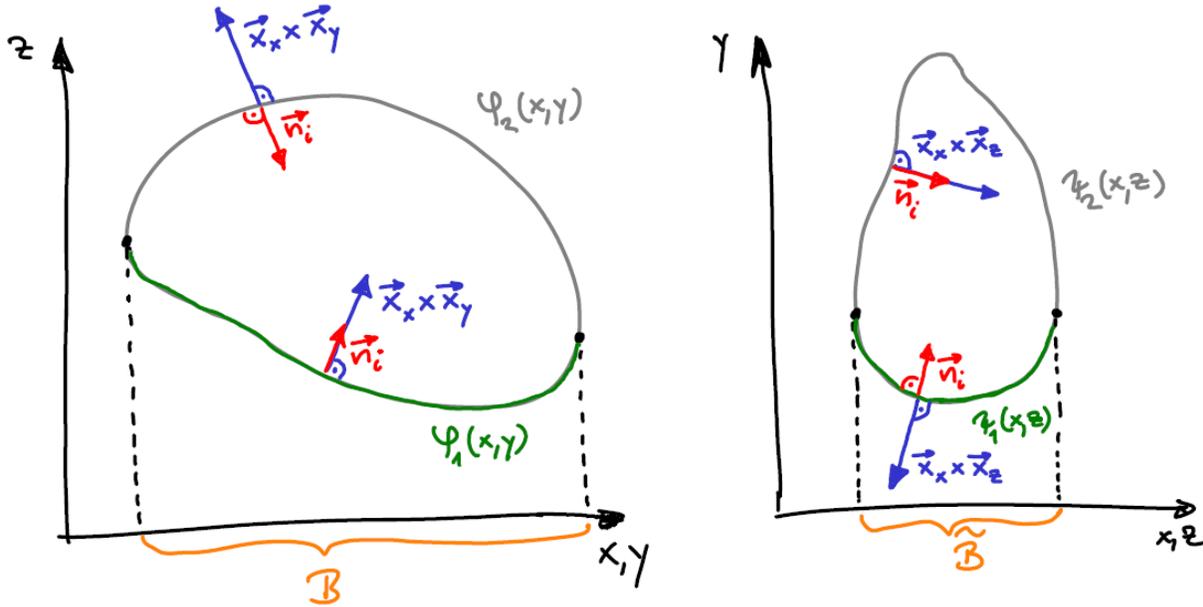
$$\partial K = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} x \\ y \end{pmatrix} \in B, z = \varphi_1(x, y) \right\} \cup \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \mid \begin{pmatrix} x \\ y \end{pmatrix} \in B, z = \varphi_2(x, y) \right\}. \quad (12.49)$$

Stückweise

$$\vec{x} = \begin{pmatrix} x \\ y \\ \varphi_j(x, y) \end{pmatrix} \Rightarrow \vec{x}_x \times \vec{x}_y = \begin{pmatrix} -\varphi_{j,x} \\ -\varphi_{j,y} \\ 1 \end{pmatrix} \quad (12.50)$$

d.h.

$$d\vec{O} = \vec{n}_i dO = \pm(\vec{x}_x \times \vec{x}_y) dx dy. \quad (12.51)$$



Schreibe $\vec{f} = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix}$ und betrachte

$$\begin{aligned} \iiint_K \frac{\partial f_3}{\partial z} dV &= \iint_B \left(\int_{\varphi_1(x,y)}^{\varphi_2(x,y)} \frac{\partial f_3}{\partial z}(x,y,z) dz \right) dx dy \\ &= \iint_B (f_3(x,y,\varphi_2(x,y)) - f_3(x,y,\varphi_1(x,y))) dx dy \\ &= - \oiint_{\partial K} f_3 (\vec{n}_i)_3 dO \end{aligned} \quad (12.52)$$

Analog für die anderen Komponenten...

z.B.

$$\vec{x} = \begin{pmatrix} x \\ \psi_j(x,z) \\ z \end{pmatrix} \Rightarrow \vec{x}_x \times \vec{x}_z = \begin{pmatrix} \psi_{j,x} \\ -1 \\ \psi_{j,z} \end{pmatrix} \quad (12.53)$$

$$\begin{aligned} \iiint_K \frac{\partial f_2}{\partial y} dV &= \iint_{\tilde{B}} \left(\int_{\psi_1(x,z)}^{\psi_2(x,z)} \frac{\partial f_2}{\partial y}(x,y,z) dy \right) dx dz \\ &= \iint_{\tilde{B}} (f_2(x,\psi_2(x,z),z) - f_2(x,\psi_1(x,z),z)) dx dz \\ &= - \oiint_{\partial K} f_2 (\vec{n}_i)_2 dO \end{aligned} \quad (12.54)$$

...Summe der drei Ergebnisse liefert Behauptung. □

Beispiel: $\vec{f}(\vec{x}) = \vec{x}$ auf Kugel

Bemerkung: Anschauliche Bedeutung von div

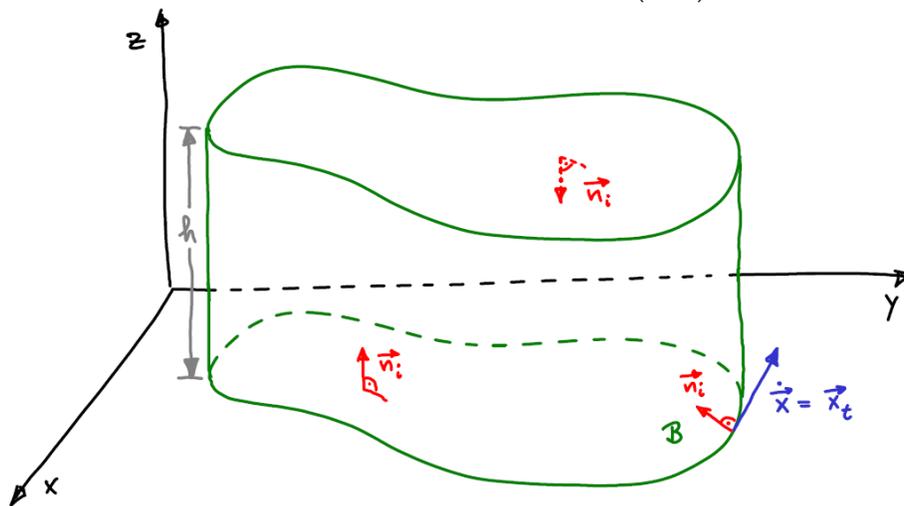
Satz 41. (Stokes'scher Integralsatz in der Ebene)

Sei $B \subset \mathbb{R}^2$ mit stückweise glatter, positiv orientierter, geschlossener Randkurve ∂B (d.h. das Innere "liegt links" der Kurve), und $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ ein stetig diffbares Vektorfeld, dann gilt

$$\iint_B \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) dV = \oint_{\partial B} \vec{v} d\vec{x}. \quad (12.55)$$

Beweis:

Betrachte Gaußschen Satz für $K := B \times [0, h]$ und $\vec{f} := \begin{pmatrix} v_2 \\ -v_1 \\ 0 \end{pmatrix}$.



Linke Seite:

$$\iiint_K \text{div} \vec{f} dx dy dz = h \iint_B \left(\frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \right) dx dy \quad (12.56)$$

Rechte Seite:

oben/unten: $\vec{n}_i = \mp \vec{e}_z \Rightarrow$ jeweils $\vec{f} \vec{n}_i = 0$; bleibt Integral über Mantelfläche.

$$\vec{x}(t, z) = \begin{pmatrix} x(t) \\ y(t) \\ z \end{pmatrix}, \quad \begin{matrix} 0 \leq t \leq T \\ 0 \leq z \leq h \end{matrix} \quad (12.57)$$

$$\begin{aligned}
\oint_{\partial K} \vec{f} \vec{n}_i \, dO &= \int_0^h \int_0^T \vec{f} \cdot (\vec{x}_z \times \vec{x}_t) \, dt \, dz \\
&= \int_0^h \int_0^T \vec{f} \cdot \left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} \dot{x} \\ \dot{y} \\ 0 \end{pmatrix} \right] \, dt \, dz \\
&= \int_0^h \int_0^T \begin{pmatrix} v_2 \\ -v_1 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} -\dot{y} \\ \dot{x} \\ 0 \end{pmatrix} \, dt \, dz \\
&= -h \int_0^T \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \cdot \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} \, dt \\
&= -h \oint_{\partial B} \vec{v} \, d\vec{x}
\end{aligned} \tag{12.58}$$

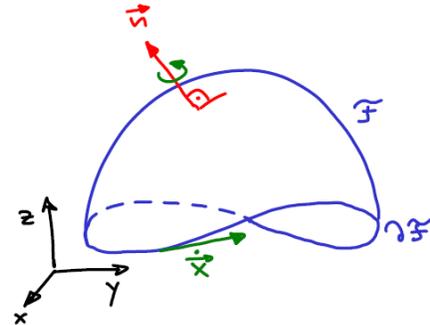
Zusammen mit (12.56) folgt die Behauptung. \square

Bemerkungen: Verallgemeinerungen auf beliebig orientierte ebene Flächenstücke in \mathbb{R}^3 und durch Stückeln auch auf gewölbte Flächen...

Satz 42. (Stokes'scher Integralsatz im \mathbb{R}^3)

Sei $\mathcal{F} \subset \mathbb{R}^3$ eine zweiseitige Fläche, regulär, mit stückweise glatter, geschlossener Randkurve $\partial\mathcal{F}$, die eine Rechtsschraube um die Flächennormale \vec{n} beschreibt, und sei $\vec{v} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ein stetig diffbares Vektorfeld, dann gilt

$$\iint_{\mathcal{F}} \text{rot } \vec{v} \cdot \vec{n} \, dO = \oint_{\partial\mathcal{F}} \vec{v} \, d\vec{x}. \tag{12.59}$$



(Beweisidee s.o.)

Bemerkungen:

1. Ebener Stokes war Spezialfall mit \mathcal{F} in xy -Ebene und $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ 0 \end{pmatrix}$,

$$\text{d.h. } \text{rot } \vec{v} = \begin{pmatrix} 0 \\ 0 \\ \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} \end{pmatrix}.$$

2. Weiterer Spezialfall: $\text{rot } \vec{v} = \vec{0} \Rightarrow \oint_{\partial\mathcal{F}} \vec{v} \, d\vec{x} = 0$ wegen Stokes.
Andererseits: $\text{rot } \vec{v} = \vec{0} \Rightarrow \exists U : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit $\vec{v} = \nabla U$, und damit

$$\oint_{\partial\mathcal{F}} \vec{v} \, d\vec{x} = \underbrace{U(\text{Endpunkt}) - U(\text{Anfangspunkt})}_{=0, \text{ da Endpkt. = Anfangspkt. für } \oint}. \tag{12.60}$$

Bemerkung: Anschauliche Bedeutung von rot

13 Wahrscheinlichkeitsrechnung & Statistik

13.1 Axiomatischer Aufbau

Sei Ω im Folgenden stets eine nicht-leere Menge.

Definition: $\mathcal{P}(\Omega) := \{A \mid A \subseteq \Omega\}$ heißt Potenzmenge von Ω (Menge aller Teilmengen).

Beispiel: $\Omega = \{1, 2, 3\}$

$$\mathcal{P}(\Omega) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \quad (13.1)$$

insbesondere: $\emptyset \in \mathcal{P}(\Omega)$ und $\Omega \in \mathcal{P}(\Omega)$ (immer)

Definition: (σ -Algebra)

Eine Teilmenge \mathcal{A} von $\mathcal{P}(\Omega)$ mit

(i) $\Omega \in \mathcal{A}$

(ii) $(A_n)_{n \in \mathbb{N}}$, Folge von Mengen aus \mathcal{A} , d.h. $A_n \in \mathcal{A} \forall n \in \mathbb{N} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

(iii) $A \in \mathcal{A} \Rightarrow A^C \in \mathcal{A}$

heißt σ -Algebra über Ω . Gilt (ii) nur für endliche Vereinigungen, so heißt \mathcal{A} Algebra.

Bemerkungen:

1. "A oder B" ("A vereinigt B")

$$A \cup B := \{\omega \in \Omega \mid \omega \in A \text{ oder } \omega \in B\} \quad (13.2)$$

2. "A und B" ("A geschnitten B")

$$A \cap B := \{\omega \in \Omega \mid \omega \in A \text{ und } \omega \in B\} \quad (13.3)$$

3. Komplementärereignis/-menge

$$A^C := \{\omega \in \Omega \mid \omega \notin A\} \quad (\Rightarrow (A^C)^C = A) \quad (13.4)$$

4. "A ohne B" / "A minus B"

$$\begin{aligned} A \setminus B &:= \{\omega \in \Omega \mid \omega \in A \text{ und } \omega \notin B\} \\ &= A \cap B^C \end{aligned} \quad (13.5)$$

5. Rechne mit (\cup, \cap) ähnlich wie mit $(+, \cdot)$, d.h. es gelten die üblichen Assoziativ- und Kommutativgesetze sowie *zwei* Distributivgesetze

$$\begin{aligned} A \cap (B \cup C) &= (A \cap B) \cup (A \cap C), \\ A \cup (B \cap C) &= (A \cup B) \cap (A \cup C). \end{aligned} \quad (13.6)$$

6. De Morgansche Regeln

$$\begin{aligned}(A \cup B)^C &= A^C \cap B^C, \\ (A \cap B)^C &= A^C \cup B^C\end{aligned}\tag{13.7}$$

Zum Beweis Bildchen malen...

Beispiele:

1. $\mathcal{P}(\Omega)$ ist (größte) σ -Algebra über Ω .
2. $\{\emptyset, \Omega\}$ ist (kleinste) σ -Algebra über Ω .

Lemma 43. Sei \mathcal{A} eine σ -Algebra über Ω , dann gilt

- (i) (A_n) Folge in $\mathcal{A} \Rightarrow \bigcap_{n=0}^{\infty} A_n \in \mathcal{A}$
- (ii) $A_1, \dots, A_n \in \mathcal{A} \Rightarrow A_1 \cup \dots \cup A_n \in \mathcal{A}$ sowie $A_1 \cap \dots \cap A_n \in \mathcal{A}$
- (iii) $A, B \in \mathcal{A} \Rightarrow A \setminus B \in \mathcal{A}$
- (iv) (A_n) Folge in $\mathcal{A} \Rightarrow \exists$ Folge (B_n) disjunkter Mengen in \mathcal{A} mit $B_n \subseteq A_n \forall n \in \mathbb{N}$ und $\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$

Beweise:

- (i) $\bigcap_{n=1}^{\infty} A_n = \left(\bigcup_{n=1}^{\infty} A_n^C \right)^C$
- (ii) $A_1 \cup \dots \cup A_n = \bigcup_{\nu=1}^{\infty} A_{\nu}$ mit $A_{\nu} := \emptyset \forall \nu > n$
 $A_1 \cap \dots \cap A_n = \bigcap_{\nu=1}^{\infty} A_{\nu}$ mit $A_{\nu} := \Omega \forall \nu > n$
- (iii) $A \setminus B = A \cap B^C$
- (iv) $B_1 := A_1, n \geq 2: B_n := A_n \setminus (A_1 \cup \dots \cup A_{n-1})$
 $\Rightarrow B_n \in \mathcal{A}$, paarweise disjunkt, $B_n \subseteq A_n, \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n$

□

Definition: (Ω, \mathcal{A}, P) heißt Wahrscheinlichkeitsraum (W-Raum), falls gilt

- (i) $\Omega \neq \emptyset$
- (ii) \mathcal{A} ist σ -Algebra über Ω
- (iii) P ist Wahrscheinlichkeitsmaß auf \mathcal{A} , d.h. $P : \mathcal{A} \rightarrow \mathbb{R}$ mit
 - (a) $P(A) \geq 0 \forall A \in \mathcal{A}$ (Nichtnegativität)
 - (b) $P(\Omega) = 1$ (Normiertheit)

(c) Ist (A_n) Folge disjunkter Mengen in \mathcal{A} , d.h. $A_j \cap A_k = \emptyset \forall j \neq k$, dann gilt

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \quad (\sigma\text{-Additivitat}). \quad (13.8)$$

Bemerkung: Es heit dann

- Ω Ergebnismenge oder Maraum
- $\omega \in \Omega$ Ergebnis oder Elementarereignis
- $A \in \mathcal{A}$ Ereignis
- $P(A)$ Wahrscheinlichkeit von A

Beispiel: idealer Wurfel, d.h. jede Augenzahl gleich wahrscheinlich

$$\Omega := \{1, 2, 3, 4, 5, 6\}, \quad \mathcal{A} := \mathcal{P}(\Omega), \quad P(A) := \frac{|A|}{6} \quad \forall A \in \mathcal{A} \quad (13.9)$$

$|A|$ ist Anzahl der Elemente von A

Satz 44. Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum, dann gilt

- (i) $P(\emptyset) = 0$
- (ii) $A, B \in \mathcal{A}, A \supseteq B \Rightarrow P(A \setminus B) = P(A) - P(B)$
- (iii) $A \in \mathcal{A} \Rightarrow P(A^C) = 1 - P(A)$
- (iv) $A, B \in \mathcal{A}, A \subseteq B \Rightarrow P(A) \leq P(B)$ (*Monotonie*)
- (v) A_1, \dots, A_n paarweise disjunkt $\Rightarrow P(A_1 \cup \dots \cup A_n) = \sum_{\nu=1}^n P(A_\nu)$
(*endliche Additivitat*)
- (vi) (A_n) Folge in $\mathcal{A} \Rightarrow P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n)$ (*Subadditivitat*)
- (vii) $A, B \in \mathcal{A} \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (*Additionsgesetz*)

Beweise: (nur ein paar...)

(iii)

$$\begin{aligned} P(A \cup A^C) &= P(\Omega) \stackrel{(b)}{=} 1 \\ &\stackrel{(c)}{=} P(A) + P(A^C) \\ \Rightarrow P(A^C) &= 1 - P(A) \quad \square \end{aligned} \quad (13.10)$$

(i)

$$P(\emptyset) \stackrel{(iii)}{=} 1 - P(\Omega) \stackrel{(b)}{=} 0 \quad \square \quad (13.11)$$

(vii)

$$\begin{aligned} P(A \cup B) &= P(A \cup (B \setminus A)) \\ &\stackrel{(c)}{=} P(A) + P(B \setminus A) + P(A \cap B) - P(A \cap B) \\ &= P(A) + P(\underbrace{(B \setminus A) \cup (A \cap B)}_{=B}) - P(A \cap B) \quad \square \end{aligned} \tag{13.12}$$

Weiteres Beispiel für einen Wahrscheinlichkeitsraum:

Sei $a < b$ und $I = [a, b]$ ein kompaktes Intervall. Für ein Teilintervall $A \subset I$ suche

$$P(A) := \text{Wahrscheinlichkeit, dass zufällig aus } I \text{ gewählte Zahl in } A \text{ liegt.} \tag{13.13}$$

Intuitiv liegt nahe: $P(A) := \frac{l(A)}{l(I)}$, wobei $l(I) = b - a$ "Länge von I ".

Konstruiere Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) :

- $\Omega := I$
- $\mathcal{A} :=$ eine σ -Algebra über Ω mit $A \in \mathcal{A} \forall$ Intervalle $A \subset I$
- $P : \mathcal{A} \rightarrow [0, 1]$ mit $P(A) := \frac{l(A)}{l(I)} \forall$ Intervalle $A \subset I$

Man kann zeigen:

Solche \mathcal{A} und P existieren, aber es ist $\mathcal{A} \neq \mathcal{P}(\Omega)$, wenn wir Translationsinvarianz verlangen, d.h.

$$\forall A \in \mathcal{A} \text{ und } \forall t \in \mathbb{R} \text{ mit } t + A \in \mathcal{A}: \quad P(t + A) = P(A), \tag{13.14}$$

wobei $t + A = \{\omega \in \Omega \mid \omega - t \in A\}$.

Definition: (Erzeugung von σ -Algebren)

Sei $\Omega \neq \emptyset$, $\mathcal{M} \subset \mathcal{P}(\Omega)$, dann heißt

$$\mathcal{A}^\sigma(\mathcal{M}, \Omega) := \bigcap_{\mathcal{B} \supset \mathcal{M}} \mathcal{B} \tag{13.15}$$

\mathcal{B} : σ -Algebra über Ω

die von \mathcal{M} über Ω erzeugte σ -Algebra (kleinste σ -Algebra, die \mathcal{M} enthält).

Beispiel: Borel- σ -Algebra über \mathbb{R}^n

$$\begin{aligned} \mathcal{M} &:= \text{Menge aller abgeschlossenen Quader } Q \subseteq \mathbb{R}^n \\ \mathcal{B}_n &:= \mathcal{A}^\sigma(\mathcal{M}, \mathbb{R}^n) \end{aligned} \tag{13.16}$$

Man kann zeigen: \mathcal{B}_n enthält alle offenen und abgeschlossenen Teilmengen des \mathbb{R}^n .

13.2 Kombinatorik

13.2.1 Laplacesche Wahrscheinlichkeitsräume

Beispiel: Beim idealen Würfeln sind alle Augenzahlen gleich wahrscheinlich.

Definition: Sei Ω eine endliche Ergebnismenge, dann heißt $(\Omega, \mathcal{P}(\Omega), P)$ mit

$$P(A) = \frac{|A|}{|\Omega|} \quad \forall A \subseteq \Omega \quad (13.17)$$

Laplacescher W-Raum.

Bemerkungen:

1. $P(A) = \frac{|A|}{|\Omega|} = \frac{\text{Anzahl der günstigen Ausgänge}}{\text{Anzahl der möglichen Ausgänge}}$
2. In einem Laplaceschen W-Raum sind alle Elementarereignisse gleich wahrscheinlich, d.h. $\omega \in \Omega: P(\{\omega\}) = \frac{1}{|\Omega|}$

Beispiel:

Zweimaliges Werfen eines idealen Würfels.

$A =$ "Summe der Augenzahlen ist = 5". $P(A) = ?$

$$\begin{aligned} \Omega &= \{1, 2, 3, 4, 5, 6\}^2 := \{1, \dots, 6\} \times \{1, \dots, 6\} \\ &:= \{(\omega_1, \omega_2) \mid \omega_j \in \{1, \dots, 6\}, j = 1, 2\} \end{aligned} \quad (13.18)$$

ω_1 (ω_2) ist Ergebnis des ersten (zweiten) Wurfs.

$$|\Omega| = 36$$

$$\begin{aligned} A &= \{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 5\} \\ &= \{(1, 4), (2, 3), (3, 2), (4, 1)\} \quad \Rightarrow \quad P(A) = \frac{4}{36} = \frac{1}{9} \end{aligned} \quad (13.19)$$

Führt auf Abzählprobleme – ein paar Formeln:

13.2.2 Permutationen

- Anzahl möglicher Anordnungen von n (unterschiedlichen) Objekten:

$$n(n-1) \cdots 1 = n!$$

n mögliche Kugeln für die erste Position, $(n-1)$ für die zweite...

Beispiel:

$n = 3$ Kugeln, $\bullet, \color{red}\bullet, \color{blue}\bullet$, lassen sich auf $3! = 6$ verschiedene Möglichkeiten anordnen:

$$(\color{red}\bullet \color{blue}\bullet \bullet) (\color{red}\bullet \bullet \color{blue}\bullet) (\color{blue}\bullet \color{red}\bullet \bullet) (\color{blue}\bullet \bullet \color{red}\bullet) (\bullet \color{red}\bullet \color{blue}\bullet) (\bullet \color{blue}\bullet \color{red}\bullet)$$

- Anzahl möglicher Anordnungen von n Objekten, die in j Gruppen mit jeweils k_1, k_2, \dots, k_j identischen Objekten vorkommen ($k_1 + k_2 + \dots + k_j = n$):

$$\frac{n!}{k_1! k_2! \cdots k_j!}$$

$n!$ mögliche Anordnungen der n Objekte dividiert durch die Anzahl $k_1!$ der möglichen Anordnungen der Objekte vom Typ 1 innerhalb einer gegebenen Anordnung. . .

Beispiel: $k_1 = 2$ rote Kugeln und $k_2 = 1$ blaue Kugel, $\bullet\bullet\bullet$, lassen sich auf $\frac{3!}{2!1!} = 3$ verschiedene Möglichkeiten anordnen:

$$(\bullet\bullet\bullet) (\bullet\bullet\bullet) (\bullet\bullet\bullet)$$

13.2.3 Urnenmodelle

- Urne mit n unterschiedlichen Kugeln
- Ziehe k Kugeln

Anzahl der möglichen Ergebnisse: (Beispiel: Urne mit $n = 3$ Kugeln, ziehe $k = 2$ -mal)

	mit Zurücklegen	ohne Zurücklegen
mit Beachtung der Reihenfolge	n^k $3^2 = 9$: $(\bullet\bullet) (\bullet\bullet) (\bullet\bullet)$ $(\bullet\bullet) (\bullet\bullet) (\bullet\bullet)$ $(\bullet\bullet) (\bullet\bullet) (\bullet\bullet)$	$\frac{n!}{(n-k)!}$ $\frac{3!}{1!} = 6$: $(\bullet\bullet) (\bullet\bullet)$ $(\bullet\bullet) (\bullet\bullet)$ $(\bullet\bullet) (\bullet\bullet)$
ohne Beachtung der Reihenfolge	$\binom{n+k-1}{k}$ $\binom{4}{2} = 6$: $(\bullet\bullet) (\bullet\bullet) (\bullet\bullet)$ $(\bullet\bullet) (\bullet\bullet)$ $(\bullet\bullet)$	$\frac{n!}{(n-k)! k!} =: \binom{n}{k}$ $\binom{3}{2} = 3$: $(\bullet\bullet) (\bullet\bullet)$ $(\bullet\bullet)$

Zum Ziehen mit Zurücklegen ohne Beachtung der Reihenfolge:

- Die Stichprobe enthält k Kugeln, davon sind jeweils k_j vom Typ j , $j = 1, \dots, n$, $\sum_{j=1}^n k_j = k$.
- Wir legen diese in n aufgereichte Eimer, jeweils die Kugeln vom Typ j in den j -ten Eimer; z.B. so: $\boxed{\bullet\bullet\bullet} \mid \mid \bullet \mid \bullet\bullet \mid \mid$ für $k=6, n=5$.
- Die Stichprobe ist jetzt eindeutig durch die Angabe der Positionen der Zwischenwände charakterisiert, d.h. wir müssen k Kugeln und $n - 1$ Zwischenwände anordnen; Anzahl der Möglichkeiten (vgl. Abschnitt 13.2.2)

$$\frac{(k+n-1)!}{k!(n-1)!} = \binom{n+k-1}{k}$$

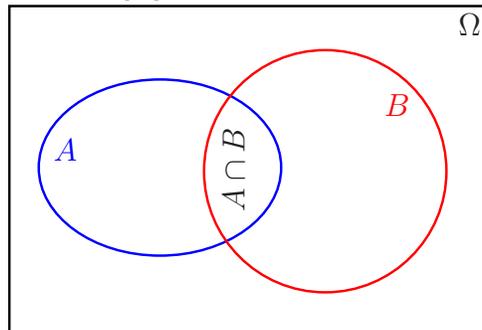
13.3 Bedingte Wahrscheinlichkeiten

Definition: (Bedingte Wahrscheinlichkeit)

Sei (Ω, \mathcal{A}, P) ein W-Raum, $A, B \in \mathcal{A}$ mit $P(B) > 0$, dann heißt

$$P(A|B) := \frac{P(A \cap B)}{P(B)} \quad (13.20)$$

bedingte Wahrscheinlichkeit von A gegeben B .



Bemerkungen:

1. Mit $\mathcal{A}_B := \{A \in \mathcal{A} \mid A \subseteq B\}$ und $P_B(A) := \frac{P(A)}{P(B)}$ für $A \in \mathcal{A}_B$ ist (B, \mathcal{A}_B, P_B) ein W-Raum.¹⁶
2. Für $P(B) = 0$ definieren wir $P(A|B) := 0 \forall A \in \mathcal{A}$.
3. $P(A \cap B) = P(A|B) P(B) \forall A \in \mathcal{A}$ (auch falls $P(B) = 0$)
4. Für $A_1, \dots, A_n \in \mathcal{A}$ mit $P(A_1 \cap \dots \cap A_n) > 0$ gilt

$$P(A_1 \cap \dots \cap A_n) = P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \dots \cap A_{n-1}) \quad (13.21)$$

(Beweis mit vollständiger Induktion nach n)

Beispiel: Wahrscheinlichkeit, dass beim Skat jeder der drei Spieler genau ein Ass ausgeteilt bekommt: Sei $A_k :=$ "Spieler k bekommt genau ein Ass ausgeteilt".

$$\begin{aligned} P(A_1) &= \frac{\binom{4}{1} \binom{28}{9}}{\binom{32}{10}} = \frac{385}{899} \\ P(A_2|A_1) &= \frac{\binom{3}{1} \binom{19}{9}}{\binom{22}{10}} = \frac{3}{7} \\ P(A_3|A_1 \cap A_2) &= \frac{\binom{2}{1} \binom{10}{9}}{\binom{12}{10}} = \frac{10}{33} \end{aligned} \quad (13.22)$$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) = \frac{50}{899} \approx 5,6\%$$

¹⁶Da B nun die Ergebnismenge ist, muss auch des Komplement bezüglich B gebildet werden, d.h. wir definieren $A^C := B \setminus A$.

Definition: (Stochastische Unabhängigkeit)

Sei (Ω, \mathcal{A}, P) ein W-Raum.

(i) Zwei Ereignisse A und B heißen (stochastisch) unabhängig, falls gilt

$$P(A \cap B) = P(A) P(B). \quad (13.23)$$

(ii) Endlich oder abzählbar viele Ereignisse $A_i \in \mathcal{A}$ heißen (stochastisch) unabhängig, falls gilt

$$P(A_{i_1} \cap \dots \cap A_{i_r}) = P(A_{i_1}) \cdots P(A_{i_r}) \quad \forall i_1, \dots, i_r \text{ mit } 1 \leq i_1 < \dots < i_r \quad (13.24)$$

Bemerkung: Falls A und B unabhängig und $P(B) \neq 0 \Rightarrow P(A|B) = P(A)$.

Satz 45. Sei (Ω, \mathcal{A}, P) ein W-Raum und B_1, B_2, \dots eine endliche oder abzählbare Folge paarweise disjunkter Ereignisse, mit $P(B_k) > 0 \forall k$ und $\bigcup_k B_k = \Omega$. Dann gilt

(i) Formel von der totalen Wahrscheinlichkeit:

$$P(A) = \sum_k P(A|B_k) P(B_k) \quad \forall A \in \mathcal{A}, \quad (13.25)$$

(ii) Satz von Bayes:

$$P(B_k|A) = \frac{P(A|B_k) P(B_k)}{\sum_j P(A|B_j) P(B_j)} \quad \forall k \text{ und } \forall A \in \mathcal{A} \text{ mit } P(A) > 0. \quad (13.26)$$

Beweis:

(i)

$$\begin{aligned} P(A) &= P(A \cap \Omega) = P\left(A \cap \bigcup_k B_k\right) \\ &= P\left(\bigcup_k \underbrace{(A \cap B_k)}_{\subseteq B_k \text{ und damit disjunkt}}\right) \\ &= \sum_k P(A \cap B_k) = \sum_k P(A|B_k) P(B_k) \end{aligned} \quad (13.27)$$

(ii)

$$P(B_k|A) = \frac{P(B_k \cap A)}{P(A)} = \frac{P(A|B_k) P(B_k)}{P(A)} \stackrel{(i)}{=} \frac{P(A|B_k) P(B_k)}{\sum_j P(A|B_j) P(B_j)} \quad (13.28)$$

Beispiel: Gefangenenparadoxon

- In einem Gefängnis sitzen drei zum Tode verurteilte Gefangene: Anton, Brigitte und Clemens.

- Genau einer von ihnen soll begnadigt werden. Dazu wird ein Los gezogen, das allen die gleiche Chance gibt, begnadigt zu werden.
- Anton, der also eine Überlebenswahrscheinlichkeit von $\frac{1}{3}$ hat, bittet den Wärter, der das Ergebnis des Losentscheids kennt, ihm einen seiner Leidensgenossen, Brigitte oder Clemens, zu nennen, der oder die sterben muss.
- Der Wärter antwortet “Brigitte”.

Wie hoch ist nun Antons Überlebenswahrscheinlichkeit?

$$\Omega = \{A, B, C\} \times \{B, C\} = \{\omega = (\omega_1, \omega_2) \mid \omega_1 \in \{A, B, C\}, \omega_2 \in \{A, B\}\} \quad (13.29)$$

in ω_1 steht, wer begnadigt wird, in ω_2 wer vom Wärter genannt wird

$$A_b := \{\omega \in \Omega \mid \omega_1 = A\} = \text{“A(nton) wird begnadigt”} \quad (13.30)$$

analog B_b und C_b .

$$B_g := \{\omega \in \Omega \mid \omega_2 = B\} = \text{“B(ritigte) wird genannt”} \quad (13.31)$$

analog C_g . Wir wissen:

$$A_b \cup B_b \cup C_b = \Omega = B_g \cup C_g \quad (13.32)$$

$$P(A_b) = P(B_b) = P(C_b) = \frac{1}{3} \quad (13.33)$$

$$P(C_g|B_b) = 1 = P(B_g|C_b) \quad (13.34)$$

$$P(C_g|C_b) = 0 = P(B_g|B_b)$$

$$P(C_g|A_b) = \frac{1}{2} = P(B_g|A_b)$$

Daraus folgt

$$\begin{aligned} P(A_b|B_g) &= \frac{P(B_g|A_b) P(A_b)}{P(B_g|A_b) P(A_b) + P(B_g|B_b) P(B_b) + P(B_g|C_b) P(C_b)} \\ &= \frac{\frac{1}{2} \cdot \frac{1}{3}}{\frac{1}{2} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3}} = \frac{\frac{1}{6}}{\frac{1}{6} + \frac{1}{3}} = \frac{1}{3}, \end{aligned} \quad (13.35)$$

und analog (bzw. als Komplement) $P(C_b|B_g) = \frac{2}{3}$.

13.4 Zufallsvariable & Verteilungsfunktionen

Oft interessieren wir uns bei der Auswertung eines Zufallsexperiments auf (Ω, \mathcal{A}, P) nicht für die Ereignisse ω selbst, sondern für eine zugeordnete Größe $X(\omega)$

Beispiel: Zweifaches Würfeln.

$$\Omega = \{(\omega_1, \omega_2) \mid \omega_1, \omega_2 \in \{1, \dots, 6\}\} \quad (13.36)$$

$X(\omega) = \omega_1 + \omega_2$, Augensumme aus beiden Würfeln.

Definition: Sei (Ω, \mathcal{A}, P) ein W-Raum.

Eine Funktion $X : \Omega \rightarrow \mathbb{R}$ heißt (reelle) Zufallsvariable (ZV), wenn gilt

$$X^{-1}((-\infty, x]) = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{A} \quad \forall x \in \mathbb{R}. \quad (13.37)$$

...ist also immer ein Ereignis

Wir schreiben

- für $\{\omega \in \Omega \mid X(\omega) \leq x\}$ kurz $X \leq x$ und
- für $P(\{\omega \in \Omega \mid X(\omega) \leq x\})$ kurz $P(X \leq x)$.

Bemerkungen:

1. Eine ZV X heißt diskret, falls die Wertemenge $X(\Omega)$ diskret, d.h. endlich oder abzählbar, ist.
2. Ist ein W-Raum (Ω, \mathcal{A}, P) diskret (d.h. Ω höchstens abzählbar), so ist jede ZV $X : \Omega \rightarrow \mathbb{R}$ diskret.
3. Sind X und Y ZV auf (Ω, \mathcal{A}, P) , dann auch $X+Y$, $X \cdot Y$ und jede Linearkombination von endlich vielen ZV.

Definition: Sei (Ω, \mathcal{A}, P) ein W-Raum und $X : \Omega \rightarrow \mathbb{R}$ eine ZV, dann heißt

$$F_X : \mathbb{R} \rightarrow [0, 1] \quad (13.38)$$

$$x \mapsto P(X \leq x)$$

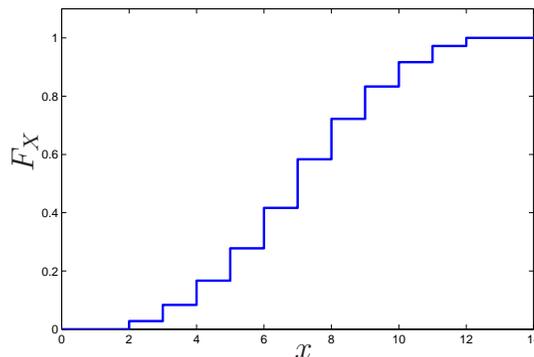
Verteilungsfunktion von X .

Beispiel: Zweifaches Würfeln (siehe oben).

$$F_X(x) = P(X \leq x) = \sum_{k \leq x} P(X=k) \quad (13.39)$$

k	2	3	4	5	6	7	8	9	10	11	12
$P(X=k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Wird gerne so geplottet:



...aber eigentlich ohne senkrechte Linien, d.h. .

Eigenschaften:

- (i) F_X ist monoton wachsend.
- (ii) F_X ist in jedem Punkt $x_0 \in \mathbb{R}$ rechtsseitig stetig.
- (iii) $\lim_{x \rightarrow \infty} F_X(x) = 1$ und $\lim_{x \rightarrow -\infty} F_X(x) = 0$
- (iv) $P(a < X \leq b) = F_X(b) - F_X(a)$ und
 $P(X = a) = F_X(a) - \lim_{\varepsilon \rightarrow 0^+} F_X(a - \varepsilon)$...linksseitiger Grenzwert

Beispiele für Verteilungen:

1. Binomialverteilung

Ein Zufallsexperiment habe 2 mögliche Ausgänge: (E)rfolg oder (M)isserfolg, d.h.

$$\Omega = \{E, M\}, \quad P(E) = p \in [0, 1], \quad P(M) = 1 - p, \quad (13.40)$$

also W-Raum $(\Omega, \mathcal{P}(\Omega), P)$. Experiment werde n -mal wiederholt, d.h.

$$\begin{aligned} \tilde{\Omega} = \Omega^n &= \{\omega = (\omega_1, \dots, \omega_n) \mid \omega_i \in \Omega, i = 1, \dots, n\} \\ \tilde{P}(\{\omega\}) &= p^k (1 - p)^{n-k}, \quad \text{wobei } k \text{ die Anzahl der } \omega_i \text{ mit } \omega_i = E \text{ ist,} \end{aligned} \quad (13.41)$$

also W-Raum $(\tilde{\Omega}, \mathcal{P}(\tilde{\Omega}), \tilde{P})$. Betrachte

$$X(\omega) = |\{\omega_i \in \omega \mid \omega_i = E\}| = k. \quad (13.42)$$

$X(\tilde{\Omega}) = \{0, \dots, n\}$, d.h. X ist diskrete Zufallsvariable, und $X(\omega)$ zählt, wie oft E in ω auftritt. $\tilde{P}(X=k) = ?$

$$\begin{aligned} |(X=k)| &\equiv |\{\omega \in \tilde{\Omega} \mid X(\omega) = k\}| \\ &= \# \text{ Möglichkeiten, } k \text{ Es und } (n - k) \text{ Ms anzuordnen} \\ &= \frac{n!}{k!(n - k)!} = \binom{n}{k} \end{aligned} \quad (13.43)$$

Damit gilt

$$\tilde{P}(X=k) = \binom{n}{k} p^k (1 - p)^{n-k} =: b(k; n, p), \quad (13.44)$$

und weiter

$$F_X(x) = \sum_{k \leq x} \binom{n}{k} p^k (1 - p)^{n-k} =: B(x; n, p). \quad (13.45)$$

Wir schreiben

$$X \sim \text{Bin}(n, p) \quad (13.46)$$

und sagen, X ist binomialverteilt mit Parametern n und p .

Normierung:

$$\sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1 \quad (13.47)$$

Beispiel:

$X := \# \text{ 6 in 10 W\u00fcrfen eines fairen W\u00fcrfels} \Rightarrow X \sim \text{Bin}(10, \frac{1}{6})$

2. Poissonverteilung

Wir schreiben

$$X \sim \text{Pois}(\lambda) \tag{13.48}$$

und sagen X ist Poisson-verteilt mit Parameter λ , falls gilt $X(\Omega) = \mathbb{N}_0$ und

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k \in \mathbb{N}_0. \tag{13.49}$$

Verteilungsfunktion:

$$F_X(x) = \sum_{k \leq x} \frac{\lambda^k}{k!} e^{-\lambda} \tag{13.50}$$

Normierung

$$\begin{aligned} \sum_{k=0}^{\infty} P(X=k) &= \lim_{x \rightarrow \infty} F_X(x) \\ &= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1 \end{aligned} \tag{13.51}$$

Definition: Eine ZV X mit Verteilungsfunktion F_X hat die (Verteilungs-)Dichte f_X , wenn eine nichtnegative, integrierbare Funktion f_X existiert mit

$$F_X(x) = \int_{-\infty}^x f_X(t) dt. \tag{13.52}$$

Folgerungen:

- (i) F_X ist stetig auf \mathbb{R} .
- (ii) $F'_X(x) = f_X(x)$ falls f_X stetig.
- (iii) $\int_{-\infty}^{+\infty} f_X(t) dt = 1$ (Normierung)
- (iv) $P(X=a) = 0 \quad \forall a \in \mathbb{R}$

Beispiele f\u00fcr Verteilungen (Forts.):**3. Gleichverteilung**

Wir sagen X ist gleichverteilt auf $[a, b]$, wenn gilt

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{f\u00fcr } x \in [a, b] \\ 0 & \text{sonst} \end{cases}. \tag{13.53}$$

Verteilungsfunktion

$$F_X(x) = \begin{cases} 0 & , \quad x \leq a \\ \frac{x-a}{b-a} & , \quad a \leq x \leq b \\ 1 & , \quad x \geq b \end{cases} \quad (13.54)$$

4. Standardnormalverteilung

Wir schreiben $X \sim \mathcal{N}(0, 1)$ und sagen X ist standardnormalverteilt, wenn gilt

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad (13.55)$$

Normierung: Siehe Beispiel für Substitutionsregel.

Verteilungsfunktion bekommt eigenen Namen,

$$\Phi(x) := F_X(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt \quad (13.56)$$

und ist tabelliert.

13.4.1 Kenngrößen von Zufallsvariablen

Definition Sei X eine ZV auf (Ω, \mathcal{A}, P) .

- (i) Ist X diskret mit Wertebereich $X(\Omega) = \{x_1, x_2, \dots\}$ und gilt $\sum_{x_i \in X(\Omega)} |x_i| P(X=x_i) < \infty$, dann nennen wir

$$E(X) := \sum_{x_i \in X(\Omega)} x_i P(X=x_i) \quad (13.57)$$

den Erwartungswert von X .

- (ii) Besitzt X die Dichte f_X mit $\int_{-\infty}^{+\infty} |t| f_X(t) dt < \infty$, so nennen wir

$$E(X) := \int_{-\infty}^{+\infty} t f_X(t) dt \quad (13.58)$$

den Erwartungswert von X .

Beispiele:

1. Faires Würfeln: $\Omega = \{\square, \blacksquare, \boxtimes, \boxplus, \boxminus, \boxdot\}$; $P(\{\omega\}) = \frac{1}{6}$, $\omega = \square, \dots, \boxdot$
 $X(\omega) :=$ Augenzahl, diskret mit $X(\Omega) = \{1, \dots, 6\}$,

$$E(X) = \sum_{k=1}^6 k P(X=k) = \frac{1}{6}(1 + 2 + \dots + 6) = 3,5. \quad (13.59)$$

2. $X \sim \text{Bin}(n, p)$

$$\begin{aligned}
 E(X) &= \sum_{k=0}^n k P(X=k) \\
 &= \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} \\
 &= \sum_{k=1}^n n \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
 &= n \sum_{k=0}^{n-1} \binom{n-1}{k} p^{k+1} (1-p)^{n-(k+1)} \\
 &= np \underbrace{\sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k}}_{=1}
 \end{aligned} \tag{13.60}$$

3. $X \sim \text{Pois}(\lambda)$

$$E(X) = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = e^{-\lambda} \lambda \underbrace{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}_{=e^\lambda} = \lambda \tag{13.61}$$

4. $X \sim \mathcal{N}(0, 1)$

$$E(X) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x e^{-\frac{x^2}{2}} dx = 0 \quad (\text{Symmetrie}) \tag{13.62}$$

5. X Cauchy-verteilt: $f_X(t) = \frac{1}{\pi(1+t^2)}$

$E(X)$ existiert nicht, da $\int_{-\infty}^{+\infty} \frac{|t|}{\pi(1+t^2)} dt$ divergent

Bem: Aber normiert, $\int_{-\infty}^{+\infty} \frac{1}{\pi(1+t^2)} dt = \frac{1}{\pi} (\arctan(\infty) - \arctan(-\infty)) = 1$.

Satz 46. Sei X eine ZV auf (Ω, \mathcal{A}, P) , für die $E(X)$ existiert. Dann gilt $\forall a, b \in \mathbb{R}$

$$E(aX + b) = aE(X) + b. \tag{13.63}$$

Beweis: Durch Nachrechnen.

Definition: Sei X eine ZV mit Erwartungswert $\mu := E(X)$. Dann heißt

- (i) $\text{Var}(X) := E((X - \mu)^2)$ Varianz von X , falls $E(X^2)$ existiert, und $\sigma_X := \sqrt{\text{Var}(X)}$ Standardabweichung von X .

(ii) $\mu_k := E(X^k)$, $k \in \mathbb{N}$, k tes Moment von X , falls $E(|X^k|)$ existiert.

(iii) X standardisiert $\stackrel{\text{Def.}}{\iff} E(X) = 0$ und $\text{Var}(X) = 1$.

Bemerkung: Allgemein kann man für stetige Funktionen $\phi : \mathbb{R} \rightarrow \mathbb{R}$ zeigen, dass, falls $E(|\phi(X)|)$ existiert, gilt:

$$\begin{aligned}
 E(\phi(X)) &= \sum_{x_i \in X(\Omega)} \phi(x_i) P(X = x_i), & \text{falls } X \text{ diskret,} \\
 E(\phi(X)) &= \int_{-\infty}^{+\infty} \phi(t) f_X(t) dt, & \text{falls } X \text{ die Dichte } f_X \text{ besitzt.}
 \end{aligned}
 \tag{13.64}$$

Beispiele:

1. X diskret mit $X(\Omega) = \{x_1, \dots, x_n\}$, $P(X=x_i) = \frac{1}{n} \forall i = 1, \dots, n$:

$$\begin{aligned}
 E(X) &= \sum_{i=1}^n x_i P(X=x_i) \\
 &= \frac{1}{n} \sum_{i=1}^n x_i =: \mu, \\
 \text{Var}(X) &= \sum_{i=1}^n (x_i - \mu)^2 P(X=x_i) \\
 &= \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2.
 \end{aligned}
 \tag{13.65}$$

2. Faires Würfeln: $E(X) = \frac{7}{2} =: \mu$ (s.o.)

$$\begin{aligned}
 \text{Var}(X) &= \sum_{k=1}^6 (k - \mu)^2 P(X=k) = \frac{(\frac{5}{2})^2 + (\frac{3}{2})^2 + (\frac{1}{2})^2 + (\frac{1}{2})^2 + (\frac{3}{2})^2 + (\frac{5}{2})^2}{6} \\
 &= \frac{2^{\frac{25+9+1}{4}}}{6} = \frac{35}{12} \approx 2,9 \\
 \sigma_X &\approx 1,7
 \end{aligned}
 \tag{13.66}$$

3. Wir schreiben $X \sim \mathcal{N}(\mu, \sigma^2)$ und sagen X ist normalverteilt mit Parametern μ und σ^2 , wenn

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.
 \tag{13.67}$$

Dann gilt

$$E(X) = \mu \quad \text{und} \quad \text{Var}(X) = \sigma^2,
 \tag{13.68}$$

Denn mit $z = \frac{x-\mu}{\sigma}$, $dz = \frac{dx}{\sigma}$, folgt

$$\begin{aligned}
 \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 1 \quad (\text{Normierung - s.o.}), \\
 E(X) &= \int_{-\infty}^{\infty} \frac{x}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{\sigma z + \mu}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \mu, \\
 \text{Var}(X) &= \int_{-\infty}^{\infty} \frac{(x-\mu)^2}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} \frac{\sigma^2 z^2}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
 &\stackrel{\text{p.I.}}{=} \sigma^2 \left(\underbrace{\left[z \cdot \frac{(-1)}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \right]_{-\infty}^{\infty}}_{=0} - \underbrace{\int_{-\infty}^{\infty} \frac{(-1)}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz}_{=1} \right) = \sigma^2.
 \end{aligned} \tag{13.69}$$

Satz 47. Sei X eine ZV. Dann gilt – falls die auftretenden Größen existieren –

- (i) $\text{Var}(aX + b) = a^2 \text{Var}(X) \quad \forall a, b \in \mathbb{R}$
- (ii) $\text{Var}(X) = E(X^2) - [E(X)]^2$

Beweis: (mit Satz 46)

(i) $E(aX + b) = aE(X) + b =: a\mu + b$

$$\begin{aligned}
 \text{Var}(aX + b) &= E\left(\left(aX + b - (a\mu + b)\right)^2\right) \\
 &= E\left(a^2(x - \mu)^2\right) \\
 &= a^2 \text{Var}(X) \quad \square
 \end{aligned} \tag{13.70}$$

(ii) $E(X) =: \mu$

$$\begin{aligned}
 \text{Var}(X) &= E\left((X - \mu)^2\right) \\
 &= E\left(X^2 - 2X\mu + \mu^2\right) \\
 &= E(X^2) - 2\mu \underbrace{E(X)}_{=\mu} + \mu^2 \\
 &= E(X^2) - \mu^2 \quad \square
 \end{aligned} \tag{13.71}$$

Bemerkung: Ist X eine ZV mit Erwartungswert μ und Varianz $\sigma^2 > 0$, so ist $\frac{X - \mu}{\sigma}$ standardisiert.

Satz 48. (Tschebyscheff'sche Ungleichung)

Sei X eine ZV, für die $\mu := E(X)$ und $\sigma^2 := \text{Var}(X)$ existieren. Dann gilt

$$P(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2} \quad \forall a > 0. \tag{13.72}$$

Bemerkung: Mit $a = \lambda\sigma$ (alle positiv) gilt auch

$$P(|X - \mu| \geq \lambda\sigma) \leq \frac{1}{\lambda^2} \quad \text{bzw.} \quad P(|X - \mu| < \lambda\sigma) \geq 1 - \frac{1}{\lambda^2}, \quad (13.73)$$

z.B.

$$\begin{aligned} P(|X - \mu| \geq 2\sigma) &\leq \frac{1}{4} = 25\% & \text{bzw.} & \quad P(|X - \mu| < 2\sigma) \geq 75\% & \text{und} \\ P(|X - \mu| \geq 3\sigma) &\leq \frac{1}{9} \approx 11\% & \text{bzw.} & \quad P(|X - \mu| < 3\sigma) \geq 89\%. \end{aligned} \quad (13.74)$$

Große Abweichungen vom Erwartungswert (gemessen in Einheiten der Standardabweichung) sind also unwahrscheinlich. Daher ist die Varianz bzw. die Standardabweichung ein Maß für die Schwankungen einer ZV um ihren Erwartungswert.

Beweis:

1. X diskret:

$$\begin{aligned} \sigma^2 &= E((X - \mu)^2) = \sum_{x_i \in X(\Omega)} (x_i - \mu)^2 P(X=x_i) \\ &\geq \sum_{\substack{x_i \in X(\Omega) \\ |x_i - \mu| \geq a}} (x_i - \mu)^2 P(X=x_i) \\ &\geq a^2 \sum_{\substack{x_i \in X(\Omega) \\ |x_i - \mu| \geq a}} P(X=x_i) = a^2 P(|X - \mu| \geq a) \end{aligned} \quad (13.75)$$

2. X kontinuierlich mit Dichte f_X :

$$\begin{aligned} \sigma^2 &= E((X - \mu)^2) = \int_{-\infty}^{+\infty} (t - \mu)^2 f_X(t) dt \\ &\geq \int_{-\infty}^{\mu-a} (t - \mu)^2 f_X(t) dt + \int_{\mu+a}^{\infty} (t - \mu)^2 f_X(t) dt \\ &\geq a^2 \underbrace{\left(\int_{-\infty}^{\mu-a} f_X(t) dt + \int_{\mu+a}^{\infty} f_X(t) dt \right)}_{=P(|X-\mu| \geq a)} \end{aligned} \quad (13.76)$$

□

Definition: (Stochastische Unabhängigkeit von ZV)

Seien $X_k : \Omega \rightarrow \mathbb{R}$, $k = 1, \dots, n$, ZV auf (Ω, \mathcal{A}, P) . X_1, \dots, X_n heißen (stochastisch) unabhängig, wenn für jedes $(a_1, \dots, a_n) \in \mathbb{R}^n$ die Ereignisse $(X_1 \leq a_1), \dots, (X_n \leq a_n)$ stochastisch unabhängig sind.

Folgerungen: Falls X_1, \dots, X_n stochastisch unabhängig sind, gilt $\forall i, j = 1, \dots, n$

- (i) $E(X_i X_j) = E(X_i) E(X_j)$ und
(ii) $\text{Var}(X_i + X_j) = \text{Var}(X_i) + \text{Var}(X_j)$.

Bemerkung: $E(X_i + X_j) = E(X_i) + E(X_j)$ gilt sogar, falls X_i und X_j nicht stochastisch unabhängig sind.

Beweis:

- (i) falls ZV diskret:

$$\begin{aligned}
E(X_i X_j) &= \sum_{x_k \in X_i(\Omega)} \sum_{x_l \in X_j(\Omega)} x_k x_l P((X_i=x_k) \cap (X_j=x_l)) \\
&\stackrel{\text{stoch. unabh.}}{=} \sum_{x_k \in X_i(\Omega)} \sum_{x_l \in X_j(\Omega)} x_k x_l P(X_i=x_k) P(X_j=x_l) \\
&= \underbrace{\sum_{x_k \in X_i(\Omega)} x_k P(X_i=x_k)}_{=E(X_i)} \underbrace{\sum_{x_l \in X_j(\Omega)} x_l P(X_j=x_l)}_{=E(X_j)}
\end{aligned} \tag{13.77}$$

analog für stetige ZV – allerdings müssen wir dazu zunächst das Konzept der gemeinsamen Dichte für mehrere ZV einführen (analog zum Konzept der Dichte für eine ZV):

$$\begin{aligned}
E(X_i X_j) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_i x_j f_{X_i X_j}(x_i, x_j) dx_i dx_j \\
&\stackrel{\text{stoch. unabh.}}{=} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_i x_j f_{X_i}(x_i) f_{X_j}(x_j) dx_i dx_j \\
&= \underbrace{\int_{-\infty}^{+\infty} x_i f_{X_i}(x_i) dx_i}_{=E(X_i)} \underbrace{\int_{-\infty}^{+\infty} x_j f_{X_j}(x_j) dx_j}_{=E(X_j)}
\end{aligned} \tag{13.78}$$

- (ii) $E(X_i) =: \mu_i$

$$\begin{aligned}
&\text{Var}(X_i + X_j) \\
&= E\left(\left((X_i + X_j) - (\mu_i + \mu_j)\right)^2\right) \\
&= E\left(\left(X_i - \mu_i\right)^2 + \left(X_j - \mu_j\right)^2 + 2(X_i - \mu_i)(X_j - \mu_j)\right) \\
&= \underbrace{E\left(\left(X_i - \mu_i\right)^2\right)}_{=\text{Var}(X_i)} + \underbrace{E\left(\left(X_j - \mu_j\right)^2\right)}_{=\text{Var}(X_j)} + 2 \underbrace{\left(\underbrace{E(X_i X_j)}_{\substack{=E(X_i)E(X_j) \\ (i)}} - E(X_i)\mu_j - \mu_i E(X_j) + \mu_i \mu_j \right)}_{=0}
\end{aligned} \tag{13.79}$$

□

Lemma 49. Seien X_1, \dots, X_n unabhängig normalverteilt mit $E(X_j) = \mu_j$ und $\text{Var}(X_j) = \sigma_j^2$, d.h. $X_j \sim \mathcal{N}(\mu_j, \sigma_j^2)$. Dann gilt für $Y := \sum_{j=1}^n X_j$:

$$Y \sim \mathcal{N}(\sum_{j=1}^n \mu_j, \sum_{j=1}^n \sigma_j^2) \quad (13.80)$$

Beweis: für $n = 2$ (Rest durch Induktion)

Verteilungsfunktion von Y :

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^y f_Y(t) dt = P(Y \leq y) = P(X_1 + X_2 \leq y) = P(X_2 \leq y - X_1) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{y-x_1} f_{X_2}(x_2) dx_2 f_{X_1}(x_1) dx_1 && \begin{aligned} z_j &= \frac{x_j - \mu_j}{\sigma_j} \\ dx_j &= \sigma_j dz_j \end{aligned} \\ &= \sigma_1 \sigma_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{y - \sigma_1 z_1 - \mu_1 - \mu_2}{\sigma_2}} f_{X_2}(\sigma_2 z_2 + \mu_2) dz_2 f_{X_1}(\sigma_1 z_1 + \mu_1) dz_1 && (13.81) \\ & && z = z_2 + \frac{\sigma_1}{\sigma_2} z_1, \quad dz = dz_2 \end{aligned}$$

$$\begin{aligned} &= \sigma_1 \sigma_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{y - \mu_1 - \mu_2}{\sigma_2}} f_{X_2}(\sigma_2 z - \sigma_1 z_1 + \mu_2) dz f_{X_1}(\sigma_1 z_1 + \mu_1) dz_1 \\ &= \sigma_1 \sigma_2 \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{y - \mu_1 - \mu_2}{\sigma_2}} \frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{\sigma_2^2 z^2 - 2\sigma_1\sigma_2 z z_1 + \sigma_1^2 z_1^2}{2\sigma_2^2} - \frac{z_1^2}{2}} dz dz_1 \end{aligned}$$

$$\begin{aligned} \text{Exponent: } & -\frac{1}{2\sigma_2^2} (\sigma_2^2 z^2 - 2\sigma_1\sigma_2 z z_1 + \sigma_1^2 z_1^2 + \sigma_2^2 z_1^2) \\ &= -\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_2^2} \left(\frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} z^2 - 2\frac{\sigma_1\sigma_2 z}{\sigma_1^2 + \sigma_2^2} z_1 + z_1^2 \right) \\ &= -\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_2^2} \left(\left(z_1 - \frac{\sigma_1\sigma_2 z}{\sigma_1^2 + \sigma_2^2} \right)^2 - \underbrace{\frac{\sigma_1^2 \sigma_2^2}{(\sigma_1^2 + \sigma_2^2)^2} z^2 + \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} z^2}_{=\frac{\sigma_2^4 z^2}{(\sigma_1^2 + \sigma_2^2)^2}} \right) && (13.82) \\ &= -\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_2^2} \left(z_1 - \frac{\sigma_1\sigma_2 z}{\sigma_1^2 + \sigma_2^2} \right)^2 - \frac{\sigma_2^2 z^2}{2(\sigma_1^2 + \sigma_2^2)} \end{aligned}$$

einsetzen, z_1 -Integral ausführen – liefert $\sqrt{\frac{2\pi\sigma_2^2}{\sigma_1^2 + \sigma_2^2}}$

$$\begin{aligned} F_Y(y) &= \int_{-\infty}^{\frac{y - \mu_1 - \mu_2}{\sigma_2}} \frac{1}{2\pi} \sqrt{\frac{2\pi\sigma_2^2}{\sigma_1^2 + \sigma_2^2}} e^{-\frac{\sigma_2^2 z^2}{2(\sigma_1^2 + \sigma_2^2)}} dz && \begin{aligned} t &= \sigma_2 z + \mu_1 + \mu_2 \\ dz &= \frac{1}{\sigma_2} dt \end{aligned} \\ &= \int_{-\infty}^y \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} e^{-\frac{(t - (\mu_1 + \mu_2))^2}{2(\sigma_1^2 + \sigma_2^2)}} dt && \square \end{aligned} \quad (13.83)$$

13.5 Exkurs: Beschreibende Statistik

Definition: Sei $\{x_1, \dots, x_n\}$ eine Stichprobe (n Zahlen, kommen irgendwo her). Dann heißen

$$\bar{x} := \frac{1}{n} \sum_{i=1}^n x_i \quad \text{und} \quad s_x^2 := \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (13.84)$$

Mittelwert (arithmetisches Mittel) und empirische Varianz (Stichprobenvarianz) der Stichprobe. Weiter heißt $s_x := \sqrt{s_x^2}$ empirische Standardabweichung.

Warum der Vorfaktor $\frac{1}{n-1}$?

Betrachte die x_i als Werte, die von Zufallsvariablen X_i angenommen werden, welche alle durch die gleiche Verteilung beschrieben werden.

Ziel: Schätze eine Kenngröße oder einen Parameter ϑ der Verteilung (z.B. den Erwartungswert μ , oder die Varianz σ^2) aus der Stichprobe mithilfe einer Schätzfunktion $\hat{\vartheta} : \mathbb{R}^n \rightarrow \mathbb{R}$ (oder $\rightarrow \mathbb{R}^m$) durch $\vartheta \approx \hat{\vartheta}(x_1, \dots, x_n)$. Das ist eigentlich bereits schließende Statistik...

Eine möglicherweise wünschenswerte Eigenschaft dieser Schätzfunktion ist die Erwartungstreue.

Definition: Seien X_1, \dots, X_n unabhängige ZV, die alle die gleiche Verteilung haben, kurz seien X_1, \dots, X_n iid (*independent and identically distributed*). Eine Schätzfunktion, $\hat{\vartheta} : \mathbb{R}^n \rightarrow \mathbb{R}$ für einen Parameter bzw. eine Kenngröße ϑ der Verteilung heißt erwartungstreu, falls gilt

$$E(\hat{\vartheta}(X_1, \dots, X_n)) = \vartheta. \quad (13.85)$$

Beispiele:

1. Kenngröße: Erwartungswert $\mu = E(X_i) \forall i = 1, \dots, n$
 Schätzfunktion: $\hat{\mu}(x_1, \dots, x_n) = \bar{x}$ (Mittelwert)
 Erwartungstreu?

$$\begin{aligned} E(\hat{\mu}(X_1, \dots, X_n)) &= E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \\ &= \frac{1}{n} \sum_{i=1}^n E(X_i) \\ &= \frac{1}{n} \sum_{i=1}^n \mu = \mu \quad \text{Ja!} \end{aligned} \quad (13.86)$$

2. Kenngröße: Varianz $\sigma^2 = E((X_i - \mu)^2) \forall i = 1, \dots, n$
 Schätzfunktion: $\hat{\sigma}^2(x_1, \dots, x_n) = s_x^2$ (empirische Varianz)

Erwartungstreu?

$$\begin{aligned}
 s_x^2 &= \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \\
 &= \frac{1}{n-1} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2 \\
 &= \frac{1}{n-1} \sum_{i=1}^n \left(x_i^2 - \frac{2}{n} \sum_{j=1}^n x_i x_j + \frac{1}{n^2} \sum_{j=1}^n \sum_{k=1}^n x_j x_k \right) \\
 &= \frac{1}{n-1} \left(\underbrace{\sum_{i=1}^n x_i^2 - \frac{2}{n} \sum_{i=1}^n \sum_{j=1}^n x_i x_j + \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^n x_j x_k}_{= -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n x_i x_j} \right) \tag{13.87} \\
 &= \frac{1}{n-1} \sum_{i=1}^n \sum_{j=1}^n \left(\frac{x_i^2}{n} - \frac{x_i x_j}{n} \right) \\
 &= \frac{1}{n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \left(\frac{x_i^2 + x_j^2}{2} - \frac{2x_i x_j}{2} \right) \\
 &= \frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2
 \end{aligned}$$

$$\begin{aligned}
 E(\hat{\sigma}^2(X_1, \dots, X_n)) &= E\left(\frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{j=1}^n (X_i - X_j)^2 \right) \\
 &= \frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \left(\underbrace{E(X_i^2) + E(X_j^2)}_{=2E(X^2)} - 2 \underbrace{E(X_i X_j)}_{= \begin{cases} E(X_i)E(X_j) & i \neq j \\ E(X^2) & i = j \end{cases}} \right) \\
 &= \frac{1}{2n(n-1)} \left(2n^2 E(X^2) - 2n E(X^2) - 2n(n-1) (E(X))^2 \right) \\
 &= E(X^2) - [E(X)]^2 = \sigma^2 \quad \text{Ja!} \tag{13.88}
 \end{aligned}$$

Bemerkung: Als Schätzung für die Standardabweichung σ verwenden wir gerne die empirische Standardabweichung s_x (aber nicht erwartungstreu).

13.6 Grenzwertsätze

Satz 50. (Schwaches Gesetz der großen Zahlen)

Sei $(X_k)_{k \in \mathbb{N}}$ eine Folge unabhängiger ZV, die alle die gleiche Verteilung haben (kurz: iid), mit $E(X_k) = \mu$ und $\text{Var}(X_k) = \sigma^2 \forall k$. Dann gilt für $Y_n := \frac{1}{n} \sum_{k=1}^n X_k$:

$$\lim_{n \rightarrow \infty} P(|Y_n - \mu| \geq \varepsilon) = 0 \quad \forall \varepsilon > 0 \quad (13.89)$$

Salopp: Der Mittelwert strebt gegen den Erwartungswert.

Relative Häufigkeiten streben gegen Wahrscheinlichkeiten.

Beweis:

$$\begin{aligned} E(Y_n) &= E\left(\frac{1}{n} \sum_{k=1}^n X_k\right) = \frac{1}{n} \sum_{k=1}^n \underbrace{E(X_k)}_{=\mu} = \mu \\ \text{Var}(Y_n) &= \text{Var}\left(\frac{1}{n} \sum_{k=1}^n X_k\right) = \frac{1}{n^2} \sum_{k=1}^n \underbrace{\text{Var}(X_k)}_{=\sigma^2} = \frac{\sigma^2}{n} \end{aligned} \quad (13.90)$$

Tschebyscheff'sche Ungleichung:

$$\forall \varepsilon > 0: \quad P(|Y_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow \infty} 0. \quad \square \quad (13.91)$$

Bemerkung: Die ZV $Z_n := \frac{Y_n - \mu}{\sigma/\sqrt{n}}$ ist standardisiert (d.h. $E(Z_n) = 0$ und $\text{Var}(Z_n) = 1$); es gilt

$$Z_n = \frac{\sqrt{n}}{\sigma} \underbrace{\frac{1}{n} \sum_{k=1}^n (X_k - \mu)}_{=Y_n - \mu} = \frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{X_k - \mu}{\sigma}. \quad (13.92)$$

Satz 51. (Zentraler Grenzwertsatz)

Sei $(X_k)_{k \in \mathbb{N}}$ eine Folge unabhängiger ZV, die alle der gleichen Verteilung folgen (d.h. iid), mit $E(X_k) = \mu$ und $\text{Var}(X_k) = \sigma^2 \forall k$, und deren 3. absolutes Moment $E(|X_k|^3)$ existiere. Dann gilt für $Z_n = \frac{1}{\sqrt{n}} \sum_{k=1}^n \frac{X_k - \mu}{\sigma}$:

$$Z_n \underset{n \rightarrow \infty}{\sim} \mathcal{N}(0, 1).$$

(ohne Beweis)

Bemerkung: Es gibt Varianten und Verschärfungen. Für den Hausgebrauch merken wir uns grob: Sind X_1, X_2, \dots iid, so ist die Summe $S_n := X_1 + X_2 + \dots + X_n$ für große n ungefähr normalverteilt, also näherungsweise $S_n \sim \mathcal{N}(E(S_n), \text{Var}(S_n))$.

13.7 Schließende Statistik: Hypothesentests

Idee: Wir möchten eine Annahme (über ein Experiment, einen Datensatz etc.) widerlegen, indem wir

- eine Größe betrachten (“Teststatistik”), die sich aus den Daten ermitteln läßt, und
- zeigen, dass der beobachtete Wert der Größe sehr unwahrscheinlich ist, falls die Annahme stimmt.

Beispiel:

- Annahme: Es liegt ein fairer Würfel vor.
- Experiment: Werfe den Würfel 100 mal.
- Teststatistik: Anzahl der geworfenen III .

Verwerfe die Annahme, falls zu viele oder zu wenige III fallen.

Vorgehensweise: (Hypothesentest)

Im Beispiel:

- | | |
|---|--|
| 1. Nullhypothese H_0 | Würfel ist fair |
| 2. Alternativhypothese H_A | Würfel ist nicht fair |
| 3. Wähle Teststatistik X (ZV) | # der geworfenen III |
| 4. Bestimme Verteilung von X , falls H_0 wahr | $X \sim \text{Bin}(100, \frac{1}{6})$ |
| 5. Wähle ein Signifikanzniveau α | $\alpha = 5\%$ |
| (Anteil der unter H_0 möglichen Werte von X , die wir für “sehr unwahrscheinlich” erklären, übliche Werte sind z.B. 10%, 5%, 1%, 0,1% ... – willkürlich!) | |
| 6. Bestimme den Verwerfungsbereich
$K := \{\text{“sehr unwahrscheinliche” } X\text{-Werte}\}$
so dass | Suche $k_1, k_2 \in \{0, 1, 2, \dots, 100\}$,
k_1 maximal, k_2 minimal,
so dass $P(X \in \{0, \dots, k_1\}) \leq 2,5\%$
und $P(X \in \{k_2, \dots, 100\}) \leq 2,5\%$.
Wir finden ¹⁷ : $k_1 = 9, k_2 = 25$, d.h. |
| <ul style="list-style-type: none"> • K diejenigen möglichen Werte von X enthält, die am stärksten für H_A sprechen und • $P(X \in K) \leq \alpha$,
und zwar möglichst groß. | $K = \{0, \dots, 9\} \cup \{25, \dots, 100\}$ |
| 7. Bestimme X aus den Daten | z.B. $X = 8$ |
| 8. Testentscheidung: Verwerfe H_0 zugunsten von H_A , falls $X \in K$. | Verwerfe H_0 . |

Wir sagen dann:

Es ist statistisch bewiesen auf $\alpha = 5\%$, dass H_0 nicht gelten kann.

oder

H_0 ist auf $\alpha = 5\%$ nicht mit Beobachtung vereinbar.

¹⁷denn $\sum_{k=0}^{9} \binom{100}{k} (\frac{1}{6})^k (\frac{5}{6})^{100-k} \approx 2,1\% \quad 4,3\%$ und $\sum_{k=25}^{100} \binom{100}{k} (\frac{1}{6})^k (\frac{5}{6})^{100-k} \approx 2,2\% \quad 3,8\%$

Bemerkungen:

1. Mit der Beobachtung $X = 11$ wäre H_0 nicht verworfen worden:
Daten widersprechen H_0 nicht. – aber auch nichts statistisch bewiesen!
2. Auf dem Signifikanzniveau $\alpha = 1\%$ hätten wir H_0 mit diesem Test nicht verwerfen können.
3. Es gibt ein kritisches α zwischen 1% und 5% , an dem sich die Testentscheidung für die Beobachtung $X = 8$ ändert; diese Zahl heißt **p-Wert**.

Genauer: Der p-Wert ist die Wahrscheinlichkeit dafür, unter H_0 etwas zu beobachten, was H_A mindestens so stark unterstützt wie die tatsächliche Beobachtung.

Im Beispiel:

$$\text{p-Wert} = 2 \sum_{k=0}^8 \binom{100}{k} \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{100-k} \approx 1,9\% \quad (13.93)$$

Faktor 2, da es sich um einen sog. beidseitigen Test handelt.

Würden wir gegen die Alternative H_A : *Würfel zeigt 1 mit gringerer Wahrscheinlichkeit als $\frac{1}{6}$* testen, so wäre der Verwerfungsbereich einseitig ($K = \{0, \dots, 10\}$ bei $\alpha = 5\%$) und der p-Wert wäre $0,95\%$

Der p-Wert hängt vom beobachteten X ab.

4. Enthält die Nullhypothese einen Parameter, so ist
 - die Menge aller Parameterwerte,
 - für der Test auf Signifikanzniveau α nicht verwirft,das **($1-\alpha$)-Vertrauensintervall** für den “wahren” Wert des Parameters.

Im Beispiel: $X \sim \text{Bin}(100, w)$

$H_0 : w = w_0, H_A : w \neq w_0$.

Bedingungen:

$$\begin{aligned} P(X \leq 8) > 2,5\% & \quad \text{und} & \quad P(X \geq 8) > 2,5\% \\ \Leftrightarrow B(8; 100, w_0) > 2,5\% & \quad \text{und} & \quad 1 - B(7; 100, w_0) > 2,5\% & \quad (13.94) \\ \Leftrightarrow w_0 < 15\% & \quad \text{und} & \quad w_0 > 3,5\% & \quad (\text{numerisch}) \end{aligned}$$

95%-Vertrauensintervall für w : (3,5%, 15%).

enthält natürlich $\frac{8}{100} = 8\%$ (Schätzung für w), aber nicht $\frac{1}{6} \approx 16,7\%$

Das Vertrauensintervall hängt (ebenso wie der p-Wert) vom beobachteten Wert der Teststatistik ab.

13.7.1 Der t-Test

Fragestellung:

Stichprobe $\{x_1, \dots, x_n\}$, betrachte als Werte von Zufallsvariablen $\{X_1, \dots, X_n\}$, iid.

- Gesucht: Erwartungswert $\mu := E(X_j)$ (gleich für alle j).
- Idee: Schätze durch Mittelwert \bar{x} .
- Genauigkeit? \rightsquigarrow Test, Vertrauensintervall,...

Annahme: Die X_j sind normalverteilt, d.h. $X_j \sim \mathcal{N}(\mu, \sigma^2)$ oder in Praxis: n ist groß und die X_j sind nicht zu "unnormale" verteilt

\Rightarrow Summe bzw. Mittelwert ist normalverteilt. (entweder laut Lemma 49 oder laut ZGS, Satz 51)

Testablauf

1. $H_0 : \mu = \mu_0$
2. $H_A : \mu \neq \mu_0$
so beidseitig – einseitig wäre $\mu > \mu_0$ oder $\mu < \mu_0$

3. Teststatistik: $T = \frac{\bar{x} - \mu_0}{s_x / \sqrt{n}}$
wobei: $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$ und $s_x^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^2$

4. Verteilung von T unter H_0 :

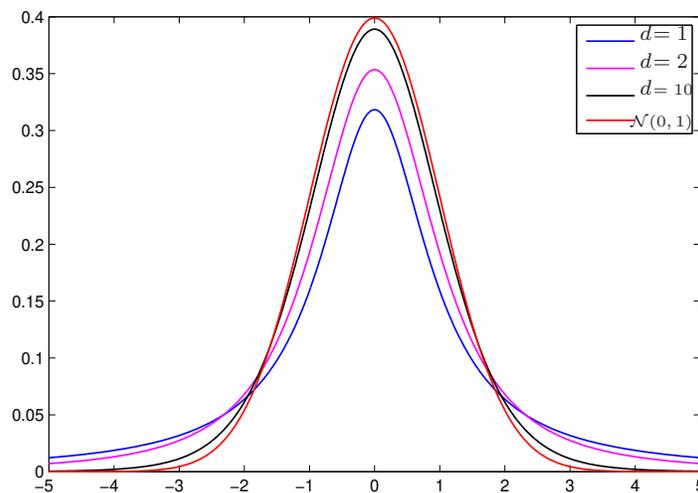
Wäre die wahre Varianz bekannt, so könnten wir $Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$ bilden,

wobei $Z \sim \mathcal{N}(0, 1)$ unter H_0 (d.h. für $\mu = \mu_0$).

Stattdessen: Schätze Varianz aus Daten.

Folge: T hängt über \bar{x} und s_x von den Daten ab und ist nur für $n \rightarrow \infty$ normalverteilt.

Man kann zeigen:¹⁸ T folgt der sogenannten (Student'schen) t-Verteilung mit $d = n - 1$ Freiheitsgraden; Dichte:



¹⁸siehe z.B. U. Krengel, *Einführung in die Wahrscheinlichkeitstheorie und Statistik*, 8. Auflage, Vieweg, Wiesbaden 2005, §14

Die Verteilungsfunktion $F_{T,d}$ ist tabelliert (siehe z.B. http://de.wikipedia.org/wiki/Studentsche_t-Verteilung).

Definieren wir $t_{d;q}$ durch $F_{T,d}(t_{d;q}) = q$, so gilt z.B.

$d = n-1$	1	3	10	20	40	60	∞ d.h. $\mathcal{N}(0,1)$
$t_{d;0,975}$	12,7	3,18	2,23	2,09	2,02	2,00	1,96

5. z.B. $\alpha = 5\%$

6. Verwerfungskriterium:

$$|T| > t_{n-1;0,975} \tag{13.95}$$

denn $P(T > t_{n-1;0,975}) = 2,5\% = P(T < -t_{n-1;0,975})$

7./8. ...

zweiseitiges 95%-Vertrauensintervall für den Erwartungswert μ

$H_0 : \mu = \mu_0$ wird angenommen, falls $|T| \leq t_{n-1;0,975}$

$$\begin{aligned} \Leftrightarrow \quad & -t_{n-1;0,975} \leq \frac{\bar{x} - \mu_0}{s_x/\sqrt{n}} \leq t_{n-1;0,975} \\ \Leftrightarrow \quad & \bar{x} - t_{n-1;0,975} \frac{s_x}{\sqrt{n}} \leq \mu_0 \leq \bar{x} + t_{n-1;0,975} \frac{s_x}{\sqrt{n}} \end{aligned} \tag{13.96}$$

d.h. das 95%-Vertrauensintervall für μ ist

$$\left[\bar{x} - t_{n-1;0,975} \frac{s_x}{\sqrt{n}}, \bar{x} + t_{n-1;0,975} \frac{s_x}{\sqrt{n}} \right] \approx \left[\bar{x} - 2 \frac{s_x}{\sqrt{n}}, \bar{x} + 2 \frac{s_x}{\sqrt{n}} \right] \quad \text{falls } n \text{ hinreichend groß.} \tag{13.97}$$