Mathematik für Physiker 2

Hermann Schulz-Baldes, Dept. Mathematik

Assistenz: Daniele Toniolo, Dept. Mathematik

Vorlesung, Sommersemester 2018

Termine

Vorlesungstermine

Mo 08:15 – 10:00 Uhr	Raum: HE	Schulz-Baldes
Mi 08:15 – 10:00 Uhr	Raum: HE	Schulz-Baldes

Videoaufzeichnung vom SS2016 über StudOn

Übungstermine (Physikum bzw. Cauerstr. 11)

Gruppe 1	Fr 12:15 – 13:45	Raum: HC	Refik Mansuroglu
Gruppe 2	Fr 14:15 – 15:45	Raum: U4	Daniele Toniolo (Eng.)
Gruppe 3	Fr 12:15 – 13:45	Raum: H13	Martin Doß
Gruppe 4	Fr 12:15 – 13:45	Raum: U4	Kevin Haas

Regeln

Anmeldung

- Melden Sie sich in Studon zu Veranstaltung und Übungen an (ab 10:00 am 9.4.). Beachten Sie Anmeldefristen für Klausur

Übungen

- Die Übungsblätter werden dienstags auf Studon bereitgestellt.
- Die Abgabe bis Dienstag 11:00 in Übungskasten (Cauerstr. 11)
- Es werden nur leserliche und ordentliche Abgaben akzeptiert
- Es sind Zweierabgaben gestattet, solange beide die gleiche Übungsgruppe besuchen
- Es sollten 50% der Gesamtpunktzahl erreicht werden

Klausur

- Die Klausur findet am 17.07.2018 um 8:00 in H12 und H13 statt
- Einsicht am 18.7.2018 von 8:00 bis 9:30 in U2 (Mathematik)
- Nachklausur 20.9.2018 um 8:00 in H12 and H13
- Hilfsmittel: beidseitig handbeschriebenes DIN A4 Blatt

Überblick

- Lineare Algebra:

  Determinanten
  Eigenwerte
  Diagonalisierung
  Skalarprodukte
  Jordansche Normalformen
  Lineare Differentialgleichungen
  Singulärwertzerlegung
  Quadratische Formen

- Analysis:

  Riemann Integral
  Grundbegriffe der Topologie
  Differentialgleichungen
  Differentialrechnung mehrerer Variablen
  Satz von Taylor
  Satz über implizite Funktionen

Literatur

Qual der Wahl: jedes Buch oder Skript zu obigen Themen!

... auch dieses Skript enthält alles Wesentliche

Lineare Algebra: Grundbegriffe und Notationen

(Oder: was Sie alles schon wissen sollten!)

Körper $\mathbb{K}$ ist hier immer $\mathbb{R}$ oder $\mathbb{C}$ (nichts Allgemeineres hier!)

$\mathbb{K}^N = \{N\text{-komponentige Vektoren mit Koeffizienten in } \mathbb{K}\}$

Struktur des Vektorraumes: Addition und Skalarmultiplikation

$$v + \lambda w \in \mathbb{K}^N \qquad v, w \in \mathbb{K}^N, \quad \lambda \in \mathbb{K}$$

Anderer Vektorraum: $V = \{\text{Polynome mit Koeffezienten in } \mathbb{K}\}$

Lineare Abhängigkeit und Unabhängigkeit....

$b_1, b_2, \ldots.$ Basis (linear unabhängig und aufspannend)

Dimension eines Vektorraumes $\dim(\mathbb{K}^M) = M$ und $\dim(V) = \infty$

$U \subset \mathbb{K}^M$ Unterraum $\iff v + \lambda w \in U \ \forall\, v, w \in U, \quad \lambda \in \mathbb{K}$

Dann hat U wieder Basis und Dimension

Matrizen

$\mathrm{Mat}(N \times M, \mathbb{K}) = \{N \times M$ Matrizen mit Einträgen in $\mathbb{K}\}$

Spezialfall Vektoren $\mathbb{K}^N = \mathrm{Mat}(N \times 1, \mathbb{K})$

Matrixprokukt $\mathrm{Mat}(N \times M, \mathbb{K}) \times \mathrm{Mat}(M \times K, \mathbb{K}) \in \mathrm{Mat}(N \times K, \mathbb{K})$

Spezialfall Matrix-Vektor Produkt $\mathrm{Mat}(N \times M, \mathbb{K}) \times \mathbb{K}^M \in \mathbb{K}^N$

Linearität von Matrixprodukt in beiden Argumenten:

$$(A + \lambda B)C = AC + \lambda BC \qquad A(B + \lambda C) = AB + \lambda AC$$

wobei $\lambda \in \mathbb{K}$ und richtige Matrixgrößen

Rangsatz für $A \in \mathrm{Mat}(N \times M, \mathbb{K})$

$\mathrm{Ker}(A) = \{v \in \mathbb{K}^M \mid Av = 0\}$ Kern von A

$\mathrm{Ran}(A) = \{Av \mid v \in \mathbb{K}^M\} = A\mathbb{K}^M$ Bild (range) von A

$\mathrm{Ker}(A)$ Unterraum von $\mathbb{K}^M$

$\mathrm{Ran}(A)$ Unterraum von $\mathbb{K}^N$

Theorem 1.1 (Rangsatz)

$\dim(\mathrm{Ker}(A)) + \dim(\mathrm{Ran}(A)) = M$

Gauss Algorithmus

Lösung von linearen Gleichungssystem nach $v \in \mathbb{K}^M$:

$$Av = b \qquad A \in \mathrm{Mat}(N \times M, \mathbb{K}), \quad b \in \mathbb{K}^N$$

Bringe Gleichungssystem auf Dreiecksgestalt

Lese Lösung(en) ab, falls welche existieren

Bestimmung des Inversen A^{-1} von quadratischen $A \in \mathrm{Mat}(N \times N, \mathbb{K})$

Mit verallgemeinerten Gauss Algorithmus (falls Inverses existiert)

$$(A, \mathbf{1}) \ \to \ (\mathbf{1}, A^{-1})$$

Hier $\mathbf{1} = \mathbf{1}_N = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & 1 \end{pmatrix}$ Einheitsmatrix

Dann $AA^{-1} = \mathbf{1} = A^{-1}A$

Und was Sie auch schon wissen sollten!

Also ist $\left(\mathrm{Mat}(N \times N, \mathbb{K}), +, \cdot\right)$ eine Algebra über $\mathbb{K}$, d.h.

- $\mathrm{Mat}(N \times N, \mathbb{K})$ Vektorraum über $\mathbb{K}$ (von Dimension NM)
- Matrixprodukt erfüllt Distributivgesetz, Assoziativität

Allgemeine lineare Gruppe unter Matrixmultiplikation

$$\mathrm{GL}(N, \mathbb{K}) = \{A \in \mathrm{Mat}(N \times N, \mathbb{K}) \text{ invertierbar}\}$$

2 Determinanten

Vorbereitung: Permutationen

Definition 2.1

Sei $N \in \mathbb{N}$. Eine bijektive Abbildung $\sigma : \{1, \ldots, N\} \to \{1, \ldots, N\}$ heißt Permutation über N Elementen. Die Menge S_N aller Permutationen über N Elementen versehen mit der Hintereinanderausführung heißt symmetrische Gruppe. Standardnotation für eine Permutation ist

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & N \\ \sigma(1) & \sigma(2) & \cdots & \sigma(N) \end{pmatrix}$$

wobei untere Zeile die der oberen zugeordneten Werte enthält

Eine Transposition (von n', m') ist eine Permutation, bei der $\sigma(n) \neq n$ für lediglich zwei Werte $n = n', m'$, bei denen $\sigma(n') = m'$

Bei benachbarter Transposition gilt zudem $m' = n' + 1 \mod N$

Beispiel 2.2

Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}$$

ist eine benachbarte Transposition

Satz 2.3

(i) *Die Gruppe S_N hat $N! = N(N-1)\cdots 1$ Elemente*

(ii) *Jede Transposition ist Hintereinanderausführung einer ungeraden Anzahl von benachbarten Transpositionen.*

(iii) *Jede Permutation ist Hintereinanderausführung von benachbarten Transpositionen.*

Satz 2.4

Jedes $\sigma \in S_N$ *hat eine darstellende Matrix* $P(\sigma) \in \mathrm{Mat}(N \times N, \mathbb{R})$*:*

$$P(\sigma)_{n,m} \;=\; \delta_{n,\sigma(m)}$$

Mit Hilfe der Standardbasisvektoren als Spaltenvektoren, kann diese Matrix auch geschrieben werden als

$$P(\sigma) \;=\; \big(e_{\sigma(1)}, \ldots, e_{\sigma(N)}\big)$$

Es gilt dann für $\sigma, \tau \in S_N$

$$P(\sigma\tau) \;=\; P(\sigma)P(\tau)$$

wobei auf der rechten Seite die Matrixmultiplikation verwandt wird

Somit: $P(S_N)$ *Untergruppe der allgemeinen linearen Gruppe* $\mathrm{Gl}(N, \mathbb{R})$

Beispiel 2.5

Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}$$

hat die darstellende Matrix:

$$P(\sigma) = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Definition 2.6

Sei $\sigma \in S_N$. Ein Fehlstand von σ ist ein Paar (n, m) mit $1 \leqslant n < m \leqslant N$, so dass $\sigma(n) > \sigma(m)$. Die Anzahl $L(\sigma)$ der Fehlstände von σ definiert das Signum von σ durch

$$\operatorname{sgn}(\sigma) \;=\; (-1)^{L(\sigma)}$$

Beispiel 2.7

Die Permutation

$$\sigma \;=\; \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$$

hat Fehlstände $(1, 3)$, $(2, 3)$ und $(2, 4)$. Also $L(\sigma) = 3$ und $\operatorname{sgn}(\sigma) = -1$.

Satz 2.8

(i) *Für* $\sigma \in S_N$,

$$\operatorname{sgn}(\sigma) = \prod_{1 \leqslant n < m \leqslant N} \frac{\sigma(n) - \sigma(m)}{n - m}$$

wobei Produkt gebildet wird über alle Paare (n, m) *mit der angegebenen Eigenschaft* $1 \leqslant n < m \leqslant N$.

(ii) *Für* $\sigma, \tau \in S_N$,

$$\operatorname{sgn}(\sigma\tau) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau)$$

(iii) *Die Abbildung* $\sigma \in S_N \mapsto \operatorname{sgn}(\sigma) \in \mathbb{Z}_2$ *ist ein surjektiver Gruppen-Homomorphismus, wobei hier* $\mathbb{Z}_2$ *die multiplikative Gruppe mit 2 Elementen ist.*

(iv) *Sei* $\sigma = \tau_K \cdots \tau_1$ *ein Produkt von Transpositionen* (*was nach* Satz 2.3(iii) *immer der Fall ist*), *so ist* $\operatorname{sgn}(\sigma) = (-1)^K$

Definition 2.9

Sei $A = (A_{n,m})_{n,m=1,\ldots,N} \in \mathrm{Mat}(N \times N, \mathbb{K})$. Seine Determinante ist

$$\det(A) = \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma) \prod_{n=1}^{N} A_{n,\sigma(n)} = \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, A_{1,\sigma(1)} A_{2,\sigma(2)} \cdots A_{N,\sigma(N)}$$

Beispiel 2.10

Sei $N = 2 \Rightarrow$ zwei Permutationen (Identität und Transposition)

$$\det(A) = A_{1,1}A_{2,2} - A_{1,2}A_{2,1}$$

Beispiel 2.11

Sei $\sigma \in S_N$ mit darstellender Matrix $P(\sigma)$. Dann nur 1 Summand:

$$\det(P(\sigma)) = \mathrm{sgn}(\sigma)$$

Definition 2.12

Seien $V_1, \ldots, V_N$ und W Vektorräume über $\mathbb{K}$. Eine Abbildung

$$F : V_1 \times \ldots \times V_N \to W$$

heißt multilinear $\Longleftrightarrow$ linear in jedem Argument ist, d.h.

$$v_n \in V_n \mapsto F(v_1, \ldots, v_N) \in W \quad \text{linear} \quad \forall\; n = 1, \ldots, N$$

und $v_m \in V_m$, $m \neq n$. Falls $N = 2$ und $V_1 = V_2$, heißt F bilinear

Beispiel 2.13

Jede lineare Abbildung ist multilinear mit $N = 1$. Sei $V = \mathbb{K}^M$ und $A \in \mathrm{Mat}(M \times M, \mathbb{K})$. Eine bilineare Abbildung ist $F : V \times V \to \mathbb{K}$

$$F(v, w) \;=\; \sum_{m,n=1}^{M} v_m A_{m,n} w_n$$

Definition 2.14

Seien V und W Vektorräume über $\mathbb{K}$

Abbildung $F : V^{\times N} \to W$ heißt alternierend $\Longleftrightarrow$ für jede benachbarte Transposition $\tau \in S_N$ (d.h. jedes Vertauschen von zwei benachbarten Argumenten) gilt

$$F(v_{\tau(1)}, \ldots, v_{\tau(N)}) \;=\; -F(v_1, \ldots, v_N) \qquad v_n \in V$$

Beispiel 2.15

Sei $A \in \mathrm{Mat}(M \times M, \mathbb{K})$ eine anti-symmetrische Matrix, d.h.

$$A^T \;=\; -A$$

Dann ist $F : \mathbb{K}^M \times \mathbb{K}^M \to \mathbb{K}$ wie oben definiert durch

$$F(v, w) \;=\; \sum_{m,n=1}^{M} v_m A_{m,n} w_n$$

alternierend

Lemma 2.16

Seien V und W Vektorräume über $\mathbb{K}$ und $F : V^{\times N} \to W$ alternierend Dann gilt für jede Permutation $\sigma \in S_N$

$$F(v_{\sigma(1)}, \ldots, v_{\sigma(N)}) \;=\; \operatorname{sgn}(\sigma)\, F(v_1, \ldots, v_N)$$

für alle $v_1, \ldots, v_N \in V$.

Beweis. Jede Permutation Hintereinanderausführung von benachbarten Transpositionen. Deren Anzahl ist genau das Signum $\square$

Korollar 2.17

$F(v_1, \ldots, v_N) = 0$, *falls* $v_n = v_m$ *für* $n \neq m$

Grund: $F(\ldots, v_m, \ldots, v_n, \ldots) = -F(\ldots, v_n, \ldots, v_m, \ldots) = 0$ $\square$

Satz 2.18

Sei $F : V^{\times N} \to W$ *eine multilineare, alternierende Abbildung*

(i) *Falls dann* $v_1, \ldots, v_N \in V$ *linear abhängig sind, so gilt*

$$F(v_1, \ldots, v_N) \; = \; 0$$

(ii) *Sei* $\dim(V) = N$ *und* $F \neq 0$*,*
d.h. es gibt $w_1, \ldots, w_N \in V$ *mit* $F(w_1, \ldots, w_N) \neq 0$
Falls dann $v_1, \ldots, v_N \in V$ *linear unabhängig sind, so gilt*

$$F(v_1, \ldots, v_N) \; \neq \; 0$$

Beweis. (i) Seien $v_1, \ldots, v_N$ linear abhängig, und $v_N = \sum_{n=1}^{N-1} \lambda_n v_n$
Dann folgt aus der Linearität im letzten Argument

$$F(v_1, \ldots, v_N) \;=\; \sum_{n=1}^{N-1} \lambda_n \, F(v_1, \ldots, v_{N-1}, v_n) \;=\; 0$$

(ii) Da $v_1, \ldots, v_N$ Basis, gibt es Zerlegung $w_n = \sum_{m=1}^{N} \lambda_{n,m} \, v_m$. Also

$$\begin{aligned}
F(w_1, \ldots, w_N) \;&=\; \sum_{m_1,\ldots,m_N=1}^{N} (\lambda_{1,m_1} \cdots \lambda_{N,m_N}) \, F(v_{m_1}, \ldots, v_{m_N}) \\
&=\; \sum_{\sigma \in S_N} (\lambda_{1,\sigma(1)} \cdots \lambda_{N,\sigma(N)}) \, F(v_{\sigma(1)}, \ldots, v_{\sigma(N)}) \\
&=\; \left(\sum_{\sigma \in S_N} (\lambda_{1,\sigma(1)} \cdots \lambda_{N,\sigma(N)}) \, \mathrm{sgn}(\sigma) \right) F(v_1, \ldots, v_N)
\end{aligned}$$

nach Lemma 2.16. Da $F(w_1, \ldots, w_N) \neq 0$, auch $F(v_1, \ldots, v_N) \neq 0$ $\square$

Satz 2.19

$A \in \mathrm{Mat}(N \times N, \mathbb{K}) \cong (\mathbb{K}^N)^{\times N} \mapsto \det(A) \in \mathbb{K}$ *multilinear, alternierend, wenn aufgefasst als Abbildung auf den N Spaltenvektoren*

Beweis. Sei $A = (a_1, \dots, a_N)$ mit Spaltenvektoren $a_n \in \mathbb{K}^N$, $1 \leqslant n \leqslant N$
Für $\lambda \in \mathbb{K}$ und $b \in \mathbb{K}^N$ gilt dann

$$\begin{aligned}
&\det(a_1, \dots, a_n + \lambda b, \dots, a_N) \\
&\quad = \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, A_{1,\sigma(1)} \cdots (A_{n,\sigma(n)} + \lambda b_{\sigma(n)}) \cdots A_{N,\sigma(N)} \\
&\quad = \det(a_1, \dots, a_N) \;+\; \lambda \det(a_1, \dots, a_{n-1}, b, a_{n+1}, \dots, a_N)
\end{aligned}$$

Für Transposition τ gilt $\tau S_N = S_N$. Weil sgn Homomorphismus

$$\begin{aligned}
\det(a_{\tau(1)}, \dots, a_{\tau(N)}) &= \sum_{\sigma \in S_N} \mathrm{sgn}(\sigma)\, A_{1,\sigma(\tau(1))} \cdots A_{N,\sigma(\tau(N))} \\
&= -\sum_{\sigma\tau \in S_N} \mathrm{sgn}(\sigma\tau)\, A_{1,\sigma(\tau(1))} \cdots A_{N,\sigma(\tau(N))} = -\det(A)
\end{aligned}$$

$\square$

Wichtigste Eigenschaften der Determinante

Außerdem rechnet man Normalisierung $\det(\mathbf{1}) = 1$. Leicht zu zeigen:

Satz 2.20 (Charakterisierung der Determinante)

Sei $F : \mathrm{Mat}(N \times N, \mathbb{K}) \cong (\mathbb{K}^N)^{\times N} \to \mathbb{K}$. *Dann gilt:*

$$F \text{ multilineare, alternierend mit } F(\mathbf{1}_N) = 1 \qquad \Longleftrightarrow \qquad F = \det$$

Satz 2.21

Sei $A \in \mathrm{Mat}(N \times N, \mathbb{K})$. *Dann gilt*

$$A \text{ invertierbar} \qquad \Longleftrightarrow \qquad \det(A) \neq 0$$

Beweis. Verwende Satz 2.18:

"$\Longrightarrow$" A invertierbar $\Rightarrow$ Zeilen linear unabhängig $\Rightarrow$ $\det(A) \neq 0$

"$\Longleftarrow$" A nicht invertierbar $\Rightarrow$ Zeilen linear abhängig $\Rightarrow$ $\det(A) = 0$ $\square$

Satz 2.22

Seien $A, B \in \mathrm{Mat}(N \times N, \mathbb{K})$*. Dann gilt*

$$\det(AB) = \det(A)\,\det(B)$$

Insbesondere gilt für invertierbares A

$$\det(A^{-1}) = \frac{1}{\det(A)} \qquad \det(A^{-1}BA) = \det(B)$$

Also ist $\det : (\mathrm{Gl}(N, \mathbb{K}), \cdot) \to (\mathbb{R}\backslash\{0\}, \cdot)$ ein Gruppenhomomorphismus

Ähnliche Matrizen in folgendem Sinne haben gleiche Determinante:

Definition 2.23

$B, C \in \mathrm{Mat}(N \times N, \mathbb{K})$ ähnlich
$\iff \exists\, A \in \mathrm{Mat}(N \times N, \mathbb{K})$ mit $C = A^{-1}BA$

Da Gauss-Algorithmus **ohne** Normieren durch Linksmultiplikation einer oberen Dreiecksmatrix mit Einserdiagonale gegeben ist, kann dies als Algorithmus zur Berechnung von det verwandt werden.

Beweis. Sei $A = (a_1, \ldots, a_N)$ mit Spaltenvektoren $a_n \in \mathbb{K}$. Dann

$$\begin{aligned}
\det(AB) &= \det\left(\sum_{n_1=1}^{N} B_{n_1,1} a_{n_1}, \ldots, \sum_{n_1=1}^{N} B_{n_N,N} a_{n_N}\right) \\
&= \sum_{n_1,\ldots,n_N=1}^{N} B_{n_1,1} \cdots B_{n_N,N} \det(a_{n_1}, \ldots, a_{n_N})
\end{aligned}$$

Es bleiben nur Summanden mit $\{n_1, \ldots, n_N\} = \{1, \ldots, N\}$, für die es eine Permutation $\sigma \in S_N$ gibt mit $n_m = \sigma(m)$. Mit Lemma 2.16

$$\begin{aligned}
\det(AB) &= \sum_{\sigma \in S_N} B_{\sigma(1),1} \cdots B_{\sigma(N),N} \det(a_{\sigma(1)}, \ldots, a_{\sigma(N)}) \\
&= \sum_{\sigma \in S_N} B_{\sigma(1),1} \cdots B_{\sigma(N),N} \operatorname{sgn}(\sigma) \det(a_1, \ldots, a_N) \\
&= \det(B^T) \det(A) = \det(B) \det(A)
\end{aligned}$$

da $\det(B^T) = \det(B)$. $\square$

Definition 2.24

Adjunkte $\widehat{A} = (\widehat{A}_{m,n})_{m,n=1,\ldots,N} \in \mathrm{Mat}(N \times N, \mathbb{K})$ von $A = (a_1, \ldots, a_N)$:

$$\widehat{A}_{m,n} \;=\; \det(a_1, \ldots, a_{m-1}, e_n, a_{m+1}, \ldots, a_N)$$

wobei a_m ersetzt durch n-ten Standardbasisvektor $e_n \in \mathbb{K}^N$.

Lemma 2.25

Sei $\check{A}^{(n,m)} \in \mathrm{Mat}((N-1) \times (N-1), \mathbb{K})$ die Matrix, die aus A durch Streichen der n-ten Zeile und der m-ten Spalte gewonnen wird. Dann

$$\widehat{A}_{m,n} \;=\; (-1)^{n+m} \det(\check{A}^{(n,m)})$$

Beweis. Durch benachbarte Transpositionen von Zeilen und Spalten

$$\begin{aligned}\widehat{A}_{m,n} \;&=\; (-1)^{N-m} \det(a_1, \ldots, a_{m-1}, a_{m+1}, \ldots, a_N, e_n) \\ &=\; (-1)^{N-m}(-1)^{N-n} \det\left(\begin{pmatrix} \check{A}^{(n,m)} & 0 \\ b & 1 \end{pmatrix} \right) \;=\; (-1)^{n+m} \det(\check{A}^{(n,m)})\end{aligned}$$

□

Satz 2.26 (Cramer's Regel für die inverse Matrix)

Sei $A \in \mathrm{Mat}(N \times N, \mathbb{K})$ *mit Adjunkter* $\widehat{A} \in \mathrm{Mat}(N \times N, \mathbb{K})$. *Dann gilt*

$$\widehat{A}\,A \;=\; \det(A)\,\mathbf{1}_N$$

Falls $\det(A) \neq 0$, *ist die Adjunkte im Wesentlichen die inverse Matrix:*

$$A^{-1} \;=\; \frac{1}{\det(A)}\,\widehat{A}$$

Beweis. Es werden die Einträge des Matrixproduktes berechnet:

$$
\begin{aligned}
(\,\widehat{A}\,A\,)_{m,n} &= \sum_{k=1}^{N} \widehat{A}_{m,k}\,A_{k,n} && (2.1)\\
&= \sum_{k=1}^{N} \det(a_1, \ldots, a_{m-1}, e_k, a_{m+1}, \ldots, a_N)\,A_{k,n}\\
&= \det(a_1, \ldots, a_{m-1}, a_n, a_{m+1}, \ldots, a_N) \;=\; \det(A)\,\delta_{n,m}
\end{aligned}
$$

□

Rückführung einer $N \times N$ Determinante auf viele $(N-1) \times (N-1)$:

Satz 2.27 (Laplace'scher Entwicklungssatz)

Entwicklung der Determinante nach der m-ten Spalte

$$\det(A) \;=\; \sum_{k=1}^{N} A_{k,m}\,(-1)^{m+k}\,\det(\check{A}^{(k,m)})$$

und der m-ten Zeile

$$\det(A) \;=\; \sum_{k=1}^{N} A_{m,k}\,(-1)^{m+k}\,\det(\check{A}^{(m,k)})$$

Beweis. In Gleichung (2.1) für $n = m$ ersetze
$\widehat{A}_{m,k} = (-1)^{m+k}\,\det(\check{A}^{(k,m)})$ gemäß Lemma 2.25

Wegen $\det(A) = \det(A^T)$ folgt Entwicklung nach m-ter Zeile aus der Entwicklung von $\det(A^T)$ nach m-ter Spalte. $\square$

3 Eigenwerte und Eigenvektoren

Definition 3.1

Sei T lineare Abbildung auf Vektorraum V über $\mathbb{K}$, d.h. $T \in \mathcal{L}(V)$

Ein Skalar $z \in \mathbb{K}$ heißt Eigenwert von T

$\Longleftrightarrow$ es gibt einen sogenannten Eigenvektor $v \in V$, $v \neq 0$, mit

$$T\,v \;=\; z\,v$$

Diese Gleichung heißt eine Eigenwertgleichung

Satz 3.2

Sei $S \in \mathcal{L}(V, W)$ eine invertierbare lineare Abbildung zwischen Vektorräumen V und W über $\mathbb{K}$. Dann gilt für $T \in \mathcal{L}(V)$:

$$z \in \mathbb{K} \text{ Eigenwert von } T \qquad \Longleftrightarrow \qquad z \in \mathbb{K} \text{ Eigenwert von } STS^{-1}$$

Beweis. "$\Longrightarrow$" Da $Tv = TS^{-1}Sv = zv$, auch $STS^{-1}(Sv) = z\,Sv$ $\square$

Wähle S als darstellende Matrix für $T \in \mathcal{L}(V)$ zu einer Basis
$\Longrightarrow$ in endlicher Dimension nur Eigenwerte von Matrizen benötigt

Satz 3.3 (Charakteristisches Polynom)

Sei $A \in \mathrm{Mat}(N \times N, \mathbb{K})$. Dann gilt

$$z \in \mathbb{K} \text{ Eigenwert von } A \qquad \Longleftrightarrow \qquad \det(A - z\,\mathbf{1}_N) \;=\; 0$$

Eigenwerte sind Nullstellen in $\mathbb{K}$ des charakteristischen Polynoms

$$z \in \mathbb{K} \;\mapsto\; p_A(z) = \det(A - z\,\mathbf{1}_N)$$

Beweis. $Av = zv \Longleftrightarrow (A - z\,\mathbf{1})v = 0$
Existenz von $v \neq 0 \Longleftrightarrow A - z\,\mathbf{1}$ nicht invertierbar

$$\Longleftrightarrow \det(A - z\,\mathbf{1}_N) = 0 \qquad \square$$

Beispiel 3.4

Sei

$$A = \begin{pmatrix} 5 & -1 \\ -1 & 5 \end{pmatrix}$$

Also

$$p_A(z) = (5-z)^2 - 1 = z^2 - 10z + 24 = (z-4)(z-6)$$

Also sind die Eigenwerte $z_1 = 4$ und $z_2 = 6$.

Satz 3.5

Das charakteristisches Polynom P_A ist N-ten Grades und erfüllt

$$\det(A - z\,\mathbf{1}_N) \;=\; (-z)^N \;+\; \mathrm{Tr}(A)\,(-z)^{N-1} \;+\; q(z) \;+\; \det(A)$$

wobei $q = \sum_{n=1}^{N-2} q_n z^n$ ein Polynom vom Grade höchstens $N-2$ ohne konstanten Term ist, und $\mathrm{Tr}(A)$ die Spur von A ist, welche definiert ist als die Summe der Diagonaleinträge von A:

$$\mathrm{Tr}(A) \;=\; \sum_{n=1}^{N} A_{n,n}$$

Beweis. Ausschreiben der Definition mit Kronecker Delta:

$$\det(A - z\,\mathbf{1}_N) \;=\; \sum_{\sigma\in S_N} \mathrm{sgn}(\sigma) \prod_{n=1}^{N} \big(A_{n,\sigma(n)} - z\,\delta_{n,\sigma(n)}\big)$$

Potenzen z^N und z^{N-1} nur ein Summand $\sigma = \mathrm{id}$

Konstanter Term ist bei $z = 0$ genau $\det(A)$ $\square$

Beispiel 3.6 (Allgemeine Formel für 2×2 Matrizen)

Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Dann ist das charakteristische Polynom

$$\begin{aligned} p_A(z) &= \det\left(\begin{pmatrix} a-z & b \\ c & d-z \end{pmatrix}\right) \\ &= (a-z)(d-z) - bc \\ &= z^2 - (a+d)z + (ad-bc) \\ &= z^2 - \mathrm{Tr}(A)\, z + \det(A) \end{aligned}$$

Nullstellen sind

$$z_\pm = \frac{\mathrm{Tr}(A)}{2} \pm \sqrt{\frac{\mathrm{Tr}(A)^2}{4} - \det(A)}$$

Möglich: A reell, aber p_A nur komplexe Nullstellen.

Bemerkung 3.7

Reelle Matrizen aufgefasst als lineare Abbildung auf dem $\mathbb{R}^N$ können evtl. keine (reellen) Eigenwerte haben. Wenn sie hingegen als lineare Abbildung auf dem komplexen Vektorraum $\mathbb{C}^N$ aufgefasst werden, so gibt es N Eigenwerte:

Satz 3.8

Charakteristisches Polynom von $A \in \mathrm{Mat}(N \times N, \mathbb{K})$ hat Faktorisierung

$$p_A(z) \;=\; (-1)^N \prod_{k=1}^{K} (z - z_k)^{\alpha(z_k)} \tag{3.1}$$

mit paarweise verschiedenen $z_k \in \mathbb{C}$ und mit $\alpha(z_k) \in \mathbb{N}$, so dass

$$\sum_{k=1}^{K} \alpha(z_k) \;=\; N \tag{3.2}$$

Falls $z_k \in \mathbb{K}$, so ist z_k Eigenwert von A und $\alpha(z_k) = \alpha_A(z_k)$ heißt algebraische Vielfachheit oder Multiplizität von z_k

Beispiel 3.9

Sei

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

Charakteristisches Polynom berechnet mit Regel von Sarrus

$$\begin{aligned} p_A(z) = \ & \det\left(\begin{pmatrix} 1-z & 2 & 0 \\ -1 & -z & 1 \\ 1 & 0 & -z \end{pmatrix} \right) \\ = \ & (1-z)z^2 + 2 - (-z)\cdot 2 \cdot (-1) \ = \ (1-z)(z^2+2) \end{aligned}$$

Also sind die Nullstellen von p_A gegeben durch

$$z_1 = 1 \quad , \quad z_2 = i\sqrt{2} \quad , \quad z_3 = -i\sqrt{2}$$

Als lineare Abbildung auf $\mathbb{R}^3$ lediglich ein Eigenwert $z_1 = 1$!

Berechnung von Eigenvektor zu Eigenwert z von A

Gauss für homogenenes Gleichungssystems $(A - z\,\mathbf{1})v = 0$

Beispiel 3.10

Sei wieder

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

Nullstellen $z_1 = 1$, $z_2 = i\sqrt{2}$ und $z_3 = -i\sqrt{2}$. Für z_1 berechne

$$A - z_1\,\mathbf{1} = \begin{pmatrix} 0 & 2 & 0 \\ -1 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix} \to \begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{array} \to \begin{array}{ccc|c} -1 & -1 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Also $\mathrm{Ker}(A - z_1\,\mathbf{1}) = \mathrm{span}(v_1)$ mit $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ ein Eigenvektor

Satz 3.11

Sei $A \in \mathrm{Mat}(N \times N, \mathbb{R})$ reell aufgefasst als Abbildung auf $\mathbb{C}^N$

Wenn $z \in \mathbb{C}$ Eigenwert, so auch sein komplex konjugiertes $\overline{z} \in \mathbb{C}$

Eigenvektoren sind auch komplex konjugiert zueinander

Beweis. Sei $Av = zv \Rightarrow A\overline{v} = \overline{z}\,\overline{v}$ $\square$

Satz 3.12

Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig

Genauer: wenn $v_1, \ldots, v_K$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $z_1, \ldots, z_K$ von T, so $v_1, \ldots, v_K$ linear unabhängig

Beweis. Induktion über K. Für $K = 1$ Aussage richtig, da $v_1 \neq 0$

Induktionsschritt von $K - 1$ zu K: Seien $\lambda_1, \dots, \lambda_K \in \mathbb{K}$, so dass

$$\sum_{k=1}^{K} \lambda_k v_k \;=\; 0$$

Multiplikation dieser Gleichung mit z_K und Anwenden von A liefert:

$$\sum_{k=1}^{K} \lambda_k z_K\, v_k \;=\; 0 \qquad , \qquad \sum_{k=1}^{K} \lambda_k z_k\, v_k \;=\; 0$$

Somit ergibt Subtraktion der beiden Gleichungen

$$\sum_{k=1}^{K-1} \lambda_k (z_K - z_k)\, v_k \;=\; 0$$

Da $z_K - z_k \neq 0$, gibt Induktionshypothese $\lambda_k = 0$ für $k = 1, \dots, K - 1$

Wird dies in $\sum_{k=1}^{K} \lambda_k v_k = 0$ eingesetzt, so folgt auch $\lambda_K = 0$ $\square$

Definition 3.13

Sei z Eigenwert von $A \in \mathrm{Mat}(N \times N, \mathbb{K})$. Eigenraum

$$E_A(z) \;=\; \mathrm{Ker}(A - z\,\mathbf{1})$$

Geometrische Vielfachheit oder Multiplizität von z als Eigenwert von A:

$$\beta(z) \;=\; \beta_A(z) \;=\; \dim(E_A(z))$$

Satz 3.14 (Geometrische Multiplizität kleiner als algebraische)

$1 \leqslant \beta(z_k) \leqslant \alpha(z_k)$ *für alle Eigenwerte* $z_k \in \mathbb{K}$

Beweis. Sei $v_1, \ldots, v_{\beta(z_k)}$ Basis des Eigenraumes $E_A(z_k)$
Sie wird vervollständigt zu geordneter Basis $(v_1, \ldots, v_N)$ von $\mathbb{K}^N$
Sei $M = (v_1, \ldots, v_N)$. Dies ist invertierbar und

$$M^{-1}AM = \begin{pmatrix} z_k \mathbf{1}_{\beta(z_k)} & B \\ 0 & C \end{pmatrix}$$

wobei B und C Matrizen geeigneter Größe sind
Somit auch $M^{-1}AM - z\,\mathbf{1}_N$ von Blockstruktur. Also

$$\begin{aligned} p_A(z) &= p_{M^{-1}AM}(z) = \det(M^{-1}AM - z\,\mathbf{1}_N) \\ &= \det\left(\begin{pmatrix} (z_k - z)\,\mathbf{1}_{\beta(z_k)} & B \\ 0 & C - z\,\mathbf{1} \end{pmatrix}\right) \\ &= (z_k - z)^{\beta(z_k)} \det(C - z\,\mathbf{1}) \end{aligned}$$

Also folgt in der Tat $\alpha(z_k) \geqslant \beta(z_k)$. $\square$

Definition 3.15

$T \in \mathcal{L}(V)$ diagonalisierbar $\Longleftrightarrow \exists$ Basis $\mathcal{B}$ aus Eigenvektoren von T

Satz 3.16

Sei $T \in \mathcal{L}(V)$ auf einem Vektorraum V über $\mathbb{K}$ der Dimension $N < \infty$. Dann sind äquivalent:

(i) *T diagonalisierbar*

(ii) *$\alpha(z_k) = \beta(z_k)$ für alle Eigenwerte $z_k \in \mathbb{K}$, und für $\mathbb{K} = \mathbb{R}$ zudem $\sum_k \alpha(z_k) = N$ mit Summe über alle Eigenwerte $z_k \in \mathbb{K}$*

(iii) *$\sum_k \beta(z_k) = N$ mit Summe über alle Eigenwerte $z_k \in \mathbb{K}$*

Falls T diagonalisierbar, so ist darstellende Matrix bez. der geordneten Basis $M = (v_1, \dots, v_N)$ von Eigenvektoren zu den Eigenwerten $z'_1, \dots, z'_N$, aufgelistet mit ihrer Multiplizität, diagonal:

$$M^{-1}TM \;=\; (z'_m\,\delta_{n,m})_{n,m=1,\dots,N} \;=\; \mathrm{diag}(z'_1, \dots, z'_N)$$

Beweis.

(i)$\Longleftrightarrow$(iii) Klar

(iii)$\Longrightarrow$(ii) Immer (auch für $\mathbb{K} = \mathbb{R}$) gilt $\sum_k \alpha(z_k) \leqslant N$ und $\beta(z_k) \leqslant \alpha(z_k)$

(ii)$\Longrightarrow$(iii) Offensichtlich

Nun zum letzten Punkt. Sei e_m der m-te Standardbasisvektor

$$
\begin{aligned}
(M^{-1}TM)_{n,m} &= (M^{-1}TMe_m)_n \\
&= (M^{-1}Tv_m)_n \\
&= (M^{-1}z'_m v_m)_n \\
&= z'_m (M^{-1}Me_m)_n \\
&= z'_m (\mathbf{1}_N e_m)_n \\
&= z'_m \, \delta_{n,m}
\end{aligned}
$$

$\square$

Beispiel 3.17

Sei A die 3×3 Matrix aus Beispiel 3.10, wo auch deren Eigenwerte und Eigenvekoren (zumindestens einer davon) berechnet wurden. Dann setze

$$M = (v_1, v_2, v_3) = \begin{pmatrix} 1 & -2 & -2 \\ 0 & 1 - i\sqrt{2} & 1 + i\sqrt{2} \\ 1 & i\sqrt{2} & -i\sqrt{2} \end{pmatrix}$$

Es gilt dann nach Satz 3.16

$$M^{-1}AM = \begin{pmatrix} 1 & & 0 \\ 0 & i\sqrt{2} & 0 \\ 0 & 0 & -i\sqrt{2} \end{pmatrix}$$

Dies kann auch explizit verifiziert werden.

Funktionalkalkül durch Diagonalisierung

Definition 3.18 (Funktionalkalkül, auch Spektralkalkül)

Sei $A \in \mathrm{Mat}(N \times N, \mathbb{C})$ und f ganze Potenzreihe (konv. auf $\mathbb{C}$) der Form

$$f(z) = \sum_{k=0}^{\infty} f_k z^k \qquad , \qquad f_n \in \mathbb{C} \tag{3.3}$$

Dann ist die Matrix $f(A)$ definiert durch die konvergente Reihe

$$f(A) = \sum_{k=0}^{\infty} f_k A^k$$

wobei $A^0 = \mathbf{1}_N$

Man kann also z.B. e^A definieren, oder $\cos(A)$, oder e^{tA}, etc.

Wieso? Z.B. Gleichung $\partial_t v = Av$ wird durch $v(t) = e^{tA} v(0)$ gelöst

Berechnung? **Nicht** durch Berechnen von Matrixpotenzen A^k!

Satz 3.19

Sei A diagonalisierbar und $M^{-1}AM = \mathrm{diag}(z_1', \ldots, z_N')$. Dann

$$f(A) \;=\; M\,\mathrm{diag}(f(z_1'), \ldots, f(z_N'))\,M^{-1}$$

Beweis. Setze $D = \mathrm{diag}(z_1', \ldots, z_N')$. Beginne mit Monom $f(z) = z^k$ Dann ist $f(A) = A^k$. Somit

$$f(A) \;=\; \left(MDM^{-1}\right)^k \;=\; MD^kM^{-1} \;=\; M\mathrm{diag}((z_1')^k, \ldots, (z_N')^k)M^{-1}$$

Für beliebige Potenzreihe gilt (Konvergenz nach Restabschätzung)

$$f(A) \;=\; \sum_{k=0}^{\infty} f_k\, MD^kM^{-1} \;=\; M\Big(\sum_{k=0}^{\infty} f_k\, D^k\Big)M^{-1} \;=\; Mf(D)M^{-1}$$

Dies ist genau die Aussage, denn $f(D)$ ist die angegebene Matrix. $\square$

Beispiel 3.20

Seien A und M die 3×3 Matrizen aus Beispiel 3.17

Es soll nun e^A berechnet werden, d.h. $f(z) = e^z$ Exponentialfunktion:

$$f(A) \;=\; Mf(D)M^{-1} \;=\; M \begin{pmatrix} e^1 & 0 & 0 \\ 0 & e^{i\sqrt{2}} & 0 \\ 0 & 0 & e^{-i\sqrt{2}} \end{pmatrix} M$$

Nun M bekannt. Hieraus M^{-1} durch Invertieren berechnen (mit Gauss Algorithmus, oder mit Adjunkten). Danach Matrixprodukt

Bemerkung 3.21 (Jordan Block)

Nach Satz 3.8 faktorisiert jedes charakteristische Polynom über $\mathbb{C}$.

Somit für $\mathbb{C}$-lineare Abbildung $\sum_k \alpha(z_k) = N$

Aber: Abbildung nicht notwendigerweise diagonalisierbar ist! Beispiel:

$$A = (z_1\, \delta_{n,m} + \delta_{n+1,m})_{n,m=1,\dots,N} = \begin{pmatrix} z_1 & 1 & 0 & & 0 \\ 0 & z_1 & 1 & 0 & \\ & \ddots & \ddots & \ddots & \\ & & 0 & z_1 & 1 \\ 0 & & & 0 & z_1 \end{pmatrix}$$

Dann $p_A(z) = (z_1 - z)^N$ mit $\alpha(z_1) = N$. Aber nur $\beta(z_1) = 1$:

$$\mathrm{Ker}(A - z_1\, \mathbf{1}) = \mathrm{Ker} \begin{pmatrix} 0 & 1 & 0 & \\ & & \ddots & \ddots \\ & & 0 & 1 \\ 0 & & & 0 \end{pmatrix} = \mathrm{span}(e_1)$$

Was immer geht über $\mathbb{C}$:

Definition 3.22

$A \in \mathrm{Mat}(N \times N, \mathbb{K})$ trigonalisierbar $\iff$ $\exists$ invertierbares $M \in \mathrm{Mat}(N \times N, \mathbb{K})$, so dass $M^{-1}AM$ obere Dreiecksmatrix

Bemerkung 3.23

Trigonalisierung von A nicht die obere Dreiecksmatrix nach Gauss!
Dort ist M, so MA obere Dreiecksmatrix, und nicht $M^{-1}AM$

Satz 3.24

Es sind äquivalent:

(i) *A trigonalisierbar*

(ii) *Charakteristische Polynom faktorisiert in $\mathbb{K}$*

(iii) $\exists$ *Folge von Unterräumen $\{0\} = V_0 \subset V_1 \subset \ldots \subset V_{N-1} \subset V_N = V$ mit $\dim(V_n) = n$ und $AV_n \subset V_n$, d.h. V_n ist A-invariant*

Beweis. (i)$\Longrightarrow$(ii) $p_A(z) = \det(A - z\mathbf{1}_N) = \det(M^{-1}AM - z\mathbf{1}_N)$
Produkt der Diagonaleinträge, also faktorisiert

(ii)$\Longrightarrow$(iii) Folge heißt auch Fahne. Existenz durch Induktion über N

Induktionsanfang $N = 1$ klar. Induktionsschritt $N - 1 \to N$

Sei $z_1 \in \mathbb{K}$ Eigenwert mit Eigenvektor $v_1 \in \mathbb{K}^N$ (nach Voraussetzung)

Vervollständige zu Basis $\Longrightarrow \; M = (v_1, \ldots, v_N) \; \Longrightarrow$

$$M^{-1}AM = \begin{pmatrix} z_1 & B \\ 0 & C \end{pmatrix} \qquad C \in \mathrm{Mat}((N-1) \times (N-1), \mathbb{K})$$

Induktion: $\exists$ invertierbares K, so dass $K^{-1}CK$ obere Dreiecksmatrix

$$\begin{pmatrix} 1 & 0 \\ 0 & K \end{pmatrix}^{-1} M^{-1}AM \begin{pmatrix} 1 & 0 \\ 0 & K \end{pmatrix} = \begin{pmatrix} z_1 & BK \\ 0 & K^{-1}CK \end{pmatrix}$$

Fahne von A ist $V_n = M \begin{pmatrix} 1 & 0 \\ 0 & K \end{pmatrix} \mathrm{span}(e_1, \ldots, e_n)$ (Check!)

(iii)$\Longrightarrow$(i) Wähle $v_n \in V_n \backslash V_{n-1} \Longrightarrow \; v_1, \ldots, v_N$ linear unabhängig

$\Longrightarrow$ mit $M = (v_1, \ldots, v_N)$ ist $M^{-1}AM$ obere Dreiecksmatrix $\square$

Korollar 3.25

$A \in \mathrm{Mat}(N \times N, \mathbb{C})$ ist über $\mathbb{C}$ trigonalisierbar,
d.h. $\exists$ invertierbares M, so dass $M^{-1}AM$ eine obere Dreiecksmatrix

weil charakteristisches Polynom faktorisiert über $\mathbb{C}$ $\square$

Anwendung der Trigonalisierung:

Theorem 3.26 (Satz von Cayley-Hamilton)

Für jedes $A \in \mathrm{Mat}(N \times N, \mathbb{C})$ gilt

$$p_A(A) \; = \; 0$$

wobei $p_A(A)$ durch Funktionalkalkül definiert ist.

Beweis. Sei $M \in \mathrm{Mat}(N \times N, \mathbb{C})$ Basiswechsel so dass $B = M^{-1}AM$ obere Dreiecksmatrix

$$p_A(A) \;=\; Mp_A(B)M^{-1} \;=\; Mp_B(B)M^{-1}$$

Also: es reicht Sachverhalt für obere Dreiecksmatrizen zu zeigen!

Induktion über N. Für $N = 1$ trivial. Für Schritt von $N-1$ nach N sei

$$B \;=\; \begin{pmatrix} \lambda & C \\ 0 & D \end{pmatrix} \qquad , \qquad \lambda \in \mathbb{C}$$

mit oberer Dreiecksmatrix $D \in \mathrm{Mat}((N-1) \times (N-1), \mathbb{K})$

Induktionsvoraussetzung: $p_D(D) = 0$. Zudem $p_B(z) = (\lambda - z)p_D(z)$

Nun folgt (Details zur Übung) für eine geeignete Matrix C':

$$p_B(B) = (\lambda \mathbf{1}_N - B)p_D(B) = \begin{pmatrix} 0 & -C \\ 0 & \lambda\mathbf{1} - D \end{pmatrix} \begin{pmatrix} p_D(\lambda) & C' \\ 0 & p_D(D) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -C \\ 0 & \lambda\mathbf{1} - D \end{pmatrix} \begin{pmatrix} p_D(\lambda) & C' \\ 0 & 0 \end{pmatrix} = 0$$

□

Weitere Anwendung der Trigonalisierung

Satz 3.27

$$\det(e^A) \;=\; e^{\mathrm{Tr}(A)} \qquad , \qquad A \in \mathrm{Mat}(N \times N, \mathbb{C})$$

Beweis. Sei $B = M^{-1}AM$ obere Dreiecksmatrix

Dann $B = D + S$ wobei D diagonal, S strikte obere Dreiecksmatrix

Dann $B^n = D^n + S_n$ und $e^B = e^D + S'$ mit strikten o.D. S_n, S'

$$\begin{aligned}
\det(e^A) &= \det(e^{MBM^{-1}}) \\
&= \det(Me^BM^{-1}) \\
&= \det(e^B) \\
&= \det(e^D) \\
&= e^{\mathrm{Tr}(D)} \\
&= e^{\mathrm{Tr}(D+S)} \;=\; e^{\mathrm{Tr}(M(D+S)M^{-1})} \;=\; e^{\mathrm{Tr}(A)}
\end{aligned}$$

$\square$

Definition 3.28 (Spektrum einer linearen Abbildung)

Das Spektrum $\mathrm{Spec}(T) \subset \mathbb{K}$ von $T \in \mathcal{L}(V)$ ist die Menge aller $z \in \mathbb{K}$, für welche $T - z\,\mathbf{1}_V$ kein Isomorphismus von V ist

Satz 3.29

Sei $T \in \mathcal{L}(V)$ auf einem Vektorraum V über $\mathbb{K}$ mit endlicher Dimension $N = \dim(V) < \infty$. Dann gilt

$$\mathrm{Spec}(T) \; = \; \{z \in \mathbb{K} \,|\, z \text{ Eigenwert von } T\}$$

Falsch falls $\dim(V) = \infty$ weil es kontinuierliches Spektrum gibt!

Beweis. "$\supset$" z Eigenwert $\Longrightarrow$ $T - z\,\mathbf{1}_V$ hat nicht-trivialen Kern

"$\subset$" $T - z\,\mathbf{1}_V$ kein Isomorphismus $\Longrightarrow$ nicht injektiv oder nicht surjektiv

Falls $T - z\,\mathbf{1}_V$ nicht surjektiv nach Rangsatz auch nicht injektiv

Falls nicht injektiv $\exists$ Eigenvektor von T zum Eigenwert z □

4 Vektorräume mit Skalarprodukt

Definition 4.1

Sei V Vektorraum über $\mathbb{K}$. Skalarprodukt ist Abbildung

$\langle\cdot|\cdot\rangle : V \times V \to \mathbb{K} \iff \forall\ v, v', w \in V$ und $\lambda \in \mathbb{K}$ gilt:

(i) $\langle w|v + \lambda v'\rangle = \langle w|v\rangle + \lambda\langle w|v'\rangle$ (Linearität im zweiten Argument)

(ii) $\langle w|v\rangle = \overline{\langle v|w\rangle}$ (Symmetrie)

(iii) $\langle v|v\rangle \geqslant 0$ (Positivität)

(iv) $\langle v|v\rangle = 0 \implies v = 0$ (Nicht-Entartung)

Dann Norm (auch Länge) definiert als $\|v\| = \sqrt{\langle v|v\rangle}$

Vektorraum mit Skalarprodukt $(V, \langle\cdot|\cdot\rangle)$ heißt

- euklidisch falls $\mathbb{K} = \mathbb{R}$
- unitär falls $\mathbb{K} = \mathbb{C}$
- Hilbert-Raum falls vollständig, z.B. wenn endlich dimensional
- Prä-Hilbertraum

Bemerkung 4.2

Im Fall $\mathbb{K} = \mathbb{R}$: Eigenschaften (i) und (ii) $\Longleftrightarrow$ Skalarprodukt bilinear

Im Fall $\mathbb{K} = \mathbb{C}$ hingegen: Skalarprodukt anti-linear im ersten Argument

$$\langle \lambda v | w \rangle = \overline{\lambda} \langle v | w \rangle$$

Mathematikliteratur: oft Antilinearität im zweiten Argument

Hier wie in Physik und insbesondere der Quantenmechanik

Dirac'sche BraKet Notation auch aus Physik:

Ein Bra ist $\langle v|$ und ein Ket ist $|w\rangle$. Zusammen *bracket=Klammer* $\langle v|w\rangle$

Ket ist Vektor in V, Bra Vektor aus Dualraum $V^* = \mathcal{L}(V, \mathbb{K})$

Andere Notationen in Mathematikliteratur: $\langle v, w\rangle$, $(v|w)$...

Beispiel 4.3

Sei $V = \mathbb{K}^N$. Das euklidische Skalarprodukt ist dann definiert als

$$\langle v|w\rangle = (\overline{v})^T w = \sum_{n=1}^{N} \overline{v_n}\, w_n \qquad , \qquad v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} , \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}$$

In der Tat gelten alle vier Axiome, z.B. (iii) ist richtig, weil

$$\langle v|v\rangle = \sum_{n=1}^{N} \overline{v_n}\, v_n = \sum_{n=1}^{N} |v_n|^2$$

als Summe nicht-negativer Zahlen nicht-negativ ist

Auch folgt (iv), da:

Summe positiv $\Longleftrightarrow$ ein Summand positiv $\Longleftrightarrow$ $v \neq 0$

Beispiel 4.4

Sei weiter $V = \mathbb{K}^N$ und $B \in \mathrm{Mat}(N \times N, \mathbb{K})$. Wir setzen

$$\langle v|w\rangle_B \;=\; (\overline{v})^T B w \;=\; \sum_{n=1}^{N} \overline{v_n}\, B_{n,m}\, w_m \qquad , \qquad v, w \in \mathbb{K}^N$$

Spezialfall $B = \mathbf{1}_N$ wieder euklidisches Skalarprodukt

Axiom (i) klar, aber Symmetrie (ii) benötigt Zusatzeigenschaft

$$(\overline{B})^T \;=\; B \qquad \text{d.h. } B \text{ selbstadjungiert}$$

In der Tat:

$$\begin{aligned} \langle v|w\rangle_B \;&=\; \sum_{n,m=1}^{N} \overline{v_n}\, B_{n,m}\, w_m \;=\; \sum_{n,m=1}^{N} \overline{v_n}\, \overline{B_{m,n}}\, w_m \\ &=\; \overline{\sum_{n=1}^{N} \overline{w_n}\, B_{n,m}\, v_m} \;=\; \overline{\langle w|v\rangle_B} \end{aligned}$$

Eigenschaften (iii),(iv) $\Longleftrightarrow$ Positivität der Matrix B (Eigenwerte positiv)

Hierzu später mehr.

Beispiel 4.5

Sei $V = \mathbb{R}_N[x]$ der reelle Vektorraum der Polynome mit reellen Koeffizienten vom Grade höchstens N. Dann $\dim(V) = N + 1$

Skalarprodukt definiert (Übung!) mit Integral:

$$\langle p|q\rangle = \int_0^1 p(x)q(x)\,dx \qquad , \qquad p, q \in \mathbb{R}_N[x]$$

Bemerkung 4.6

$\dim(V) < \infty \implies \exists$ Skalarprodukt

Zu beliebiger Basis $(b_1, \ldots, b_N)$ definiere Skalarprodukt durch

$$\langle b_n|b_m\rangle = \delta_{n,m}$$

und anschließend antilineare bzw. lineare Fortsetzung auf $V \times V$, d.h.

$$\langle v|w\rangle = \sum_{n=1}^{N} \overline{v_n}\, w_n \qquad , \qquad v = \sum_{n=1}^{N} v_n b_n\,, \quad w = \sum_{n=1}^{N} w_n b_n$$

Definition 4.7 (Begriffe in Prä-Hilbertraum $(V, \langle\cdot|\cdot\rangle)$)

(i) $v, w \in V$ orthogonal (Kurzschreibweise $v \perp w$) $\iff$ $\langle v|w\rangle = 0$

(ii) $v \in V$ normiert $\iff$ $\|v\| = \sqrt{\langle v|v\rangle} = 1$

(iii) $(u_i)_{i\in I}$ orthogonale Familie $\iff$ $u_i \perp u_j \quad \forall\ i,j \in I,\ i \neq j$

(iv) $(u_i)_{i\in I}$ orthonormierte Familie

$\iff$ $(u_i)_{i\in I}$ orthogonale Familie und $\|u_i\| = 1\ \forall\ i \in I$

$\iff$ $\langle u_i|u_j\rangle = \delta_{i,j} \ \ \forall\ \ i,j \in I$

(v) Orthonormalbasis ist Basis, die auch orthonormierte Familie ist

Beispiel 4.8

In $\mathbb{K}^N$ mit dem euklidischen Skalarproduk, ist die Standardbasis eine Orthonormalbasis (Übung)

Satz 4.9

Seien $(u_k)_{k=1,\dots,K}$ *orthonormiert. Dann gilt für alle* $v \in V$*:*

$$\|v\|^2 \;=\; \sum_{k=1}^{K} |\langle v|u_k\rangle|^2 + \|v - \sum_{k=1}^{K}\langle u_k|v\rangle u_k\|^2 \qquad \text{(Pythagoras)}$$

und somit

$$\|v\|^2 \;\geqslant\; \sum_{k=1}^{K} |\langle v|u_k\rangle|^2 \qquad \text{(Bessel Ungleichung)}$$

Korollar 4.10 (Cauchy-Schwarz Ungleichung)

Sei V *ein Vektorraum mit Skalarprodukt und seien* $v, w \in V$*. Dann*

$$|\langle v|w\rangle| \;\leqslant\; \|v\|\|w\|$$

Beweis: Fall $w = 0$ trivial. Sonst Bessel Ungleichung für $\{\frac{w}{\|w\|}\}$. $\square$

Beweis: Setze $w = \sum_{k=1}^{K} \langle u_k | v \rangle u_k$. Dann gilt $(v - w) \perp w$, da

$$\begin{aligned}
\langle v - w | w \rangle &= \sum_{k=1}^{K} \langle u_k | v \rangle \langle v - w | u_k \rangle \\
&= \sum_{k=1}^{K} \langle u_k | v \rangle \left(\langle v | u_k \rangle - \sum_{l=1}^{K} \overline{\langle u_l | v \rangle} \underbrace{\langle u_l | u_k \rangle}_{\delta_{l,k}} \right) = 0
\end{aligned}$$

Geometrisch: w orthogonale Projektion von v auf $\operatorname{span}\{u_1, \ldots, u_K\}$

Somit:

$$\begin{aligned}
\|v\|^2 &= \langle v | v \rangle \\
&= \langle v - w + w | v - w + w \rangle \\
&= \langle v - w | v - w \rangle + \langle w | w \rangle
\end{aligned}$$

und $\langle w | w \rangle = \|w\|^2 = \sum_{k=1}^{K} |\langle u_k | v \rangle|^2$ $\square$

Bemerkung 4.11 (Winkel zwischen zwei Vektoren)

Zu $v, w \in V$ ist Winkel $\sphericalangle(v, w) \in [0, \frac{\pi}{2}]$ definiert durch

$$\cos(\sphericalangle(v, w)) \;=\; \frac{|\langle v|w\rangle|}{\|v\|\|w\|}$$

Cauchy-Schwarz: rechte Seite in $[0, 1]$ und $\cos : [0, \frac{\pi}{2}] \to [0, 1]$ bijektiv

Satz 4.12

$(V, \langle\cdot|\cdot\rangle)$ Vektorraum mit Skalarprodukt. Norm $\|v\| = \sqrt{\langle v|v\rangle}$ erfüllt:

(i) $\|\lambda v\| \;=\; |\lambda|\, \|v\|$ (Homogenität)

(ii) $\|v + w\| \leqslant \|v\| + \|w\|$ (Dreiecksungleichung)

(iii) $\|v\| = 0 \implies v = 0$ (Nicht-Entartung)

Definition 4.13

Ein Vektorraum V mit $\|.\| : V \to \mathbb{R}_{\geqslant 0}$ und (i),(ii),(iii) heißt normiert

Beweis: $\|\lambda v\| = |\lambda|\|v\|$ klar, ebenso Nicht-Entartung

Nun zur Dreiecksungleichung: Mit Cauchy-Schwarz Ungleichung

$$\begin{aligned}\|v+w\|^2 &= \langle v|v\rangle + \langle v|w\rangle + \langle w|v\rangle + \langle w|w\rangle \\ &= \|v\|^2 + 2\,\Re e\,\langle v|w\rangle + \|w\|^2 \\ &\leqslant \|v\|^2 + 2\,\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2\end{aligned}$$

□

Korollar 4.14

Für alle $v \in V$

$$\|v\| = \sup_{\|w\|=1} |\langle w|v\rangle|$$

Beweis: Nach Cauchy-Schwarz $|\langle w|v\rangle| \leqslant \|w\|\,\|v\| = \|v\|$

Somit Supremum auch $\leqslant \|v\|$

Andererseits führt Wahl $w = \frac{v}{\|v\|}$ zur Gleichheit □

Nun zu weiteren geometrischen Informationen.

Satz 4.15 (Parallelogrammregel)

Sei $(V, \langle\cdot|\cdot\rangle)$ *ein Vektorraum mit Skalarprodukt. Dann gilt*

$$\|v+w\|^2 \; + \; \|v-w\|^2 \; = \; 2(\|v\|^2 \; + \; \|w\|^2)$$

Beweis: Rein algebraische Rechnung:

$$\begin{aligned}
&\|v+w\|^2 + \|v-w\|^2 \\
&\quad = \; \langle v+w|v+w\rangle \; + \; \langle v-w|v-w\rangle \\
&\quad = \; \langle v|v\rangle + \langle v|w\rangle + \langle w|v\rangle + \langle w|w\rangle + \langle v|v\rangle - \langle v|w\rangle - \langle w|v\rangle + \langle w|w\rangle \\
&\quad = \; 2\langle v|v\rangle \; + \; 2\langle w|w\rangle
\end{aligned}$$

□

Satz 4.16 (Polarisationsidentitäten = Skalarprodukt durch $\|.\|$)

Sei $(V, \langle\cdot|\cdot\rangle)$ *ein Vektorraum mit Skalarprodukt. Für* $\mathbb{K} = \mathbb{C}$

$$\langle v|w\rangle = \frac{1}{4}\left[(\|v+w\|^2 - \|v-w\|^2) - i(\|v+iw\|^2 - \|v-iw\|^2)\right]$$

und für $\mathbb{K} = \mathbb{R}$*:*

$$\langle v|w\rangle = \frac{1}{4}\,[\|v+w\|^2 - \|v-w\|^2]$$

Beweis wieder nur algebraische Rechnung. Schwieriger ist:

Satz 4.17 (Charakterisierung nach John von Neumann)

Norm eines normierten Vektorraumes durch Skalarprodukt induziert

$\iff$ *Norm erfüllt Parallelogrammregel*

Das Skalarprodukt ist dann durch Polarisationsidentität gegeben

Satz 4.18

$(V, \langle\cdot|\cdot\rangle)$ endlich-dimensionaler Vektorraum mit Skalarprodukt $(u_1, \ldots, u_N)$ Orthonormalbasis. Koordinatenabbildung $M : V \to \mathbb{K}^N$

$$M(v) \;=\; \begin{pmatrix} \langle u_1|v\rangle \\ \vdots \\ \langle u_N|v\rangle \end{pmatrix} \quad \Longleftrightarrow \quad v = \sum_{n=1}^{N} \langle u_n|v\rangle\, u_n$$

Außerdem

$$\langle v|w\rangle \;=\; \langle M(v)|M(w)\rangle \qquad , \qquad v, w \in V$$

wobei rechts euklidisches Skalarprodukt im $\mathbb{K}^N$ steht

Kein Gauss für Zerlegung nach ONB, nur Skalarprodukte!

Umschreiben:

$$v \;=\; |v\rangle \;=\; \sum_{n=1}^{N} |u_n\rangle\langle u_n|v\rangle \quad \Longleftrightarrow \quad \mathbf{1} \;=\; \sum_{n=1}^{N} |u_n\rangle\langle u_n|$$

Beweis: Bilde Skalarprodukt von $v = \sum_{k=1}^{N} \lambda_k \, u_k$ mit u_n:

$$\langle u_n | v \rangle \; = \; \sum_{k=1}^{N} \lambda_k \, \langle u_n | u_k \rangle \; = \; \lambda_n$$

Zweite Behauptung:

$$\begin{aligned}
\langle v | w \rangle \; &= \; \langle \sum_{k=1}^{N} \langle u_k | v \rangle \, u_k \; | \; \sum_{n=1}^{N} \langle u_n | w \rangle \, u_n \rangle \\
&= \; \sum_{k,n=1}^{N} \overline{\langle u_k | v \rangle} \langle u_n | w \rangle \, \delta_{n,k} \\
&= \; \sum_{n=1}^{N} \langle v | u_n \rangle \langle u_n | w \rangle \\
&= \; \langle M(v) | M(w) \rangle
\end{aligned}$$

$\square$

Standardverfahren zur Erzeugung von Orthonormalbasen in $(V, \langle\cdot|\cdot\rangle)$:

Satz 4.19 (Gram-Schmidt Verfahren)

$v_1, \ldots, v_N$ linear unabhängiger Vektoren in V
$\Longrightarrow \exists$ orthonormierte Familie $u_1, \ldots, u_N$ mit

$$\operatorname{span}(\{v_1, \ldots, v_n\}) = \operatorname{span}(\{u_1, \ldots, u_n\}) \quad , \quad n = 1, \ldots, N$$

Die orthonormierte Familie kann iterativ konstruiert werden durch

$$u_n = \frac{v_n - \sum_{k=1}^{n-1} \langle u_k | v_n \rangle \, u_k}{\| v_n - \sum_{k=1}^{n-1} \langle u_k | v_n \rangle \, u_k \|}$$

Zusammen mit dem Basiserweiterungssatz erhält man:

Korollar 4.20

Jede orthonormierte Familie kann zu Orthonormalbasis erweitert werden

Beweis: Setze $w_n = \sum_{k=1}^{n-1} \langle u_k | v_n \rangle \, u_k$

w_n ist Projektion von v_n auf von $u_1, \ldots, u_{n-1}$ aufgespannten Unterraum weil $(v_n - w_n) \perp u_k$ für alle $k = 1, \ldots, n-1$:

$$\begin{aligned} \langle u_k | v_n - w_n \rangle &= \langle u_k | v_n \rangle - \sum_{l=1}^{n-1} \langle u_l | v_n \rangle \langle u_k | u_l \rangle \\ &= \langle u_k | v_n \rangle - \sum_{l=1}^{n-1} \langle u_l | v_n \rangle \, \delta_{k,l} \;=\; 0 \end{aligned}$$

Also bilden $u_1, \ldots, u_{n-1}, v_n - w_n$ eine orthogonale Familie

Normieren von $v_n - w_n$ ergibt u_n und somit orthonormierte Familie □

Bemerkung 4.21

Mit Hilfe des Gram-Schmidt Verfahrens werden sämtliche Familien orthogonaler Polynome erzeugt (Legendre, Hermite, Laguerre, etc.)

Beispiel 4.22 (in euklidischen $V = \mathbb{R}^3$)

$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ Basis. Bilde $u_1 = \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

Dann $w_2 = v_2 - \langle v_2 \mid u_1 \rangle \, u_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$

und $u_2 = \frac{w_2}{\|w_2\|} = \frac{1}{\sqrt{\frac{1}{4} + \frac{1}{4} + 1}} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix} = \frac{2}{\sqrt{6}} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{pmatrix}$

Analog: $w_3 = v_3 - \langle v_3 \mid u_1 \rangle \, u_1 - \langle v_3 \mid u_2 \rangle \, u_2 = \dots = \begin{pmatrix} -\frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$

und $u_3 = \frac{w_3}{\|w_3\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ Dann ist (u_1, u_2, u_3) ONB

Definition 4.23 (Orthogonale Projektionen in $(V, \langle\cdot|\cdot\rangle)$)

(i) Eine lineare Abbildung $P \in \mathcal{L}(V)$ heißt idempotent $\Longleftrightarrow P^2 = P$

(ii) Ein idempotentes $P \in \mathcal{L}(V)$ heißt (orthogonale) Projektion $\Longleftrightarrow \langle v|Pw\rangle = \langle Pv|w\rangle$ für alle $v, w \in V$

(iii) Dimension einer Projektion ist die Dimension des Bildes $\mathrm{Ran}(P)$

(iv) Orthogonales Komplement zu P ist die Projektion $\mathbf{1} - P$

Beispiel 4.24

Eine 2×2-Matrizen, die idempotent, aber nicht Projektion ist:

$$P = \begin{pmatrix} \frac{1}{2} - \sqrt{\frac{1}{4} - bc} & b \\ c & \frac{1}{2} + \sqrt{\frac{1}{4} - bc} \end{pmatrix} , \qquad |bc| \leqslant \frac{1}{4}$$

Satz 4.25 (Eigenschaften von Projektion P)

(i) *Für* $w \in \mathrm{Ran}(P)$ *gilt* $Pw = w$

(ii) *Für* $w \perp \mathrm{Ran}(P)$*, d.h.* $w \perp v$ *für alle* $v \in \mathrm{Ran}(P)$*, gilt* $Pw = 0$

(iii) *Sei* $u_1, \dots, u_M$ *Orthonormalbasis von* $\mathrm{Ran}(P)$*. Dann*

$$P = \sum_{m=1}^{M} |u_m\rangle\langle u_m|$$

Beweis: (i) Sei $w = Pv$. Dann gilt (nur idempotent verwandt!)

$$Pw - w = P^2 v - Pv = Pv - Pv = 0$$

(ii) Da $\langle Pw|v\rangle = \langle w|Pv\rangle = \langle w|v\rangle = 0$ gilt $Pw \perp \mathrm{Ran}(P)$

Zudem $Pw \in \mathrm{Ran}(P)$, also $Pw = 0$

(iii) Rechte Seite erfüllt (i) und (ii) und stimmt somit mit P überein $\square$

Umformulierung:

Satz 4.26 (Gram-Schmidt Verfahren)

$v_1, \dots, v_N$ linear unabhängiger Vektoren in V
$\Longrightarrow \exists$ orthonormierte Familie $u_1, \dots, u_N$ mit

$$\mathrm{span}(\{v_1, \dots, v_n\}) \;=\; \mathrm{span}(\{u_1, \dots, u_n\}) \qquad , \qquad n = 1, \dots, N$$

Die orthonormierte Familie kann iterativ konstruiert werden durch

$$u_n \;=\; \frac{v_n - P_{n-1} v_n}{\|v_n - P_{n-1} v_n\|}$$

wobei P_{n-1} die Projektion auf den von $u_1, \dots, u_{n-1}$ augespannten UR:

$$P_{n-1} \;=\; \sum_{k=1}^{n-1} |u_k\rangle\langle u_k|$$

Satz 4.27 (Riesz Lemma für endlich dimensionales $(V, \langle\cdot|\cdot\rangle)$)

Zu jedem linearen Funktional $T \in \mathcal{L}(V, \mathbb{K})$ existiert eindeutiger Vektor $w_T \in V$ mit

$$T(v) = \langle w_T | v \rangle \quad , \quad \forall \ v \in V$$

Also ist Dualraum $V^* = \mathcal{L}(V, \mathbb{K})$ isomorph zu V

Beweis: T durch Werte auf ONB $u_1, \ldots, u_N$ festgelegt. Wähle

$$w_T = \sum_{k=1}^{N} \overline{T(u_k)}\, u_k$$

Einsetzen zeigt dann $\langle w_T | u_n \rangle = T(u_n)$ für $n = 1, \ldots, N$

Nun sind beide Seiten in $T(v) = \langle w_T | v \rangle$ linear in v

Übereinstimmung auf Basis $\implies$ gleich als lineare Abbildungen $\square$

Satz 4.28 (Adjungierte lineare Abbildung)

Seien V, W, U endlich dimensionale Vektorräume mit Skalarprodukt

Zu linearem $T \in \mathcal{L}(V, W)$ existiert eindeutiges $T^* \in \mathcal{L}(W, V)$ mit

$$\langle T^* w | v \rangle = \langle w | T v \rangle \qquad , \qquad v \in V \, , \quad w \in W$$

Falls $S \in \mathcal{L}(W, U)$, dann $(ST)^* = T^* S^*$

Satz 4.29 (Adjungierte Matrix ist adjunigierte lineare Abbildung)

Spezialfall: $V = \mathbb{K}^M$ und $W = \mathbb{K}^N$ euklidisch

$T = A \in \mathrm{Mat}(N \times M, \mathbb{K})$ lineare Abbildung durch Matrixmultiplikation

$\implies \quad T^* = A^* \in \mathrm{Mat}(M \times N, \mathbb{K})$ wobei adjungierte Matrix gegeben als

$$(A^*)_{m,n} = \overline{A_{n,m}} \qquad \iff \qquad A^* = \overline{A^T} = (\overline{A})^T$$

Für Matrix $B \in \mathrm{Mat}(M \times K, \mathbb{K})$ gilt $(AB)^* = B^* A^*$

Beweis von Satz 4.28: Nur Existenz von T^*, Eindeutigkeit Übung

Für festes w ist Abbildung $v \in V \mapsto \langle w|Tv\rangle$ lineares Funktional

Nach Riesz Lemma existiert $v_w \in V$, so dass

$$\langle w|Tv\rangle = \langle v_w|v\rangle$$

Setze $T^*w = v_w$. So definiertes T^* ist tatsächlich linear:

$$\begin{aligned}\langle T^*(w+\lambda w')|v\rangle &= \langle v_{w+\lambda w'}|v\rangle = \langle w+\lambda w'|Tv\rangle = \langle w|Tv\rangle + \overline{\lambda}\langle w'|Tv\rangle \\ &= \langle v_w|v\rangle + \overline{\lambda}\langle v_{w'}|v\rangle = \langle T^*w|v\rangle + \overline{\lambda}\langle T^*w'|v\rangle\end{aligned}$$

Also

$$\langle T^*(w+\lambda w') - T^*w - \lambda T^*w'|v\rangle = 0 \quad , \quad \forall \ v \in V$$

Also Linearität von T^* wegen Lemma. Regel $(ST)^* = T^*S^*$ dann klar

Lemma 4.30

Falls $\langle w|v\rangle = 0$ für alle $v \in V$, so gilt $w = 0$

Beweis: Insbesondere $\langle w|w\rangle = 0$ und somit $w = 0$ □

Beweis von Satz 4.29:

Regel $(ST)^* = T^*S^*$ folgt nach Ausschreiben

Spezialfall: sei $v = \sum_{m=1}^{M} v_m e_m \in \mathbb{K}^M$ und $w = \sum_{n=1}^{N} w_n e_n \in \mathbb{K}^N$

$$\begin{aligned}
\langle w|Av\rangle &= \sum_{n=1}^{N}\sum_{m=1}^{M} \overline{w_n}\, v_m \langle e_m|Ae_n\rangle = \sum_{n=1}^{N}\sum_{m=1}^{M} \overline{w_n}\, v_m\, A_{m,n} \\
\langle A^* w|v\rangle &= \sum_{n=1}^{N}\sum_{m=1}^{M} \overline{w_n}\, v_m \langle A^* e_m|e_n\rangle = \sum_{n=1}^{N}\sum_{m=1}^{M} \overline{w_n}\, v_m\, \overline{(A^*)_{n,m}}
\end{aligned}$$

Also $\langle w|Av\rangle = \langle A^* w|v\rangle$ für alle v, w.

Schließe mit Eindeutigkeit der adjungierten Abbildung $\square$

Definition 4.31

V und W endlich dimensionale Prä-Hilbert-Räume über $\mathbb{K}$

(i) $U \in \mathcal{L}(V, W)$ heißt unitär $\Longleftrightarrow U^* U = \mathbf{1}_V$ und $UU^* = \mathbf{1}_W$

(ii) Falls $\mathbb{K} = \mathbb{R}$, heißen unitäre Abbildungen auch orthogonal

(iii) $T \in \mathcal{L}(V) = \mathcal{L}(V, V)$ heißt normal $\Longleftrightarrow T^* T = TT^*$

(iv) $T \in \mathcal{L}(V)$ heißt selbstadjungiert oder hermitisch $\Longleftrightarrow T^* = T$

Bei Matrizen wird euklidisches Skalarprodukt verwandt, d.h.

(i) $U \in \mathrm{Mat}(N \times N, \mathbb{K})$ unitär $\Longleftrightarrow U^* U = \mathbf{1}$ und $UU^* = \mathbf{1}$

(ii) Falls $\mathbb{K} = \mathbb{R}$, heißen unitäre Abbildungen auch orthogonal

(iii) $A \in \mathrm{Mat}(N \times N, \mathbb{K})$ normal $\Longleftrightarrow A^* A = AA^*$

(iv) $A \in \mathrm{Mat}(N \times N, \mathbb{K})$ selbstadjungiert oder hermitisch $\Longleftrightarrow A^* = A$

Bemerkung 4.32

1. Jede selbstadjungierte Abbildung ist normal
2. Jede unitäre Abbildung ist normal
3. Für jede unitäre Abbildung $U \in \mathcal{L}(V, W)$ gilt $U^* = U^{-1}$
   Insbesondere: $U^* U = \mathbf{1}_V$ impliziert $UU^* = \mathbf{1}_W$
4. Jede unitäre Abbildung $U \in \mathcal{L}(V, W)$ erhält das Skalarprodukt:
   $$\langle Uv | Uw \rangle = \langle v | w \rangle \qquad , \qquad \forall \; v, w \in V$$
   Umgekehrt: $U \in \mathcal{L}(V, W)$ erhält Skalarprodukte $\Longrightarrow$ U unitär
5. $(u_1, \ldots, u_N)$ ONB von $\mathbb{K}^N \Longrightarrow U = (u_1, \ldots, u_N)$ unitär, da
   $$\langle e_n | U^* U e_m \rangle = \langle U e_n | U e_m \rangle = \langle u_n | u_m \rangle = \delta_{n,m}$$
6. $U(N) = \{U \in \mathrm{Mat}(N \times N, \mathbb{C}) \text{ unitär}\}$ unitäre Gruppe
   da: U, U' unitär $\Longrightarrow (UU')^*(UU') = (U')^* U^* U U' = (U')^* U' = \mathbf{1}$
   $O(N) = \{U \in \mathrm{Mat}(N \times N, \mathbb{R}) \text{ orthogonal}\}$ orthogonale Gruppe

Zusammenhang zwischen hermitisch und unitär

Satz 4.33

Sei $A \in \mathrm{Mat}(N \times N, \mathbb{C})$ selbstadjungiert, d.h. $A^* = A$. Dann

$$e^{iA} \quad \text{unitär}$$

Beweis: $(e^{iA})^* = e^{-iA}$ und da A normal ist

$$(e^{iA})^* e^{iA} = e^{-iA} e^{iA} = e^{-iA+iA} = e^0 = \mathbf{1}$$

□

Umgekehrt: $A = -i \log(U)$ selbstadjungiert wenn U unitär

Die unitären bilden $\mathrm{U}(N)$, eine sogenannte Lie-Gruppe

Lie-Algebra hierzu sind die hermitischen Matrizen

Beispiel 4.34

$e^{i\theta\sigma_2} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ wobei $\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ Generator

Satz 4.35 (Unitäre Trigonalisierung nach Schur)

Sei $A \in \mathrm{Mat}(N \times N, \mathbb{C})$. *Dann existiert unitäres* $U \in \mathrm{Mat}(N \times N, \mathbb{C})$*, so dass*

$$U^* A U \quad \text{obere Dreiecksmatrix}$$

Beweis ist ganz analog zu (ii)$\Longrightarrow$(iii) in Satz 3.24 (Trigonalisierung)

Nur noch Unitarität als Zusatzeigenschaft

Wiederholung des Arguments!

Beweis: Induktion über N. Fall $N = 1$ klar, also Schritt $N - 1 \to N$

Sei z_1 Eigenwert mit normiertem Eigenvektor u_1

Vervollständige zu ONB $u_1, \ldots, u_N$ und setze $V = (u_1, \ldots, u_N)$

$$\begin{aligned} V^* A V &= (u_1, \ldots, u_N)^* (A u_1, \ldots, A u_N) \\ &= (u_1, \ldots, u_N)^* (z_1 u_1, A u_2 \ldots, A u_N) = \begin{pmatrix} z_1 & B \\ 0 & C \end{pmatrix} \end{aligned}$$

Voraussetzung: $\exists$ unitäres W, so dass $W^* C W$ obere Dreiecksmatrix

$$\begin{pmatrix} 1 & 0 \\ 0 & W^* \end{pmatrix} V^* A V \begin{pmatrix} 1 & 0 \\ 0 & W \end{pmatrix} = \begin{pmatrix} z_1 & BW \\ 0 & W^* C W \end{pmatrix}$$

ist eine obere Dreiecksmatrix

Aber Basiswechsel $U = V \begin{pmatrix} 1 & 0 \\ 0 & W \end{pmatrix}$ unitär als Produkt von unitären $\square$

Satz 4.36 (Spektralsatz für normale Matrizen)

Sei $A \in \mathrm{Mat}(N \times N, \mathbb{C})$*. Dann gilt*

$$A \text{ normal} \qquad \Longleftrightarrow \qquad A \text{ unitär diagonalisierbar}$$

Genauer bedeutet Letzteres: es gibt Eigenwerte $z'_1, \ldots, z'_N \in \mathbb{C}$ (*nicht notwendigerweise paarweise verschieden*) *und eine Unitäre* U*, so dass*

$$U^* A U = D \quad , \quad D = \begin{pmatrix} z'_1 & & \\ & \ddots & \\ & & z'_N \end{pmatrix}$$

Für normales A *gelten zudem die Äquivalenzen:*

$$\begin{aligned} A \text{ selbstadjungiert} \quad &\Longleftrightarrow \quad z'_1, \ldots, z'_N \in \mathbb{R} \\ A \text{ unitär} \quad &\Longleftrightarrow \quad 1 = |z'_1| = \ldots = |z'_N| \end{aligned}$$

Beweis: "$\Longleftarrow$" Weil Diagonalmatrizen vertauschen gilt dann

$$A^*A = UD^*U^*UDU^* = UD^*DU^* = UDD^*U^* = UDU^*UD^*U^* = AA^*$$

"$\Longrightarrow$" Sei U unitär mit $B = U^*AU$ obere Dreiecksmatrix (Satz 4.35)

$$B^*B = U^*A^*UU^*AU = U^*A^*AU = U^*AA^*U = U^*AUU^*A^*U = BB^*$$

d.h. B auch normal

Behauptung: normale obere Dreiecksmatrizen sind diagonal

Induktion über N. Anfang klar. Schritt von $N-1 \to N$:

$$|B_{1,1}|^2 = (B^*B)_{1,1} = (BB^*)_{1,1} = |B_{1,1}|^2 + \sum_{n=2}^{N} |B_{1,n}|^2$$

so dass $B_{1,n} = 0$ für $n = 2, \ldots, N$. Somit

$$B = \begin{pmatrix} B_{1,1} & 0 \\ 0 & C \end{pmatrix}$$

wobei C normale obere Dreiecksmatrix der Größe $N-1$,

also nach Induktionsvoraussetzung diagonal!

Nun zum Zusatz

Sei also A selbstadjungiert und $Av = zv$ mit $\|v\| = 1$. Dann

$$\begin{aligned} z &= z\langle v|v\rangle \\ &= \langle v|zv\rangle \\ &= \langle v|Av\rangle \\ &= \langle A^* v|v\rangle \\ &= \langle Av|v\rangle \\ &= \langle zv|v\rangle \\ &= \overline{z} \end{aligned}$$

Umgekehrt, wenn $z'_1, \ldots, z'_N \in \mathbb{R}$, dann gilt

$$A^* = (UDU^*)^* = UD^* U^* = UDU^* = A$$

Ähnlich wird bei unitärem A verfahren $\square$

Beispiel 4.37

Sei $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ mit $a, b \in \mathbb{C}$. Rechnung: $AA^* = A^*A$, d.h. A normal

$p_A(z) = (a - z)^2 - b^2 \implies$ Eigenwerte $z_1 = a + b$, $z_2 = a - b$

Normalisierte Eigenvektoren $v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Tatsächlich auch orthogonal und somit $U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ unitär

Dann ist $U^{-1}AU = U^*AU = D$ diagonal

Eigenvektoren nur bis auf Phase eindeutig

$\implies \quad U' = \frac{1}{\sqrt{2}} \begin{pmatrix} e^{i\theta_1} & e^{i\theta_2} \\ e^{i\theta_1} & -e^{i\theta_2} \end{pmatrix}$ mit $\theta_1, \theta_2 \in \mathbb{R}$ andere Wahl

Satz 4.38 (Spektralsatz für reelle symmetrische Matrizen)

Sei $A = A^T \in \mathrm{Mat}(N \times N, \mathbb{R})$ *symmetrisch (d.h. reell hermitisch). Dann*

$$O^T A O \;=\; D \quad , \quad D \;=\; \begin{pmatrix} z_1' & & \\ & \ddots & \\ & & z_N' \end{pmatrix} \quad , \quad O \text{ orthogonal}$$

Beweis: Als komplexe Matrix gilt $A^* = (\overline{A})^T = A^T = A$

Nach Spektralsatz $\exists$ unitäre (komplexe) Matrix U mit $U^* A U = D$

Spaltenvektoren u_n von $U = (u_1, \ldots, u_N)$ Eigenvektoren von A

Jeder Eigenraum $E_A(z_k)$ invariant unter komplexer Konjugation $v \mapsto \overline{v}$:

$$Av \;=\; z_k v \qquad \Longleftrightarrow \qquad A\overline{v} \;=\; z_k \overline{v}$$

Nach unten stehendem Lemma existiert eine reelle ONB von $E_A(z_k)$

$\Longrightarrow$ zusammen reelle ONB $O = (v_1, \ldots, v_N)$ von $\mathbb{R}^N$ $\square$

Lemma 4.39

Sei Unterraum $V \subset \mathbb{C}^N$ invariant unter komplexer Konjugation ($\overline{V} = V$)
Dann existiert reelle Orthonormalbasis von V

Beweis: Die Basis wird iterativ konstruiert

Sei $V' \subset V$ Unterraum, der schon reelle ONB $v_1, \ldots, v_n$ besitzt

Wähle Einheitsvektor $u \in V$, orthogonal auf V'. Dann sind

$$\Re e(u) = \frac{1}{2}(u + \overline{u}) \qquad , \qquad \Im m(u) = \frac{1}{2i}(u - \overline{u})$$

beide in V und beide orthogonal auf V', da

$$\langle v_k | \overline{u} \rangle = \overline{\langle v_k | u \rangle} = 0 \qquad , \qquad k = 1, \ldots, n$$

Nun $\Re e(u)$ oder $\Im m(u)$ ungleich Nullvektor

In ersterem Fall setze $v_{n+1} = \Re e(u) / \|\Re e(u)\|$, im anderen analog $\square$

Verallgemeinerung in abstrakten Rahmen

Satz 4.40 (Spektralsatz für normale Abbildungen)

Sei $(V, \langle\,.\,|\,.\,\rangle)$ *endlich dimensionaler* $\mathbb{C}$*-Vektorraum mit Skalarprodukt*
Zu normalen $T \in \mathcal{L}(V)$ *existiert ONB* $(u_1, \ldots, u_N)$ *so dass darstellende Matrix diagonal:*

$$U^* T U = D \qquad , \qquad U = (u_1, \ldots, u_N)$$

Insbesondere: algebraische = geometrische Vielfachheiten

Beweis: Sei $W = (v_1, \ldots, v_N)$ beliebige ONB von V
Dann ist $A = W^* T W \in \mathrm{Mat}(N \times N, \mathbb{C})$ normal, weil

$$A^* A = (W^* T W)^* W^* T W = W^* T^* T W = W^* T T^* W = A A^*$$

Spektralsatz: $\exists$ unitäres $V \in \mathrm{Mat}(N \times N, \mathbb{C})$ mit $V^* A V = D$ diagonal
Somit erfüllt $U = W V = (u_1, \ldots, u_N)$ die Behauptung $\square$

Vorbereitungen für weitere Formulierung des Spektralsatzes:

Definition 4.41

Seien $W, W' \subset V$ Unterräume von $(V, \langle . | . \rangle)$

Dann ist orthogonales Komplement von W definiert durch

$$W^{\perp} = \{v \in V \mid v \perp w \quad \forall \quad w \in W\}$$

Zudem:

$$W' \perp W \qquad \Longleftrightarrow \qquad w \perp w' \quad \forall\, w \in W,\ w' \in W'$$

Satz 4.42

Sei $\dim(V) < \infty$ *und* $W \subset V$ *Unterraum*

(i) $W^{\perp}$ *ist Unterraum*

(ii) $W \oplus W^{\perp} = V$ *direkte orthogonale Summe,*
d.h. $\mathrm{span}(W, W^{\perp}) = V$*, aber* $W \cap W^{\perp} = \{0\}$

(iii) $(W^{\perp})^{\perp} = W$

Beweis: (i) Seien $v, v' \in W^\perp$ und $\lambda \in \mathbb{K}$. Dann $v + \lambda v' \in W^\perp$ weil

$$\langle v + \lambda v' | w \rangle \;=\; \langle v | w \rangle \;+\; \lambda \langle v' | w \rangle \;=\; 0 + \lambda \cdot 0 \;=\; 0$$

(ii) Sei $u_1, \ldots, u_M$ eine ONB von W
Vervollständige zu ONB $u_1, \ldots, u_N$ von V
Dann sind $u_{M+1}, \ldots, u_N$ in $W^\perp$ und spannen $W^\perp$ auch auf
(iii) folgt aus Argument in (ii) $\square$

Satz 4.43

Sei $V = W \oplus W^\perp$ Zerlegung in Unterraum W und sein Komplement.
Zu jedem $v \in V$ gibt es Zerlegung $v = w + u$ mit $w \in W$ und $u \in W^\perp$
Definiere $P : V \to V$ durch $Pv = w$.
Dann ist P eine Projektion mit $\mathrm{Ran}(P) = W$.

Merke Korrespondenz: Unterraum $\leftrightarrow$ Projektion auf Unterraum

Beweis: Zu zeigen: P linear, idempotent, selbstadjungiert

Für die Linearität von P, seien $v, v' \in V$ und $\lambda \in \mathbb{K}$

Zerlegungen: $v = w + u$ und $v' = w' + u'$ mit $w, w' \in W$, $u, u' \in W^\perp$

Also $w + \lambda w' \in W$ und $u + \lambda u' \in W^\perp$ (weil W und $W^\perp$ Unterräume)

Zudem $v + \lambda v' = (w + \lambda w') + (u + \lambda u')$

Also $P(v + \lambda v') = w + \lambda w' = Pv + \lambda Pv'$

Idempotent klar nach Definition der Abbildung P

Für $P = P^*$, verwende vervollständigte ONB von W:

$$v = \sum_{n=1}^{N} v_n u_n \qquad , \qquad v' = \sum_{n=1}^{N} v'_n u_n$$

Dann gilt

$$Pv = \sum_{n=1}^{M} v_n u_n \qquad , \qquad Pv' = \sum_{n=1}^{M} v'_n u_n$$

Einsetzen zeigt $\langle v' | Pv \rangle = \langle Pv' | v \rangle$ $\square$

Das orthogonale Komplement tritt auch in Zusammenhang mit Abbildungseigenschaften auf

Satz 4.44

Sei V ein Vektorraum mit Skalarprodukt und $T \in \mathcal{L}(V)$. Dann gilt

$$\mathrm{Ker}(T^*) \; = \; \mathrm{Ran}(T)^{\perp}$$

Beweis. In der Tat:

$$\begin{aligned} v \in \mathrm{Ker}(T^*) \quad & \Longleftrightarrow \quad \langle w | T^* v \rangle = 0 \quad \forall \; w \in V \\ & \Longleftrightarrow \quad \langle Tw | v \rangle = 0 \quad \forall \; w \in V \\ & \Longleftrightarrow \quad v \in \mathrm{Ran}(T)^{\perp} \end{aligned}$$

was den Beweis schon beendet □

Satz 4.45

Seien $W_1, \dots, W_K$ *Unterräume endlich dimensionalen* $(V, \langle\,.\,|\,.\,\rangle)$ $P_1, \dots, P_K$ *zugehörige Projektionen. Dann:*

$$W_1, \dots, W_K \ \text{paarweise orthogonal} \quad \Longleftrightarrow \quad P_k P_n = \delta_{k,n} P_k$$

Beweis: "$\Longrightarrow$" Wenn $k = n$, dann $P_n P_k = (P_k)^2 = P_k = \delta_{k,n} P_k$
Sei nun $k \neq n$ und $u_1(k), \dots, u_{M(k)}(k)$ ONB von W_k. Dann

$$P_k P_n = \sum_{l=1}^{M(k)} |u_l(k)\rangle\langle u_l(k)| \sum_{l'=1}^{M(n)} |u_{l'}(n)\rangle\langle u_{l'}(n)| = 0$$

"$\Longleftarrow$" Seien $k \neq n$ und $w_k \in W_k$ und $w_n \in W_n$
Dann gilt $P_k w_k = w_k$ und $P_n w_n = w_n$. Mit $(P_k)^* = P_k$ folgt

$$\langle w_k | w_n \rangle = \langle P_k w_k | P_n w_n \rangle = \langle w_k | (P_k)^* P_n w_n \rangle = \langle w_k | P_k P_n w_n \rangle = 0$$

Also $W_k \perp W_n$ $\square$

Satz 4.46 (Projektionsoperatorversion des Spektralsatzes)

V endlich dimensionaler komplexer Vektorraum mit Skalarprodukt
Zu $T \in \mathcal{L}(V)$ seien $z_1, \ldots, z_K$ paarweise verschiedene Eigenwerte und $P_1, \ldots, P_K$ die Projektionen auf die zugehörigen Eigenräume
Dann ist $T \in \mathcal{L}(V)$ normal genau dann, wenn Folgendes gilt:

$$T = \sum_{k=1}^{K} z_k P_k$$

mit

$$\mathbf{1}_V = \sum_{k=1}^{K} P_k \qquad , \qquad P_k P_n = \delta_{k,n} P_k$$

Beweis: Sei $U = (u_1, \dots, u_N)$ die in Satz 4.40 gegebene ONB, d.h.

$$T = UDU^* \quad , \quad D = \begin{pmatrix} z_1' & & \\ & \ddots & \\ & & z_N' \end{pmatrix} = \sum_{n=1}^{N} z_n' \, |e_n\rangle\langle e_n|$$

wobei $e_1, \dots, e_N$ die Standardbasis von $\mathbb{C}^N$ ist. Somit gilt

$$T = \sum_{n=1}^{N} z_n' \, U|e_n\rangle\langle e_n|U^* = \sum_{k=1}^{K} z_k \sum_{n,\, z_n' = z_k} |u_n\rangle\langle u_n|$$

Hier wurde Summe umgeordnet, so dass lediglich paarweise verschiedene Eigenwerte z_k auftreten. Aber:

$$P_k = \sum_{n,\, z_n' = z_k} |u_n\rangle\langle u_n|$$

und somit die Darstellung von T sowie andere Eigenschaften

Die Umkehrung sei als Übung verifiziert $\square$

5 Jordansche Normalform

Erinnerung: Nicht jede Matrix ist diagonalisierbar, so wie z.B.:

Definition 5.1

Ein Jordan Block der Größe N mit Eigenwert $z_1 \in \mathbb{C}$ ist

$$J_N(z_1) \;=\; (z_1\, \delta_{n,m} + \delta_{n+1,m})_{n,m=1,\dots,N} \;=\; \begin{pmatrix} z_1 & 1 & 0 & & 0 \\ 0 & z_1 & 1 & 0 & \\ & \ddots & \ddots & \ddots & \\ & & 0 & z_1 & 1 \\ 0 & & & 0 & z_1 \end{pmatrix}$$

"Problem" ist: $p_{J_N(z_1)}(z) = (z_1 - z)^N$ aber nur ein Eigenvektor e_1

algebraische Mult. $\alpha(z_1) \;=\; N \;>\; 1 \;=\; \beta(z_1)$ geometrische Mult.

Jordansche Normalform: in geeigneter Basis Jordan Blöcke typisch

Satz 5.2 (Jordansche Normalform von Matrizen)

Sei $A \in \mathrm{Mat}(N \times N, \mathbb{C})$. Dann existiert sogenannte Jordan Basis $M = (b_1, \dots, b_N)$ und Jordan Blöcke $J_{N_1}(z'_1), \dots, J_{N_K}(z'_K)$, so dass

$$M^{-1}AM = \begin{pmatrix} J_{N_1}(z'_1) & & \\ & \ddots & \\ & & J_{N_K}(z'_K) \end{pmatrix}$$

Bemerkung 5.3

1. Satz überträgt sich auf $T \in \mathcal{L}(V)$, wenn $\dim(V) < \infty$
2. Struktur des Skalarproduktes wird nicht verwandt
3. In Jordan'scher Normalform können Jordan Blöcke gleiche Eigenwerte haben
4. Für reelles A gilt Jordan Zerlegung mit reellen M nur, wenn P_A über $\mathbb{R}$ faktorisiert. Immer gilt es mit komplexem M!

Beweis des Satzes: Konstruktiv

Wichtige Begriffsbildungen:

Definition 5.4

Der verallgemeinerte Eigenraum von A zu z der Stufe $j \geqslant 1$ ist

$$E_{A,j}(z) \;=\; \mathrm{Ker}\big((A - z\mathbf{1})^j\big)$$

Vektoren aus $E_{A,j}(z)$ heißen auch verallgemeinerte Eigenvektoren

Bemerkung 5.5

1. $E_{A,j}(z)$ ist als Kern einer linearen Abbildung ein Unterraum
2. Eigenraum $E_A(z)$ von A zu z ist gleich $E_{A,1}(z)$
3. Offensichtlich gilt $E_{A,j}(z) \subset E_{A,j+1}(z)$. Also

$$E_{A,1}(z) \;\subset\; E_{A,2}(z) \;\subset\; E_{A,3}(z) \;\subset\; \ldots$$

   Aber diese Folge ist endlich!

Beispiel 5.6

Für einen Jordan Block $A = J_N(z)$ wie oben gilt für $j = 1, \ldots, N$

$$E_{A,j}(z) \;=\; \mathrm{span}(e_1, \ldots, e_j)$$

In der Tat,

$$(A - z\mathbf{1})^j \;=\; \begin{pmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & 0 & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 0 & 1 \\ 0 & & & 0 & 0 \end{pmatrix}^j \;=\; \begin{pmatrix} \ldots & 0 & 1 & 0 & \ldots \\ & \ddots & \ddots & \ddots & \\ & & 0 & 0 & 1 \\ & & & 0 & 0 \\ 0 & & & 0 & \vdots \end{pmatrix}$$

wobei die ersten j Spalten gleich Nullvektoren sind

Definition 5.7

Der Hauptraum von A zu z ist

$$H_A(z) \;=\; \bigcup_{j\geqslant 1} E_{A,j}(z)$$

Beispiel 5.8

Für einen Jordan Block $A = J_N(z)$ ist der Hauptraum $H_A(z) = \mathbb{K}^N$

Außerdem ist

$$J_N(0) \;=\; J_N(z) - z\mathbf{1}_N \;=\; A - z\mathbf{1}_N$$

nilpotent.

Definition 5.9

Eine Matrix $S \in \mathrm{Mat}(N \times N, \mathbb{C})$ heißt nilpotent $\Longleftrightarrow \exists \;\; n \in \mathbb{N}$ mit $S^n = 0$

Lemma 5.10 (Lemma von Fitting)

Abkürzung $E_j = E_{A,j}(z) = \mathrm{Ker}((A - z\mathbf{1})^j)$ *zu Eigenwert* $z \in \mathbb{K}$
Definiere den Fitting Index zu z

$$f \;=\; \min\{j \in \mathbb{N} \mid E_{j+1} = E_j\}$$

und setze $R_j = \mathrm{Ran}((A - z\mathbf{1})^j)$*. Dann gilt*

(i) $f = \min\{j \in \mathbb{N} \,|\, R_j = R_{j+1}\}$

(ii) $A - z\mathbf{1} : R_f \to R_f$ *ist ein Isomorphismus*

(iii) $E_{f+i} = E_f$ *und* $R_{f+i} = R_f$ *für alle* $i \geqslant 0$*. Insbesondere* $E_f = H_A(z)$

(iv) $(A - z\mathbf{1})^f(E_f) = \{0\}$, *d.h.* $A - z\mathbf{1}$ *ist nilpotent auf* E_f

(v) E_f *und* R_f *sind invariant unter* $A - z\mathbf{1}$, *d.h.*

$$(A - z\mathbf{1})(E_f) \subset E_f \qquad , \qquad (A - z\mathbf{1})(R_f) \subset R_f$$

(vi) $\mathbb{K}^N = E_f \oplus R_f = H_A(z) \oplus R_f$

(vii) $\dim(E_f) = \dim(H_A(z)) =$ *algebraische Multiplizität von z*

Beweis: ersetze $A - z\mathbf{1}$ durch $A \implies$ o.B.d.A. $z = 0$

Oben gezeigt: $E_j = \mathrm{Ker}(A^j)$ erfüllt $E_j \subset E_{j+1}$

Zudem $R_j \supset R_{j+1}$, da $R_{j+1} = A^{j+1} V = A^j R_1$ wobei $V = \mathbb{K}^N$. Also:

$$\begin{array}{ccccc} E_j & \subset & V & \xrightarrow{A^j} & R_j \\ \cap & & \| & & \cup \\ E_{j+1} & \subset & V & \xrightarrow{A^{j+1}} & R_{j+1} \end{array}$$

Nach Dimensionssatz angewandt auf $A^j : V \to V$ und $A^{j+1} : V \to V$:

$$\dim(V) \;=\; \dim(E_j) + \dim(R_j) \;=\; \dim(E_{j+1}) + \dim(R_{j+1})$$

Somit

$$\begin{aligned} R_{j+1} = R_j \;&\iff\; \dim(R_{j+1}) = \dim(R_j) \quad &&\text{(rechte Spalte des Diagramms)} \\ &\iff\; \dim(E_{j+1}) = \dim(E_j) \quad &&\text{(Dimensionssatz)} \\ &\iff\; E_{j+1} = E_j \quad &&\text{(linke Spalte des Diagramms)} \end{aligned}$$

Dies impliziert (i)

(ii) folgt da $R_f = R_{f+1}$, so dass $A : R_f \to R_{f+1} = R_f$ surjektiv
Somit $A : R_f \to R_f$ Isomorphismus

(iii) Hieraus folgt $R_{f+i} = R_f$ für alle $i \geqslant 1$, mit Äquivalenz auch $E_{f+i} = E_f$

(iv) ist offensichtlich

(v) Lediglich hinzuzufügen ist $AE_f \subset E_{f-1} \subset E_f$

(vi) Nach Dimensionssatz für A^f gilt: $\dim(E_f) + \dim(R_f) = \dim(V)$
Noch zu zeigen: $E_f \cap R_f = \{0\}$
Sei $v \in E_f \cap R_f$. Dann $A^f v = 0$ und $v = A^f w$ für ein $w \in V$
Somit $A^{2f} w = 0$, d.h. $w \in E_{2f} = E_f$. Also $v = A^f w = 0$

(vii) Sei $d = \dim(E_f)$. Setze $S = A|_{E_f}$. Dann $S^f = 0$ nach (iv)
Also nur 0 Eigenwert und $p_S(z) = (-z)^d$
Sei nun $M = (M_1, M_2) =$ (Basis von E_f,Basis von R_f). Nach (v)

$$M^{-1}AM = \begin{pmatrix} M_1^{-1} S M_1 & 0 \\ 0 & M_2^{-1}(A|_{R_f})M_2 \end{pmatrix}$$

Also $p_A(z) = p_S(z)\, p_{A|_{R_f}}(z) = (-z)^d\, p_{A|_{R_f}}(z)$
(ii) $\implies$ $A|_{R_f}$ Isomorphismus $\implies$ $p_{A|_{R_f}}(0) \neq 0 \implies d = \alpha(0)$ $\square$

Satz 5.11 (Hauptraumzerlegung)

Gegeben Faktorisierung des charakteristischen Polynomes:

$$p_A(z) \;=\; (z_1 - z)^{\alpha(z_1)} \cdots (z_K - z)^{\alpha(z_K)}$$

Haupträume $H_k = H_A(z_k)$, $k = 1, \ldots, K$, *mit* $\dim(H_k) = \alpha(z_k)$. *Dann*

(i) $A(H_k) \subset H_k$, *d.h. die* H_k *sind A-invariante Unterräume*

(ii) $V = H_1 \oplus \ldots \oplus H_K$

(iii) $\exists$ *Basis* $(b_1, \ldots, b_N)$ *so dass*

$$M^{-1}AM \;=\; \begin{pmatrix} z_1\,\mathbf{1}_{\alpha(z_1)} + S_1 & & \\ & \ddots & \\ & & z_K\,\mathbf{1}_{\alpha(z_K)} + S_K \end{pmatrix}$$

wobei $S_k : \mathbb{K}^{\alpha(z_k)} \to \mathbb{K}^{\alpha(z_k)}$ *nilpotent ist*

(iv) $A = D + S$, *wobei D diagonalisierbar ist, S nilpotent und* $DS = SD$

Beweis: (i) und (ii) nach iterativer Anwendung des Lemmas von Fitting möglich wegen Faktorisierung des char. Polynoms (vgl. Satz 3.24)
(iii) M ist Hintereinanderreihung von Basen der H_k und wieder Fitting
Die Zerlegung in (iv) ist nach (iii) klar und $DS = SD$ nachrechnen $\square$

Verbleibend: Wahl der Basis in Haupträumen

Satz 5.12 (Jordansche Normalform für nilpotente Abbildung)

Sei $S \in \mathrm{Mat}(N \times N, \mathbb{K})$ nilpotent (also nur 0 Eigenwert)
Dann $\exists$ invertierbares M und $n_1, \ldots, n_K \in \mathbb{N}$ mit $n_1 + \ldots + n_K = N$ und

$$M^{-1}SM = \begin{pmatrix} J_{n_1}(0) & & \\ & \ddots & \\ & & J_{n_K}(0) \end{pmatrix}$$

Der technische Beweis zeigt, dass folgendes Verfahren funktioniert:

(Für Details siehe Bücher, oder Skript LA)

Konstruktion der Basis: Sei $E_j = \mathrm{Ker}(S^j)$. Mit Fitting Index f von S

$$\{0\} = E_0 \subset E_1 \subset E_2 \ldots \subset E_{f-1} \subset E_f = V = \mathbb{K}^N$$

Alle Inklusionen echt und

$$S(E_j) \subset E_{j-1} \qquad , \qquad j = 1, \ldots, f$$

Setze

$$r_j = \dim(E_j) - \dim(E_{j-1})$$

Dann zeigt man

$$r_f \leqslant r_{f-1} \leqslant r_{f-2} \leqslant \ldots \leqslant r_1$$

Wähle r_f lin. unabh. $b_1, \ldots, b_{r_f} \in E_f$ mit $\mathrm{span}\{b_1, \ldots, b_{r_f}\} \cap E_{f-1} = \{0\}$
Für jedes $r = 1, \ldots, r_f$ gibt es sogenannte Jordan Kette

$$b_r,\, Sb_r,\, \ldots,\, S^{f-1} b_r$$

Da $S^f b_r = 0$, ist $M^{-1} S M = J_f(0)$ für $M = (S^{f-1} b_r, \ldots, b_r)$

Falls $r_f = r_1$ ist Konstruktion beendet (da dann $fr_f = fr_1 = N$)

Sonst existiert j mit $r_f = r_{j+1} < r_j$

Vervollständige $S^{r_f - j} b_1, \dots, S^{r_f - j} b_{r_f} \in E_j \backslash E_{j-1}$ durch $s_j = r_{j-1} - r_j$ weitere linear unabhängige Vektoren $b_{r_{j+1}}, \dots, b_{r_j} \in E_j \backslash E_{j-1}$, so dass der gesamte Spann nur trivialen Schnitt mit E_{j-1} hat

Nun s_j zusätzliche Jordan Ketten der Länge j:

$$b_{r_{j+1}+s}, Sb_{r_{j+1}+s}, \dots, S^{j-1} b_{r_{j+1}+s} \qquad , \qquad s = 1, \dots, s_j$$

Dies gibt Jordan Blöcke der Größe j

Itereriere die Prozedur. Wenn $r_j = r_1$ fertig, sonst beginne von Neuem

Kochrezept für Konstruktion einer Jordan Basis:

- Berechne die Potenzen S^j
- Berechne $E_j = \text{ker}(S^j)$ mit Hilfe des Gauss Algorithmus
- Wähle sukzessive zusätzliche linear unabh. Vektoren in $E_j \backslash E_{j-1}$,
  nicht im Spann längerer Jordan Ketten
  und deren Spann E_{j-1} nur trivial schneidet
- Bilde zugehörige Jordan Ketten
- Jordan Basis durch Umkehrung der Reihenfolge

Beachte: viele Freiheiten, nicht eindeutig!

Graphische Darstellung dieser Prozedur

$$
\begin{array}{ccccccccccc}
E_f\setminus E_{f-1} & \xrightarrow{S} & E_{f-1}\setminus E_{f-2} & \xrightarrow{S} & E_{f-2}\setminus E_{f-3} & \xrightarrow{S} & \dots & \xrightarrow{S} & E_2\setminus E_1 & \xrightarrow{S} & E_1 \\
\hline
b_1 & & Sb_1 & & S^2 b_1 & & & & S^{f-2} b_1 & & S^{f-1} b_1 \\
b_2 & & Sb_2 & & S^2 b_2 & & & & S^{f-2} b_2 & & S^{f-1} b_2 \\
\vdots & & \vdots & & \vdots & & & & \vdots & & \vdots \\
b_{r_f} & & Sb_{r_f} & & S^2 b_{r_f} & & & & S^{f-2} b_{r_f} & & S^{f-1} b_{r_f} \\
 & & b_{r_f+1} & & Sb_{r_f+1} & & & & S^{f-3} b_{r_f+1} & & S^{f-2} b_{r_f+1} \\
 & & \vdots & & \vdots & & & & \vdots & & \vdots \\
 & & b_{r_{f-1}} & & Sb_{r_{f-1}} & & & & S^{f-3} b_{r_{f-1}} & & S^{f-2} b_{r_{f-1}} \\
 & & & & & \ddots & & & & & \\
 & & & & & & \ddots & & & & \\
 & & & & & & & & b_{r_3+1} & & Sb_{r_3+1} \\
 & & & & & & & & \vdots & & \vdots \\
 & & & & & & & & b_{r_2} & & Sb_{r_2} \\
 & & & & & & & & & & b_{r_2+1} \\
 & & & & & & & & & & \vdots \\
 & & & & & & & & & & b_{r_1}
\end{array}
$$

Beispiel 5.13

Sei

$$S = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix}$$

Dann ist

$$S^2 = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 1 \\ 1 & -1 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Also S nilpotent mit Fitting Index ist 2

Zunächst Berechnung von $\mathrm{Ker}(S) = \mathrm{Ker}(S^{f-1}) = E_1$

Beispiel (Fortsetzung)

$$E_1 = \mathrm{Ker}(S) = \mathrm{span}\left\{\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\-1\end{pmatrix}\right\}$$

Da $\dim(E_2) = \dim(\mathbb{C}^4) = 4$ und $\dim(E_1) = 2$, gibt es
$r - f = 4 - 2 = 2$ Jordan Blöcke der Größe $f = 2$, keine weiteren
Mögliche Wahl der zwei Vektoren in $E_2 \backslash E_1 = \mathbb{C}^4 \backslash E_1$ ist

$$b_1 = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \ b_2 = \begin{pmatrix}0\\1\\0\\0\end{pmatrix} \implies Sb_1 = \begin{pmatrix}0\\-1\\1\\1\end{pmatrix}, \ Sb_2 = \begin{pmatrix}1\\1\\0\\-1\end{pmatrix}$$

$M^{-1}SM = \mathrm{diag}(J_2(0), J_2(0))$ mit Basiswechsel $M = (Sb_1, b_1, Sb_2, b_2)$

Beispiel 5.14

Sei

$$S = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & a \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad a \in \mathbb{C}$$

Jordan Basis in Abhängigkeit von a. Zunächst:

$$S^2 = \begin{pmatrix} 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad , \quad S^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Fallunterscheidung:

$a \neq 0 \Longrightarrow$ Fitting Index $f = 3 \Longrightarrow$ eine Jordan Kette der Länge 3

$a = 0 \Longrightarrow$ Fitting Index $f = 2 \Longrightarrow$ Jordan Ketten der Längen 2 und 1

Beispiel 5.14 (Fortsetzung)

Für $a \neq 0$

$$E_2 = \mathrm{Ker}(S^2) = \mathrm{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$$

Wahl

$$b_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in E_3 \backslash E_2 \implies Sb_1 = \begin{pmatrix} 1 \\ a \\ 0 \end{pmatrix}, \quad S^2 b_1 = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix}$$

Somit

$$M = \begin{pmatrix} a & 1 & 0 \\ 0 & a & 0 \\ 0 & 0 & 1 \end{pmatrix} \implies M^{-1} S M = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Beispiel 5.14 (Fortsetzung)

Fall $a = 0$, also Fitting Index $f = 2$. Dann $E_2 = \mathbb{C}^3$ und

$$E_1 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right\}$$

Wahl

$$b_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in E_2 \backslash E_1 \implies Sb_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Dann wähle $b_2 \in E_1$ linear unabhängig von Sb_1, z.B.

$$b_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

$M^{-1}SM = \operatorname{diag}(J_2(0), J_1(0))$ mit Basiswechsel $M = (Sb_1, b_1, b_2)$

Beweis vom Hauptsatz (Satz 5.2):

$\mathbb{K} = \mathbb{C} \Longrightarrow$ charakteristisches Polynom faktorisiert

$\Longrightarrow$ Hauptraumzerlegung nach Satz 5.11

In jedem Hauptraum H_k wende Satz 5.12 auf nilpotententes S_k an $\square$

Zusammenfassung zur Berechnung einer Jordan Basis:

- Berechne und faktorisiere charakteristisches Polynom
- Zu jedem Eigenwert z_j bestimme den Hauptraum $H_A(z_j)$
- Innerhalb jedes Hauptraumes bestimme eine Jordan Basis zur nilpotenten Abbildung $(A - z_j\mathbf{1})|_{H_A(z_j)}$

Beispiel 5.15

Sei

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

Bestimme Jordan Basis!

Hier $p_A(z) = (2-z)(1-z)^2 \Longrightarrow z_1 = 2$ und $z_2 = 1$ Eigenwerte

Da $z_1 = 2$ algebraische Vielfachheit 1 hat, ist Hauptraum = Eigenraum

$$E_{A,1}(2) = \mathrm{Ker}(A - 2\,\mathbf{1}) = \mathrm{Ker}\begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & -1 \\ 0 & 0 & 0 \end{pmatrix} = \mathrm{span}\left\{\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}\right\}$$

Hingegen hat z_2 algebraische Vielfachheit 2. Also ggf. Jordan Block

Beispiel 5.16

$$E_{A,1}(1) = \mathrm{Ker}(A - 1\,\mathbf{1}) = \mathrm{Ker}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix} = \mathrm{span}\left\{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\}$$

Also $\beta(1) = 1 < \alpha(1) = 2$. Hauptraum $H_A(1) = E_{A,2}(1)$:

$$H_A(1) = \mathrm{Ker}\big((A - 1\,\mathbf{1})^2\big) = \mathrm{Ker}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \mathrm{span}\left\{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right\}$$

Wahl

$$b_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \in E_{A,2}(1) \backslash E_{A,1}(1)$$

Beispiel 5.17

Dann

$$(A - \mathbf{1})b_2 \;=\; \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

Somit führt

$$M \;=\; (b_1, (A - \mathbf{1})b_2, b_2) \;=\; \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

also zu

$$M^{-1}AM \;=\; \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Erste Anwendung der Jordan Form

Erinnerung (Definition 3.28 und Satz 3.29):
Spektrum von $A \in \mathrm{Mat}(N \times N, \mathbb{C})$

$$\begin{aligned}\mathrm{Spec}(A) &= \{z \in \mathbb{C} \,|\, z \text{ Eigenwert von } A\} \\ &= \{z \in \mathbb{C} \,|\, z\mathbf{1} - A \text{ nicht invertierbar}\}\end{aligned}$$

Wegen Letzterem gilt $\mathrm{Spec}(A) = \mathrm{Spec}(M^{-1}AM)$

Satz 5.18 (Spektraler Abbildungssatz)

$f : \mathbb{C} \to \mathbb{C}$ *analytische Funktion (konvergente Potenzreihe). Dann*

$$f(\mathrm{Spec}(A)) = \mathrm{Spec}(f(A))$$

Für diagonalisierbare A ist das sofort klar!

Beweis. Jordan Zerlegung $A = M(D+J)M^{-1}$ von A

Also o.B.d.A. $A = D + J$ (nach obiger Bemerkung)

Sei $f(z) = \sum_{n \geqslant 0} f_n z^n$ mit $f_n \in \mathbb{C}$. Dann

$$f(D+J) \ = \ \sum_{n \geqslant 0} f_n (D+J)^n \ = \ f(D) + S$$

wobei S wieder nilpotente obere Dreiecksmatrix. Also

$$f(D+J) \ = \ \begin{pmatrix} f(z_1') & * & * \\ 0 & \ddots & * \\ 0 & 0 & f(z_N') \end{pmatrix} \qquad , \qquad D \ = \ \begin{pmatrix} z_1' & & \\ & \ddots & \\ & & z_N' \end{pmatrix}$$

Somit charakteristisches Polynom von $f(D+J)$ berechenbar

Behauptung folgt. $\square$

6 Lineare Differentialgleichungen

Lineares Differentialgleichungssystem erster Ordnung

$$\dot{x}(t) = A\,x(t) \qquad , \qquad x(0) = x_0$$

wobei $A \in \mathrm{Mat}(N \times N, \mathbb{C})$ konstante Koeffizienten

$t \in \mathbb{R} \mapsto x(t) \in \mathbb{C}^N$ differenzierbar mit Ableitung $\dot{x}(t) = \partial_t x(t)$

und $x_0 \in \mathbb{C}^N$ eine sogenannte Anfangsbedingung

Wegen $\partial_t e^{tA} = Ae^{tA}$ ist Lösung gegeben durch

$$x(t) = e^{tA}\,x_0$$

Außerdem Lösung eindeutig

(weil $y(t) = e^{-tA}x(t)$ erfüllt $\dot{y}(t) = 0$, so dass $y(t) =$ konst,

und diese Konstante kann bei $t = 0$ berechnet werden)

Somit: auf Berechnung der Exponentialfunktion e^{tA} reduziert.

Klar für A diagonalisierbar, jetzt aber für alle A!

Berechnung von e^{tA} mit Jordanscher Normalform

$$A = M^{-1}(D+J)M \qquad , \qquad [D,J] = DJ - JD = 0$$

wobei D diagonal und J nilpotent, Summe von Jordan Blöcken

$$\begin{aligned}
e^{tA} &= \sum_{k\geqslant 0} \frac{t^k}{k!} A^k \\
&= \sum_{k\geqslant 0} \frac{t^k}{k!} \big(M^{-1}(D+J)M\big)^k \\
&= M^{-1} \sum_{k\geqslant 0} \frac{t^k}{k!} (D+J)^k M \\
&= M^{-1} e^{t(D+J)} M
\end{aligned}$$

Außerdem kommutieren D und J, also:

$$e^{tA} = M^{-1} e^{tD} e^{tJ} M$$

Nun e^{tD} **und** e^{tJ} beide leicht berechenbar!

Da J block-diagonal mit nilpotenten Jordan Blöcken:

$$J_n(0) = \begin{pmatrix} 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & 0 & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 0 & 1 \\ 0 & & & 0 & 0 \end{pmatrix} \implies J_n(0)^2 = \begin{pmatrix} 0 & 0 & 1 & & 0 \\ 0 & 0 & 0 & 1 & \\ & \ddots & \ddots & \ddots & \ddots \\ & & 0 & 0 & 0 \\ 0 & & & 0 & 0 \end{pmatrix}$$

Also

$$e^{tJ_n(0)} = \sum_{k=0}^{n-1} \frac{t^k}{k!} J_n(0)^k = \begin{pmatrix} 1 & \frac{t}{1!} & \frac{t^2}{2!} & & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & \frac{t}{1!} & \frac{t^2}{2!} & \\ & \ddots & \ddots & \ddots & \\ & & 0 & 1 & \frac{t}{1!} \\ 0 & & & 0 & 1 \end{pmatrix}$$

Nur noch Einsetzen in $e^{tA} = M^{-1} e^{tD} e^{tJ} M$

Beispiel 6.1

Es soll die Lösung des Anfangswertproblemes

$$\dot{x}(t) = A\,x(t) \qquad , \qquad A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix}, \quad x(0) = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

berechnet werden. In Beispiel 5.15 wurde gezeigt, dass

$$\begin{aligned} A &= \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \left[\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \right] \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{aligned}$$

Beispiel 6.1 (Fortsetzung)

Also

$$e^{tA} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} e^{2t} & 0 & 0 \\ 0 & e^t & 0 \\ 0 & 0 & e^t \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} e^t & 0 & 0 \\ te^t & e^t & e^t - e^{2t} \\ 0 & 0 & e^{2t} \end{pmatrix}$$

und die Lösung ist

$$x(t) = e^{tA}\,x(0) = \begin{pmatrix} e^t \\ te^t - e^{2t} \\ e^{2t} \end{pmatrix}$$

was durch Einsetzen nochmals überprüft werden kann.

Systeme höherer Ordnung

Homogenes lineares Differentialgleichungssystem
mit konstanten Koeffizienten und k-ter Ordnung:

$$y^{(k)} - A_{k-1} y^{(k-1)} - \ldots - A_1 y' - A_0 y = 0$$

wobei $A_{k-1}, \ldots, A_0 \in \mathrm{Mat}(N \times N, \mathbb{C})$ zeitunabhängig
$t \in \mathbb{R} \mapsto y(t) \in \mathbb{C}^N$ ist C^k (komponentenweise, Real- & Imaginärteil)
mit Ableitungen $y^{(k)}(t) = \partial_t y^{(k-1)}(t)$

Anfangsbedingungen:

$$y^{(j)}(0) = b_j \in \mathbb{C}^N \qquad , \qquad j = 0, \ldots, k-1$$

Reduktion auf System erster Ordnung ist möglich!

Reduktion von Ordnung

Setze:

$$A = \begin{pmatrix} 0 & 1 & 0 & \\ & 0 & \ddots & 0 \\ & & 0 & 1 \\ A_0 & A_1 & \cdots & A_{k-1} \end{pmatrix}, \quad x = \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(k-2)} \\ y^{(k-1)} \end{pmatrix}, \quad x_0 = \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{k-2} \\ b_{k-1} \end{pmatrix}$$

Dann gilt nach Ausmultiplizieren:

$$\dot{x}(t) = A\,x(t) \qquad , \qquad x(0) = x_0$$

Beispiel 6.2

Gedämpfte Schwingungsgleichung $my'' + ry' + ky = 0$ mit $m, r, k \in \mathbb{R}$

$$\partial_t \begin{pmatrix} y \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{r}{m} \end{pmatrix} \begin{pmatrix} y \\ y' \end{pmatrix}$$

7 Positive Matrizen und Singulärwertzerlegung

Definition 7.1

$A \in \mathrm{Mat}(N \times N, \mathbb{C})$ positiv semidefinit $\Longleftrightarrow A = A^*$ und $\mathrm{Spec}(A) \subset \mathbb{R}_{\geqslant 0}$

A positiv definit, falls zudem $0 \notin \mathrm{Spec}(A)$

Analog definiert sind negativ definite und semidefinite Matrizen

Kurzschreibweisen: $A \geqslant 0$ und $A > 0$, bzw. $A \leqslant 0$ und $A < 0$

Beispiel 7.2

$A > 0$ bedeutet **nicht** positive Matrix-Einträge:

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \quad \Longrightarrow \quad z_\pm = \frac{1}{2}(3 \pm \sqrt{5}) > 0 \quad \Longrightarrow \quad A > 0$$

Bemerkung 7.3

Es gilt: $A < 0 \Longleftrightarrow -A > 0$

Weil: $p_{-A}(z) = (-1)^N p_A(-z) \Longrightarrow$ Nullstellen an 0 gespiegelt

Satz 7.4

Sei $A = A^*$. *Dann*

$$A \geqslant 0 \qquad \Longleftrightarrow \qquad \langle v|Av\rangle \geqslant 0 \quad \forall\ v \in \mathbb{C}^N$$

Analog: $A > 0$, $A \leqslant 0$, $A < 0$

Beweis. "$\Longrightarrow$" Nach Spektralsatz $A = U^* DU$ mit $D \geqslant 0$. Also

$$w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} = Uv \quad \Longrightarrow \quad \langle v|Av\rangle = \langle w|Dw\rangle = \sum_{n=1}^{N} d_n\,|w_n|^2 \geqslant 0$$

"$\Longleftarrow$" Sei $u_1, \ldots, u_N$ ONB zu Eigenwerten $z'_1, \ldots, z'_N$ (Spektralsatz)
Nach Voraussetzung

$$z'_n = \langle u_n|Au_n\rangle \geqslant 0$$

$\square$

Korollar 7.5

Für jedes $B \in \mathrm{Mat}(N \times M, \mathbb{C})$ *gilt* $B^* B \geqslant 0$

Falls $N = M$ *und* B *invertierbar ist, gilt* $B^* B > 0$

Beweis. $\langle v | B^* B v \rangle = \langle Bv | Bv \rangle \geqslant 0$, und > 0 falls B invertierbar

Schließe mit Satz 7.4. □

Korollar 7.6

Für $A = A^* \in \mathrm{Mat}(N \times N, \mathbb{C})$ *und* $M \in \mathrm{Mat}(N \times N, \mathbb{C})$ *invertierbar:*

$$A \geqslant 0 \qquad \Longleftrightarrow \qquad M^* A M \geqslant 0$$

Falls M *nicht invertierbar ist, gilt nur Implikation "*$\Longrightarrow$*"*

Analoge Aussagen für positive Definitheit, und negative Semidefinitheit und Definitheit

Beweis. Wieder nach Satz 7.4, da $\langle v | M^* A M v \rangle = \langle Mv | A(Mv) \rangle$ □

Korollar 7.7

Die positiv definiten Matrizen bilden einen Kegel, d.h.

$$A > 0 \,, B > 0 \,, \lambda > 0 \quad \Longrightarrow \quad A + \lambda B > 0$$

Beweis. In der Tat, für jeden Vektor v gilt

$$\langle v|(A + \lambda B)v\rangle \;=\; \langle v|Av\rangle + \lambda\langle v|Bv\rangle \;>\; 0$$

$\square$

Bemerkung 7.8

Es ist möglich eine partielle Ordnung auf der Menge der selbstadjungierten Matrizen einzuführen:

$$A \;>\; B \quad \Longleftrightarrow \quad A - B \;>\; 0$$

“Partiell” bedeutet, dass nicht alle selbstadjungierten vergleichbar sind.

Erstes Kriterium für positive Definitheit

Satz 7.9 (Descartes' Vorzeichenregel)

Sei $A = A^* \in \mathrm{Mat}(N \times N, \mathbb{C})$ *mit charakteristischem Polynom*

$$p_A(z) \;=\; \det(A - z\,\mathbf{1}) \;=\; \sum_{n=0}^{N} p_n z^n$$

Dann sind $p_n \in \mathbb{R}$ *und es gilt*

$$(-1)^n p_n > 0 \;\text{ für }\; n = 0, \ldots, N \qquad \Longleftrightarrow \qquad A > 0$$

Beweis. $p_n \in \mathbb{R}$ folgt aus $p_A(z) = p_{A*}(z) = \overline{p_A}(z)$
Hauptaussage gilt immer für Polynome mit reellen Koeffizienten.
Trickreicher Beweis (z.B. Fischer, Lineare Algebra) $\square$

Zweites Kriterium für positive Definitheit

Satz 7.10 (Hurwitz' Hauptminorenkriterium)

Sei $A = A^* \in \mathrm{Mat}(N \times N, \mathbb{C})$. *Für* $n = 1, \ldots, N$ *seien Untermatrizen*

$$A_n = (A_{k,l})_{1 \leqslant k,l \leqslant n} \in \mathrm{Mat}(n \times n, \mathbb{C})$$

Ihre Determinanten heißen Hauptminoren. Dann gilt

$$A > 0 \qquad \Longleftrightarrow \qquad \det(A_n) > 0 \quad \forall\, n = 1, \ldots, N$$

Beweis. "$\Longrightarrow$" Ergänze $v \in \mathbb{C}^n$ zu $\hat{v} \in \mathbb{C}^N$ durch Nullen

Dann $\langle v | A_n v \rangle = \langle \hat{v} | A \hat{v} \rangle > 0$ nach Voraussetzung

Nach Satz 7.4 folgt $A_n > 0$

Nun $\det(A_n) > 0$ als Produkt der positiven Eigenwerte

"$\Longleftarrow$" Aussage $A_n > 0$ durch Induktion über n. Fall $n = N$ dann $A > 0$

$A_1 > 0$ klar (weil Determinante Diagonaleintrag). Schritt $(n-1) \to n$

Sei $A_{n-1} > 0$ und $u_1, \ldots, u_{n-1} \in \mathbb{C}^{n-1}$ ONB aus Eigenvektoren zu A_{n-1} zu Eigenwerten $z_1, \ldots, z_{n-1}$

Ergänzen durch je eine 0 liefert $\hat{u}_1, \ldots, \hat{u}_{n-1} \in \mathbb{C}^n$

$\Longrightarrow \hat{u}_1, \ldots, \hat{u}_{n-1}, e_n \in \mathbb{C}^n$ ONB von $\mathbb{C}^n$. Ersetze e_n durch

$$v \;=\; e_n \;-\; \sum_{k=1}^{n-1} \frac{1}{z_k} \langle \hat{u}_k | A_n e_n \rangle \, \hat{u}_k$$

Dann ist $M = (\hat{u}_1, \ldots, \hat{u}_{n-1}, v) \in \mathrm{Mat}(n \times n, \mathbb{C})$ invertierbar und

$$\begin{aligned}
\langle \hat{u}_l | A_n v \rangle &= \langle \hat{u}_l | A_n e_n \rangle - \sum_{k=1}^{n-1} \frac{1}{z_k} \langle \hat{u}_k | A_n e_n \rangle \langle \hat{u}_l | A_n \hat{u}_k \rangle \\
&= \langle \hat{u}_l | A_n e_n \rangle - \sum_{k=1}^{n-1} \langle \hat{u}_k | A_n e_n \rangle \langle \hat{u}_l | \hat{u}_k \rangle \;=\; 0
\end{aligned}$$

Weil $M^* A_n M$ selbstadjungiert, folgt

$$M^* A_n M = \begin{pmatrix} D & 0 \\ 0 & \langle v | A_n v \rangle \end{pmatrix}$$

Nach Voraussetzung $D > 0$ und $\det(M^* A_n M) = |\det(M)|^2 \det(A_n) > 0$

$\Longrightarrow$ letzter Diagonaleintrag $\langle v | A_n v \rangle > 0$

$\Longrightarrow$ $M^* A_n M > 0$ und nach Korollar 7.6 somit $A_n > 0$ $\square$

Beispiel 7.11

Sei

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

Dann ist

$$p_A(z) = (2-z)^3 - (2-z) - (2-z) = -z^3 + 6z^2 - 10z + 4$$

Also liegen alternierende Vorzeichen vor und somit $A > 0$

Auch das Hurwitz Kriterium kann angewandt werden, denn

$$\det(A_3) = 4\,, \quad \det(A_2) = \det\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = 3\,, \quad \det(A_1) = 2$$

so dass alle Hauptminoren positiv sind.

Definition 7.12

Betrag von $A \in \mathrm{Mat}(N \times M, \mathbb{C})$ ist $0 < |A| = (A^*A)^{\frac{1}{2}} \in \mathrm{Mat}(M \times M, \mathbb{C})$, wobei Wurzel durch Spektralkalkül von $A^*A > 0$ definiert ist:

$$A^*A = U^* \begin{pmatrix} (\mu_1)^2 & & \\ & \ddots & \\ & & (\mu_M)^2 \end{pmatrix} U \implies |A| = U^* \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_M \end{pmatrix} U$$

Eigenwerte μ_m von $|A|$ heißen Singulärwerte von A. Ordnung

$$0 \leqslant \mu_M \leqslant \mu_{M-1} \leqslant \ldots \leqslant \mu_1$$

Bemerkung 7.13

Auch wenn $N = M$ gilt nicht immer $|A| = |A^*|$, es sei denn $N = 1$

Tatsächlich: $|A| = |A^*| \iff A$ normal

Aber positive Singulärwerte von A und A^* sind gleich (später)

Satz 7.14

Sei $A \in \mathrm{Mat}(N \times M, \mathbb{C})$. *Dann gilt*

$$\mathrm{Ker}(A) \;=\; \mathrm{Ran}(A^*)^{\perp}$$

wobei orthogonales Komplement in $\mathbb{C}^M$ *bez. Standardskalarprodukt*
Zudem

$$\mathrm{Ker}(A) \;=\; \mathrm{Ker}(A^*A) \qquad , \qquad \mathrm{Ran}(A) \;=\; \mathrm{Ran}(AA^*)$$

Die Ränge von A, A^*, A^*A, AA^*, $|A|$ *und* $|A^*|$ *sind allesamt gleich, d.h.*

$$\begin{aligned}\dim(\mathrm{Ran}(A)) &= \dim(\mathrm{Ran}(A^*)) \\ &= \dim(\mathrm{Ran}(A^*A)) \\ &= \dim(\mathrm{Ran}(AA^*))\end{aligned}$$

Bemerkung: Erste Aussage verallgemeinert Satz 4.44

Beweis. Mit Lemma 4.30

$$\begin{aligned} v \in \mathrm{Ker}(A) \quad &\Longleftrightarrow \quad \langle w|Av\rangle = 0 \quad \forall \;\; w \in \mathbb{C}^N \\ &\Longleftrightarrow \quad \langle A^* w|v\rangle = 0 \quad \forall \;\; w \in \mathbb{C}^N \\ &\Longleftrightarrow \quad v \in \mathrm{Ran}(A^*)^\perp \end{aligned}$$

somit erste Behauptung. Klar ist $\mathrm{Ker}(A) \subset \mathrm{Ker}(A^*A)$,
aber $\mathrm{Ker}(A^*A) \subset \mathrm{Ker}(A)$ gilt auch, weil für $v \in \mathrm{Ker}(A^*A)$

$$\|Av\|^2 \;=\; \langle Av|Av\rangle \;=\; \langle v|A^*Av\rangle \;=\; 0 \quad \Longrightarrow \quad Av = 0$$

Hiermit

$$\mathrm{Ran}(A^*A) \;=\; \mathrm{Ker}(A^*A)^\perp \;=\; \mathrm{Ker}(A)^\perp \;=\; \mathrm{Ran}(A^*)$$

also $\dim(\mathrm{Ran}(A^*A)) = \dim(\mathrm{Ran}(A^*))$, und analog für AA^*. Nun:

$$\begin{aligned} \dim(\mathrm{Ran}(A^*)) &= \; M - \dim(\mathrm{Ran}(A^*)^\perp) \;=\; M - \dim(\mathrm{Ker}(A)) \\ &= \; M - \big(M - \dim(\mathrm{Ran}(A))\big) \;=\; \dim(\mathrm{Ran}(A)) \end{aligned}$$

wobei im vorletzten Schritt der Rangsatz für A verwandt wird $\square$

Satz 7.15 (Singulärwertzerlegung)

Sei $A \in \mathrm{Mat}(N \times M, \mathbb{C})$*. Dann gibt es:*

- $W \in \mathrm{Mat}(N \times N, \mathbb{C})$ *unitär*
- $U \in \mathrm{Mat}(M \times M, \mathbb{C})$ *unitär*
- $D \in \mathrm{Mat}(N \times M, \mathbb{C})$ *Diagonalmatrix (nicht quadratisch)*

so dass

$$A = WDU$$

Auf der Diagonale von D stehen die Singulärwerte von A.

Beweis.

Zunächst Fall $N = M$ und A invertierbar

Nach Definition $|A| = U^* D U$ mit U unitär und $D > 0$ invertierbar

Also $D = U|A|U^*$

Also setze $W = AU^* D^{-1}$, so dass $A = WDU$ wie gewünscht

Noch zu zeigen ist, dass W unitär ist:

$$\begin{aligned} W^* W &= (AU^* D^{-1})^* AU^* D^{-1} \\ &= D^{-1} UA^* AU^* D^{-1} \\ &= D^{-1} D^2 D^{-1} \\ &= \mathbf{1} \end{aligned}$$

Da W invertierbar, gilt auch $W^* = W^{-1}$ und $WW^* = \mathbf{1}$

Fall $M > N$ ($M < N$ analog) $\Longrightarrow$ $\mathrm{rk}(A) \leqslant N$ $\Longrightarrow$ $\mu_{N+1} = 0$. Setze

$$D = \begin{pmatrix} \mu_1 & & & 0 & \dots & 0 \\ & \ddots & & \vdots & & \vdots \\ & & \mu_N & 0 & \dots & 0 \end{pmatrix} \in \mathrm{Mat}(N \times M, \mathbb{C})$$

Sei $r = \mathrm{rk}(A)$. Also $\mu_r > 0$ und $\mu_{r+1} = 0$. Mit Standardbasis $e_n \in \mathbb{C}^M$

$$A\, U^*\, e_n = 0 \qquad , \qquad n = r+1, \dots, M \tag{7.1}$$

Dann

$$w_n = \frac{1}{\mu_n} A\, U^*\, e_n \qquad , \qquad n = 1, \dots, r \tag{7.2}$$

wohldefiniert und

$$\langle w_n | w_m \rangle = \frac{1}{\mu_n} \frac{1}{\mu_m} \langle AU^* e_n | AU^* e_m \rangle = \frac{1}{\mu_n} \frac{1}{\mu_m} \langle e_n | UA^* AU^* e_m \rangle = \delta_{n,m}$$

Also $w_1, \dots, w_r \in \mathbb{C}^N$ orthonormal. Ergänze zu ONB $W = (w_1, \dots, w_N)$
Zusammen ergeben dann (7.1) und (7.2) nach Auswertung auf e_n

$$WD = AU^*$$

□

Algorithmus (Kochrezept) zur Bestimmung der Singulärwertzerlegung:

- Berechne A^*A
- Führe unitäre Diagonalisierung $A^*A = U^*\widehat{D}U$ durch
- Bilde D (von gleicher Größe wie A) aus $\widehat{D}$
- Berechne die Bildvektoren $w_n = \frac{1}{\mu_n} AU^* e_n$ für $n = 1, \ldots, r = \mathrm{rk}(A)$
- Vervollständige $w_1, \ldots, w_r$ zu einer Orthonormalbasis
- Bilde W

Alternativ: erst Singulärwertzerlegung von A^*

Durch Adjungieren dann die Singulärwertzerlegung von A

Dies kann weniger Rechenaufwand bedeuten, wenn $M > N$

8 Hermitische Formen und Quadriken

Erinnerung (Definition 2.12): $B : V \times V \to \mathbb{K}$ Bilinearform $\iff$

$$B(v+\lambda w, v'+\lambda' w') \;=\; B(v,v')+\lambda\, B(w,v')+\lambda'\, B(v,w')+\lambda\lambda'\, B(w,w')$$

Dann: B symmetrisch $\iff$ $B(v,w) = B(w,v)$ für alle $v, w \in V$

B antisymmetrisch $\iff$ $B(v,w) = -B(w,v)$ für alle $v, w \in V$

Beispiel 2.13 betrachtet Bilinearformen $B : \mathbb{C}^N \times \mathbb{C}^N \to \mathbb{C}$ der Gestalt

$$B(v,w) \;=\; \sum_{m,n=1}^{N} v_m A_{m,n} w_n \qquad , \qquad v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}$$

wobei $A \in \mathrm{Mat}(N \times N, \mathbb{K})$. Bez. Basis sieht jede Bilinearform so aus.

Dann: B symmetrisch $\iff$ $A^T = A$ symmetrisch

B antisymmetrisch $\iff$ $A^T = -A$ antisymmetrisch

Definition 8.1

Eine Abbildung $Q : V \times V \to \mathbb{K}$ heißt Sesquilinearform auf V, falls Q antilinear im ersten Argument und linear im zweiten Argument:

$$Q(v+\lambda w, v'+\lambda' w') \;=\; Q(v,v')+\overline{\lambda}\,Q(w,v')+\lambda'\,Q(v,w')+\overline{\lambda}\lambda'\,Q(w,w')$$

(i) Q hermitisch $\iff$ $Q(v,w) = \overline{Q(w,v)}$ für $v, w \in V$

(ii) Q anti-hermitisch $\iff$ $Q(v,w) = -\overline{Q(w,v)}$ für $v, w \in V$

(iii) Q positiv definit $\iff$ $Q(v,v) > 0$ für $v \in V \backslash \{0\}$

(iv) Q positiv semi-definit $\iff$ $Q(v,v) \geqslant 0$ für $v \in V$

(v) Analog: negative Definitheit und Semidefinitheit von Q

(vi) Q nicht ausgeartet $\iff$ $\forall\; v \in V\backslash\{0\}\; \exists\; w \in V$ mit $Q(v,w) \neq 0$

(vii) Wenn $\mathcal{B} = (b_1, \ldots, b_N)$ Basis von V, so darstellende Matrix von Q

$$Q_{\mathcal{B}} \;=\; \big(Q(b_n, b_m)\big)_{n,m=1,\ldots,N}$$

Bemerkung 8.2

1. Falls $\mathbb{K} = \mathbb{R}$ gilt: Bilinearform $=$ Sequilinearform
2. Skalarprodukte sind immer auch hermitische Sesquilinearformen
   Hinzu kommen noch die Positivität und Nichtentartung
3. Bezeichnung: *sesqui = semis que = anderthalb*
4. Menge der Sesquilinearformen auf V ist $\mathbb{K}$-Vektorraum

Beispiel 8.3

Sei $V = \mathbb{K}^N$ und $A \in \mathrm{Mat}(N \times N, \mathbb{K})$. Definiere $Q : V \times V \to \mathbb{K}$ durch

$$Q(v, w) = \langle v | A w \rangle = \sum_{n,m=1}^{N} \overline{v_n} A_{n,m} w_m \,, \qquad v = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix}, \; w = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix}$$

Q ist Sesquilinearform. Nach folgendem Satz ist dies Standardform

Satz 8.4

Sei $\mathcal{B} = (b_1, \ldots, b_N)$ Basis von V und sei zugehörige Koordinatenabbildung $K = (b_1, \ldots, b_N)^{-1} : V \to \mathbb{K}^N$. Dann

$$Q(v, w) \;=\; \langle K(v) | Q_{\mathcal{B}} K(w) \rangle$$

wobei rechts Standardskalarprodukt auf $\mathbb{K}^N$ steht. Es gilt:

(i) *Q hermitisch $\iff$ $Q_{\mathcal{B}} = (Q_{\mathcal{B}})^*$ selbstadjungiert (auch hermitisch)*

(ii) *Q anti-hermitisch $\iff$ $Q_{\mathcal{B}} = -(Q_{\mathcal{B}})^*$ antiselbstadjungiert*

(iii) *Q positiv definit $\iff$ $Q_{\mathcal{B}} > 0$*

(iv) *Q positiv semi-definit $\iff$ $Q_{\mathcal{B}} \geqslant 0$*

(v) *Wenn $\mathcal{C}$ weitere Basis, $K' = \mathcal{C}^{-1}$ und $M = K(K')^{-1}$, dann*

$$Q_{\mathcal{C}} \;=\; M^* Q_{\mathcal{B}} M$$

Beweis. Koordinatenabbildung ist

$$K(v) = \begin{pmatrix} v_1 \\ \vdots \\ v_N \end{pmatrix} \quad , \qquad \text{für} \quad v = \sum_{n=1}^{N} v_n b_n$$

und analog für $K(w)$. Einsetzen ergibt

$$\langle K(v) | Q_{\mathcal{B}} K(w) \rangle = \sum_{n,m=1}^{N} \overline{v_n} (Q_{\mathcal{B}})_{n,m} w_m = \sum_{n,m=1}^{N} \overline{v_n} Q(b_n, b_m)\, w_m = Q(v, w)$$

Zu (i):

$$\overline{Q(v, w)} = \langle Q_{\mathcal{B}} K(w) | K(v) \rangle = \langle K(w) | (Q_{\mathcal{B}})^* K(v) \rangle$$

was gleich $Q(w, v) = \langle K(w) | Q_{\mathcal{B}} K(v) \rangle$ für alle $v, w \iff Q_{\mathcal{B}} = (Q_{\mathcal{B}})^*$

Analog folgt (ii). (iii) und (iv) sind Umformulierungen von Satz 7.4. Nun:

$$\begin{aligned} \langle K(v) | Q_{\mathcal{B}} K(w) \rangle &= \langle K(K')^{-1} K'(v) | Q_{\mathcal{B}} K(K')^{-1} K'(w) \rangle \\ &= \langle K'(v) | M^* Q_{\mathcal{B}} M K'(w) \rangle \implies \text{(v)} \end{aligned}$$

$\square$

Definition 8.5

Quadratische Matrizen $A, B \in \mathrm{Mat}(N \times N, \mathbb{K})$ hermitisch kongruent
$\iff$ $\exists$ invertierbares $M \in \mathrm{Mat}(N \times N, \mathbb{K})$ mit

$$A = M^* BM$$

$A, B \in \mathrm{Mat}(N \times N, \mathbb{K})$ kongruent $\iff$ $A = M^T BM$ für invertierbares M

Bemerkung 8.6

Kongruenz und hermitische Kongruenz sind Äquivalenzrelationen

Satz 8.7 (Normalform für hermitische Sesquilinearformen)

Zu hermitischer Sesquilinearform Q existiert Basis $\mathcal{B}$ mit

$$Q_{\mathcal{B}} = \begin{pmatrix} \mathbf{1}_{N_+} & 0 & 0 \\ 0 & 0_{N_0} & 0 \\ 0 & 0 & -\mathbf{1}_{N_-} \end{pmatrix} \quad , \qquad N_+ + N_0 + N_- = N$$

Beweis. Zunächst $\mathbb{K} = \mathbb{C}$ und $\mathcal{C} = (c_1, \ldots, c_N)$ beliebige Basis

Nach Satz 8.4 ist $Q_\mathcal{C} = (Q_\mathcal{C})^*$ und mit Spektralsatz

$$U^* Q_\mathcal{C} U \;=\; \begin{pmatrix} z_1' & & \\ & \ddots & \\ & & z_N' \end{pmatrix} \qquad , \qquad U \text{ unitär}\,, \quad z_n' \in \mathbb{R}$$

Ordnung $z_n' \geqslant z_{n+1}'$. Nun $N_\pm$ Anzahl positiver/negativer Eigenwerte

Mit $N_0 = N - N_+ - N_-$ gilt $z_{N_++1}' = \ldots = z_{N_++N_0}' = 0$. Definiere $D > 0$

$$\begin{pmatrix} z_1' & & \\ & \ddots & \\ & & z_N' \end{pmatrix} = D \begin{pmatrix} \mathbf{1}_{N_+} & 0 & 0 \\ 0 & 0_{N_0} & 0 \\ 0 & 0 & -\mathbf{1}_{N_-} \end{pmatrix} D$$

Dann $\mathcal{B} = \mathcal{C}M$ mit Basiswechsel $M = UD^{-1}$ weil

$$(UD^{-1})^* Q_\mathcal{C} (UD^{-1}) \;=\; \begin{pmatrix} \mathbf{1}_{N_+} & 0 & 0 \\ 0 & 0_{N_0} & 0 \\ 0 & 0 & -\mathbf{1}_{N_-} \end{pmatrix}$$

Falls $\mathbb{K} = \mathbb{R}$, kann U orthogonal gewählt werden. $\square$

Erster Teil obigen Beweises zeigt:

Satz 8.8 (Hauptachsentransformation)

$(V, \langle\,.\,|\,.\,\rangle)$ $\mathbb{K}$-Vektorraum mit Skalarprodukt und Q hermitische Sesquilinearform $\Longrightarrow$ $\exists$ ONB $\mathcal{B}$ mit

$$Q_{\mathcal{B}} = \begin{pmatrix} z_1' & & \\ & \ddots & \\ & & z_N' \end{pmatrix}$$

Satz 8.9 (Trägheitssatz von Sylvester)

Zahlen N_+, N_0 und N_- aus Satz 8.7 *sind*

$$N_+ = \max\{\dim(U) \mid U \text{ Unterraum mit } Q(u,u) > 0 \ \forall\, u \in U\backslash\{0\}\ \}$$
$$N_0 = \max\{\dim(U) \mid U \text{ Unterraum mit } Q(v,u) = 0 \ \forall\, v \in V \text{ und } u \in U\ \}$$
$$N_- = \max\{\dim(U) \mid U \text{ Unterraum mit } Q(u,u) < 0 \ \forall\, u \in U\backslash\{0\}\ \}$$

Insbesondere: N_+, N_0 und N_- von $M^* Q_{\mathcal{B}} M$ unabhängig von M

Beweis mit Variationsprinzipien (siehe Literatur) □

Korollar 8.10

Sei Q eine hermitische Sesquilinearform

(i) *Q positiv definit* $\iff N_0 = N_- = 0$

(ii) *Q positiv semi-definit* $\iff N_- = 0$

(iii) *Q nicht ausgeartet* $\iff N_0 = 0$

Definition 8.11

Sei Q eine hermitische Sesquilinearform

(i) Trägheit (auf Englisch *inertia*) von Q ist Tupel (N_+, N_0, N_-)

(ii) Signatur von Q ist $N_+ - N_-$

(iii) Witt Index von Q ist $N_0 + \min\{N_+, N_-\}$

Bemerkung 8.12

Dies sind sogenannte Invarianten der hermitischen Sesquilinearform

Beispiel 8.13 (Indefinite Metrik)

Sei $V = \mathbb{R}^4$ Minkovski Raum-Zeit versehen mit Minkovski Produkt

$$Q(v, v') = xx' + yy' + zz' - tt' \quad , \quad v = \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}, \; v' = \begin{pmatrix} x' \\ y' \\ z' \\ t' \end{pmatrix}$$

Dies ist Bilinearform und hier im Reellen auch Sesquilinearform

Wenn $\mathcal{B}$ Standardbasis,

$$Q_{\mathcal{B}} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}$$

Also Trägheit $(3, 1)$

Beispiel 8.14 (Anti-hermitische Bilinearform)

Sei $V = \mathbb{R}^{2N}$. Vektoren $v \in V$ werden mit $v = \binom{u}{w}$ bezeichnet, wobei $u, w \in \mathbb{R}^N$. Wenn $v' = \binom{u'}{w'} \in \mathbb{R}^{2N}$, dann

$$Q(v, v') \ = \ \langle w|u'\rangle - \langle u|w'\rangle$$

anti-hermitische Bilinearform. Bez. Standardbasis $\mathcal{B}$ des $\mathbb{R}$ ist

$$Q_{\mathcal{B}} \ = \ \begin{pmatrix} 0 & -\mathbf{1}_N \\ \mathbf{1}_N & 0 \end{pmatrix}$$

Diese Bilinearform heißt auch symplektische Form

Definition 8.15 (Quadriken)

Reelles quadratisches Polynom P in N Variablen $x_1, \dots, x_N$:

$$P(x_1, \dots, x_N) \;=\; \sum_{n,m=1}^{N} p_{n,m}\, x_n x_m \;+\; \sum_{n=1}^{N} p_n\, x_n \;+\; p \qquad (8.1)$$

wobei $p_{n,m}, p_n, p \in \mathbb{R}$. Reelle Quadrik $\mathcal{Q}$ im $\mathbb{R}^N$ ist Nullstellenmenge:

$$\mathcal{Q} \;=\; \{(x_1, \dots, x_N) \in \mathbb{R}^N \mid P(x_1, \dots, x_N) = 0\}$$

Satz 8.16

$\exists\; A = A^* = (a_{n,m})_{n,m=1,\dots,N}$, $a_0 \in \mathbb{R}^N$ und $a_{0,0} \in \mathbb{R}$ mit

$$P(x_1, \dots, x_N) \;=\; \langle x | Ax \rangle \;+\; 2\langle a_0 | x \rangle \;+\; a_{0,0} \qquad , \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}$$

Beweis. Ausgehend von Darstellung (8.1) setze

$$a_{n,m} = \frac{1}{2}(p_{n,m} + p_{m,n}) \quad , \quad a_{0,n} = \frac{1}{2}p_n \quad , \quad a_{0,0} = p$$

Da $x_n x_m = x_m x_n$, gilt dann die Darstellungsformel $\square$

Korollar 8.17

Umzuschreiben mit Hilfe von $A' = (a_{n,m})_{n,m=0,\dots,N}$*:*

$$P(x_1, \dots, x_N) = \langle x' | A' x' \rangle \quad , \quad A' = \begin{pmatrix} a_{0,0} & (a_0)^T \\ a_0 & A \end{pmatrix} , \quad x' = \begin{pmatrix} 1 \\ x \end{pmatrix}$$

Definition 8.18 (Affine Abbildung)

Abbildung $F : V \to V$ heißt affin

$\iff \exists$ lineare Abbildung $T \in \mathcal{L}(V)$ und $b \in V$ mit

$$F(v) = Tv + b$$

(i) Falls $T = \mathbf{1}_V$, heißt F eine Translation

(ii) Eine Affinität ist eine bijektive affine Abbildung

(iii) Falls V Skalarprodukt hat und T orthogonal, heißt F euklidische Bewegung

Zugehörig zu affiner Abbildung wird $F' \in \mathcal{L}(\mathbb{K} \oplus V)$ definiert via

$$F' = \begin{pmatrix} 1 & 0 \\ b & T \end{pmatrix}$$

Hierbei $\mathbb{K} \oplus V$ der $\mathbb{K}$-Vektorraum (die direkte Summe), gegeben durch Menge $\mathbb{K} \times V$ versehen mit natürlicher Vektorraumstruktur

Bemerkung 8.19

Inverse der Affinität $F(v) = Tv + b$ ist $F^{-1}(v) = T^{-1}v - T^{-1}b$

Also: Affinitäten bilden Gruppe mit Translationen als Untergruppe.

Satz 8.20 (affine Abbildungen $\cong$ lineare auf $\mathbb{K} \oplus V$)

F, G affin auf V und zu $v \in V$ setze $v' = \binom{1}{v} \in \mathbb{K} \oplus V$

Dann gelten folgende Identitäten:

(i) $F'v' = F(v)'$, *wobei $F'v'$ Matrix-Vektormultiplikation in $\mathbb{K} \oplus V$*

(ii) $(F \circ G)' = F'G'$

(iii) $(F^{-1})' = (F')^{-1}$, *falls F eine Affinität ist*

Bild der Zuordnung $F \mapsto F'$ ist Untergruppe von $\mathrm{Gl}(\mathbb{K} \oplus V)$,

genauer gegeben durch obere Dreiecksmatrizen

$$\begin{pmatrix} 1 & 0 \\ b & T \end{pmatrix}$$

Beweis. In der Tat,

$$F'v' = \begin{pmatrix} 1 & 0 \\ b & T \end{pmatrix} \begin{pmatrix} 1 \\ v \end{pmatrix} = \begin{pmatrix} 1 \\ b + Tv \end{pmatrix} = F(v)'$$

und ähnlich

$$F'G' = \begin{pmatrix} 1 & 0 \\ b & T \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & S \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ b + Tc & TS \end{pmatrix} = (F \circ G)'$$

Hieraus folgt auch (iii). $\square$

Satz 8.21

Sei $\mathcal{Q}$ Quadrik im $\mathbb{R}^N$ und F Affinität auf dem $\mathbb{R}^N$
$\Longrightarrow$ *Bild $F(\mathcal{Q})$ ist eine Quadrik*

Beweis. Nach Satz 8.16 ist Quadrik von Gestalt

$$\mathcal{Q} \; = \; \{x \in \mathbb{R}^N \mid \langle x' | A' x' \rangle = 0\}$$

wobei $A' = (A')^T$ quadratische Matrix der Größe $N+1$. Nach Satz 8.20

$$\begin{aligned} F(\mathcal{Q}) &= \{F(x) \in \mathbb{R}^N \mid \langle x' | A' x' \rangle = 0\} \\ &= \{y \in \mathbb{R}^N \mid \langle (F^{-1}y)' | A' (F^{-1}y)' \rangle = 0\} \\ &= \{y \in \mathbb{R}^N \mid \langle (F^{-1})' y' | A' (F^{-1})' y' \rangle = 0\} \\ &= \{y \in \mathbb{R}^N \mid \langle y' | ((F')^{-1})^T A' (F')^{-1} y' \rangle = 0\} \end{aligned}$$

Da $((F')^{-1})^T A' (F')^{-1}$ wieder symmetrisch, liegt Quadrik vor. $\square$

Satz 8.22 (Affine Normalformen von reellen Quadriken)

Sei $\mathcal{Q}$ eine Quadrik im $\mathbb{R}^N$ von der Form

$$\begin{aligned}\mathcal{Q} &= \left\{x \in \mathbb{R}^N \,\middle|\, \langle x|Ax\rangle + 2\langle a_0|x\rangle + a_{0,0} = 0\right\} \\ &= \left\{x \in \mathbb{R}^N \,\middle|\, \langle x'|A'x'\rangle = 0\right\}\end{aligned}$$

Trägheit von A sei $(N_+, N_0, N_-) = (k, N-k-K, K)$
Zudem seien Ränge $r = \mathrm{rk}(A)$ und $r' = \mathrm{rk}(A')$, d.h. $k + K = r$
$\Longrightarrow$ $\exists$ *Affinität F auf $\mathbb{R}^N$ mit $F(\mathcal{Q})$ in einer der Normalformen (Typen):*

$$\begin{array}{llll}
\text{(I)} & y_1^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 = 0 & & \textit{falls } r = r' \\
\text{(II)} & y_1^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 = \pm 1 & & \textit{falls } r+1 = r' \\
\text{(III)} & y_1^2 + \ldots + y_k^2 - y_{k+1}^2 - \ldots - y_r^2 + 2\,y_{r+1} = 0 & & \textit{falls } r+2 = r'
\end{array}$$

Beweis. Konstruktiv! Siehe Literatur □

Liste der Normalformen von Quadriken in Dimension $N = 2$

Typ	r	r'	k	Gleichung	Quadrik
I	0	0	0	$0 = 0$	Ebene $\mathbb{R}^2$
I	1	1	1	$x_1^2 = 0$	Gerade
I	2	2	1	$x_1^2 - x_2^2 = 0$	Geradenpaar mit Schnitt
I	2	2	2	$x_1^2 + x_2^2 = 0$	Punkt
II	1	2	1	$x_1^2 = 1$	paralleles Geradenpaar
II	2	3	1	$x_1^2 - x_2^2 = 1$	Hyperbel
II	2	3	2	$x_1^2 + x_2^2 = 1$	Kreis
III	1	3	1	$x_1^2 + 2x_2 = 0$	Parabel

Beachte: Normalform der Ellipse im Satz 8.22 ist Kreis
(affine Abbildungen erlauben es, die Hauptachsen zu normieren)
Andere Situation wenn nur euklidische Bewegungen

Dimension $N = 3$: Liste der Normalformen noch überschaubar:

Typ	r	r'	k	K	Gleichung	Quadrik
I	0	0	0	0	$0 = 0$	$\mathbb{R}^3$
I	1	1	1	0	$x_1^2 = 0$	Ebene
I	2	2	1	1	$x_1^2 - x_2^2 = 0$	Ebenenpaar mit Schnitt
I	2	2	2	0	$x_1^2 + x_2^2 = 0$	Gerade
I	3	3	2	1	$x_1^2 + x_2^2 - x_3^2 = 0$	Kreiskegel
I	3	3	3	0	$x_1^2 + x_2^2 + x_3^2 = 0$	Punkt
II	1	2	1	0	$x_1^2 = 1$	paralleles Ebenenpaar
II	2	3	1	1	$x_1^2 - x_2^2 = 1$	hyperbolischer Zylinder
II	2	3	2	0	$x_1^2 + x_2^2 = 1$	Kreiszylinder
II	3	4	1	2	$x_1^2 - x_2^2 - x_3^2 = 1$	zweischaliges Hyperboloid
II	3	4	2	1	$x_1^2 + x_2^2 - x_3^2 = 1$	einschaliges Hyperboloid
II	3	4	3	0	$x_1^2 + x_2^2 + x_3^2 = 1$	Kugel
III	1	3	1	0	$x_1^2 + 2x_2 = 0$	parabolischer Zylinder
III	2	4	1	1	$x_1^2 - x_2^2 + 2x_3 = 0$	hyperbolisches Paraboloid
III	2	4	2	0	$x_1^2 + x_2^2 + 2x_3 = 0$	elliptisches Paraboloid

Geometrische Begriffe zu Sesquilinearformen

Definition 8.23

Q hermitische/anti-hermitische Sesquilinearform auf V

(i) Vektoren $u, v \in V$ heißen Q-orthogonal $\iff Q(u, v) = 0$

(ii) Unterräume U, U' Q-orthogonal (Kurzschreibweise $U \perp_Q U'$) $\iff Q(u, u') = 0 \; \forall \, u \in U, \; u' \in U'$

(iii) Q-orthogonales Komplement eines Unterraumes U

$$U^{\perp_Q} = \{v \in V \mid Q(v, u) = 0 \quad \text{für alle } u \in U\}$$

(iv) Unterraum U ist Q-isotrop $\iff Q(u, u) = 0$ für alle $u \in U$.

(v) U ist Q-Lagrange $\iff$ U maximal als Q-isotroper Unterraum

(vi) Unterraum U Q-positiv $\iff Q(u, u) > 0$ für alle $u \in U \setminus \{0\}$

Bemerkung 8.24

Wort *isotrop* Erfindung der Mathematiker des 19ten Jahrhundert

Satz 8.25

Q hermitisch und U Unterraum von V

(i) $U \subset (U^{\perp_Q})^{\perp_Q}$

(ii) *U ist Q-isotrop genau dann, wenn $U \subset U^{\perp_Q}$.*

(iii) *U Q-Lagrange* $\implies \dim(U) = \textit{Witt-Index} = N_0 + \min\{N_+, N_-\}$.

Wenn Q zudem nicht entartet, so

(iv) $\dim(U) + \dim(U^{\perp_Q}) = \dim(V)$

(v) $U = (U^{\perp_Q})^{\perp_Q}$

Beweis zum Teil nicht schwer (siehe Literatur) □

Bemerkung 8.26

Analoge Aussagen für anti-hermitische Bilinearformen,
also insbesondere symplektische Form.

Definition 8.27

Q eine hermitische oder anti-hermitische Sesquilinearform auf V

Lineare Abbildung $T \in \mathcal{L}(V)$ heißt Q-unitär $\iff$

$$Q(Tv, Tw) = Q(v, w) \qquad \forall \ v, w \in V$$

Satz 8.28

Die Q-unitären bilden Gruppe bez. Hintereinanderausführung

Beweis. Seien $S, T \in \mathcal{L}(V)$ beide Q-unitär. Dann für alle $v, w \in V$:

$$Q(STv, STw) = Q(Tv, Tw) = Q(v, w)$$

□

Bemerkung 8.29

Weyl's 10 sogenannten Klassische Gruppen von dieser Gestalt.

Hier zwei Beispiele.

Beispiel 8.30

Sei $V = \mathbb{C}^{N+M}$ und Lorentzform der Signatur (N, M)

$$Q(v, w) \;=\; \langle v | J\, w \rangle \qquad J \;=\; \begin{pmatrix} \mathbf{1}_N & 0 \\ 0 & -\mathbf{1}_M \end{pmatrix}$$

Lorentzgruppe $\mathrm{U}(N, M)$ der Signatur (N, M) zugehörige Gruppe

Für Matrix A der Größe $N + M$

$$Q(Av, Aw) \;=\; \langle Av | JAw \rangle \;=\; \langle v | (A^* JA) w \rangle \,, \qquad Q(v, w) \;=\; \langle v | Jw \rangle$$

Also

$$\mathrm{U}(N, M) \;=\; \{A \in \mathrm{Mat}((N + M) \times (N + M), \mathbb{C}) \,|\, A^* JA \;=\; J\}$$

Wenn A reell und $(N, M) = (3, 1)$, heißt A Lorentztransformation

Beispiel 8.31

Sei $V = \mathbb{R}^{2N}$ und symplektische Form

$$Q(v, w) = \langle v | I\, w \rangle \,, \qquad I = \begin{pmatrix} 0 & -\mathbf{1}_N \\ \mathbf{1}_N & 0 \end{pmatrix}$$

Zugehörige Gruppe $\mathrm{SP}(2N, \mathbb{R})$ heißt symplektische Gruppe

Matrix $A \in \mathrm{Mat}(2N \times 2N, \mathbb{R})$ symplektisch $\Leftrightarrow$

$$A^T I A = I$$

analog wie oben (Unterschied ist I und $\mathbb{R}$ statt J und $\mathbb{C}$)

Satz 8.32 (Spektraltheorie in Lorentzgruppe)

Sei $A \in \mathrm{U}(N, M)$ *und* J *wie oben. Dann gilt:*

(i) *Spektrum hat Spiegelsymmetrie am Einheitskreis:*

$$\mathrm{Spec}(A) \;=\; \left(\overline{\mathrm{Spec}(A)}\right)^{-1}$$

(ii) *Wenn* $Av = zv$ *und* $Aw = \zeta w$ *mit* $\bar{z}\zeta \neq 1$*, dann* $v \perp_J w$*.*

(iii) *Die gesamte algebraische Multiplizität aller Eigenwerte auf Einheitskreis* $\mathbb{S}^1$ *ist mindestens Signatur* $|N - M|$

Beweis. (i) Aus $J^{-1} = J$ folgt $A^{-1} = JA^*J$. Also

$$p_A(z) \;=\; \det(z\,\mathbf{1} - A) \;=\; \overline{\det(J)\det(\bar{z}\,\mathbf{1} - A^*)\det(J)} \;=\; \overline{p_{A^{-1}}(\bar{z})}$$

Somit

$$\mathrm{Spec}(A) \;=\; \overline{\mathrm{Spec}(A^{-1})} \;=\; \overline{\mathrm{Spec}(A)^{-1}}$$

nach spektralem Abbildungssatz (Satz 5.18)

(ii) Aussage $v \perp_J w$ äquivalent zu $\langle v|Jw\rangle = 0$. Nun

$$\langle v|Jw\rangle \;=\; \frac{1}{\overline{z}\zeta}\langle Av|JAw\rangle \;=\; \frac{1}{\overline{z}\zeta}\langle v|A^*JAw\rangle \;=\; \frac{1}{\overline{z}\zeta}\langle v|Jw\rangle$$

so dass aus Annahme $\overline{z}\zeta \neq 1$ tatsächlich $\langle v|Jw\rangle = 0$ folgt.

(iii) Vereinfachende Annahme: keine nicht-trivialen Jordan Blöcke

U Unterraum aufgespannt von Eigenvektoren zu Eigenwerten $|\lambda| > 1$

Nach (ii) ist U isotrop

Nach Satz 8.25 ist $\dim(U)$ beschränkt durch den Witt-Index $\min\{N, M\}$

Analog: Unterraum zu Eigenwerten $|\lambda| < 1$ hat Dimension $\min\{N, M\}$

Also liegen $N + M - 2\min\{N, M\} = |N - M|$ Eigenwerte auf $\mathbb{S}^1$ $\square$

9 Riemann Integral für Funktionen einer Variablen

Integral = Fläche zwischen Funktion $f : [a, b] \to \mathbb{R}$ und x-Achse

$$\approx \sum_n (x_n - x_{n-1}) f(\xi_n) \text{ mit Zwischenpunkten } \xi_n \in [x_{n-1}, x_n]$$

Alternative: Lebesgue Integral durch Zerlegung der y-Achse

Definition 9.1

(i) Zerlegung Z des endlichen Intervalls $[a, b]$ ist gegeben durch Punkte $a = x_0 < x_1 < \ldots < x_N = b$ mit endlichem $N \in \mathbb{N}$

(ii) Zerlegung Z' feiner als Zerlegung Z (Schreibweise $Z \leqslant Z'$) $\iff$ jeder Punkt in Z ist auch in Z'

(iii) Gegeben Zerlegungen Z und Z', entsteht Verfeinerung $Z \cup Z'$ durch deren Überlagerung

(iv) $\mathcal{F}(Z) = \max_{n=1,\ldots,N} \Delta x_n$ mit $\Delta x_n = x_n - x_{n-1}$ ist Feinheit von Z

Definition 9.2 (Riemann Integral)

$f : [a, b] \to \mathbb{R}$ Funktion und Z Zerlegung von $[a, b]$

(i) Untersumme $U_Z(f)$ und Obersumme $O_Z(f)$ sind

$$U_Z(f) = \sum_{n=1}^{N} \Delta x_n \cdot m_n\,, \qquad m_n = \inf\{f(x) \mid x \in [x_{n-1}, x_n]\}$$

$$O_Z(f) = \sum_{n=1}^{N} \Delta x_n \cdot M_n \qquad M_n = \sup\{f(x) \mid x \in [x_{n-1}, x_n]\}$$

(ii) Unteres und oberes Riemann Integral:

$$U(f) = \sup_Z U_Z(f)\,, \qquad O(f) = \inf_Z O_Z(f)$$

(iii) f Riemann integrierbar $\iff$ $U(f) = O(f) \in \mathbb{R}$. Dann

$$\int_a^b dx\, f(x) = U(f) = O(f) = \int_a^b f(x)\, dx = \int_{[a,b]} dx\, f(x)$$

Bemerkung 9.3

$f : [a,b] \to \mathbb{R}$ unbeschränkt $\implies O_Z(f) = \infty$ für alle Z

Riemann integrierbare Funktionen sind also immer beschränkt

Lemma 9.4

Z, Z' Zerlegungen von $[a,b] \implies U_Z(f) \leqslant O_{Z'}(f)$

Somit $U(f) \leqslant O(f)$

Beweis: Behauptung klar für $Z = Z'$

Sei zunächst $Z \leqslant Z'$ und z.B. $x_{n-1} = x'_{j-1} < x'_j < x'_{j+1} = x_n$. Dann

$$\begin{aligned} \Delta x_n m_n &= \Delta x'_j m_n + \Delta x'_{j+1} m_n \\ &\leqslant \Delta x'_j m'_j + \Delta x'_{j+1} m'_{j+1} \quad \text{(weil inf über kleineres Intervall)} \end{aligned}$$

Somit $U_Z(f) \leqslant U_{Z'}(f)$. Analog: $O_Z(f) \geqslant O_{Z'}(f)$

Jetzt $Z \leqslant Z \cup Z'$ und $Z' \leqslant Z \cup Z'$, so dass

$$U_Z(f) \leqslant U_{Z \cup Z'}(f) \leqslant O_{Z \cup Z'}(f) \leqslant O_{Z'}(f)$$

□

Satz 9.5 (Riemannsches Integrabilitätskriterium)

f integrierbar $\iff \forall\, \varepsilon > 0\ \exists$ *Zerlegung Z mit* $O_Z(f) - U_Z(f) < \varepsilon$

Beweis: "$\Longrightarrow$" $O(f) = U(f) \implies Z, Z'$ mit $O_Z(f) - U_{Z'}(f) < \varepsilon$
$\implies O_{Z\cup Z'}(f) - U_{Z\cup Z'}(f) \leqslant O_Z(f) - U_{Z'}(f) < \varepsilon$ (Beweis Lemma 9.4)
"$\Longleftarrow$" $0 \leqslant O(f) - U(f) \leqslant O_Z(f) - U_Z(f) < \varepsilon$ für geeignetes Z □

Satz 9.6 (Verschärftes Riemannkriterium)

f integrierbar
$\iff \forall\, \varepsilon > 0\ \exists\, \delta > 0$ *mit* $O_Z(f) - U_Z(f) < \varepsilon\ \forall\ Z$ *mit* $\mathcal{F}(Z) < \delta$

Beweis: (Argument f in O_Z, U_Z unterdrückt) "$\Longleftarrow$" klar nach Satz 9.5

"$\Longrightarrow$" Sei $\varepsilon > 0$. Nach Satz 9.5 $\exists$ Zerlegung Z' mit $O_{Z'} - U_{Z'} < \frac{\varepsilon}{3}$

Sei Z beliebige Zerlegung.

Behauptung: $\exists$ von Z unabhängige Konstante $C = C(f, Z')$ mit

$$O_Z - O_{Z \cup Z'} \leqslant C\,\mathcal{F}(Z)\,, \qquad U_{Z \cup Z'} - U_Z \leqslant C\,\mathcal{F}(Z)$$

Begründung: Seien x_n, x_{n+1} in Z.

Betrachte zugehörigen Beitrag B_n zu $O_Z - O_{Z \cup Z'}$.

Falls kein Punkt von Z' in $[x_n, x_{n+1}]$, ist $B_n = 0$.

Falls ein Punkt x' von Z' in $[x_n, x_{n+1}]$, gilt:

$$\begin{aligned} B_n &= (x_{n+1} - x_n) \sup_{[x_n, x_{n+1}]} f - (x_{n+1} - x') \sup_{[x', x_{n+1}]} f - (x' - x_n) \sup_{[x_n, x']} f \\ &= (x_{n+1} - x') \left(\sup_{[x_n, x_{n+1}]} f - \sup_{[x', x_{n+1}]} f \right) + (x' - x_n) \left(\sup_{[x_n, x_{n+1}]} f - \sup_{[x_n, x']} f \right) \\ &\leqslant \mathcal{F}(Z) \left(\sup_{[a,b]} f - \inf_{[a,b]} f \right) \cdot 2 \end{aligned}$$

Falls k Punkte von Z' in $[x_n, x_{n+1}]$, gilt analog:

$$B_n \leqslant \mathcal{F}(Z) \left(\sup_{[a,b]} f - \inf_{[a,b]} f \right) (k+1)$$

Summieren über Beitrag zeigt erste Behauptung. Zweite analog
Somit

$$\begin{aligned} O_Z - U_Z &\leqslant |O_Z - O_{Z \cup Z'}| + |O_{Z \cup Z'} - U_{Z \cup Z'}| + |U_{Z \cup Z'} - U_Z| \\ &\leqslant C\,\mathcal{F}(Z) + \frac{\varepsilon}{3} + C\,\mathcal{F}(Z) \\ &< \varepsilon \end{aligned}$$

für $\mathcal{F}(Z) < \delta = \frac{\varepsilon}{3C}$. $\square$

Definition 9.7

$f : [a, b] \to \mathbb{R}$ beschränkt

Z Zerlegung von $[a, b]$ mit Punkten $x_0 < x_1 < \ldots < x_N$

Seien $\xi_n \in [x_{n-1}, x_n]$ Zwischenpunkte und setze $\xi = (\xi_1, \ldots, \xi_N)$

Die zugehörige Riemannsche Zwischensumme ist definiert als

$$S_{Z,\xi}(f) \;=\; \sum_{n=1}^{N} \Delta x_n \cdot f(\xi_n)$$

Satz 9.8 (Zwischensummenbeschreibung des Integrals)

$f : [a, b] \to \mathbb{R}$ *beschränkt. Dann:* $\int_a^b dx\, f(x) = I$

$$\iff \forall\, \varepsilon > 0\ \exists\, \delta > 0 \quad \textit{mit}$$
$$|S_{Z,\xi}(f) - I| < \varepsilon \quad \forall\, Z \textit{ mit } \mathcal{F}(Z) < \delta \textit{ und } \forall\, \xi$$

Beweis "$\Longrightarrow$" Es gilt

$$U_Z - O_Z \leqslant U_Z - I \leqslant S_{Z,\xi} - I \leqslant O_Z - I \leqslant O_Z - U_Z$$

Somit $|S_{Z,\xi} - I| < O_Z - U_Z < \varepsilon \quad \forall\, \mathcal{F}(Z) < \delta$ (nach Satz 9.6)

"$\Longleftarrow$" Zu $\varepsilon > 0$ sei $\delta > 0$, so dass

$$|S_{Z,\xi} - I| < \frac{\varepsilon}{4} \qquad \text{für } \mathcal{F}(Z) < \delta \quad \forall \text{ Zwischenpunkte } \xi$$

Bestimme ξ und ξ', so dass

$$|S_{Z,\xi} - O_Z| < \frac{\varepsilon}{4} \quad \text{und} \quad |S_{Z,\xi'} - U_Z| \leqslant \frac{\varepsilon}{4}$$

Dann

$$O_Z - U_Z \leqslant |O_Z - S_{Z,\xi}| + |S_{Z,\xi} - I| + |S_{Z,\xi'} - I| + |S_{Z,\xi'} - U_Z| < \varepsilon$$

Schließe mit Satz 9.5 oder 9.6 $\square$

Korollar 9.9

$f : [a,b] \to \mathbb{R}$ *Riemann integrierbar*

$\Longrightarrow \int_a^b dx\, f(x) = \lim_{N\to\infty} \frac{b-a}{N} \sum_{n=1}^N f\left(a + \frac{n-t}{N}(b-a)\right) \quad \forall\, t \in [0,1]$

Beweis: Spezielle Zwischensumme mit $\xi_n = a + \frac{n-t}{N}(b-a)$ □

Beispiel 9.10 (Berechnung eines Integrals)

$$
\begin{aligned}
\int_a^b dx\, x^2 &= \lim_{N\to\infty} \frac{b-a}{N} \sum_{n=1}^{N} \left(a + \tfrac{n}{N}(b-a)\right)^2 \\
&= \lim_{N\to\infty} \frac{b-a}{N} \sum_{n=1}^{N} \left(a^2 + n\tfrac{2}{N}(b-a)a + \tfrac{1}{N^2}(b-a)^2 n^2\right) \\
&= \lim_{N\to\infty} \frac{b-a}{N} \left(Na^2 + \frac{2(b-a)a}{N}\frac{N(N+1)}{2} + \frac{(b-a)^2}{N^2}\frac{N(N+1)(2N+1)}{6}\right) \\
&= (b-a)\left(a^2 + (b-a)a + \tfrac{1}{3}(b-a)^2\right) \\
&= (b-a)\tfrac{1}{3}(b^2 + ab + a^2) \;=\; \tfrac{1}{3}(b^3 - a^3)
\end{aligned}
$$

Länglich. Bessere Alternative später (Hauptsatz)

Satz 9.11

Jede stetige Funktion ist (Riemann) integrierbar

Beweis: Da $[a, b]$ kompakt, $f : [a, b] \to \mathbb{R}$ gleichmäßig stetig (Ana 1), d.h. $\forall\, \varepsilon > 0\ \exists\, \delta > 0$, so dass $|f(x) - f(x')| < \frac{\varepsilon}{b-a}\ \ \forall\, |x - x'| < \delta$
Für Zerlegung Z mit $\mathcal{F}(Z) < \delta$ gilt also

$$O_Z(f) - U_Z(f) \ < \ \sum_{n=1}^{N} \Delta x_n \cdot \frac{\varepsilon}{b-a} \ = \ \varepsilon$$

Schließe mit Satz 9.6 □

Definition 9.12

$f : [a, b] \to \mathbb{R}$ (Riemann'sche) Treppenfunktion
$\iff\ \exists$ Zerlegung Z von $[a, b]$ mit f konstant auf $(x_{n-1}, x_n)\ \forall\ n$

Es gibt unstetige, integrierbare Funktionen:

Satz 9.13

Treppenfunktionen sind integrierbar

Beweis: $O_Z(f) - U_Z(f)$ beliebig klein, denn Beitrag der endlich vielen Sprungstellen wird klein $\square$

Beispiel 9.14 (Eine nicht Riemann integrierbare Funktion)

$$f(x) = \begin{cases} 1 & x \in \mathbb{Q} \cap [0,1] \\ 0 & x \in \mathbb{R} \backslash \mathbb{Q} \cap [0,1] \end{cases}$$

Dann $\forall$ Zerlegungen Z gilt $O_Z(f) = 1 > U_Z(f) = 0$

Aber: f Lebesgue integrierbar und $\int_0^1 dx\, f(x) = 0$, weil $\mathbb{Q}$ "nur dünn"

Satz 9.15

(i) *Monotone Funktionen sind integrierbar*

(ii) f, g *integrierbar*, $\lambda \in \mathbb{R} \implies f + \lambda g$ *und* $f \cdot g$ *integrierbar*
(d.h. integrierbare Funktionen bilden eine Algebra)

(iii) f *integrierbar* $\implies |f|$ *integrierbar*

(iv) f *integrierbar*, $h \in C(\mathbb{R}, \mathbb{R}) \implies h \circ f$ *integrierbar*

Beweis: (i) Sei f monoton steigend. Dann

$$\begin{aligned} O_Z(f) - U_Z(f) &= \sum_{n=1}^{N} \Delta x_n \underbrace{(f(x_n) - f(x_{n-1}))}_{\geqslant 0} \\ &\leqslant \mathcal{F}(Z) \sum_{n=1}^{N} (f(x_n) - f(x_{n-1})) \\ &= \mathcal{F}(Z)(f(b) - f(a)) < \varepsilon \end{aligned}$$

Letzteres falls $\mathcal{F}(Z) < \frac{\varepsilon}{f(b)-f(a)} = \delta$.

(ii) Sei $\lambda > 0$.

$$\begin{aligned} O_Z(f+\lambda g) &= \sum_{n=1}^{N} \Delta x_n \sup_{[x_{n-1},x_n]} (f+\lambda g) \\ &\leqslant \sum_{n=1}^{N} \Delta x_n \left(\sup_{[x_{n-1},x_n]} f + \lambda \sup_{[x_{n-1},x_n]} g \right) \\ &= O_Z(f) + \lambda O_Z(g) \end{aligned}$$

Analog $U_Z(f+\lambda g) \geqslant U_Z(f) + \lambda U_Z(g)$

Somit nach Satz 9.5

$$O_Z(f+\lambda g) - U_Z(f+\lambda g) \leqslant (O_Z(f) - U_Z(f)) + \lambda(O_Z(g) - U_Z(g)) < \frac{\varepsilon}{2} + \frac{\lambda\varepsilon}{2}$$

Also ist $f + \lambda\, g$ integrierbar

Fall $\lambda < 0$ ist eine Übung

Hilfsmittel lokale Variation $v_n(f) = \sup_{x,x' \in [x_{n-1},x_n]} |f(x) - f(x')|$. Dann:

$$O_Z(f) - U_Z(f) \;=\; \sum_{n=1}^{N} \Delta x_n \sup_{x,x' \in [x_{n-1},x_n]} |f(x) - f(x')| \;=\; \sum_{n=1}^{N} \Delta x_n v_n(f)$$

Da

$$f(x)g(x) - f(x')g(x') \;=\; (f(x) - f(x'))g(x) + f(x')(g(x) - g(x'))$$

gilt

$$v_n(fg) \;\leqslant\; \|g\|_\infty v_n(f) \;+\; \|f\|_\infty v_n(g)$$

wobei

$$\|f\|_\infty \;=\; \sup_{x \in [a,b]} |f(x)|$$

Somit

$$O_Z(fg) - U_Z(fg) \;\leqslant\; \|g\|_\infty (O_Z(f) - U_Z(f)) \;+\; \|f\|_\infty (O_Z(g) - U_Z(g))$$

und *fg* integrierbar weil *f* und *g* beide integrierbar.

(iii) Variation erfüllt $v_n(|f|) \leqslant v_n(f)$. Dann wie in (ii)

(iv) h gleichmäßig stetig auf Kompaktum $[-\|f\|_\infty, \|f\|_\infty]$ (Ana 1),

d.h. $\forall\, \varepsilon > 0\ \exists\, \delta > 0$ mit $|h(y) - h(y')| < \frac{\varepsilon}{2(b-a)} \quad \forall\, |y - y'| < \delta$

Da f integrierbar, wähle Z mit $O_Z(f) - U_Z(f) < \frac{\varepsilon \cdot \delta}{4\|h\|_\infty}$

Seien $\sum_n'$ und $\sum_n''$ Summen über Terme mit $v_n(f) < \delta$ bzw. $v_n(f) \geqslant \delta$

Dann $O_Z(f) - U_Z(f) \geqslant \delta \sum_n'' \Delta x_n$, und somit $\sum_n'' \Delta x_n < \frac{\varepsilon}{4\|h\|_\infty}$. Also

$$\begin{aligned}
O_Z(h \circ f) - U_Z(h \circ f) &= \sum_{n=1}^{N}{}' \Delta x_n v_n(h \circ f) + \sum_{n=1}^{N}{}'' \Delta x_n v_n(h \circ f) \\
&\leqslant \sum_{n=1}^{N}{}' \Delta x_n \frac{\varepsilon}{2(b-a)} + \sum_{n=1}^{N}{}'' \Delta x_n 2\|h\|_\infty \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

$\square$

Jetzt: Aussagen für Ober- und Untersummen gelten für Integral:

Satz 9.16 (Rechenregeln fürs Integral)

$f, g : [a, b] \to \mathbb{R}$ *integrierbar,* $c \in (a, b)$ *und* $\lambda \in \mathbb{R}$

(i) *Linearität des Integrals:*

$$\int_a^b dx(f(x) + \lambda g(x)) = \left(\int_a^b dx\, f(x)\right) + \lambda \left(\int_a^b dx\, g(x)\right)$$

(ii) *Intervalladditivität:*

$$\int_a^b dx\, f(x) = \int_a^c dx\, f(x) + \int_c^b dx\, f(x)$$

(iii) *Monotonie:*

$$f(x) \leqslant g(x) \ \forall\, x \in [a, b] \quad \Longrightarrow \quad \int_a^b dx\, f(x) \leqslant \int_a^b dx\, g(x)$$

(iv) *Standardabschätzung:*

$$\left|\int_a^b dx\, f(x)\right| \leqslant \int_a^b dx |f(x)| \leqslant (b - a)\|f\|_\infty$$

Satz 9.17

$(f_n)_{n\geqslant 1}$ *Folge integrierbarer Funktionen auf* $[a,b]$
Folge konvergiere gleichmäßig gegen $f = \lim_{n\to\infty} f_n(x)$
$\Longrightarrow$ *f integrierbar und* $\lim_{n\to\infty}\int_a^b dx\, f_n(x) = \int_a^b dx\, f(x)$ ($\int\lim = \lim\int$)
Entsprechendes gilt für gleichmäßig konvergente Reihen: $\sum_n \int = \int \sum_n$

Beweis: Wie im Beweis von Satz 9.15(ii):

$$\begin{aligned} O_Z(f) - U_Z(f) &= (O_Z(f) - O_Z(f_n)) - (U_Z(f) - U_Z(f_n)) + O_Z(f_n) - U_Z(f_n) \\ &\leqslant O_Z(f - f_n) - U_Z(f - f_n) + O_Z(f_n) - U_Z(f_n) \\ &\leqslant 2(b-a)\|f - f_n\|_\infty + O_Z(f_n) - U_Z(f_n) \end{aligned}$$

Wähle n mit $\|f - f_n\|_\infty < \frac{\varepsilon}{4(b-a)}$, dann Z mit $O_Z(f_n) - U_Z(f_n) < \frac{\varepsilon}{2}$
Nach Riemannkriterium ist f also integrierbar. Mit Rechenregeln

$$\begin{aligned} \left|\int_a^b dx\, f(x) - \int_a^b dx\, f_n(x)\right| &= \left|\int_a^b dx(f(x) - f_n(x))\right| \\ &\leqslant (b-a)\|f - f_n\|_\infty \xrightarrow{n\to\infty} 0 \end{aligned}$$

$\square$

Definition 9.18

$f, F : [a, b] \to \mathbb{R}$ Funktionen

F stetig auf $[a, b]$ und differenzierbar auf (a, b)

F Stammfunktion von f (Bezeichnung $F(x) = \int dx\, f(x)$) $\iff F' = f$

Bemerkung 9.19

F Stammfunktion von $f \iff G = F + c$ Stammfunktion von $f \ \forall\, c \in \mathbb{R}$

Theorem 9.20 (Hauptsatz der Differential- und Integralrechnung)

$f : [a, b] \to \mathbb{R}$ *Riemann integrierbar*

(i) *Für jede Stammfunktion F von f gilt*

$$\int_a^b dx\, f(x) = F(b) - F(a)$$

(ii) *Unbestimmtes Integral $G(x) = \int_a^x dt\, f(t)$ Lipschitz-stetig auf $[a, b]$*

(iii) *f stetig auf $(a, b) \implies G$ differenzierbar und $G'(x) = f(x)$*

Beweis: (i) Für Zerlegung Z von $[a, b],\ a = x_0 < x_1 < \ldots < x_N = b$:

$$\begin{aligned} F(b) - F(a) &= \sum_{n=1}^{N} (F(x_n) - F(x_{n-1})) \\ &= \sum_{n=1}^{N} (x_n - x_{n-1}) f(\xi_n) \quad (\text{MWS } x_{n-1} < \xi_n < x_n) \\ &= S_{Z,\xi}(f) \end{aligned}$$

Nun wähle Folge $(Z_j)_{j\in\mathbb{N}}$ mit $\lim_{j\to\infty} \mathcal{F}(Z_j) = 0$. Dann nach Satz 9.8:

$$F(b) - F(a) = \lim_j S_{Z_j,\xi_j}(f) = \int_a^b dx\, f(x)$$

(ii) Lipschitz-Stetigkeit: Seien $x, x' \in [a, b],\ x < x'$ und $L = \|f\|_\infty$:

$$|G(x') - G(x)| = \left| \int_x^{x'} dt\, f(t) \right| \leqslant \int_x^{x'} dt\, |f(t)| \leqslant |x - x'| \sup_{[a,b]} |f| \leqslant L|x - x'|$$

Zuletzt zur Differenzierbarkeit:

$$\begin{aligned}\left|\frac{G(x')-G(x)}{x'-x}-f(x)\right| &= \frac{1}{|x'-x|}\left|\int_x^{x'} dt(f(t)-f(x))\right| \\ &\leqslant \frac{1}{|x'-x|}\int_x^{x'} dt|f(t)-f(x)| \\ &\leqslant \sup_{t\in[x,x']}|f(t)-f(x)| < \varepsilon\end{aligned}$$

für $|x-x'|<\delta$ nach Stetigkeit von f in x. $\square$

Bemerkung 9.21

Umschreiben: wenn $F \in C^1([a,b])$, dann F' Riemann integrierbar und

$$\int_a^x dt\, F'(t) = F(x)-F(a)$$

Aber $f(x)=\mathrm{sgn}(x)$ hat unbestimmtes Integral $G(x)=\int_0^x dt\, f(t)=|x|$, was nicht differenzierbar in 0 ist, dort wohl aber Lipshitz-stetig.

Beispiel 9.22 (Einige Stammfunktionen)

All die folgenden Formeln erhält man durch Ableiten:

1. $\alpha \neq 1,\ \int dx\, x^\alpha = \frac{x^{\alpha+1}}{\alpha+1} + c$ für entweder $\alpha \in \mathbb{N}$ oder $x > 0$
2. $\int dx \frac{1}{x} = \ln|x| + c$ für $x \in \mathbb{R}\setminus\{0\}$
3. $\int dx\, e^x = e^x + c$
4. $\int dx\, a^x = \int dx\, e^{\ln(a)x} = \frac{a^x}{\ln(a)} + c$ für $a > 0,\ a \neq 1$
5. $\int dx\, \cos(x) = \sin(x) + c$
6. $\int dx\, \sin(x) = -\cos(x) + c$
7. $\int dx\, \tan(x) = -\ln|\cos(x)| + c$ für $x \in \mathbb{R}\setminus\left\{\left(n+\frac{1}{2}\right)\pi \mid n \in \mathbb{Z}\right\}$
8. $\int dx\, \cot(x) = \ln|\sin(x)| + c$ für $x \in \mathbb{R}\setminus\{n\pi \mid n \in \mathbb{Z}\}$
9. $\int dx \frac{1}{1+x^2} = \arctan(x) + c$
10. $\int dx \frac{1}{\sqrt{1-x^2}} = \arcsin(x) + c$ für $x \in (-1, 1)$

Aber:

Stammfunktionen elementarer Funktionen nicht immer elementar!

Definition 9.23

Elementare Funktionen sind rationale Funktionen, trigonometrische Funktionen, Exponentialfunktionen, all deren Umkehrungen und sämtliche Kombinationen (Summen, Produkte, Kompositionen) davon.

Beispiel 9.24

Folgende Stammfunktionen definieren nicht elementare Funktionen:

Integrallogarithmus	$\int dx \, \frac{1}{\ln(x)}$
Integralsinus	$\int dx \, \frac{\sin(x)}{x}$
Gauss'sche Fehlerfunktion	$\int dx \, e^{-x^2}$

Integrationstechniken

Satz 9.25 (Partielle Integration)

$f, g : [a, b] \to \mathbb{R}$ *differenzierbar mit integrierbaren Ableitungen* f', g'

$$\int_a^b dx\ f(x)g'(x) \ = \ f(b)g(b) - f(a)g(a) - \int_a^b dx\ f'(x)g(x)$$

und analog für unbestimmte Integrale.

Beweis: Produktregel $(fg)' = f'g + fg'$, und dann Fundamentalsatz. □

Beispiel 9.26

1. $\int dx\ \ln(x) \ = \ x\ln(x) - \int dx\ \frac{1}{x}x \ = \ x\ln(x) - x + c$
2. $\int dx\ x^n \cos(x) \ = \ x^n \sin(x) - n\int dx\ x^{n-1}\sin(x)$ dann Iteration
3. $\int dx\ x^3 \arctan(x) \ = \ \left(\frac{x^4}{4} + c'\right)\arctan(x) - \int dx\ \left(\frac{x^4}{4} + c'\right)\frac{1}{1+x^2}$
   $= \frac{1}{4}(x^4 - 1)\arctan(x) - \frac{1}{4}\int dx (x^2 - 1)$ nach Wahl $c' = \frac{1}{4}$
   $= \frac{1}{4}(x^4 - 1)\arctan(x) - \frac{1}{12}x^3 + \frac{1}{4}x + c$

Satz 9.27 (Substitutionsregel)

$\phi : [\alpha, \beta] \to \mathbb{R}$ differenzierbar mit $\phi([\alpha, \beta]) \subset [a, b]$ und ϕ' integrierbar
Für $f \in C([a, b], \mathbb{R})$

$$\int_{\phi(\alpha)}^{\phi(\beta)} dx\ f(x) = \int_{\alpha}^{\beta} dt\ f(\phi(t))\, \phi'(t)$$

$$\left(\int dx\ f(x)\right) \circ \phi = \int dt\ f(\phi(t))\phi'(t)$$

Beweis: Stammfunktion F von f stetig differenzierbar (Hauptsatz)
$\implies (F \circ \phi)' = (F' \circ \phi)\phi' = (f \circ \phi)\phi'$ integrierbar, dann Hauptsatz □

Korollar 9.28 (Variablentransformation)

Zudem $\phi : [\alpha, \beta] \to [a, b]$ bijektiv. Dann

$$\int_{a}^{b} dx\ f(x) = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} dt\ f(\phi(t))\phi'(t)$$

Beispiel 9.29

1. $\int_a^b dt\, f(\alpha t+\beta) = \frac{1}{\alpha}\int_{\alpha a+\beta}^{\alpha b+\beta} dx\, f(x)$
2. $\int dx \frac{\phi'(x)}{\phi(x)} = \log|\phi(x)| + c,$ vorausgesetzt $\phi(x) \neq 0$
3. $R(x,y) = \frac{P(x,y)}{Q(x,y)}$ rationale Funktion in zwei Variablen

$$\int dx\, R(x, \sqrt[n]{ax+b}) = \int dt\, R\left(\frac{t^n - b}{a}, t\right) \frac{n}{a} t^{n-1} + c$$

weil: $t = \sqrt[n]{ax+b}$, $ax+b = t^n$, $x = \frac{t^n-b}{a}$, $dx = \frac{nt^{n-1}}{a}dt$

Rechte Seite jetzt rational in t (Methode unten). Beispiel:

$$\int_0^1 dx \frac{x}{\sqrt{x+1}} = \int_1^{\sqrt{2}} dt\, \frac{t^2-1}{t} 2t = \frac{2}{3}t^3 - 2t\Big|_1^{\sqrt{2}} = \frac{4}{3} - \frac{2}{3}\sqrt{2}$$

4. $\int dx\, R(\cos x, \sin x) = \int dt\, \frac{2}{1+t^2} R\left(\frac{1-t^2}{1+t^2}, \frac{2t}{1+t^2}\right) + c$ (rational in t)

weil: setze $t = \tan\left(\frac{x}{2}\right)$, $\frac{1-t^2}{1+t^2} = \cos(x)$, $x = 2\arctan(t)$, $dx = \frac{2\,dt}{1+t^2}$

Methode für Integration rationaler Funktionen

1. Schritt: Polynomdivision mit Rest

$R(x) = \frac{P(x)}{Q(x)} = \text{Polynom}(x) + \frac{p(x)}{q(x)}$ wobei $\deg(p) < \deg(q)$, teilerfremd

2. Schritt: Faktorisierung von q (Fundamentalsatz der Algebra)

$q(x) = (x - z_1)^{\ell_1} \cdot (x - z_2)^{\ell_2} \cdot \ldots \cdot (x - z_n)^{\ell_n}$

$z_j \in \mathbb{C}$ ist ℓ_j-fache Nullstelle von q, nicht Nullstelle von p

3. Schritt: Partialbruchzerlegung (eindeutig)

$$\frac{p(x)}{q(x)} = \sum_{j=1}^{n} \sum_{\ell=1}^{\ell_j} \frac{A_{j,\ell}}{(x - z_j)^{\ell}} \qquad \text{mit } A_{j,\ell} \in \mathbb{C}$$

Berechnung der $A_{j,\ell}$:

(i) multipliziere mit q

(ii) Koeffizientenvergleich

(iii) löse lineare Gleichungen

Nachweis der Existenz der Partialbruchzerlegung

Induktion über $m = \deg(q)$. Anfang OK, also Schritt $(m-1) \to m$

Sei z Nullstelle von q der Ordnung $\ell \geqslant 1$, d.h.

$$q(x) = (x-z)^\ell r(x) \qquad \text{mit Polynom } r \text{ mit } r(z) \neq 0$$

Dann

$$\frac{p(x)}{r(x)} - \frac{p(z)}{r(z)} = \frac{p(x)r(z) - p(z)r(x)}{r(x)r(z)} = \frac{(x-z)s(x)}{r(x)}$$

mit Polynom s (weil Zähler z als Nullstelle hat). Somit

$$\frac{p(x)}{q(x)} = \frac{p(x)}{(x-z)^\ell r(x)} = \frac{p(z)}{r(z)} \frac{1}{(x-z)^\ell} + \frac{s(x)}{(x-z)^{\ell-1} r(x)}$$

d.h. neuer Summand und Term mit $\deg((x-z)^{\ell-1} r(x)) = m-1$

Eindeutigkeit: Gegeben zweite Zerlegung mit Koeffizienten $B_{j,\ell}$

multipliziere mit $(x-z_j)^{\ell_j}$ und setze $x = z_j \implies A_{j,\ell_j} = B_{j,\ell_j}$

Dann multipliziere mit $(x-z_j)^{\ell_j - 1}$, etc. $\square$

4. Schritt: Termweise Integrieren

$$\int dx \frac{P(x)}{Q(x)} = \int dx \text{ Polynom}(x) + \sum_{j=1}^{n} \sum_{\ell=1}^{\ell_j} A_{j,\ell} \int dx \frac{1}{(x - z_j)^\ell}$$

Integrand ist komplexwertig

Definition 9.30

$f : [a,b] \to \mathbb{C}$ integrierbar $\iff$ $\Re e(f), \Im m(f)$ integrierbar. Dann

$$\int_a^b dx\, f(x) = \int_a^b dx\, \Re e\big(f(x)\big) + i \int_a^b dx\, \Im m\big(f(x)\big)$$

Hauptsatz gilt für Real- und Imaginärteil separat.

Gegeben $F : [a,b] \longrightarrow \mathbb{C}$ mit $F' = f$ (Ableitung von $\Re e(F), \Im m(F)$)

$$\implies \int_a^b dx\, f(x) = F(b) - F(a)$$

In Schritt 4 werden nun folgende Integrale benötigt:

Lemma 9.31

Seien $\ell \in \mathbb{N}$ *und* $z \in \mathbb{C}$

(i) $\int dx\,(x-z)^\ell \;=\; \frac{1}{\ell+1}(x-z)^{\ell+1} \;+\; c$ *(falsch für* $\ell \in \mathbb{R}\setminus\mathbb{Z}$ *!)*

(ii) $\int dx \frac{1}{(x-z)^\ell} \;=\; -\frac{1}{\ell-1}\frac{1}{(x-z)^{\ell-1}} \;+\; c$ *für* $\ell > 1$

(iii) $\int dx \frac{1}{x-z} = \frac{1}{2}\ln((x-u)^2+v^2) + i\arctan\left(\frac{x-u}{v}\right) \;+\; c$ *mit* $z = u + iv$

Beweis: (i) Sei mit $z = u + iv$,

$$
\begin{aligned}
\frac{d}{dx}(x-z)^{\ell+1} &= \frac{d}{dx}\sum_{k=0}^{\ell+1}\binom{\ell+1}{k}(x-u)^k(-iv)^{\ell+1-k} \\
&= \sum_{k=1}^{\ell+1}(\ell+1)\frac{\ell!}{(\ell+1-k)!k!}\cdot k(x-u)^k(-iv)^{\ell-(k-1)} \\
&= (\ell+1)(x-z)^\ell
\end{aligned}
$$

Dann komplexer Hauptsatz.

Für (ii) überprüfe $\frac{d}{dx}\frac{1}{(x-z)^{\ell-1}} = -(\ell-1)\frac{1}{(x-z)^{\ell}}$

Zu (iii):

$$\int dx\,\frac{1}{x-z} = \int dx\,\frac{x-u}{(x-u)^2+v^2} + i\int dx\frac{v}{(x-u)^2+v^2}$$

Substitutionen im ersten Integral: $y=(x-u)^2$, $dy = 2(x-u)dx$

Substitutionen im zweiten Integral: $\tilde{y} = \frac{x-u}{v}$, $d\tilde{y} = \frac{dx}{v}$

Also

$$\begin{aligned}\int dx\,\frac{1}{x-z} &= \frac{1}{2}\int dy\,\frac{1}{y-v^2} + i\int d\tilde{y}\frac{1}{\tilde{y}^2+1}\\ &= \frac{1}{2}\ln|y-v^2| + i\arctan(\tilde{y}) + c\end{aligned}$$

□

Satz 9.32

Die Stammfunktion einer rationalen Funktion ist elementar.

Sie kann algorithmisch mit Partialbruchzerlegung berechnet werden

Beispiel 9.33

$$
\begin{aligned}
\int dx \frac{4x^2-4x}{(x^2+1)^2} &= \int dx \frac{4x^2-4x}{(x-i)^2(x+i)^2} \quad (2.Schritt) \\
&= \int dx \left(\frac{A}{x-i} + \frac{B}{(x-i)^2} + \frac{C}{x+i} + \frac{D}{(x+i)^2} \right) \quad (3.Schritt) \\
&= \int dx \left(\frac{-i}{x-i} + \frac{1+i}{(x-i)^2} + \frac{i}{x+i} + \frac{1-i}{(x+i)^2} \right) \quad (Nebenrechnung) \\
&= -\frac{i}{2}\ln(x^2+1) + \arctan(x) + (1+i)\frac{-1}{x-i} \\
&\quad + \frac{i}{2}\ln(x^2+1) - \arctan(-x) + (1-i)\frac{-1}{x+i} \; + \; c \quad (4.Schritt) \\
&= 2\arctan(x) + 2\frac{1-x}{1+x^2} \; + \; c
\end{aligned}
$$

Beispiel 9.33 (Nebenrechnung:)

$$\begin{aligned} 4x^2 - 4x &= A(x^2+1)(x+i) + B(x+i)^2 + C(x^2+1)(x-i) + D(x-i)^2 \\ &= A(x^3 + ix^2 + x + i) + B(x^2 + 2ix - 1) \\ &\quad + C(x^3 - ix^2 + x - i) + D(x^2 - 2ix - 1) \end{aligned}$$

Koeffizientenvergleich und LA:

$$\begin{aligned} O(x^3):&\ 0 = A + C \quad \Longrightarrow \quad C = -A \\ O(x^2):&\ 4 = iA + B - iC + D = 2iA + B + D \\ O(x^1):&\ -4 = A + 2iB + C - 2iD = 2i(B - D) \\ O(x^0):&\ 0 = iA - B + C - D = 2iA - (B + 0) \end{aligned}$$

$$\begin{aligned} O(x^2) + O(x^0):&\ 4 = 4iA \Longrightarrow A = -i \Longrightarrow C = i \\ O(x^2) - O(x^0):&\ 4 = 2(B + D) \Longrightarrow B + D = 2 \\ O(x):&\ B - D = 2i \Longrightarrow B = 1 + i \text{ und } D = 1 - i \end{aligned}$$

Definition 9.34

$a, b \in \mathbb{R} \cup \{\pm\infty\}$ und $f : (a, b) \to \mathbb{R}$ (oder $\mathbb{C}$)

f uneigentlich integrierbar $\iff$ f integrierbar auf allen $[c, d] \subset (a, b)$ und $\lim_{c\downarrow a} \int_c^d dx\, f(x)$ und $\lim_{d\uparrow b} \int_c^d dx\, f(x)$ existieren. Dann

$$\int_a^b dx\, f(x) = \lim_{c\downarrow a} \lim_{d\uparrow b} \int_c^d dx\, f(x)$$

Bemerkung 9.35

Eine uneigentlich integrierbare Funktion auf (a, b) kann bei a und b Singularitäten haben, d.h. z.B. $\lim_{c\downarrow a} f(c) = \infty$ möglich

Integrierbarkeit besagt, dass die Fläche der unendlich langen Spitze dennoch endlich ist

Beispiel 9.36

Für $\alpha < 1$

$$\int_0^1 dx\,\frac{1}{x^\alpha} = \lim_{c \downarrow 0} \frac{-1}{(\alpha-1)x^{\alpha-1}}\bigg|_c^1 = \frac{1}{1-\alpha}$$

Für $\alpha > 1$ ist $\frac{1}{x^\alpha}$ nicht (uneigentlich) integrierbar bei 0. Analog:

$$\int_1^\infty dx\,\frac{1}{x^\alpha} = \lim_{d \to \infty} \frac{-1}{(\alpha-1)x^{\alpha-1}} - \frac{-1}{(\alpha-1)}\bigg|_1^d = \frac{1}{\alpha-1} \qquad \text{falls } \alpha > 1$$

Für $\alpha < 1$ ist $\frac{1}{x^\alpha}$ nicht integrierbar bei ∞

Beispiel 9.37 (Cauchy-Integral)

$$\begin{aligned}\int_{-\infty}^\infty dx\,\frac{1}{1+x^2} &= \lim_{a \to -\infty} \lim_{b \to +\infty} \int_a^b dx\,\frac{1}{1+x^2} \\ &= \lim_{a \to -\infty} \lim_{b \to \infty} (\arctan(b) - \arctan(a)) = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi\end{aligned}$$

Beispiel 9.38 (Ein rationales Integral in $\sin(x)$)

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} dx\, \frac{2y}{y^2 - \sin^2(x)} &= \int_0^{\frac{\pi}{2}} dx\, \frac{2y}{y^2 - \frac{1}{2}(1 - \cos(2x))} \\
&= \int_0^{\pi} dx\, \frac{2y}{2y^2 - 1 + \cos(x)} \quad \text{mit } t = \tan\left(\tfrac{x}{2}\right) \\
&= \int_{\tan\left(\frac{0}{2}\right)}^{\tan\left(\frac{\pi}{2}\right)} dt\, \frac{2}{1+t^2}\, \frac{2y}{2y^2 - 1 + \frac{1-t^2}{1+t^2}} \\
&= 2\int_0^{\infty} dt\, \frac{2y}{(1+t^2)(2y^2-1) + 1 - t^2} \\
&= 2\int_0^{\infty} dt\, \frac{2y}{t^2(2y^2-2) + 2y^2} \\
&= \frac{2}{y}\int_0^{\infty} dt\, \frac{1}{t^2\left(1 - \frac{1}{y^2}\right) + 1} \\
&= \frac{2}{y}\frac{1}{\sqrt{1 - \frac{1}{y^2}}}\int_0^{\infty} dt\, \frac{1}{t^2+1} = \frac{2}{\sqrt{y^2-1}}\frac{\pi}{2}
\end{aligned}$$

Satz 9.39 (Majorantenkriterium)

$f, g : (a, b) \to \mathbb{R}$, *mit* $|f| \leqslant g$ *und* f *stetig*

$$\int_a^b dx\, g(x) \text{ existiert} \quad \Longrightarrow \quad \int_a^b dx\, f(x) \text{ existiert}$$

Beweis:

$$\left| \int_c^d dx\, f(x) \right| \leqslant \int_c^d dx\, |f(x)| \leqslant \int_c^d dx\, g(x)$$

Dann bilde Limes $c \downarrow a$ und $d \uparrow b$. □

Beispiel 9.40

$$\int_0^\infty dx\, \frac{e^{-\frac{1}{1+x}}}{1+x^2} \leqslant \int_0^\infty dx\, \frac{1}{1+x^2} = \frac{\pi}{2}$$

Beispiel 9.41 (Euler'sche Γ-Funktion)

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

Integral uneigentlich bei 0 und ∞. Spalte auf $\int_0^\infty = \int_0^1 + \int_1^a + \int_a^\infty$

$$\int_0^1 dt\, e^{-1} t^{x-1} \;\leqslant\; \int_0^1 dt\, t^{x-1} \;<\; \infty \qquad \text{für } x-1 > -1 \iff x > 0$$

$$\int_a^\infty dt\, e^{-t} t^{x-1} \;=\; \int_a^\infty dt\, e^{-t+(x-1)\ln(t)} \;\leqslant\; \int_a^\infty dt\, e^{-t+\frac{1}{2}t} \;<\; \infty$$

für a ausreichend groß. Zusammenfassend: $\Gamma(x)$ definiert $\forall\; x > 0$

$$\int_\varepsilon^R dt\, e^{-t} t^x \;=\; -e^{-t} t^x \big|_{t=\varepsilon}^{t=R} + x \int_\varepsilon^R dt\, e^{-t} t^{x-1} \quad \text{(partielle Integration)}$$

Im Limes $\varepsilon \downarrow 0$ und $R \uparrow \infty$ folgt Funktionalgleichung: $\Gamma(x+1) = x\Gamma(x)$

Da $\Gamma(1) = \int_0^\infty dt\, e^{-t} = 1$, folgt $\Gamma(n) = (n-1)!$

Satz 9.42 (Integralkriterium für Konvergenz von Reihen)

Sei $f : [0, \infty) \to \mathbb{R}$ *positiv und monoton fallend mit* $\int_0^\infty dx\, f(x) < \infty$

Sei $(a_n)_{n\in\mathbb{N}}$ *Folge in* $\mathbb{R}$ *mit* $|a_n| \leqslant f(n)$

$\implies \sum_{n\geqslant 0} a_n$ *absolut konvergent.*

Beweis:

$$\sum_{n=1}^{\infty} |a_n| \leqslant \sum_{n=1}^{\infty} \int_{n-1}^{n} dx\, f(x) = \int_0^\infty dx\, f(x) < \infty$$

□

Beispiel 9.43 (Riemann'sche ζ-Funktion)

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}\,, \qquad s > 1,\ s \in \mathbb{R}$$

Tatsächlich ist dies wohldefiniert für $s > 1$, wie Vergleich mit der integrierbaren Funktion $f(x) = \min\left\{1, \frac{1}{x^s}\right\}$ zeigt

Satz 9.44 (Parameterabhängige Integrale)

Seien $a < b,\ c < d$ und $f : [a,b] \times [c,d] \to \mathbb{R}$ stetig.

Dann existiert $F(y) = \int_a^b dx\, f(x,y)$ und ist stetig auf $[c,d]$

Zudem: $f(x,y)$ stetig differenzierbar in $y \Longrightarrow F$ auch und

$$\frac{d}{dy} \int_a^b dx\, f(x,y) \;=\; \int_a^b dx\, \frac{d}{dy}\, f(x,y)$$

Beweis: Sei $y \in [c,d]$ und $(y_n)_{n\geqslant 1}$ Folge mit $\lim y_n = y$

Für die Stetigkeit von F zeigen wir $\lim_{n\to\infty} F(y_n) = F(y)$

Dann $f_n(x) := f(x, y_n)$ integrierbar weil stetig. Zudem

$$\lim_{n\to\infty} f_n(x) \;=\; g(x) \qquad \text{mit}\ \ g(x) = f(x,y)$$

Sogar gleichmässige (uniforme) Konvergenz $\|f_n - g\|_\infty \to 0$ auf $[a,b]$

Nach Satz 9.17

$$\lim_{n\to\infty} F(y_n) \;=\; \lim_{n\to\infty} \int_a^b dx\, f_n(x) \;=\; \int_a^b dx\, \lim_{n\to\infty} f_n(x) \;=\; F(y)$$

Beweis zweiter Aussage verwendet gleichmäßige Stetigigkeit von

$$(x, y) \in [a, b] \times [c, d] \;\mapsto\; \frac{d}{dy} f(x, y)$$

Jetzt:

$$\begin{aligned}
&\left| \frac{F(y+\varepsilon) - F(y)}{\varepsilon} - \int_a^b dx\, \frac{d}{dy} f(x, y) \right| \\
&= \left| \int_a^b dx \left(\frac{f(x, y+\varepsilon) - f(x, y)}{\varepsilon} - \frac{d}{dy} f(x, y) \right) \right| \qquad \text{(Linearität von } \textstyle\int\text{)} \\
&\leqslant \int_a^b dx \left| \frac{d}{dy} f(x, \xi) - \frac{d}{dy} f(x, y) \right| \qquad \text{(MWS mit } \xi \in [y, y+\varepsilon]\text{)} \\
&\leqslant \int_a^b dx \sup_{a \leqslant x' \leqslant b} \sup_{y \leqslant \xi \leqslant y+\varepsilon} \left| \frac{d}{dy} f(x', \xi) - \frac{d}{dy} f(x', y) \right| \;\leqslant\; \delta(\varepsilon)
\end{aligned}$$

mit $\lim_{\varepsilon \to 0} \delta(\varepsilon) = 0$ (nach gleichmäßiger Stetigkeit von $\frac{d}{dy} f$)

Somit $\frac{d}{dy} F(y) = \int_a^b dx\, \frac{d}{dy} f(x, y)$ $\square$

Beispiel 9.45

Berechnung von

$$F(y) \;=\; \int_0^{\frac{\pi}{2}} dx\; \ln(y^2 - \sin^2(x))\;, \qquad y > 1$$

Direkt schwierig, aber mit Satz 9.44:

$$\frac{d}{dy} F(y) \;=\; \int_0^{\frac{\pi}{2}} dx\, \frac{2y}{y^2 - \sin^2(x)} \;=\; \frac{\pi}{\sqrt{y^2 - 1}}$$

wobei Letzteres Beispiel 9.38 ist. Jetzt:

$$F(y) \;=\; \int dy \frac{\pi}{\sqrt{y^2 - 1}} \;=\; \pi \ln(y + \sqrt{y^2 - 1}) \;+\; c$$

entweder nach Ableiten oder mit Substitution $y = \cosh(t)$

Zur Bestimmung der Konstante: Verhalten für $y \to \infty$

Definition 9.46 (Asymptotik und Landau Symbole $\mathcal{O}$ und o)

$f:(a,\infty)\to\mathbb{C}$ und Vergleichsfunktion $g:(a,\infty)\to\mathbb{R}_{>0}$
(meist $g(x)=x^\alpha$ oder $g(x)=\ln(x)$, etc.)

(i) f ist von Ordnung g für $x\to\infty$ (Schreibweise $f=\mathcal{O}(g)$)
$\iff \quad \exists\, b>a$ mit $\sup_{x\geqslant b}\frac{|f(x)|}{g(x)} \leqslant C < \infty$

(ii) f von Ordnung klein o von g für $x\to\infty$ (Schreibweise $f=o(g)$)
$\iff \quad \lim_{x\to\infty}\frac{|f(x)|}{g(x)}=0$

(iii) Analoge Definitionen für Verhalten bei a und Singularitäten

Beispiel 9.47

1. $\frac{1}{x^{2\gamma}+2+\cos(x)}=\mathcal{O}(x^{-2\gamma})$ für $x\to\infty$ wobei $\gamma>\frac{1}{2}$
2. $\sin(x)-x=o(x)$ für $x\to 0$
3. $\ln(1-x)=\mathcal{O}(x)$ für $x\to 0$
4. $\ln(1-x)-x=\mathcal{O}(x^2)$ für $x\to 0$

Beispiel (Bestimmung der Konstante in Beispiel 9.45)

$$F(y) = \int_0^{\frac{\pi}{2}} dx \; \ln(y^2 - \sin^2(x)) = \int_0^{\frac{\pi}{2}} dx \left(\ln(y^2) + \ln\left(1 - \frac{\sin^2(x)}{y^2}\right)\right)$$

$$= \frac{\pi}{2}\ln(y^2) \; + \; \mathcal{O}\left(\frac{1}{y^2}\right) \qquad \text{(da } \ln(1-x) \; = \; \mathcal{O}(x) \text{ für } x \text{ klein)}$$

$$= \pi \ln(y) \; + \; \mathcal{O}\left(\frac{1}{y^2}\right)$$

Vergleich mit

$$\begin{aligned} F(y) \; &= \; \pi \ln(y + \sqrt{y^2 - 1}) + c \\ &= \; \pi \ln(2y) \; + \; \pi \ln\left(\frac{1}{2} \; + \; \sqrt{\frac{1}{4} - \frac{1}{y^2}}\right) \; + \; c \\ &= \; \pi \ln(2) \; + \; \pi \ln(y) \; + \; c \; + \; \mathcal{O}\left(\frac{1}{y^2}\right) \end{aligned}$$

zeigt $c = -\pi \ln(2)$

10 Topologische Grundlagen

Schlagwörter: Metrische, normierte, topologische, kompakte Räume

Definition 10.1

$d : X \times X \to \mathbb{R}_{\geqslant 0} = [0, \infty)$ heißt Metrik (oder Abstand) auf Menge X, falls $\forall\ x, y, z \in X$ gilt, dass

(i) $d(x, y) = 0 \iff x = y$ (Nichtentartung)

(ii) $d(x, y) = d(y, x)$ (Symmetrie)

(iii) $d(x, y) \leqslant d(x, y) + d(z, y)$ (Dreiecksungleichung)

Dann heißt (X, d) metrischer Raum

Beispiel 10.2

$X = \mathbb{C}^N$ versehen mit euklidischer Metrik

$$d(x, y) = \|x - y\| = \Big(\sum_{n=1}^{N} |x_n - y_n|^2\Big)^{\frac{1}{2}}$$

Dies ist ein Spezialfall von Folgendem:

Satz 10.3

Jeder normierte Vektorraum $(V, \|\,.\,\|)$ *ist ein metrischer Raum mit induzierter Metrik*

$$d(x,y) \;=\; \|x-y\|$$

Beweis: Erinnerung an Definition 4.13: $\|.\| : V \to \mathbb{R}_{\geqslant 0}$ Norm, wenn

(i) $\|\lambda v\| \;=\; |\lambda|\,\|v\| \qquad \text{(Homogenität)}$

(ii) $\|v+w\| \leqslant \|v\| + \|w\| \qquad \text{(Dreiecksungleichung)}$

(iii) $\|v\| = 0 \implies v = 0 \qquad \text{(Nicht-Entartung)}$

Dann $d(x,y) = \|x-y\| = 0 \iff x-y=0 \iff x=y$

Symmetrie klar nach Homogenität und Dreiecksungleichung:

$$\begin{aligned} d(x,y) &= \|x-y\| \\ &= \|x-z+z-y\| \\ &\leqslant \|x-z\| + \|z-y\| \\ &= d(x,z) + d(z,y) \end{aligned}$$

$\square$

Beispiel 10.4

$X = C([a,b]) = \{f : [a,b] \to \mathbb{R} \text{ stetige Funktion}\}$

reeller Vektorraum der Dimension ∞ mit Norm

$$\|f\|_\infty = \sup_{x \in [a,b]} |f(x)|$$

Die induzierte Metrik ist also

$$d(f,g) = \|f-g\|_\infty = \sup_{x \in [a,b]} |f(x) - g(x)|$$

Beispiel 10.5

Beispiel ganz anderer Natur: Für Menge X setze

$$d(x,y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$$

Dies ist eine Metrik

Definition 10.6

(X, d) metrischer Raum, $x \in X, \ r > 0, \ A \subset X$ Teilmenge

(i) Die (offene) Kugel mit Radius r und Mittelpunkt x ist

$$B_r(x) \ = \ \{y \in X \mid d(x, y) < r\}$$

(ii) A offen $\iff \ \forall\, a \in A \,\exists\, r > 0$ mit $B_r(a) \subset A$

Satz 10.7

(X, d) metrischer Raum, $r > 0, \ x \in X$

(i) $(A_i)_{i \in I}$, *I Indexmenge, $A_i \subset X$ offen*
$\implies \bigcup_{i \in I} A_i = \{x \in X \mid x \in A_i$ *für ein* $i \in I\}$ *ist offen*

(ii) $B_r(x)$ *offen*

(iii) $A \subset X$ *offen* $\iff$ *A ist Vereinigung von Kugeln*

(iv) $A, B \subset X$ *offen* $\implies$ $A \cap B$ *offen*

Beweis:

(i) Setze $A = \bigcup_{i \in I} A_i$. Sei $a \in A$

$\Longrightarrow \ \exists\, i \in I$ mit $a \in A_i$

$\Longrightarrow \ \exists\, r > 0$ mit $B_r(a) \subset A_i$ (weil A_i offen)

$\Longrightarrow \ B_r(a) \subset A$, somit A offen

(ii) Sei $a \in B_r(x)$. Sei $\delta = d(a, x) < r$. Dann $B_{r-\delta}(a) \subset B_r(x)$

(iii) "$\Longleftarrow$" klar nach (i) und (ii)

"$\Longrightarrow$" Zu $a \in A$ wähle $r_a > 0$ mit $B_{r_a}(a) \subset A$.

Dann $A = \bigcup_{a \in A} B_{r_a}(a)$ (hier ist A Indexmenge!)

(iv) $a \in A \cap B$. Da sowohl A als auch B offen

$\Longrightarrow \ \exists\, r_A > 0$ und $r_B > 0$ mit $B_{r_A}(a) \subset A$ und $B_{r_B}(a) \subset B$

Setze $r = \min\{r_A, r_B\}$. Dann $B_r(a) \subset A \cap B$. Somit $A \cap B$ offen $\square$

Beispiel 10.8

1. Offene Mengen in $\mathbb{R}$ sind Vereinigungen offener Intervalle (a, b).
2. Offene Mengen in $\mathbb{R}^d$ sind Vereinigungen offener Kugeln

Verallgemeinerung des metrischen Raumes: topologischer Raum

Notation: $\mathcal{P}(X) = \{A \mid A \subset X\}$ Potenzmenge von X

Definition 10.9

X Menge, $\mathcal{O} \subset \mathcal{P}(X)$ Mengensystem von Teilmengen von X

Dann heißt $\mathcal{O}$ Topologie auf X, falls

(i) $\emptyset \in \mathcal{O}, \;\; X \in \mathcal{O}$

(ii) $(A_i)_{i \in I}, \;\; A_i \in \mathcal{O} \implies \bigcup_{i \in I} A_i \in \mathcal{O}$ ($\mathcal{O}$ vereinigungsstabil)

(iii) Für $N \in \mathbb{N}$ und $A_1, \ldots, A_N \in \mathcal{O} \implies \bigcap_{i=1}^N A \in \mathcal{O}$
($\mathcal{O}$ endlich durchschnittsstabil)

Dann heißt $(X, \mathcal{O})$ topologischer Raum und Elemente von $\mathcal{O}$ offen

Bemerkung 10.10

(X, d) metrischer Raum. Setze

$$\begin{aligned} \mathcal{O} &= \{\text{beliebige Vereinigungen von Kugeln in } X\} \cup \{\varnothing\} \\ &= \left\{ \bigcup_{i \in I} B_{r_i}(x_i) \,\middle|\, r_i > 0,\ x_i \in X,\ i \in I \right\} \cup \{\varnothing\} \end{aligned}$$

Nach Satz 10.7 ist dies Topologie, genannt induzierte Topologie

Beispiel 10.11

Aber: nicht jede Topologie wird von einer Metrik induziert!

X habe $\geqslant 2$ Punkte. Betrachte "Klumpentopologie" $\mathcal{O} = \{\varnothing, X\}$

Annahme: $\exists$ Metrik d auf X, welche $\mathcal{O}$ induziert

$\implies$ einzige Kugel ist X $\implies$ $\forall\, r > 0$ gilt $d(x, y) < r \quad \forall\, x, y \in X$

$\implies$ $d(x, y) = 0 \quad \forall\, x, y \in X$ Widerspruch zur Definition von d $\lightning$

Bemerkung 10.12

Metriken d und d' auf X können gleiche Topologie erzeugen

Dies gilt insbesondere, falls ein $C > 0$ existiert mit

$$\frac{1}{C}\, d(x,y) \;\leqslant\; d'(x,y) \;\leqslant\; C\, d(x,y) \qquad \forall\, x,y \in X$$

weil dann (mit der Bezeichnung $B'_r(x)$ für Kugeln bezüglich d')

$$B_{\frac{r}{C}}(x) \;\subset\; B'_r(x) \;\subset\; B_{Cr}(x)$$

Also $B_r(x) = \bigcup_{y \in B_r(x)} B_{r_y}(y) = \bigcup_{y \in B_r(x)} B'_{\frac{r_y}{C}}(y)$ offen in Topologie zu d'

Beispiel 10.13

$\mathbb{R}^d$ mit euklidischer Metrik d und Maximumsmetrik $d' = d_\infty$:

$$\frac{1}{\sqrt{d}} \left(\sum_{i=1}^{d} |x_i - y_i|^2 \right)^{\frac{1}{2}} \;\leqslant\; \max_{i=1,\dots,d} |x_i - y_i| \;\leqslant\; \left(\sum_{i=1,\dots,d} |x_i - y_i|^2 \right)^{\frac{1}{2}}$$

Konstruktion neuer topologischer Räume aus bekannten:

Definition 10.14

$(X, \mathcal{O})$ topologischer Raum und $A \subset X$

$\mathcal{O}_A = \{B \cap A \mid B \in \mathcal{O}\}$ Unterraumtopologie auf A

$(A, \mathcal{O}_A)$ heißt topologischer Unterraum von $(X, \mathcal{O})$

Satz 10.15

$\mathcal{O}_A$ Topologie auf A

Beweis: Nachweis der Eigenschaften aus Definition 10.9:

(i) $\emptyset, A \in \mathcal{O}_A$

(ii) $C_i \in \mathcal{O}_A \implies \exists\ B_i \in \mathcal{O}$ mit $C_i = A \cap B_i \implies \bigcup_{i\in I} B_i \in \mathcal{O}$

$\implies \bigcup_{i\in I} B_i \cap A = \bigcup_{i\in I} C_i \in \mathcal{O}_A$

(iii) Übung

$\square$

Notation: $A^c = \{x \in X \mid x \notin A\} = X \setminus A$ Komplement von $A \subset X$ in X

Definition 10.16

$(X, \mathcal{O})$ topologischer Raum und $A \subset X$. Dann:

A abgeschlossen $\iff$ A^c offen

Satz 10.17

$(X, \mathcal{O})$ topologischer Raum

(i) *$\emptyset, X$ abgeschlossen*

(ii) *$(A_i)_{i \in I}$ Familie abgeschlossener Mengen $\implies$ $\bigcap_{i \in I} A_i$ abgeschl.*

(iii) *$A_1, \ldots, A_N$ abgeschlossen $\implies$ $\bigcup_{i=1}^N A_i$ abgeschlossen*

Beweis: (i) $(\varnothing)^c = X$ und$X^c = \varnothing$ offen

(ii) Mengentheoretische Identität:

$$\begin{aligned}
\left(\bigcap_{i\in I} A_i\right)^c &= \left\{x \in X \mid x \notin \bigcap_{i\in I} A_i\right\} \\
&= \{x' \in X \mid x \notin A_i \text{ für ein } i \in I\} \\
&= \bigcup_{i\in I} \{x \in X \mid x \notin A_i\} = \bigcup_{i\in I} A_i^c
\end{aligned}$$

Nun ist A_i^c offen $\forall\ i \in I$

$\implies \bigcup_{i\in I} A_i^c$ offen (nach Definition 10.9) $\implies \bigcap_{i\in I} A_i$ abgeschlossen

(iii) $\left(\bigcup_{i=1,\ldots,N} A_i\right)^c = \bigcap_{i=1}^N A_i^c$ offen weil A_i^c offen und Definition 10.9 $\square$

Definition 10.18

$(X, \mathcal{O})$ topologischer Raum, $\quad x \in X\,, \quad U \subset X$

$U = U(x)$ heißt Umgebung von $x \iff \exists\, A \in \mathcal{O}$ mit $x \in A \subset U$

Bemerkung 10.19

In (X, d) ist U Umgebung von $x \iff \exists$ Kugel $B_r(x) \subset U$

Beispiel 10.20

$a, b \in \mathbb{R}$ und $a < b$

1. $[a, b] = \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$ abgeschlossen in $\mathbb{R}$
2. $[a, b)$ weder offen noch abgeschlossen
3. $\mathbb{Q} \subset \mathbb{R}$ weder offen noch abgeschlossen

   Begründung $\quad q \in \mathbb{Q}$, dann ist kein $B_r(q) \subset \mathbb{Q}$ wenn $r > 0$
4. $[a, b)$ Umgebung von allen $x \in (a, b)$, aber nicht von a und b
5. $\bigcap_{n \geqslant 1} \left(0, 1 + \frac{1}{n}\right) = (0, 1]$ nicht offen (unendlicher Durchschnitt)

Satz 10.21

$(X, \mathcal{O})$ *topologischer Raum,* $A \subset X$

A *offen* $\iff$ A *Umgebung all seiner Punkte*

Beweis:

"$\Longrightarrow$" $x \in A \implies x \in A \subset A$ mit A offen $\implies$ A Umgebung von x

"$\Longleftarrow$" A Umgebung von $x \ \ \forall\, x \in A$

$\implies \forall \ x \in A \ \exists$ offenes B_x mit $x \in B_x \subset A$

$\implies \bigcup_{x \in A} B_x = A$ offen, nach Vereinigungsstabilität $\square$

Definition 10.22

$(X, \mathcal{O})$ topologischer Raum, $x \in X$ und $A \subset X$

(i) x Berührungspunkt (BP) von A

$\iff \forall$ Umgebungen U von x gilt $U \cap A \neq \emptyset$

(ii) x Häufungspunkt (HP) von A

$\iff \forall$ Umgebungen U von x gilt $A \cap (U \setminus \{x\}) \neq \emptyset$

Bemerkung 10.23

$A \subset X$ eines topologischen Raums und $x \in X$

1. $x \in A \implies x$ BP von A, aber nicht immer HP von A

   Zum Beispiel ist 1 BP von $A = \{1\} \subset \mathbb{R}$, aber 1 nicht HP von A
2. Jeder HP von A ist auch BP von A (nicht umgekehrt!)
3. $A \subset \mathbb{R}$ von oben beschränkt $\implies a = \sup(A)$ BP von A

   Sonst gäbe es $r > 0$ mit $B_r(a) \cap A = \emptyset$

   und $a - \frac{r}{2}$ wäre kleinere obere Schranke. Widerspruch $\lightning$

Satz 10.24

$(X, \mathcal{O})$ topologischer Raum und $A \subset X$

Äquivalent sind:

(i) *A abgeschlossen*

(ii) *Jeder BP von A gehört zu A*

(iii) *Jeder HP von A gehört zu A*

Beweis: (i) $\Longrightarrow$ (ii): A abgeschlossen, $x \notin A$

$\Longrightarrow$ $x \in A^c$ offen $\Longrightarrow$ A^c Umgebung von x (nach Satz 10.21)

Da aber $A^c \cap A = \emptyset$, ist x nicht BP von A

Somit Negation: x BP von $A \Longrightarrow x \in A$

(ii) $\Longrightarrow$ (iii): klar nach Bemerkung 10.23

(iii) $\Longrightarrow$ (i): Jeder HP von A ist in A, also ist $x \in A^c$ nicht HP von A

$\Longrightarrow$ $\exists$ offene Umgebung $U(x)$ mit $A \cap (U(x) \backslash \{x\}) = \emptyset$, d.h. $U(x) \subset A^c$

Somit $A^c = \bigcup_{x \in A^c} U(x)$ offen $\Longrightarrow$ A abgeschlossen $\square$

Definition 10.25

Topologischer Raum $(X, \mathcal{O})$ heißt Hausdorff-Raum

$\iff \forall \;\; x \neq y \in X \;\; \exists$ Umgebungen $U(x),\; U(y)$ mit $U(x) \cap U(y) = \varnothing$

Man sagt auch, es gilt die Trennungseigenschaft T_2 in X

Satz 10.26

Jeder metrische Raum (X, d) ist ein Hausdorff-Raum

(wenn versehen mit der induzierten Topologie)

Beweis: $x, y \in X$, $x \neq y \implies r = d(x, y) > 0$ (Nichtentartung)

Wähle $U(x) = B_{\frac{r}{2}}(x)$ und $U(y) = B_{\frac{r}{2}}(y)$. $\square$

Satz 10.27

$(X, \mathcal{O})$ Hausdorff und x Häufungspunkt von $A \subset X$
$\implies$ in jeder Umgebung U von x liegen unendlich viele Punkte von A

Beweis:

U Umgebung von x

$\implies \exists\, x_1 \in U \setminus \{x\} \cap A$, insbesondere $x_1 \neq x$

$\implies \exists$ Umgebungen $U_1(x) \subset U$ und $V_1(x_1)$ mit $U_1 \cap V_1 = \varnothing$

$\implies \exists\, x_2 \in U_1(x) \setminus \{x\} \cap A$ und $x_2 \neq x_1$ sowie $x_2 \neq x$

Dann iteriere. $\square$

Definition 10.28

Folge $(x_n)_{n\in\mathbb{N}}$ in Hausdorff Raum $(X, \mathcal{O})$ konvergiert gegen $x \in X$

$\iff$ zu jeder Umgebung U von x $\exists$ $N \in \mathbb{N}$ mit $x_n \in U \quad \forall\, n \geqslant N$

Schreibweise: $\lim_{n\to\infty} x_n = \lim_n x_n = x$

Bemerkung 10.29

(X, d) metrischer Raum. Dann

$$\lim_{n\to\infty} x_n = x \iff \Big(\forall\, \epsilon > 0\ \exists\, N \text{ mit } x_n \in B_\epsilon(x) \quad \forall\, n \geqslant N \Big)$$

Für $X = \mathbb{R}^d$ oder $X = \mathbb{C}^d$ stimmt Konvergenzbegriff mit Ana 1 überein!

Definition 10.30

$x \in X$ Häufungspunkt (HP) von Folge $(x_n)_{n\in\mathbb{N}}$ in Hausdorff $(X, \mathcal{O})$

$\iff$ zu jeder Umgebung U von x $\exists$ unendlich viele n mit $x_n \in U$

Achtung! HP von Menge $\{x_n \mid n \in \mathbb{N}\} \implies$ HP von Folge $(x_n)_{n\in\mathbb{N}}$

aber Umkehrung falsch (z.B. konstante Folgen)

Satz 10.31

$(x_n)_{n\in\mathbb{N}}$ *Folge in Hausdorff-Raum* $(X, \mathcal{O})$

$x = \lim_{n\to\infty} x_n \implies x$ *einziger HP von* $(x_n)_{n\in\mathbb{N}}$ *(Grenzwert eindeutig)*

Beweis: Sei y zweiter HP von $(x_n)_{n\in\mathbb{N}}$ mit $y \neq x$

$\implies \exists$ Umgebungen $U(x)$ und $U(y)$ mit $U(x) \cap U(y) = \varnothing$

Nach Definition der Konvergenz $\exists\ N$ mit $x_n \in U(x)\ \ \forall\ n \geqslant N$

$\implies$ nur endlich viele x_n in $U(y)$. Widerspruch $\lightning$ $\square$

Satz 10.32

(X, d) *metrischer Raum und* $x \in X$ *HP von Folge* $(x_n)_{n\in\mathbb{N}}$ *in* X

$\implies \exists$ *Teilfolge, die gegen* x *konvergiert.*

Beweis: Bestimme iterativ $n_k > n_{k-1}$, so dass $d(x, x_{n_k}) < \frac{1}{k}$

Dann $\lim_{k\to\infty} x_{n_k} = x$ $\square$

Metrische Räume sind natürlicher Kontext für folgende Begriffe.

Definition 10.33

(X, d) metrischer Raum

(i) $A \subset X$ beschränkt $\iff \exists\, C \in \mathbb{R}$ mit $d(x, y) \leqslant C \quad \forall\, x, y \in A$

(ii) $(x_n)_{n \in \mathbb{N}}$ Cauchy-Folge in X

$\iff \forall\, \epsilon > 0\ \exists\, N$ mit $d(x_n, x_m) \leqslant \epsilon \quad \forall\, n, m \geqslant N$

(iii) (X, d) vollständig $\iff$ jede Cauchy-Folge in X konvergent

Also: X vollständig $\implies$ X metrisch $\implies$ X Hausdorff $\implies$ X topologisch

Definition 10.34

$(V, \|\,.\,\|)$ normierter Vektorraum, also auch metrischer Raum

V vollständig $\quad\iff\quad$ V Banachraum

Definition 10.35

$(V, \langle\,|\,\rangle)$ Vektorraum mit Skalarprodukt, also auch metrischer Raum

V vollständig $\quad\iff\quad$ V Hilbertraum

Satz 10.36

(X, d) *metrischer Raum*

(i) $(x_n)_{n\in\mathbb{N}}$ *Cauchy-Folge* $\Longrightarrow$ $\{x_n \mid n \in \mathbb{N}\}$ *beschränkt*

(ii) $(x_n)_{n\in\mathbb{N}}$ *konvergent* $\Longrightarrow$ $(x_n)_{n\in\mathbb{N}}$ *Cauchy-Folge*

Beweis: (analog zu $\mathbb{R}$)

(i) $(x_n)_{n\in\mathbb{N}}$ Cauchy-Folge

$\Longrightarrow$ $\exists$ N mit $d(x_n, x_m) \leqslant 1 \quad \forall\ n, m \geqslant N$

Setze $r = \max\{d(x_1, x_N), \dots, d(x_{N-1}, x_N)\} + 1$

Dann $x_k \in B_r(x_N) \quad \forall\ k \in \mathbb{N}$

(ii) $\lim_n x_n = x$. Sei $\epsilon > 0$. Bestimme N, so dass

$$d(x, x_n) < \frac{\epsilon}{2} \qquad \forall\ n \geqslant N$$

Dann $\forall\ n, m \geqslant N$

$$d(x_n, x_m) \leqslant d(x_n, x) + d(x, x_m) \leqslant \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \qquad \square$$

Beispiele vollständiger Räume

Beispiel 10.37

$\mathbb{R}^N$ und $\mathbb{C}^N$ sind vollständig (Ana 1)

Beispiel 10.38

$C([a,b])$ versehen mit der Norm $\|\,.\,\|_\infty$ (Beispiel 10.4) ist wegen folgendem Resultat vollständig (Verallgemeinerung Satz 11.25)

Satz 10.39 (Uniforme Limites stetiger Funktionen sind stetig)

$f_n \in C([a,b])$ *und* $g : [a,b] \to \mathbb{R}$ *mit* $\lim_{n\to\infty} \|f_n - g\|_\infty = 0$

$\Longrightarrow \quad g \in C([a,b])$

Bemerkung 10.40

Also ist $(C([a,b]), \|\,.\,\|_\infty)$ ein Banachraum

Die von $\|\,.\,\|_\infty$ induzierte Metrik heißt auch die Metrik der uniformen Konvergenz, manchmal auch der gleichmäßigen Konvergenz

Beweis von Satz 10.39: Sei $x \in [a, b]$ und $\epsilon > 0$

$\implies \exists\, N$ mit $\|f_n - g\|_\infty < \frac{\epsilon}{3} \quad \forall\, n \geqslant N$

f_N stetig bei $x \implies \exists$ Umgebung $\delta > 0$, so dass

$$|f_N(x) - f_N(x')| < \frac{\epsilon}{3} \qquad \forall\, x' \in B_\delta(x)$$

Somit für $x' \in B_\delta(x)$ gilt

$$\begin{aligned} |g(x) - g(x')| &\leqslant |g(x) - f_N(x)| + |f_N(x) - f_N(x')| + |f_N(x') - g(x')| \\ &< \|g - f_N\|_\infty + \frac{\epsilon}{3} + \|f_N - g\|_\infty < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} < \epsilon \end{aligned}$$

□

Beispiel 10.41

$L^2([a, b])$ Menge der quadrat-integrierbaren Funktionen mit Skalarprodukt

$$\langle f | g \rangle = \int_a^b dx\, \overline{f(x)}\, g(x)$$

ist vollständig, also Hilbert-Raum (Riesz-Fischer, Beweis später)

Fakt: Es gibt unvollständige metrische Räume,

aber jeder metrische Raum (X, d) kann vervollständigt werden:

Setze $Y = \{(x_n)_{n\in\mathbb{N}}$ Cauchy-Folge in $X\}$ und $\tilde{X} = Y/\sim$, wobei

$$(x_n)_{n\in\mathbb{N}} \sim (y_n)_{n\in\mathbb{N}} \qquad \Longleftrightarrow \qquad \lim_{n\to\infty} d(x_n, y_n) = 0$$

Definiere $\tilde{d}((x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}) = \lim_{n\to\infty} d(x_n, y_n)$ (Limes existiert!)

Einbettung: $I : X \to \tilde{X} \qquad I(x) = (x)_{n\in\mathbb{N}}$ konstante Folge

Satz 10.42

$(\tilde{X}, \tilde{d})$ ist vollständig und $X \subset \tilde{X}$ dicht,

d.h. zu jedem $\tilde{x} \in \tilde{X}$ und Umgebung $\tilde{U}$ von $\tilde{x}$ existiert $x \in X$ mit $x \in \tilde{U}$

Ohne detaillierten Beweis, aber im Prinzip genauso wie bei $\mathbb{R}$

Beispiel 10.43

Zu $X = \mathbb{Q}$ ist $\tilde{X} = \mathbb{R}$

Satz 10.44 (Fixpunktsatz von Banach)

(X, d) *vollständiger metrischer Raum*

$f : X \to X$ *Lipshitz-stetig mit Lipshitzkonstante* $L < 1$, *d.h.*

$$d(f(x), f(y)) \leqslant L\, d(x, y)$$

$\Longrightarrow$ f *hat genau einen Fixpunkt* $x \in X$, *d.h.* $f(x) = x$

Beweis: Sei $x_n = f(x_{n-1})$ Orbit von beliebigem Startpunkt $x_0 \in X$

$$d(x_n, x_{n+1}) \;=\; d(f(x_{n-1}), f(x_n)) \;\leqslant\; L\, d(x_{n-1}, x_n) \;\leqslant \ldots \leqslant\; L^n\, d(x_0, x_1)$$

Zudem nach der Dreiecksungleichung für $n \leqslant m$:

$$\begin{aligned} d(x_n, x_m) \;&\leqslant\; d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \ldots + d(x_{m-1}, x_m) \\ &\leqslant\; L^n \left(\sum_{k=0}^{m-n} L^k \right) d(x_0, x_1) \;\leqslant\; L^n \frac{1}{1-L} d(x_0, x_1) \end{aligned}$$

Somit ist $(x_n)_{n \geqslant 0}$ eine Cauchy-Folge in X

Wegen Vollständigkeit von X existiert Limespunkt $x = \lim_{n\to\infty} x_n$

Nun ist $\lim_{n\to\infty} d(f(x), f(x_n)) \leqslant \lim_{n\to\infty} L\, d(x, x_n) = 0$ und somit

$$f(x) \;=\; \lim_{n\to\infty} f(x_n) \;=\; \lim_{n\to\infty} x_{n+1} \;=\; x$$

Also ist x Fixpunkt. Sei x' ein zweiter Fixpunkt. Dann

$$d(x, x') \;=\; d(f(x), f(x')) \;\leqslant\; L\, d(x, x')$$

was wegen $L < 1$ impliziert, dass $d(x, x') = 0$, d.h. $x = x'$ ist □

Definition 10.45

$(X, \mathcal{O})$ topologischer Raum

(i) $(A_i)_{i\in I}$ offene Überdeckung von $X \iff A_i \in \mathcal{O}$ und $\bigcup_{i\in I} A_i = X$

(ii) $(A_i)_{i\in I_0}$ Teilüberdeckung von $(A_i)_{i\in I} \iff I_0 \subset I$ und $X = \bigcup_{i\in I_0} A_i$

(iii) X kompakt $\iff$ X Hausdorff und jede offene Überdeckung von X besitzt eine endliche Teilüberdeckung

(iv) $K \subset X$ kompakt $\iff$ K kompakt bez. Unterraumtopologie
$\iff \forall\ (B_i)_{i\in I}$ offen in Hausdorff X und Überdeckung von K $\exists$ endliche Teilüberdeckung von K

Beispiel 10.46

1. $(0,1)$ nicht kompakte Teilmenge von $\mathbb{R}$
   weil: $(0,1) = \bigcup_{n\geqslant 1} \left(\frac{1}{n}, 1\right)$ ohne endliche Teilüberdeckung
2. Später: Satz von Heine-Borel zeigt $[0,1]$ kompakt
3. Endliche Mengen sind immer kompakt

Satz 10.47

$(X, \mathcal{O})$ Hausdorff und $K \subset X$ kompakt. Dann

(i) *K abgeschlossen*

(ii) *$A \subset K$ abgeschlossen in $K \implies A$ kompakt*

Beweis: (i) Sei $y \in K^c$. Verwende die Trennungseigenschaft

Zu jedem $x \in K$ bestimme offene Umgebungen

$U(x)$ von x und $V_x(y)$ von y mit $U(x) \cap V_x(y) = \varnothing$

$\implies \bigcup_{x \in K} U(x)$ offene Überdeckung von kompakter Menge K

$\implies \exists$ endliche Teilüberdeckung $(U(x_n))_{n=1,\ldots,N}$ von K

$\implies \bigcap_{n=1}^{N} V_{x_n}(y)$ offene Umgebung von y mit $\bigcap_{n=1}^{N} V_{x_n}(y) \subset K^c$

Somit ist K^c Umgebung all seiner Punkte

$\implies K^c$ offen nach Satz 10.21

$\implies K$ abgeschlossen

(ii) $(A_i)_{i\in I}$ offene Überdeckung von A

$\Longrightarrow\ \forall\ i \in I\ \exists\ B_i \subset K$ offen in K mit $A_i = B_i \cap K$

$\Longrightarrow\ ((B_i)_{i\in I}, A^c = K\backslash A)$ offene Überdeckung von K (weil A^c offen!)

$\Longrightarrow\ \exists$ endliche Teilüberdeckung $(B_{i_1}, \dots, B_{i_N}, A^c)$ von K

$\Longrightarrow\ (A_{i_1}, \dots, A_{i_N})$ endliche Teilüberdeckung von A $\square$

Satz 10.48

(X, d) metrischer Raum , $K \subset X$ kompakt

$\Longrightarrow$ *K beschränkt, d.h.* $\mathrm{diam}(K) = \sup_{x,y\in K} d(x, y) < \infty$

Beweis: $x \in K,\ (B_n(x))_{n\in\mathbb{N}}$ ist offene Überdeckung von K

$\Longrightarrow\ \exists$ endliche Teilüberdeckung $(B_{n_i}(x))_{i=1,\dots,N}$

$\Longrightarrow\ K \subset B_{n_N}(x)$

$\Longrightarrow\ \mathrm{diam}\, K < n_N < \infty$ $\square$

Definition 10.49

$(X, \mathcal{O})$ Hausdorff-Raum , $A \subset X$

(i) X folgenkompakt $\Longleftrightarrow$ jede Folge in X besitzt konvergente Teilfolge

(ii) A folgenkompakt

$\Longleftrightarrow$ A versehen mit Unterraumtopologie folgenkompakt

Satz 10.50

(X, d) metrischer Raum. Äquivalent sind

(i) *X kompakt*

(ii) *X folgenkompakt*

Lemma 10.51 (Lebesguesches Überdeckungslemma)

(X, d) folgenkompakt, $(A_i)_{i\in I}$ offene Überdeckung von X

$\implies \exists$ Lebesgue'sche Zahl $\delta > 0$, so dass $\forall\, x \in X \;\exists\, i \in I$ mit $B_\delta(x) \subset A_i$

Beweis: Gegenannahme: $\nexists$ solches $\delta > 0$

$\implies \forall\, n \geqslant 1 \;\exists\, x_n \in X$ mit $B_{\frac{1}{n}}(x_n) \not\subset A_i \quad \forall\, i \in I$

Sei x_0 HP von $(x_n)_{n\geqslant 1}$ (nach Voraussetzung)

Sei $i_0 \in I$, so dass $x_0 \in A_{i_0}$ (Überdeckung)

Sei $\epsilon > 0$, so dass $B_\epsilon(x_0) \subset A_{i_0}$ (A_{i_0} offen)

Wähle $k \geqslant \frac{2}{\epsilon}$ mit $x_k \in B_{\frac{\epsilon}{2}}(x_0)$ (x_0 ist HP)

$\implies B_{\frac{1}{k}}(x_k) \subset B_{\frac{\epsilon}{2}}(x_k) \overset{\text{Dreieck}}{\subset} B_\epsilon(x_0) \subset A_{i_0}$ Widerspruch $\lightning$ $\square$

Lemma 10.52

(X, d) *folgenkompakt*

$\implies \forall\, \delta > 0\ \exists$ *endlich viele* $x_1, \ldots, x_r$ *mit* $X = \bigcup_{j=1}^r B_\delta(x_j)$

Beweis: Gegenannahme:

$\exists\, \delta > 0$ mit $X \neq \bigcup_{j=1}^r B_\delta(x_j)$ für jede Wahl von $x_1, \ldots, x_r$ und $r \in \mathbb{N}$

Sei y_0 beliebig $\implies \exists\, y_1$ mit $d(y_1, y_0) \geqslant \delta$ (Aussage für Fall $r = 1$)

$\implies \exists\, y_2$ mit $d(y_2, y_1) \geqslant \delta$ und $d(y_2, y_0) \geqslant \delta$ (Fall $r = 2$)

Iteration: $\forall\, n \in \mathbb{N}$

$$\exists\, y_n \ \text{ mit } \ d(y_k, y_n) \geqslant \delta \qquad \forall\, k < n$$

Diese Folge $(y_n)_{n \in \mathbb{N}}$ erfüllt also $d(y_n, y_m) \geqslant \delta \quad \forall\, n, m \in \mathbb{N}$

Also kann diese Folge keinen HP oder konvergente Teilfolge haben

Widerspruch zur Folgenkompaktheit $\lightning$ $\square$

Beweis von Satz 10.50:

(ii)$\Longrightarrow$(i), d.h. folgenkompakt $\Longrightarrow$ kompakt (schwierigerer Teil)

Sei $(A_i)_{i \in I}$ gegebene offene Überdeckung

Sei $\delta > 0$ zugehörige Lebesguezahl

Nach Lemma 10.52 wähle $x_1, \ldots, x_r$ mit $X = \bigcup_{j=1}^r B_\delta(x_j)$

Da $B_\delta(x_j) \subset A_{i_j}$ für geeignetes i_j nach Lemma 10.52 gilt $X = \bigcup_{j=1}^r A_{i_j}$

d.h. es gibt endliche Teilüberdeckung

(i)$\Longrightarrow$(ii), d.h. kompakt $\Longrightarrow$ folgenkompakt

Sei $(x_n)_{n\in\mathbb{N}}$ beliebige Folge in X. Ziel: Konstruktion von HP

Setze

$$\begin{aligned} A_n &= \{x \in X \mid \forall\, \epsilon > 0\ \exists\, k \geqslant n \text{ mit } x_k \in B_\epsilon(x)\} \\ &= \{x \in X \mid x \text{ BP von } \{x_k \mid k \geqslant n\}\} \ \subset\ A_{n-1} \end{aligned}$$

A_n abgeschlossen nach Satz 10.24 (da alle BP enthalten)

Behauptung: $\bigcap_{n\geqslant 1} A_n \neq \emptyset$

Begründung: Sonst wäre $\bigcup_{n\geqslant 1} A_n^c = X$ offene Überdeckung

$\overset{X \text{ komp.}}{\Longrightarrow}\ \bigcup_{n=1}^N A_n^c = X$ endliche Teilüberdeckung

$\Longrightarrow\ \emptyset = \bigcap_{n=1}^N A_n = A_N$, aber $A_N \neq \emptyset$. Widerspruch $\lightning$

Also $\exists\, x \in \bigcap_{n\geqslant 1} A_n$, d.h. x BP von $\{x_k \mid k \geqslant n\}\ \forall\, n \in \mathbb{N}$

$\Longrightarrow\ x$ HP von $(x_n)_{n\in\mathbb{N}}$ $\quad(\forall\, \epsilon > 0\ \exists$ unendlich viele n mit $d(x, x_n) < \epsilon)$

Nach Satz 10.32 existiert zu diesem HP eine konvergente Teilfolge $\square$

Satz 10.53 (Satz von Heine-Borel)

Sei $\mathbb{R}^d$ versehen mit der euklidische Metrik und $A \subset \mathbb{R}^d$. Dann:

A kompakt $\Longleftrightarrow$ A beschränkt und abgeschlossen

Beweis:

"$\Longrightarrow$" Satz 10.47 und Satz 10.48

"$\Longleftarrow$"

Bolzano-Weierstrass: Jede beschränkte Folge in $\mathbb{R}^d$ besitzt einen HP

Also hat Folge $(x_n)_{n\in\mathbb{N}}$ in A einen HP x, der auch BP von A ist

(entweder $x = x_n$ für unendlich viele n, oder x HP von $\{x_n \mid n \in \mathbb{N}\}$)

Jetzt:

A abgeschlossen $\Longrightarrow$ $x \in A$ nach Satz 10.24

Somit hat jede Folge in A einen HP in A

$\Longrightarrow$ A kompakt nach Satz 10.50

□

11 Stetige Funktionen

Definition 11.1 (Grenzwerte von Funktionen)

Seien X, Y Hausdorff-Räume, $\emptyset \neq A \subset X$ und $a \in X$ BP von A

Für eine Funktion $f : A \to Y$ und $b \in Y$ sei definiert:

$\lim_{x \to a} f(x) = b$

$\iff$ $\forall$ Umgebungen $V(b)$ $\exists$ Umgebung $U(a)$ mit $f(U(a) \cap A) \subset V(b)$

Lemma 11.2

Grenzwerte von Funktionen sind eindeutig

Beweis: (wie Satz 10.31) Seien $b_1 \neq b_2 \in Y$ zwei Grenzwerte

$\implies$ $\exists$ Umgebungen $V(b_1)$, $V(b_2)$ mit $V(b_1) \cap V(b_2) = \emptyset$

$\overset{\text{Def.}}{\implies}$ $\exists$ Umgebungen $U_1(a)$, $U_2(a)$ mit $f(U_j(a) \cap A) \subset V(b_j)$, $j = 1, 2$

$\implies$ $f(U_1(a) \cap U_2(a) \cap A) \subset \bigcap_{j=1,2} f(U_j(a) \cap A) \subset V(b_1) \cap V(b_2) = \emptyset$

$\implies$ $U_1(a) \cap U_2(a) \cap A = \emptyset$, aber $U_1(a) \cap U_2(a)$ Umgebung von a

$\implies$ a nicht BP von A. Widerspruch $\lightning$ $\square$

Satz 11.3

(X, d), (Y, d') metrische Räume, $f : A \subset X \to Y$ und a BP von A. Äquivalent sind:

(i) $\lim_{x \to a} f(x) = b$

(ii) $\forall\, \epsilon > 0\ \exists\, \delta > 0$ *mit* $f(B_\delta(a) \cap A) \subset B'_\epsilon(b)$

(iii) $\forall\, \epsilon > 0\ \exists\, \delta > 0$, *so dass für* $x \in A$ *mit* $d(x, a) < \delta$ *gilt* $d'(f(x), b) < \epsilon$

(iv) $\forall$ *Folgen* $(x_n)_{n \in \mathbb{N}}$ *in* A *mit* $\lim_n x_n = a$ *gilt* $\lim_n f(x_n) = b$

Beweis: (i)$\Longrightarrow$(ii) Sei $\epsilon > 0$. Dann $B'_\epsilon(b)$ Umgebung von b

$\Longrightarrow\ \exists$ Umgebung $U(a)$ mit $f(U(a) \cap A) \subset B'_\epsilon(b)$

$\Longrightarrow\ \exists\, \delta > 0$ mit $B_\delta(a) \subset U(a)$ (da $U(a)$ Umgebung von a)

Somit $f(B_\delta(a) \cap A) \subset f(U(a) \cap A) \subset B'_\epsilon(b)$

(i)$\Longleftarrow$(ii) Sei $V(b)$ Umgebung von b

$\Longrightarrow\ \exists\, \epsilon > 0$ mit $B'_\epsilon(b) \subset V(b)$

$\overset{\text{Voraus.}}{\Longrightarrow}\ \exists\, \delta > 0$ mit $f(B_\delta(a) \cap A) \subset B'_\epsilon(b) \subset V(b)$

Zudem ist $B_\delta(a)$ Umgebung von a, so dass (i) gilt.

(ii)$\Longleftrightarrow$(iii) ist lediglich Umformulierung

(iii)$\Longrightarrow$(iv) Sei $(x_n)_{n\in\mathbb{N}}$ Folge in A mit $\lim_n x_n = a$
Zu beliebigem $\epsilon > 0$ wähle $\delta > 0$ wie in (iii)
$\Longrightarrow \ \exists\, N$ mit $d(x_n, a) < \delta \quad \forall\, n \geqslant N$
$\overset{\text{(iii)}}{\Longrightarrow} \ d'(f(x_n), b) < \epsilon \quad \forall\, n \geqslant N$
Somit $\lim_n f(x_n) = b$

(iv)$\Longrightarrow$(iii) Gelte Negation von (iii)
$\Longrightarrow \ \exists\, \epsilon > 0 \quad \forall\, \delta > 0\ \exists\, x \in A$ mit $d(x, a) < \delta$, so dass $d'(f(x), b) \geqslant \epsilon$
Insbesondere für $\delta = \frac{1}{n}\ \exists\, x_n \in A$ mit $d(x_n, a) < \frac{1}{n}$ und $d'(f(x_n), b) \geqslant \epsilon$
Also $\lim_n x_n = a$, aber $\lim_n f(x_n) \neq b$, d.h. Negation von (iv) $\square$

Beispiel 11.4

$X = Y = \mathbb{C}$ mit euklidischer Metrik $d(z, z') = |z - z'|$
Wir zeigen:

$$\lim_{z \to 0} \frac{e^z - 1}{z} = 1$$

Beachte $\frac{e^0 - 1}{0} = \frac{0}{0}$ nicht definiert, aber 0 BP des Definitionsbereiches!
In der Tat, für $z \in \mathbb{C}$ mit $|z| < 1$ gilt

$$\begin{aligned} d\left(\frac{e^z - 1}{z}, 1\right) &= \left|\frac{e^z - 1}{z} - 1\right| = \left|\sum_{n \geqslant 2} \frac{1}{n!} z^{n-1}\right| \\ &\leqslant \sum_{n \geqslant 2} \frac{1}{n!} |z| = c\,|z| = c\,d(z, 0) \end{aligned}$$

wobei $c = \sum_{n \geqslant 2} \frac{1}{n!}$
Zu $\epsilon > 0$ wähle $\delta = \frac{\epsilon}{c}$. Dann: $d(z, 0) < \delta$ gibt $d(f(z), 1) < \epsilon$

Satz 11.5 (Cauchy-Kriterium)

(X, d), (Y, d') *metrische Räume und sei* (Y, d') *vollständig*

Ferner sei a BP von $A \subset X$ *und* $f : A \to Y$ *Abbildung. Dann:*

$\lim_{x\to a} f(x)$ *existiert* $\iff$ $\forall\, \epsilon > 0\ \exists\, \delta > 0$, *so dass für* $x, x' \in A$ *mit* $d(x, a) < \delta$, $d(x', a) < \delta$ *gilt* $d'(f(x),\ f(x')) < \epsilon$

Bemerkung 11.6

Beachte, dass Grenzwert $b = \lim_{x\to 0} f(x)$ hier nicht benötigt wird, um Konvergenz zu untersuchen (aber auch nicht berechnet wird)

Beweis: "$\Longrightarrow$" (ohne Vollständigkeit von Y)

$\lim_{x\to a} f(x) = b$

$\Longrightarrow\ \forall\, \epsilon > 0\ \exists\, \delta > 0$ mit $d(x, a) < \delta$ gilt $d'(f(x), b) < \frac{\epsilon}{2}$

$\Longrightarrow\ \forall\, x, x' \in A$ mit $d(x, a) < \delta$, $d(x', a) < \delta$ gilt

$$d'(f(x), f(x')) \leqslant d'(f(x), b) + d'(f(x'), b) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

"$\Longleftarrow$" Verwende Kriterium (iv) von Satz 11.3

Sei $(x_n)_{n\in\mathbb{N}}$ Folge in A mit $\lim x_n = a$.

Wir zeigen, dass $(f(x_n))_{n\in\mathbb{N}}$ Cauchy-Folge in (Y, d')

In der Tat, sei $\epsilon > 0$. Wähle $\delta > 0$ gemäß Voraussetzung

Nach Konvergenz $\lim x_n = a \; \exists \; N$ mit $d(x_n, a) < \delta \quad \forall \; n \geqslant N$

also nach Vorausetzung $d'(f(x_n), f(x_m)) < 2\epsilon \;\; \forall \; n, m \geqslant N$, d.h. Cauchy

Weiter: Y vollständig $\Longrightarrow$ $\lim f(x_n) = b$ existiert

Gegeben andere Folge $(x_n')_{n\in\mathbb{N}}$ mit $\lim x_n' = a$ und $\lim f(x_n') = b'$

Wir zeigen $b = b'$, was dann den Beweis beendet

In der Tat, betrachte $(x_n'')_{n\in\mathbb{N}} = (x_0, x_0', x_1, x_1', \ldots)$

Dann $\lim x_1'' = a$ und auch $\lim f(x_n'') = b''$ existiert

Aber $b'' = b'$ und $b'' = b$ $\square$

Beispiel 11.7 (Funktionen ohne Grenzwerte)

Zu $X, Y = \mathbb{R}$ und $A = \mathbb{R}\backslash\{0\}$ betrachte $f(x) = \sin\left(\frac{1}{x}\right)$ für $x \in A$

$\lim_{x\to 0} \sin\left(\frac{1}{x}\right)$ existiert nicht

In der Tat, zu $x_n = \frac{1}{2\pi n + \frac{\pi}{2}} \to 0$ gilt $f(x_n) = 1$

aber zu $x'_n = \frac{1}{2\pi n + \frac{3\pi}{2}} \to 0$ gilt $f(x'_n) = -1$

Tatsächlich:

$\forall\ b \in [-1, 1]\ \exists$ Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim x_n = 0$ und $\lim f(x_n) = b$

Beispiel 11.8

$$X, Y \in \mathbb{R} \text{ und } f(x) = \operatorname{sgn}(x) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \end{cases}, x \in \mathbb{R}\backslash\{0\}$$

$\lim_{x\to 0} f(x)$ existiert nicht

Hier ist jedoch Limes von rechts und links sinnvoll

Definition 11.9

$a \in \mathbb{R}$ Berührungspunkt von $A \subset \mathbb{R}$ und $f : A \to (Y, d')$

(i) $\lim_{x \downarrow a} f(x) = b$ (rechtsseitiger Limes)

$\iff$ $\forall$ Folgen $(x_n)_{n \in \mathbb{N}}$ in A mit $x_n > a$, $\lim x_n = a$ gilt $\lim f(x_n) = b$

(ii) $\lim_{x \to \infty} f(x) = b$ (uneigentlicher Limes)

$\iff$ $\forall$ Folgen $(x_n)_{n \in \mathbb{N}}$ in A mit $\lim x_n = \infty$ gilt $\lim f(x_n) = b$

(iii) Analog linksseitiger Limes $\lim_{x \uparrow a} f(x)$ und $\lim_{x \to -\infty} f(x)$

Beispiele

1. $\lim_{x \downarrow 0} \mathrm{sgn}(x) = 1$
2. $\lim_{x \uparrow 0} \mathrm{sgn}(x) = -1$
3. $\lim_{x \to \infty} \sin\left(\frac{1}{x}\right) = 0$
4. $\lim_{x \to \infty} \sin(x)$ existiert nicht

Satz 11.10 (Regeln für Limites komplexwertiger Funktionen)

(X, d) *metrischer Raum, a BP von* $A \subset X$, $f, g : A \to \mathbb{C}$ *und* $\lambda \in \mathbb{C}$.

Jeweils wenn rechte Seiten existieren, gilt:

(i) $\lim_{x\to a}(f(x) + \lambda g(x)) = (\lim_{x\to a} f(x)) + \lambda(\lim_{x\to a} g(x))$

(ii) $\lim_{x\to a}(f(x) \cdot g(x)) = (\lim_{x\to a} f(x))(\lim_{x\to a} g(x))$

(iii) $\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$, *falls* $\lim_{x\to a} g(x) \neq 0$

Falls $X = \mathbb{R}$, *gilt alles auch für einseitige und uneigentliche Limites*

Die Linearität in (i) gilt auch für vektorwertige Funktionen

Beweis: Mit Satz 11.3(iv) übertragen sich die Regeln für konvergente Zahlenfolgen (Ana 1) $\square$

Jetzt: Im Allgemeinen hat der Grenzwert $b = \lim_{x\to a} f(x)$ nichts mit dem Funktionswert $f(a)$ zu tun (falls a im Definitionsbereich)

Definition 11.11

X, Y Hausdorff-Raum, $f : X \to Y$ Abbildung und $x_0 \in X$

(i) f stetig in $x_0 \iff \lim_{x \to x_0} f(x) = f(x_0)$

(auf der rechten Seite: Limes existiert und gleich $f(x_0)$)

(ii) f stetig (im Großen) $\iff$ f stetig in allen Punkten $x_0 \in X$

Satz 11.12

Äquivalent sind

(i) *f stetig in x_0*

(ii) *$\forall$ Umgebungen V von $f(x_0)$ $\exists$ Umgebung U von x_0 mit $f(U) \subset V$*

(iii) *$\forall$ Umgebungen V von $f(x_0)$ ist $f^{-1}(V)$ Umgebung von x_0*

Beweis:

(i)$\iff$(ii) nach Definition des Grenzwertes (iii)$\implies$(ii) klar

(ii)$\implies$(iii) $f(U) \subset V \implies U \subset f^{-1}(f(U)) = f^{-1}(V)$

Da U Umgebung von x_0 ist, ist auch $f^{-1}(V)$ Umgebung von x_0. $\square$

Satz 11.13

$(X, \mathcal{O}_X)$, $(Y, \mathcal{O}_Y)$ Hausdorff-Räume und $f : X \to Y$. Äquivalent sind:

(i) *f stetig*

(ii) *Urbilder offener Mengen sind offen, d.h. $\forall\ A \in \mathcal{O}_Y$ gilt $f^{-1}(A) \in \mathcal{O}_X$*

(iii) *Urbilder abgeschlossener Mengen sind abgeschlossen*

Beweis: (i)$\Longrightarrow$(ii) $A \subset Y$ offen. Sei $x \in f^{-1}(A)$

$\Longrightarrow$ $f(x) \in A$, d.h. A offene Umgebung von $f(x)$

$\Longrightarrow$ $f^{-1}(A)$ Umgebung von x nach Satz 11.12. Dies gilt $\forall\ x \in X$

$\Longrightarrow$ $f^{-1}(A)$ Umgebung all seiner Punkte, also offen (Satz 10.21)

(ii)$\Longrightarrow$(i) V Umgebung von $f(x)$

$\Longrightarrow$ $\exists$ offene Umgebung $V' \subset V$ von $f(x)$

$\overset{\text{(ii)}}{\Longrightarrow}$ $f^{-1}(V')$ offene Umgebung von $x \in f^{-1}(V')$ mit $f(f^{-1}(V')) \subset V$

(ii)$\Longleftrightarrow$(iii) $A \subset Y$ abgeschlossen $\Longleftrightarrow$ Komplement A^c in Y offen

$\Longleftrightarrow$ $f^{-1}(A^c) = f^{-1}(A)^c$ offen $\Longleftrightarrow$ $f^{-1}(A)$ abgeschlossen $\square$

Satz 11.14 (Hintereinanderausführung stetiger Abbildungen)

$f : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ und $g : (Y, \mathcal{O}_Y) \to (Z, \mathcal{O}_Z)$ stetig
$\implies g \circ f : (X, \mathcal{O}_X) \to (Z, \mathcal{O}_Z)$ stetig

Beweis: $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A)) \in \mathcal{O}_X$ da $g^{-1}(A) \in \mathcal{O}_Y$ für $A \in \mathcal{O}_Z$ $\square$

Satz 11.15 (Hölder-Kriterium für Stetigkeit)

(X, d), (Y, d') metrische Räume, $f : (X, d) \to (Y, d')$ Hölder-stetig, d.h. es existiere Hölder-Exponent $\alpha > 0$ und $L \in \mathbb{R}$ mit

$$d'(f(x), f(x')) \leqslant L\, d(x, x')^\alpha \qquad \forall\, x, x' \in X$$

$\implies f$ stetig

Beweis: Sei $x_0 \in X$ und $\epsilon > 0$. Setze $\delta = (\frac{\epsilon}{L})^{\frac{1}{\alpha}}$
Dann für alle $x \in X$ mit $d(x, x_0) < \delta$

$$d(f(x), f(x_0)) \leqslant L\, d(x, x_0)^\alpha < L \frac{\epsilon}{L} = \epsilon$$

Somit $\lim_{x \to x_0} f(x) = f(x_0)\ \ \forall\, x_0 \in X$ $\square$

Nun zu strukturerhaltenden Abbildungen der Topologie

Definition 11.16

Abbildung $f : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ heißt Homöomorphismus

$\iff$ f bijektiv und f sowie f^{-1} stetig

Falls so eine Abbildung existiert sind X und Y homöomorph

Beispiel 11.17

Seien $I = (0, 1)$, $J = (0, \infty)$ versehen mit Unterraumtopologoie von $\mathbb{R}$

Dann sind I und J homöomorph

Ein Homöomorphismus ist $f(x) = \tan(\frac{\pi}{2}x)$

Bemerkung 11.18

Homöomorphe Räume sind mit Mitteln der Topologie ununterscheidbar

Natürlich haben I und J unterschiedliche Länge,

aber dies ist eine metrische Eigenschaft

Satz 11.19 (Stetige Bilder kompakter Mengen sind kompakt)

$f : (X, \mathcal{O}_X) \to (Y, \mathcal{O}_Y)$ stetig zwischen Hausdorff Räumen
$K \subset X$ kompakt $\Longrightarrow$ $f(K)$ kompakt

Beweis: $(B_i)_{i\in I}$ offene Überdeckung von $f(K) \subset \bigcup_{i\in I} B_i$. Also

$$K \subset f^{-1}(f(K)) \subset \bigcup_{i\in I} \underbrace{f^{-1}(B_i)}_{\text{offen, da } f \text{ stetig}} \qquad \text{offene Überdeckung}$$

$\Longrightarrow \exists$ endliches $I_0 \subset I$ mit $K \subset \bigcup_{i\in I_0} f^{-1}(B_i)$

$\Longrightarrow f(K) \subset f\left(\bigcup_{i\in I_0} f^{-1}(B_i)\right) = \bigcup_{i\in I_0} B_i$

also $f(K)$ kompakt $\square$

Satz 11.20 (Extrema stetiger Funktion auf Kompaktum)

K kompakt und $f : K \to \mathbb{R}$ stetig

$\Longrightarrow \; \exists \, x_0 \in K$ *mit* $f(x_0) = \sup_{x \in K} f(x) = \max_{x \in K} f(x)$

Beweis: $f(K) \subset \mathbb{R}$ kompakt nach Satz 11.19

$\Longrightarrow \; f(K)$ beschränkt und abgeschlossen (nach Heine-Borel)

Sei $M = \sup f(K)$, dann M Berührungspunkt von $f(K)$

Da $f(K)$ abgeschlossen, ist $M \in f(K)$

$\Longrightarrow \; \exists \, x_0 \in K$ mit $f(x_0) = M$ □

Achtung: x_0 im Allgemeinen nicht eindeutig!

Eine Anwendung von Satz 11.20 ist Beweis von

Theorem 11.21 (Fundamentalsatz der Algebra)

Sei $P(z) = \sum_{n=0}^{N} a_n z^n$ *mit* $a_n \in \mathbb{C},\ a_N \neq 0$
Dann existieren $z_1, \ldots, z_N \in \mathbb{C}$ *(eventuell gleich), so dass*

$$P(z) = a_N(z - z_1)(z - z_2)\cdots(z - z_N)$$

Beweis: Es reicht zu zeigen, dass eine Nullstelle z_N existiert (danach betrachte das Polynom $\frac{P(z)}{z-z_N}$)

Setze $\mu = \inf_{z\in\mathbb{C}} |P(z)|$

Für $|z| = R \in \mathbb{R}$ gilt

$$|P(z)| \ \geqslant\ R^N \left(|a_N| - |a_{N-1}|\frac{1}{R} - \ldots - |a_0|\frac{1}{R^N} \right)$$

Somit existiert ein R_c, so dass

$$|P(z)| > \mu \qquad \forall\, z \in \mathbb{C},\ |z| \geqslant R_c$$

Auf dem Kompaktum $\overline{B_{R_c}(0)} = \{z \in \mathbb{C} \mid |z| \leqslant R_c\}$ nimmt stetige Funktion $|P|$ ihr Infimum an (Satz 11.20), d.h. $\exists\ z_N \in \overline{B_{R_c}(0)}$ mit

$$|P(z_N)| = \mu$$

Behauptung: $\mu = 0$

Gegenannahme $\mu > 0$. Definiere Polynom $Q(z) = \frac{P(z+z_N)}{P(z_N)}$

Q ist nicht konstant und erfüllt $Q(0) = 1$ und $|Q(z)| \geqslant 1$

Somit existiert $k \in \{1, \dots, N\}$, so dass

$$Q(z) \ = \ 1 + b_k z^k + \ldots + b_N z^N \quad \text{mit} \quad b_k \neq 0\ , \quad b_n \in \mathbb{C}$$

Sei $\theta \in \left[0, \frac{2\pi}{k}\right)$ definiert durch $e^{ik\theta} \ = \ -\frac{|b_k|}{b_k}$

Für $r > 0$ und $r^k|b_k| < 1$ folgt $|1 + b_k(re^{i\theta})^k| = 1 - r^k|b_k|$ und somit

$$\begin{aligned} |Q(re^{i\theta})| \ &\leqslant \ 1 - r^k|b_k| + r^{k+1}|b_{k+1}| + \ldots + r^N|b_N| \\ &= \ 1 - r^k(|b_k| - r|b_{k+1}| - \ldots - r^{N-k}|b_N|) \\ &\leqslant \ 1 - r^k|b_k| \cdot \tfrac{1}{2} \ < \ 1 \end{aligned}$$

Letzteres für r ausreichend klein. Widerspruch $\lightning$ $\square$

Satz 11.22 (Stetige Funktion auf Kompaktum)

X kompakt und $f : (X, d) \to (Y, d')$ stetig

$\Longrightarrow$ f gleichmäßig stetig auf X, d.h. $\forall\ \epsilon > 0\ \exists\ \delta > 0$ mit

$$d(x, x') < \delta \quad \Longrightarrow \quad d'(f(x), f(x')) < \epsilon$$

Beweis: Gegenannahme: $\exists\ \epsilon > 0 : \forall\ n \geqslant 1\ \exists\ x_n, x'_n$ mit

$$d(x_n, x'_n) \ < \ \frac{1}{n} \qquad d'(f(x_n), f(x'_n)) \ \geqslant \ \epsilon$$

X kompakt $\Longrightarrow$ $(x_{n_k})_{k\in\mathbb{N}}$ konvergente Teilfolge und $x = \lim x_{n_k}$

Dann

$$d(x, x'_{n_k}) \ \leqslant \ d(x, x_{n_k}) + d(x_{n_k}, x'_{n_k}) \ \leqslant \ d(x, x_{n_k}) + \frac{1}{n_k} \ \xrightarrow{k\to\infty} \ 0$$

Somit: $\lim x'_{n_k} = x = \lim x_{n_k}$

f stetig $\Longrightarrow$ $y = f(x) = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} f(x'_{n_k})$

Aber $\epsilon \leqslant d'(f(x_{n_k}), f(x'_{n_k})) \leqslant d'(f(x_{n_k}), y) + d'(y, f(x_{n_k})) \to 0$

Widerspruch $\lightning$ $\square$

Definition 11.23

$(X, \mathcal{O}_X)$ topologischer Raum und (Y, d') metrischer Raum

$\mathcal{F}(X, Y) = \{f : X \to Y \text{ beschränkte Abbildung}\}$

$C_b(X, Y) = \{f : X \to Y \text{ beschränkte stetige Abbildung}\} \subset \mathcal{F}(X, Y)$

Auf $\mathcal{F}(X, Y)$ definiere

$$D(f, g) = \sup_{x \in X} d'(f(x), g(x))$$

Konvergenz einer Folge $(f_n)_{n \in \mathbb{N}}$ in $\mathcal{F}(X, Y)$ bez. Metrik D heißt uniforme oder gleichmäßige Konvergenz **auf** X

Bezeichnung: $g = \text{u-lim}\, f_n$.

Bemerkung 11.24

Spezialfall $X = [a, b]$, $Y = \mathbb{R}$

ergibt $(C([a, b]), \|\,.\,\|_\infty)$, wobei $D(f, g) = \|f - g\|_\infty$

Satz 10.39 über Vollständigkeit wird nun verallgemeinert:

Satz 11.25 (Uniforme Limites stetiger Funktionen sind stetig)

X Hausdorff-Raum und (Y, d') metrischer Raum

$f_n \in C_b(X, Y)$ und $g \in \mathcal{F}(X, Y)$ mit $\lim_{n\to\infty} D(f_n, g) = 0$

$\implies g \in C_b(X, Y)$

Beweis: (Genau wie Satz 10.39) Sei $x \in X$ und $\epsilon > 0$

$\implies \exists\, N$ mit $D(g, f_n) < \frac{\epsilon}{3} \quad \forall\, n \geqslant N$

f_N stetig bei $x \implies \exists$ Umgebung U von x, so dass

$$d'(f_N(x), f_N(x')) < \frac{\epsilon}{3} \qquad \forall\, x' \in U$$

Somit für $x' \in U$ gilt

$$\begin{aligned} d'(g(x), g(x')) &\leqslant d'(g(x), f_N(x)) + d'(f_N(x), f_N(x')) + d'(f_N(x'), g(x')) \\ &\leqslant D(g, f_N) + \frac{\epsilon}{3} + D(f_N, g) < \epsilon \end{aligned}$$

□

Satz 11.26 (Cauchy-Kriterium für uniforme Konvergenz)

(Y, d') vollständig und $f_n : X \to Y$ stetig und beschränkt. Dann:

$(f_n)_{n\in\mathbb{N}}$ uniform konvergent auf X $\iff$ $(f_n)_{n\in\mathbb{N}}$ Cauchy-Folge bez. D

Also $(C_b(X, Y), D)$ vollständig

Beweis: Rechte Seite: $\forall\ \epsilon > 0\ \exists\ N$ mit $D(f_n, f_m) < \epsilon \quad \forall\ n, m \geqslant N$

"$\Longrightarrow$" klar nach Satz 10.36

"$\Longleftarrow$" $\forall\ x \in X$ gilt $d'(f_n(x), f_m(x)) \leqslant D(f_n, f_m)$

$\Longrightarrow \quad (f_n(x))_{n\in\mathbb{N}}$ Cauchy-Folge in Y

$\Longrightarrow \quad g(x) = \lim_{n\to\infty} f_n(x)$ existiert für alle $x \in X$ (da Y vollständig)

Noch zu zeigen: Konvergenz uniform (dann g stetig nach Satz 11.25)

Sei $\epsilon > 0$. Wähle $N = N(\epsilon)$, so dass

$$d'(f_n(x), f_m(x)) \; < \; \frac{\epsilon}{2} \qquad \forall\, n, m \geqslant N \qquad \forall\, x \in X$$

Außerdem wähle $m = m(x) \geqslant N$, so dass

$$d'(g(x), f_{m(x)}(x)) \; < \; \frac{\epsilon}{2}$$

Dann für $n > N$:

$$\begin{aligned}
D(f_n, g) &= \sup_{x \in X} d'(f_n(x), g(x)) \\
&\leqslant \sup_{x \in X} d'(f_n(x), f_{m(x)}(x)) + \sup_{x \in X} d'(f_{m(x)}(x), g(x)) \\
&\leqslant \frac{\epsilon}{2} + \frac{\epsilon}{2} \; = \; \epsilon
\end{aligned}$$

$\square$

12 Stetige lineare Operatoren

Ziel: Untersuchung der Stetigkeit von linearen Abbildungen

Wieso? Intrinsisch interessant und

wird benötigt z.B. für Analysis mehrdimensionaler Ableitungen

Erinnerung: Taylor-Formel für C^2-Funktion $f : D \subset \mathbb{R} \to \mathbb{R}$

$$f(x_0 + v) = f(x_0) + f'(x_0)v + \frac{1}{2}f''(x_0)v\,v + \mathcal{O}(v^3)$$

für $x_0 \in D$ und kleines $v \in \mathbb{R}$

Für Funktion $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ wird gleiche Formel gelten, wobei:

- $x_0 \in D$ und $v \in \mathbb{R}^n$
- $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ lineare Abbildung
- $f''(x_0) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ bilineare symmetrische Abbildung
- Analog $f^{(k)}(x_0)$ k-lineare symmetrisch von $\mathbb{R}^n \times \ldots \times \mathbb{R}^n$ nach $\mathbb{R}^m$

Seien $(V, \|.\|_V)$ und $(W, \|.\|_W)$ normierte Vektorräume über $\mathbb{K}$
Index an der Norm wird unterdrückt (aus Zusammenhang meist klar)

Definition 12.1 (Erinnerung)

Abbildung $T : V \to W$ heißt $\mathbb{K}$-linear oder $\mathbb{K}$-linearer Operator
$\iff \forall\, v, w \in V,\ \lambda \in \mathbb{K}$ gilt: $T(v + \lambda w) = T(v) + \lambda T(w)$
Notation: $\mathcal{L}(V, W) =$ Menge linearer Abbildungen von V nach W
Zudem $\mathcal{L}(V) = \mathcal{L}(V, V)$

Theorem 12.1

Für eine lineare Abbildung $T : (V, \|.\|) \to (W, \|.\|)$ sind äquivalent

(i) *T stetig in 0*
(ii) *T stetig*
(iii) *T gleichmäßig stetig*
(iv) *T beschränkt, d.h. $\exists\, M > 0$ mit $\|Tv\| \leqslant M\|v\|\ \forall\, v \in V$*

Beweis:

(i) $\Longrightarrow$ (iv) Gegenannahme: $\forall\ n \in \mathbb{N}\ \exists\ v_n \in V$ mit $\|Tv_n\| > n\|v_n\|$

Setze $w_n = \frac{v_n}{n\|v_n\|}$. Dann $\|w_n\| = \frac{1}{n}$ und $\lim_{n\to\infty} w_n = 0 \in V$

Aber $\|Tw_n\| = \frac{\|Tv_n\|}{n\|v_n\|} > 1$, d.h. Tw_n konvergiert **nicht** gegen $0 \in W$

Also wäre T unstetig bei 0 $\lightning$

(iv) $\Longrightarrow$ (iii) Nach Linearität gilt

$$\|Tv - Tw\| \ = \ \|T(v-w)\| \ \leqslant \ M\|v-w\|$$

Somit T global Lipshitz-stetig mit Lipshitz-Konstante M

Insbesondere also ist T gleichmäßig stetig.

(iii) $\Longrightarrow$ (ii) $\Longrightarrow$ (i) klar $\square$

Definition 12.2

Für stetiges $T \in \mathcal{L}(V, W)$ ist Operatornorm definiert als

$$\|T\| \ = \ \sup_{v\neq 0} \frac{\|Tv\|}{\|v\|} \ = \ \sup_{\|v\|=1} \|Tv\|$$

Satz 12.3

Sei $\mathbb{K}^n$ versehen mit Maximumsnorm $\|.\|_\infty$ und $(V, \|.\|)$ normierter VR
Jede lineare Abbildung $T : (\mathbb{K}^n, \|.\|_\infty) \to (V, \|.\|)$ ist stetig

Beweis: Sei $e_1, \ldots, e_n$ die Standardbasis von $\mathbb{K}^n$
Dann $v = \sum_{j=1}^n v_j e_j \in \mathbb{K}^n$ und

$$\begin{aligned} \|Tv\| &= \|\sum_{j=1}^n v_j Te_j\| \\ &\leqslant \max\{\|Te_1\|, \ldots, \|Te_n\|\} \sum_{j=1}^n |v_j| \\ &\leqslant C\, n \|v\|_\infty \end{aligned}$$

Also $\|T\| < \infty$ und T ist stetig (Anderer Beweis mit Satz 11.20) $\square$

Beachte: Schranke explodiert im Limes unendlicher Dimensionen

Satz 12.4

Seien $(V, \|.\|)$ *und* $(W, \|.\|)$ *normierte Vektorräume*
Definiere Menge der linearen und beschränkten Operatoren:

$$\mathcal{B}(V, W) \;=\; \{T \in \mathcal{L}(V, W) \mid \|T\| < \infty\}$$

$(\mathcal{B}(V, W), \|.\|)$ *ist ein normierter Vektorraum, wobei*
$\|.\| : \mathcal{B}(V, W) \to \mathbb{R}_{\geqslant 0}$ *Opertornorm und* $(T + \lambda S)(v) = Tv + \lambda Sv$

Beweis: Wenn T und S linear sind, dann auch $T + \lambda S$. Zudem

$$\|(T + \lambda S)v\| \;\leqslant\; \|Tv\| + |\lambda|\|Sv\| \;\leqslant\; (\|T\| + |\lambda|\|S\|)\|v\|$$

so dass $T + \lambda S \in \mathcal{B}(V, W)$. Insbesondere also

$$\|T + S\| \;\leqslant\; \|T\| + \|S\| \qquad \|\lambda T\| \;=\; |\lambda|\|T\|$$

Außerdem $\|T\| = 0 \iff T = 0$. Also $\|\,.\,\|$ Norm auf $\mathcal{B}(V, W)$ $\square$

Definition 12.5 (Vergleiche Definition 10.34)

Ein vollständiger normierter Vektorraum heißt Banachraum

Satz 12.6

$(V, \|\,.\,\|)$ *normierter Raum und* $(W, \|\,.\,\|)$ *Banachraum*
$\Longrightarrow$ $(\mathcal{B}(V, W), \|\,.\,\|)$ *Banachraum*

Beweis: Nach Satz 12.4 bleibt noch die Vollständigkeit zu zeigen
Sei also $(T_n)_{n\in\mathbb{N}}$ Cauchy-Folge in $\mathcal{B}(V, W)$
Für jedes $v \in V$ ist dann

$$\|T_n v - T_m v\| = \|(T_n - T_m)v\| \leqslant \|T_n - T_m\|\|v\|$$

d.h. $(T_n v)_{n\in\mathbb{N}}$ Cauchy-Folge in W
Nach Vollständigkeit von W wird $T : V \to W$ definiert durch

$$Tv = \lim_{n\to\infty} T_n v$$

Zu zeigen: $T \in \mathcal{B}(v, w)$ und $(T_n)_{n\in\mathbb{N}}$ konvergiert gegen T

Sei $\varepsilon > 0$. Dann $\exists\, N$ mit $\|T_n - T_m\| < \frac{\varepsilon}{2} \;\; \forall\, n, m \geqslant N$
Zudem: $\forall\, v \in V \;\exists\, m \geqslant N$ mit $\|Tv - T_m v\| < \frac{\varepsilon}{2}\|v\|$
Weiter

$$\|(T - T_n)v\| \;\leqslant\; \|(T - T_m)v\| + \|(T_m - T_n)v\| \;\leqslant\; \left(\tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2}\right)\|v\| \;=\; \varepsilon\,\|v\|$$

$\implies \;\; \|T - T_n\| \leqslant \varepsilon \;\forall\, n \geqslant N$
$\implies \;\; (T_n)_{n\in\mathbb{N}}$ gegen T. Weiter, weil T_n linear,

$$\begin{aligned}
&\|T(v + \lambda w) - Tv - \lambda Tw\| \\
&= \;\|(T - T_n)(v + \lambda w) - (T - T_n)v - \lambda(T - T_n)w\| \\
&\leqslant\; \|T - T_n\|(\|v + \lambda w\| + \|v\| + |\lambda|\|w\|) \;+\; 0 \;\xrightarrow{n\to\infty}\; 0
\end{aligned}$$

d.h. T linear. Zuletzt für n ausreichend groß,
$\|T\| \leqslant \|T - T_n\| + \|T_n\| \leqslant \varepsilon + \|T_n\| < \infty$, also T beschränkt $\square$

Satz 12.7

Seien U, V, W normierte Vektorräume und $T \in \mathcal{B}(U, V)$, $S \in \mathcal{B}(V, W)$
Dann ist $ST = S \circ T \in \mathcal{B}(U, W)$ und $\|ST\| \leqslant \|S\| \|T\|$

Beweis: Linearität von ST klar, und $\forall\, u \in U$ gilt

$$\|STu\| \leqslant \|S\| \|Tu\| \leqslant \|S\| \|T\| \|u\|$$

Somit

$$\|ST\| = \sup_{u \neq 0} \frac{\|STu\|}{\|u\|} \leqslant \|S\| \|T\|$$

□

Definition 12.8

V normierter Vektorraum und sei $\mathcal{B}(V) = \mathcal{B}(V, V)$
$(\mathcal{B}(V), +, \cdot, \circ, \|\,.\,\|)$ ist eine normierte Algebra
Falls V vollständig, so auch $\mathcal{B}(V)$ vollständig (nach Satz 12.7)
und dann ist $\mathcal{B}(V)$ eine so genannte Banachalgebra

Definition 12.9

Linearer Operator $T : V \to W$ invertierbar $\iff$ $T : V \to W$ bijektiv
Dann existiert Inverses $T^{-1} : W \to V$

Bemerkung 12.10 (Tiefliegender Satz der inversen Abbildung)

Inverses eines beschränkten invertierbaren Operators ist beschränkt

Satz 12.11

V ein Banachraum und $T \in \mathcal{B}(V)$ Kontraktion, d.h. $\|T\| < 1$
$\Longrightarrow \mathbf{1} - T$ invertierbar mit Inversen gegeben durch konvergente Neumann Reihe (bez. Operatornorm):

$$(\mathbf{1} - T)^{-1} = \sum_{n \geqslant 0} T^n$$

Zudem

$$\|(\mathbf{1} - T)^{-1}\| \leqslant \frac{1}{1 - \|T\|}$$

Beweis: Nach Satz 12.7 gilt $\|T^n\| \leqslant \|T\|^n$. Somit

$$\|\sum_{n=0}^{N} T^n\| \;\leqslant\; \sum_{n=0}^{N} \|T^n\| \;\leqslant\; \frac{1}{1-\|T\|}$$

und $\left(\sum_{n=0}^{N} T^n\right)_{N\in\mathbb{N}}$ Cauchy-Folge bez. Operatornorm
Nach Satz 12.7 konvergiert Reihe also gegen einen Operator $S \in \mathcal{B}(V)$
Es gilt:

$$(\mathbf{1}-T)S \;=\; \sum_{n=0}^{\infty} T^n - \sum_{n=1}^{\infty} T^n \;=\; \mathbf{1}$$

und analog $S(\mathbf{1}-T) = \mathbf{1}$. Also $S = (\mathbf{1}-T)^{-1}$. $\square$

Satz 12.12

Seien V, W Banachräume und $T \in \mathcal{B}(V, W)$ invertierbar
Dann sind alle Elemente der offenen Kugel um T mit Radius $r = \frac{1}{\|T^{-1}\|}$

$$B_r(T) \;=\; \{S \in \mathcal{B}(V, W) \mid \|T - S\| < r\}$$

invertierbar. Zudem gilt für $S \in B_r(T)$

$$\|S^{-1}\| \;\leqslant\; \frac{\|T^{-1}\|}{1 - \|T^{-1}\|\|T - S\|}$$

$$\|T^{-1} - S^{-1}\| \;\leqslant\; \frac{\|T^{-1}\|^2\|T - S\|}{1 - \|T^{-1}\|\|T - S\|}$$

Bemerkung 12.13

Sprachgebrauch: Invertierbarkeit ist eine offene Bedingung

Beweis: Da $S = T - (T - S) = T(\mathbf{1} - T^{-1}(T - S))$ und

$$\|T^{-1}(T - S)\| \leqslant \|T^{-1}\| \|T - S\| < 1 \qquad \text{für } S \in B_{\frac{1}{\|T^{-1}\|}}(T)$$

folgt konvergente Neumann-Reihe (Satz 12.11)

$$(\mathbf{1} - T^{-1}(T - S))^{-1} = \sum_{n=0}^{\infty} (T^{-1}(T - S))^n \in \mathcal{B}(V)$$

und

$$\|(\mathbf{1} - T^{-1}(T - S))^{-1}\| \leqslant \frac{1}{1 - \|T^{-1}\| \|T - S\|}$$

$\Longrightarrow S^{-1} = (\mathbf{1} - T^{-1}(T - S))^{-1} T^{-1}$ existiert und Schranke an $\|S^{-1}\|$

Weiter

$$\begin{aligned} S^{-1} - T^{-1} &= \left((\mathbf{1} - T^{-1}(T - S))^{-1} - \mathbf{1}\right) T^{-1} \\ &= (\mathbf{1} - (\mathbf{1} - T^{-1}(T - S)))(\mathbf{1} - T^{-1}(T - S))^{-1} T^{-1} \\ &= T^{-1}(T - S)(\mathbf{1} - T^{-1}(T - S))^{-1} T^{-1} \end{aligned}$$

so dass $\|S^{-1} - T^{-1}\| \leqslant \|T^{-1}\| \|T - S\| \frac{1}{1 - \|T^{-1}\| \|T - S\|} \|T^{-1}\|$ $\square$

13 Differentialrechnung mehrerer Veränderlicher

Definition 13.1

V, W normierte Vektorräume über $\mathbb{R}$, z.B. $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$

Sei $D \subset V$ offene Teilmenge und $x_0 \in D$

Abbildung $f : D \to W$ heißt (Fréchet) differenzierbar in Punkt x_0

$\iff \exists\ T \in \mathcal{B}(V, W)$, so dass

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - T(x - x_0)\|}{\|x - x_0\|} = 0$$

$\iff \exists\ T \in \mathcal{B}(V, W)$, so dass $\forall\ \varepsilon > 0\ \exists\ \delta > 0$ mit

$$\|f(x) - f(x_0) - T(x - x_0)\| < \varepsilon\|x - x_0\| \qquad \forall\ x \neq x_0 \ \text{ mit } \ \|x - x_0\| < \delta$$

Dann heißt T Ableitung von f in x_0 und wird mit $f'(x_0)$ bezeichnet

Alternative Notationen: $\partial f(x_0)$, $Df(x_0)$, $df(x_0)$, $\nabla f(x_0)$ ...

Bemerkung 13.2

1. Im Fall $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ ist $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ linear, also gegeben durch Matrix aus $\mathrm{Mat}(n \times m, \mathbb{R})$
2. Im Fall $V = W = \mathbb{R}$ stimmt die Definition mit Ana 1 überein
   In der Tat: lineares $T : \mathbb{R} \to \mathbb{R}$ ist Multiplikation mit reeller Zahl, d.h. $\mathcal{B}(\mathbb{R}, \mathbb{R}) \cong \mathbb{R}$.
3. Definition 13.1 überträgt sich auch auf komplexe Vektorräume, wenn $\mathcal{B}(V, W) = \{$komplex lineare Abbildungen$\}$
   Differenzierbare Funktionen heißen dann holomorph
   Für $V = W = \mathbb{C}$ (vgl. Funktionentheorie) ist $f'(z_0) \in \mathcal{B}(\mathbb{C}, \mathbb{C}) \cong \mathbb{C}$

Lemma 13.3

Wenn Ableitung existiert, dann ist sie eindeutig

Beweis:

Seien T, T' zwei Ableitungen

Für alle $\varepsilon > 0 \; \exists \; \delta > 0$ (für T, T' gleichzeitig gewählt), so dass

$\forall \; x \neq x_0, \; \|x - x_0\| < \delta$ gilt

$$\begin{aligned} \frac{\|(T - T')(x - x_0)\|}{\|x - x_0\|} &\leqslant \frac{\|f(x) - f(x_0) - T(x - x_0)\|}{\|x - x_0\|} \\ &\quad + \frac{\|f(x) - f(x_0) - T'(x - x_0)\|}{\|x - x_0\|} \\ &\leqslant \varepsilon + \varepsilon = 2\varepsilon \end{aligned}$$

□

Beispiel 13.4

(i) $f = w$ konstant $\Longrightarrow f'(x_0) = 0 \quad \forall\, x_0 \in D$

(ii) $f : V \to V,\ f = \mathrm{id}_V \ \Longrightarrow\ f'(x_0) = \mathrm{id}_V$

(iii) $f : V \to W$ linear, d.h. $f \in \mathcal{B}(V, W)$. Dann $f'(x_0) = f \ \ \forall\, x_0 \in V$, weil

$$\frac{\|f(x) - f(x_0) - f(x - x_0)\|}{\|x - x_0\|} = \frac{0}{\|x - x_0\|} = 0$$

Lemma 13.5

f differenzierbar in $x_0 \Longrightarrow f$ stetig in x_0

Beweis: Für $\varepsilon > 0$ existiert $\delta > 0$, so dass für $\|x - x_0\| < \delta$ gilt:

$$\begin{aligned} \|f(x) - f(x_0)\| &\leqslant \|f(x) - f(x_0) - f'(x_0)(x - x_0)\| + \|f'(x_0)(x - x_0)\| \\ &\leqslant \varepsilon\|x - x_0\| + \|f'(x_0)\|\|x - x_0\| \end{aligned}$$

Also ist f stetig in x_0 $\square$

Satz 13.6 (Linearität, Produktregel, Kettenregel)

V, W, U normierte Vektorräume und $D \subset V$ offen

$f, g : D \to W$ differenzierbar in $x_0 \in D$ und $\lambda \in \mathbb{R}$

(i) *$f + \lambda g : D \to W$ differenzierbar in x_0 und*

$$(f + \lambda g)'(x_0) \;=\; f'(x_0) + \lambda g'(x_0)$$

(ii) *Sei zudem W normierte Algebra (z.B. $W = \mathbb{R}, \mathbb{C}, \mathrm{Mat}(n \times n, \mathbb{R})$)*

Dann ist $f \cdot g : D \to W$ in x_0 differenzierbar und

$$(f \cdot g)'(x_0) \;=\; f'(x_0)g(x_0) + f(x_0)g'(x_0)$$

(iii) *$h : f(D) \subset W \to U$ differenzierbar in $f(x_0)$*

Dann ist $h \circ f$ differenzierbar in x_0 und

$$(h \circ f)'(x_0) \;=\; h'(f(x_0))f'(x_0)$$

Beweis:

(i) Nach der Dreiecksungleichung

$$\begin{aligned}
&\frac{\|(f+\lambda g)(x)-(f+\lambda g)(x_0)-(f'(x_0)+\lambda g'(x_0))(x-x_0)\|}{\|x-x_0\|}\\
&= \frac{\|f(x)+\lambda g(x)-f(x_0)-\lambda g(x_0)-f'(x_0)(x-x_0)-\lambda g'(x_0)(x-x_0)\|}{\|x-x_0\|}\\
&\leqslant \frac{1}{\|x-x_0\|}\Big[\|f(x)-f(x_0)-f'(x_0)(x-x_0)\|\\
&\qquad + |\lambda|\,\|g(x)-g(x_0)-g'(x_0)(x-x_0)\|\Big]\\
&\xrightarrow{x\to x_0} 0
\end{aligned}$$

(ii)

$$\frac{\|f(x)g(x) - f(x_0)g(x_0) - f'(x_0)(x-x_0)g(x_0) - f(x_0)g'(x_0)(x-x_0)\|}{\|x-x_0\|}$$

$$\begin{aligned}
&\leqslant \frac{1}{\|x-x_0\|}\Big[\|f(x)g(x) - f(x_0)g(x) - f'(x_0)(x-x_0)g(x)\| \\
&\qquad + \|f(x_0)g(x) - f(x_0)g(x_0) - f(x_0)g'(x_0)(x-x_0)\| \\
&\qquad + \|f'(x_0)(x-x_0)(g(x)-g(x_0))\|\Big] \\
&\leqslant \frac{1}{\|x-x_0\|}\Big[\|f(x) - f(x_0) - f'(x_0)(x-x_0)\|\|g(x)\| \\
&\qquad + \|f(x_0)\|\|g(x) - g(x_0) - g'(x_0)(x-x_0)\| \\
&\qquad + \|f'(x_0)\|\|x-x_0\|\|g(x)-g(x_0)\|\Big] \\
&\xrightarrow{x\to x_0} 0
\end{aligned}$$

weil g stetig in x_0, siehe Lemma 13.5

Hierbei wurde Operatorungleichung $\|TS\| \leqslant \|T\|\|S\|$ verwandt

(iii) Es gilt

$$
\begin{aligned}
&\frac{1}{\|x-x_0\|}\,\|h\circ f(x)-h\circ f(x_0)-h'(f(x_0))f'(x_0)(x-x_0)\| \\
&\leqslant \frac{\|h(f(x))-h(f(x_0))-h'(f(x_0))(f(x)-f(x_0))\|}{\|x-x_0\|}\,\frac{\|f(x)-f(x_0)\|}{\|f(x)-f(x_0)\|} \\
&\quad+\frac{\|h'(f(x_0))(f(x)-f(x_0)-f'(x_0)(x-x_0))\|}{\|x-x_0\|} \\
&\leqslant \frac{\|h(f(x))-h(f(x_0))-h'(f(x_0))(f(x)-f(x_0))\|}{\|f(x)-f(x_0)\|}\cdot \\
&\quad\cdot\left(\frac{\|f(x)-f(x_0)-f'(x_0)(x-x_0)\|}{\|x-x_0\|}+\frac{\|f'(x_0)(x-x_0)\|}{\|x-x_0\|}\right) \\
&\quad+\frac{\|h'(f(x_0))\|\cdot\|f(x)-f(x_0)-f'(x_0)(x-x_0)\|}{\|x-x_0\|} \\
&\xrightarrow{x\to x_0} 0\cdot C+C\cdot 0 \;=\; 0
\end{aligned}
$$

$\square$

Satz 13.7 (Mittelwertsatz für reellwertige Funktionen)

V normierter Vektorraum und $D \subset V$ offen

$f : D \to \mathbb{R}$ differenzierbar auf ganz D

Ferner $x, y \in D$, so dass Intervall $[x, y] = \{(1-t)\,x + ty \mid t \in [0,1]\}$ in D

$\Longrightarrow \exists\; t \in (0,1)$ mit

$$f(y) \;=\; f(x) \;+\; f'((1-t)\,x + ty)(y - x)$$

Beweis: Definiere $\gamma : [0,1] \to V$ durch $\gamma(t) = (1-t)x + ty$

Dann ist $f \circ \gamma : [0,1] \to \mathbb{R}$ stetig und differenzierbar auf $(0,1)$

Nach eindimensionalem Mittelwertsatz $f \circ \gamma(1) = f \circ \gamma(0) + (f \circ \gamma)'(t)$

Nach Kettenregel gilt $(f \circ \gamma)'(t) = f'(\gamma(t))\gamma'(t) = f'(\gamma(t))(y - x)$ $\square$

Bemerkung 13.8

$f'(x_0) : V \to \mathbb{R}$ ist lineare und stetige Abbildung

Falls $V = \mathbb{R}^n$ ist sie durch Skalarprodukt mit Vektor gegeben,

der mit $\nabla f(x_0) \in \mathbb{R}^n$ bezeichnet wird, d.h.

$$f'(x_0)(x - y) \;=\; \langle \nabla f(x_0) \mid x - y \rangle_{\mathbb{R}^n}$$

Bemerkung 13.9

Für Funktionen, die nicht reellwertig sind, gilt der MWS nicht!

Dies zeigt schon folgendes zweidimensionale Beispiel:

$$f : [0, 2\pi] \to \mathbb{C} \qquad f(x) = e^{ix}$$

$$f(0) = f(2\pi) = 1 \qquad f'(x) = ie^{ix} \neq 0 \quad \forall\, x \in [0, 2\pi]$$

Aber $f(2\pi) \neq f(0) + f'(t) \cdot (2\pi - 0) \quad \forall\, t \in [0, 2\pi]$

In vielen Situationen ist folgender Satz ein guter Ersatz:

Satz 13.10 (Schrankensatz)

$f : D \to W$ differenzierbar auf offenem $D \subset V$ in normiertem Vektorr. Ferner $x, y \in D$, so dass $[x, y] \subset D$. Setze

$$M \;=\; \sup_{t\in[0,1]} \|f'((1-t)x + ty)\| \;=\; \sup_{x'\in[x,y]} \|f'(x')\|$$

Dann

$$\|f(x) - f(y)\| \;\leqslant\; M\,\|x - y\|$$

Beweis: Gegenannahme: $\exists\, \varepsilon > 0$ mit $\|f(x) - f(y)\| \geqslant (M+\varepsilon)\|x - y\|$
Unterteile $[x, y] = [x, x_0] \cup [x_0, y]$ mit Mittelpunkt $x_0 = \frac{x+y}{2}$. Nun:
$\|f(x) - f(x_0)\| \geqslant (M+\varepsilon)\|x - x_0\|$ oder $\|f(x_0) - f(y)\| \geqslant (M+\varepsilon)\|x_0 - y\|$
weil sonst Widerspruch zur Gegenannahme

$$\begin{aligned} \|f(x) - f(y)\| &\leqslant \|f(x) - f(x_0)\| + \|f(x_0) - f(y)\| \\ &< (M+\varepsilon)(\|x - x_0\| + \|x_0 - y\|) \\ &= (M+\varepsilon)\|x - y\| \qquad \text{(weil } x, x_0, y \text{ auf Gerade liegen)} \end{aligned}$$

Sei $[x_1, y_1]$ Intervall mit $\geqslant$ (entweder $[x, x_0]$ oder $[x_0, y]$)
Iteration dieser Konstruktion gibt Folge $[x_n, y_n] \subset [x_{n-1}, y_{n-1}]$ mit

$$\|x_n - y_n\| = \frac{\|x - y\|}{2^n} \qquad \|f(x_n) - f(y_n)\| \geqslant (M + \varepsilon)\|x_n - y_n\|$$

Nun existiert $x' = \lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n$
Wiederum Ungleichung $\geqslant$ auf einem Teil von $[x_n, y_n] = [x_n, x'] \cup [x', y_n]$
Ohne Einschränkung für unendlich viele x_n. Dann

$$\begin{aligned}
\|f'(x')\| &\geqslant \frac{\|f'(x')(x_n - x')\|}{\|x_n - x'\|} \\
&\geqslant \frac{\|f(x_n) - f(x')\|}{\|x_n - x'\|} - \frac{\|f(x_n) - f(x') - f'(x')(x_n - x')\|}{\|x_n - x'\|} \\
&\geqslant (M + \varepsilon) - \frac{\|f(x_n) - f(x') - f'(x')(x_n - x')\|}{\|x_n - x'\|} \xrightarrow{n\to\infty} (M + \varepsilon)
\end{aligned}$$

Widerspruch zur Voraussetzung $\lightning$ $\square$

Erinnerung:

$V_1 \times V_2$ kartesisches Produkt zweier normierter Vektorräume V_1, V_2

Vektorraumstruktur: $(x_1, x_2) + \lambda(y_1, y_2) = (x_1 + \lambda y_1, x_2 + \lambda y_2)$

(Eine) Norm: $\|(x_1, x_2)\| = \|x_1\| + \|x_2\|$

Definition 13.11 (Partielle Ableitungen)

$D \subset V_1 \times V_2$ offen und $f : D \to W$

f heißt partiell differenzierbar in $(x_1, x_2) \in D$

$\iff x \in V_1 \mapsto f(x, x_2)$ differenzierbar in x_1 und

$x \in V_2 \mapsto f(x_1, x)$ differenzierbar in x_2

Zugehörige lineare Abbildungen $\partial_{x_1} f(x_1, x_2) \in \mathcal{B}(V_1, W)$ und

$\partial_{x_2} f(x_1, x_2) \in \mathcal{B}(V_2, W)$ heißen partielle Ableitungen

Gilt dies allen Punkten von D, heißt f partiell differenzierbar auf D

Analog $\partial_{x_k} f$ für $f : D \subset V_1 \times \ldots \times V_n \to W$ und $k = 1, \ldots, n$

Wichtigster Spezialfall:

$V = \mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R}$ und $W = \mathbb{R}^m = \mathbb{R} \times \ldots \times \mathbb{R}$

$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} : D \subset \mathbb{R}^n \to \mathbb{R}^m$ mit Komponentenfunktion $f_\ell : D \subset \mathbb{R}^n \to \mathbb{R}$

Für $k = 1, \ldots, n$ und $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ist $\partial_{x_k} f(x) : \mathbb{R} \to \mathbb{R}^m$ lineare Abbildung

Diese partiellen Ableitungen existieren

$\iff$ alle $f_\ell : D \subset \mathbb{R}^n \to \mathbb{R}$ nach x_k eindimensional differenzierbar

$\partial_{x_k} f_\ell(x)$ Ableitung von $y \in \mathbb{R} \mapsto f_\ell(\ldots, x_{k-1}, y, x_{k+1}, \ldots) \in \mathbb{R}$ bei $y = x_k$

Lineare Abbildung $\partial_{x_k} f(x) \in \mathcal{B}(\mathbb{R}, \mathbb{R}^m) \cong \mathbb{R}^m$ ist ein Vektor

$$\partial_{x_k} f(x) = \begin{pmatrix} \partial_{x_k} f_1(x) \\ \vdots \\ \partial_{x_k} f_m(x) \end{pmatrix}$$

Definition 13.12 (Jacobi Matrix)

$f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ partiell differenzierbar

Jacobi Matrix $\partial f : D \to \mathrm{Mat}(m \times n, \mathbb{R})$ ist definiert als

$$\partial f \;=\; (\partial_{x_1} f, \dots, \partial_{x_n} f) = \begin{pmatrix} \partial_{x_1} f_1 & \cdots & \partial_{x_n} f_1 \\ \vdots & & \vdots \\ \partial_{x_1} f_m & \cdots & \partial_{x_n} f_m \end{pmatrix}$$

Im Fall $m = 1$ ist $\partial f(x)$ Zeilenvektor $=$ Vektor des Dualraums

Dieser Vektor $\nabla f(x) = (\partial f(x))^T \in \mathbb{R}^n$ wird der Gradient genannt. Dann

$$f'(x)v \;=\; \langle \nabla f(x) | v \rangle_{\mathbb{R}^n}$$

$\nabla f(x)$ ist dann ein Vektorfeld im Sinne folgender Definition:

Definition 13.13 (Vektorfeld)

Ein Vektorfeld auf $D \subset \mathbb{R}^n$ ist eine Funktion $g : D \to \mathbb{R}^n$

g heißt Gradientenfeld $\iff \exists\ f : D \subset \mathbb{R}^n \to \mathbb{R}$ mit $g = \nabla f(x)$

Weitere Begriffsbildungen aus partiellen Ableitungen:

Zu partiell differenzierbarem Vektorfeld $g : D \subset \mathbb{R}^n \to \mathbb{R}^n$ ist Divergenz

$$\mathrm{div}(g) = \nabla \cdot g : D \to \mathbb{R} \qquad \mathrm{div}(g) = \partial_{x_1} g_1 + \partial_{x_2} g_2 + \ldots + \partial_{x_n} g_n$$

Falls $n = 3$ ist Rotation des Vektorfeldes ein neues Vektorfeld

$$\mathrm{rot}(g) = \begin{pmatrix} \partial_{x_2} g_3 - \partial_{x_3} g_2 \\ \partial_{x_3} g_1 - \partial_{x_1} g_3 \\ \partial_{x_1} g_2 - \partial_{x_2} g_1 \end{pmatrix}$$

Zuletzt sei $f : D \subset \mathbb{R}^n \to \mathbb{R}$ zweimal partiell differenzierbar,

d.h. $\nabla f : D \to \mathbb{R}^n$ sei auch partiell differenzierbar

Dann ist Laplaceoperator Δ angewandt auf f definiert durch

$$\Delta f = \mathrm{div}\, \nabla f = \nabla \cdot \nabla f = \partial^2_{x_1} f + \ldots + \partial^2_{x_n} f \ : \ D \to \mathbb{R}$$

Satz 13.14 (Darstellung der Ableitung durch partielle Ableitung)

$f : D \subset V_1 \times V_2 \to W$ differenzierbar in $x = (x_1, x_2) \in D$
$\Longrightarrow f$ partiell differenzierbar in x und für $v_1 \in V_1,\ v_2 \in V_2$

$$f'(x)(v_1, v_2) \;=\; \partial_{x_1} f(x) v_1 \;+\; \partial_{x_2} f(x) v_2 \;=\; (\partial_{x_1} f(x), \partial_{x_2} f(x)) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

Lemma 13.15

Abbildung $\varphi : \mathcal{B}(V_1, W) \times \mathcal{B}(V_2, W) \to \mathcal{B}(V_1 \times V_2, W)$ gegeben durch

$$\varphi(T_1, T_2)(x_1, x_2) \;=\; T_1 x_1 + T_2 x_2$$

ist linear, stetig und bijektiv mit stetigem Inversen φ^{-1}

Hierbei ist $\mathcal{B}(V_1, W) \times \mathcal{B}(V_2, W)$ wieder ein kartesisches Produkt versehen mit Norm $\|(T_1, T_2)\| = \|T_1\| + \|T_2\|$

Beweis: Umkehrabbildung $\varphi^{-1} = ((\varphi^{-1})_1, (\varphi^{-1})_2)$ ist

$$(\varphi^{-1})_1(T)(x_1) = T(x_1, 0) \qquad (\varphi^{-1})_2(T)(x_2) = T(0, x_2)$$

wobei $T \in \mathcal{B}(V_1 \times V_2, W)$
Linearität von φ offensichtlich, Stetigkeit folgt aus

$$\begin{aligned} \|\varphi(T_1, T_2)\| &= \sup_{\|(x_1,x_2)\|=1} \|\varphi(T_1, T_2)(x_1, x_2)\| \\ &= \sup_{\|x_1\|+\|x_2\|=1} \|T_1 x_1 + T_2 x_2\| \\ &\leqslant \|T_1\| + \|T_2\| \end{aligned}$$

und Stetigkeit von φ^{-1} aus

$$\begin{aligned} \|\varphi^{-1}(T)\| &= \|(\varphi^{-1})_1(T)\| + \|(\varphi^{-1})_2(T)\| \\ &= \sup_{\|x_1\|=1} \|T(x_1, 0)\| + \sup_{\|x_2\|=1} \|T(0, x_2)\| \\ &\leqslant \|T\| + \|T\| = 2\,\|T\| \end{aligned}$$

□

Beweis von Satz 13.14:

Sei $T = f'(x) \in \mathcal{B}(V_1 \times V_2, W)$

Setze $T_1 = (\varphi^{-1})_1(T) \in \mathcal{B}(V_1, W)$ und $T_2 = (\varphi^{-1})_2(T) \in \mathcal{B}(V_2, W)$

Nach Lemma 13.15 sind T_1 und T_2 linear und stetig. Außerdem

$$\begin{aligned} &\frac{\|f(y, x_2) - f(x_1, x_2) - T_1(y - x_1)\|}{\|y - x_1\|} \\ &= \frac{\|f(y, x_2) - f(x_1, x_2) - T(y - x_1, 0)\|}{\|y - x_1\|} \\ &= \frac{\|f(y, x_2) - f(x_1, x_2) - T((y, x_2) - (x_1, x_2))\|}{\|(y, x_2) - (x_1, x_2)\|} \\ &\xrightarrow{y \to x_1} 0 \end{aligned}$$

Letzteres nach Differenzierbarkeit von f in (x_1, x_2)

Somit gilt $\partial_{x_1} f(x) = T_1$

Analog erhält man $\partial_{x_2} f(x) = T_2$ $\square$

Beispiel 13.16 (Umkehrung von Satz 13.14 gilt nicht)

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad f(x,y) = \begin{cases} \frac{2xy}{x^2+y^2} & , \quad (x,y) \neq 0 \\ 0 & , \quad (x,y) = 0 \end{cases}$$

f besitzt partielle Ableitungen (nachrechnen!)

$$\partial_x f(x,y) = \begin{cases} \frac{2y}{x^2+y^2} - \frac{4x^2y}{(x^2+y^2)^2} & , \quad (x,y) \neq 0 \\ 0 & , \quad (x,y) = 0 \end{cases}$$

$$\partial_y f(x,y) = \begin{cases} \frac{2x}{x^2+y^2} - \frac{4y^2x}{(x^2+y^2)^2} & , \quad (x,y) \neq 0 \\ 0 & , \quad (x,y) = 0 \end{cases}$$

Aber

$$f(r\cos\varphi, r\sin\varphi) = \frac{2r^2\cos\varphi\sin\varphi}{r^2(\cos^2\varphi + \sin^2\varphi)} = \sin(2\varphi)$$

so dass $\lim_{(x,y)\to 0} f(x,y)$ nicht existiert

f also nicht stetig ist bei 0 und somit auch nicht differenzierbar

Definition 13.17

V, W normierte Vektorräume und $D \subset V$ offen

$f : D \to W$ stetig differenzierbar auf D

$\iff \forall\, x_0 \in D$ ist f differenzierbar in x_0

und Abbildung $x \in D \mapsto f'(x) \in \mathcal{B}(V, W)$ ist stetig

Für stetig differenzierbare Funktionen gilt Umkehrung von Satz 13.14:

Satz 13.18

$f : D \subset V_1 \times V_2 \to W$ partiell differenzierbar. Äquivalent sind:

(i) $\partial_{x_1} f : D \to \mathcal{B}(V_1, W)$ *und* $\partial_{x_2} f : D \to \mathcal{B}(V_2, W)$ *stetig*

(ii) *f auf D stetig differenzierbar*

Dann gilt $f' = \varphi(\partial_{x_1} f, \partial_{x_2} f)$, wobei φ die Abbildung aus Lemma 13.15 ist

Analoges gilt für Abbildungen $f : D \subset V_1 \times \ldots \times V_n \to W$

Beweis: (ii) $\Longrightarrow$ (i) Da $f' : D \subset V_1 \times V_2 \to \mathcal{B}(V_1 \times V_2, W)$ stetig und

$$\partial_{x_j} f \;=\; (\varphi^{-1})_j \circ f' : D \to \mathcal{B}(V_j, W) \qquad j = 1, 2$$

folgt nach Lemma 13.15 die Stetigkeit der partiellen Ableitungen

(i) $\Longrightarrow$ (ii) Wir zeigen, dass für $(x_1, x_2) \in D$ gilt

$$f'(x_1, x_2)(y_1 - x_1, y_2 - x_2) = \partial_{x_1} f(x_1, x_2)(y_1 - x_1) + \partial_{x_2} f(x_1, x_2)(y_2 - x_2)$$

Die Stetigkeit folgt dann wieder aus Lemma 13.15. Tatsächlich $\forall\, \varepsilon > 0$:

$$\begin{aligned}
&\|f(y_1, y_2) - f(x_1, x_2) - \partial_{x_1} f(x_1, x_2)(y_1 - x_1) - \partial_{x_2} f(x_1, x_2)(y_2 - x_2)\| \\
&\qquad \leqslant \|f(y_1, y_2) - f(x_1, y_2) - \partial_{x_1} f(x_1, y_2)(y_1 - x_1)\| \\
&\qquad\quad + \|[\partial_{x_1} f(x_1, y_2) - \partial_{x_1} f(x_1, x_2)](y_1 - x_1)\| \\
&\qquad\quad + \|f(x_1, y_2) - f(x_1, x_2) - \partial_{x_2} f(x_1, x_2)(y_2 - x_2)\| \\
&\qquad \leqslant \varepsilon\|y_1 - x_1\| + \varepsilon\|y_1 - x_1\| + \varepsilon\|y_2 - x_2\|
\end{aligned}$$

für $\|(y_1, y_2) - (x_1, x_2)\| < \delta$, nach Def. & Stetigkeit partieller Ableitungen

Da ε beliebig klein, ist die Ableitung berechnet $\square$

Bemerkung 13.19

Wenn $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ differenzierbar, so ist $f'(x) \in \mathcal{B}(\mathbb{R}^n, \mathbb{R}^m)$ durch Jacobi Matrix $\partial f = (\partial_{x_i} f_j)_{j=1,\dots,m,\ i=1,\dots,n}$ gegeben (Satz 13.14), d.h. für $x \in D$ und $v \in \mathbb{R}^n$

$$f'(x)v \;=\; \partial f\, v \;=\; \begin{pmatrix} \sum_{i=1}^n \partial_{x_i} f_1 v_i \\ \vdots \\ \sum_{i=1}^n \partial_{x_i} f_m v_i \end{pmatrix} = \begin{pmatrix} \langle \nabla f_1 \mid v \rangle_{\mathbb{R}^n} \\ \vdots \\ \langle \nabla f_m \mid v \rangle_{\mathbb{R}^n} \end{pmatrix}$$

Insbesondere gilt im Fall $m = 1$, d.h. $f : D = \mathbb{R}^n \to \mathbb{R}$, dass

$$f'(x)v \;=\; \langle \nabla f(x) | v \rangle_{\mathbb{R}^n} \qquad v \in \mathbb{R}^n$$

Umgekehrt (Satz 13.18), wenn Jacobi Matrix existiert und *stetige* Einträge hat, so stellt sie die Ableitung dar.

Beispiel 13.20 (Differenzierbare Abbildung, die nicht stetig partiell differenzierbar ist)

$$f : \mathbb{R}^2 \to \mathbb{R} \qquad f(x,y) = \begin{cases} x^2 y \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Dann

$$\partial_x f(x,y) = \begin{cases} 2xy \sin\left(\frac{1}{x}\right) - y \cos\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Letzteres, weil $\forall\ y \in \mathbb{R}$

$$\frac{f(\varepsilon, y) - f(0,y) - 0 \cdot (\varepsilon - 0)}{\varepsilon} = \varepsilon y \sin\left(\frac{1}{\varepsilon}\right) \xrightarrow{\varepsilon \to 0} 0$$

Außerdem

$$\partial_y f(x,y) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Jetzt ist $\partial_y f$ stetig auf $\mathbb{R}^2$, aber $\partial_x f$ ist unstetig auf $S = \{(0,y) \mid y \neq 0\}$

Beispiel (Fortsetzung)

Dennoch ist f differenzierbar!

f' nach Satz 13.14 dann durch partiellen Ableitungen gegeben

In der Tat gilt für $y \neq 0$ und $(\varepsilon_x, \varepsilon_y) \neq (0,0)$

$$\frac{|f(\varepsilon_x, y+\varepsilon_y) - f(0,y) - \partial_x f(0,y)\varepsilon_x - \partial_y f(0,y)\varepsilon_y|}{\|(\varepsilon_x, \varepsilon_y)\|}$$

$$= \frac{|f(\varepsilon_x, y+\varepsilon_y)|}{|\varepsilon_x| + |\varepsilon_y|} = \frac{1}{|\varepsilon_x| + |\varepsilon_y|} \begin{cases} \left|\varepsilon_x^2 (y+\varepsilon_y) \sin\left(\frac{1}{\varepsilon_x}\right)\right| & \varepsilon_x \neq 0 \\ 0 & \varepsilon_x = 0 \end{cases}$$

$$\leqslant \begin{cases} |\varepsilon_x||y+\varepsilon_y| \left|\sin\left(\frac{1}{\varepsilon_x}\right)\right| & \varepsilon_x \neq 0 \\ 0 & \varepsilon_x = 0 \end{cases} \xrightarrow{(\varepsilon_x,\varepsilon_y)\to 0} 0$$

Es gibt noch ein weiteres Konzept von Ableitung:

Definition 13.21 (Richtungsableitungen)

V, W normierte Vektorräume, $D \subset V$ offen, $x \in D$ und $v \in V$

$f : D \to W$ hat Richtungsableitung $(\partial_v f)(x) \in W$ in Richtung v bei x

$\iff (\partial_v f)(x) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}(f(x + \varepsilon v) - f(x))$ existiert in $(W, \|.\|)$

Alternativ $(\partial_v f)(x) = \frac{d}{dt} f(x + tv) \,|_{t=0}$

Falls f in x Richtungsableitungen in alle Richtungen $v \in V$ besitzt, heißt f in x Gâteaux-differenzierbar

Bemerkung 13.22

1. Partielle Ableitungen sind Spezialfälle von Richtungsableitungen, z.B. für $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ gilt $\partial_{x_j} f = \partial_{e_j} f$
2. f differenzierbar $\Longrightarrow$ f Gâteaux-differenzierbar, weil (Kettenregel):

$$(\partial_v f)(x) \;=\; f'(x) \frac{d}{dt}(x + tv)|_{t=0} \;=\; f'(x)v$$

Beispiel 13.23

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & , \quad (x,y) \neq 0 \\ 0 & , \quad (x,y) = 0 \end{cases}$$ nicht differenzierbar in 0 (oben)

Aber f ist Gâteaux-differenzierbar, weil für $v = \binom{v_x}{v_y}$

$$(\partial_v f)(0,0) \;=\; \frac{d}{dt} \frac{v_x v_y}{v_x^2 + v_y^2} \,\Big|_{t=0} = \; 0$$

Beispiel 13.24

$$f(x,y) \;=\; \begin{cases} \frac{x^2 y}{x^4+y^2} \sqrt{x^2 + y^2} & , \quad (x,y) \neq 0 \\ 0 & , \quad (x,y) = 0 \end{cases} \qquad v = \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

f Gâteaux-differenzierbar in 0 da

$$\partial_v f(0) \;=\; \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \frac{\varepsilon^3 v_x^2 v_y}{\varepsilon^4 v_x^4 + \varepsilon^2 v_y^2} \, \varepsilon \sqrt{v_x^2 + v_y^2} \;=\; 0$$

Aber f nicht differenzierbar in 0, weil $\lim_{(x,y)\to 0} \frac{x^2 y}{x^4+y^2}$ nicht existiert, denn z.B. mit $x = t,\; y = at^2$ folgt $\lim_{t\to 0} \frac{t^2 a t^2}{t^4 + a^2 t^4} = \frac{a}{1-a^2}$

Zusammenfassung zu Ableitungsbegriffen:

Folgende echte Inklusionen gelten

(erste davon wird in Übung diskutiert):

$$\begin{aligned}\{\text{partiell differenzierbar}\} &\supsetneqq \{\text{Gâteaux-differenzierbar}\}\\ &\supsetneqq \{\text{(Fréchet) differenzierbar}\}\\ &\supsetneqq \{\text{stetig differenzierbar}\}\\ &= \{\text{stetig partiell differenzierbar}\}\end{aligned}$$

Definition 13.25 (Zweite Ableitung)

V, W normierte Vektorräume und $D \subset V$

$f : D \to W$ zweimal in $x \in D$ differenzierbar

$\Longleftrightarrow$ f differenzierbar auf D und

$\quad f' : D \to \mathcal{B}(V, W)$ differenzierbar in x

(wobei Vektorraum $\mathcal{B}(V, W)$ mit der Operatornorm versehen ist)

Zweite Ableitung ist dann $f''(x) \in \mathcal{B}(V, \mathcal{B}(V, W))$

Satz 13.26 (Satz von Schwarz)

f zweimal in $x \in D$ differenzierbar

$\Longrightarrow$ *$f''(x) \in \mathcal{B}(V, \mathcal{B}(V, W))$ ist symmetrisch, d.h. $\forall\ u, v \in V$*

$$(f''(x)u)v \;=\; (f''(x)v)u$$

Beweis: Für $u, v \in V$ ausreichend klein, definiere $g : [0,1] \to W$ durch

$$g(t) = f(x + tu + v) - f(x + tu)$$

Nach der Kettenregel ist g differenzierbar und

$$\begin{aligned} g'(t) &= f'(x + tu + v)u - f'(x + tu)u \\ &= (f'(x + tu + v) - f'(x))u - (f'(x + tu) - f'(x))u \end{aligned}$$

Weil f' differenzierbar ist, existiert $\forall\ \varepsilon > 0$ ein $\delta > 0$ mit

$$\|f'(x + tu + v) - f'(x) - f''(x)(tu + v)\| \leqslant \varepsilon\|tu + v\|$$

und

$$\|f'(x + tu) - f'(x) - f''(x)(tu)\| \leqslant \varepsilon\,\|tu\|$$

für $\|tu + v\| < \delta$ und $\|tu\| < \delta$. Somit

$$\begin{aligned} \|g'(t) - (f''(x)v)u\| &\leqslant \|(f'(x + tu + v) - f'(x) - f''(x)(tu + v))u\| \\ &\quad + \|(f'(x + tu) - f'(x) - f''(x)(tu))u\| \\ &\leqslant \varepsilon\|tu + v\|\|u\| + \varepsilon\|tu\|\|u\| \leqslant \varepsilon\|u\|(2\|u\| + \|v\|) \end{aligned}$$

Nun wenden wir Schrankensatz auf $t \in [0,1] \mapsto g(t) - t(f''(x)v)u$ an:

$$\|g(1)-(f''(x)v)u-g(0)\| \ \leqslant \ \sup_{t\in[0,1]} \|g'(t)-(f''(x)v)u\| \ \leqslant \ \varepsilon\,\|u\|(2\|u\|+\|v\|)$$

Da

$$g(1) - g(0) \ = \ f(x+u+v) - f(x+u) - f(x+v) + f(x)$$

symmetrisch in u und v ist, gilt ebenso

$$\|g(1) - g(0) - (f''(x)u)v\| \leqslant \varepsilon\|v\|(2\|v\| + \|u\|)$$

Somit folgt mit der Dreiecksungleichung:

$$\|(f''(x)v)u - (f''(x)u)v\| \ \leqslant \ \varepsilon\,2(\|u\|^2 + \|v\|^2 + \|u\|\|v\|)$$

Da ε beliebig klein, folgt $(f''(x)v)u = (f''(x)u)v$ zunächst für kleine u, v
Wegen der Linearität folgt es dann aber für alle u, v □

Bemerkung 13.27

Zweite Ableitung $T = f''(x) \in \mathcal{B}(V, \mathcal{B}(V, W))$ kann als bilineare Abbildung $\widetilde{T} : V \times V \to W$ aufgefasst werden, indem man definiert:

$$\widetilde{T}(v, w) \;=\; T(v)w$$

In der Tat, gilt dann

$$\begin{aligned}\widetilde{T}(v + \lambda v', w) \;&=\; \widetilde{T}(v, w) + \lambda \widetilde{T}(v', w)\\ \widetilde{T}(v, w + \lambda w') \;&=\; \widetilde{T}(v, w) + \lambda \widetilde{T}(v, w')\end{aligned}$$

Satz von Schwarz besagt, dass $\widetilde{T}$ symmetrisch ist, d.h.

$$\widetilde{T}(v, w) \;=\; \widetilde{T}(w, v)$$

Bezeichnung $\mathcal{B}(V, V; W) \cong \mathcal{B}(V, \mathcal{B}(V, W))$ mit Norm

$$\|\widetilde{T}\| = \sup_{\|v\|=1} \sup_{\|w\|=1} \|\widetilde{T}(v, w)\|$$

Meist: T und $\widetilde{T}$ identifiziert, d.h. auch $f''(x) = \widetilde{f''(x)} \in \mathcal{B}(V, V; W)$

Multilineare Abbildungen: (vergleiche Definition 2.12)

Analog wird

$$T \in \mathcal{B}(V, \mathcal{B}(V, \mathcal{B}(V, \ldots, \mathcal{B}(V, W)) \ldots)$$

mit k Argumenten aus V mit k-multilinearer Abbildung identifiziert

$$\widetilde{T} \in \mathcal{B}(V, \ldots, V; W) \;=\; \mathcal{B}(V^{\times k}; W)$$

Diese erfüllt dann für $\ell = 1, \ldots, k$:

$$\widetilde{T}(v_1, \ldots, v_\ell + \lambda w_\ell, \ldots, v_k) \;=\; \widetilde{T}(v_1, \ldots, v_k) + \lambda \widetilde{T}(v_1, \ldots, w_\ell, \ldots, v_k)$$

Die Norm auf den k-linearen Abbildungen ist wie oben definiert:

$$\|\widetilde{T}\| \;=\; \sup_{\|v_1\|=1} \cdots \sup_{\|v_k\|=1} \|\widetilde{T}(v_1, \ldots, v_k)\|$$

Dies ist wieder gleich der Norm $\|T\|$

Zuletzt: auch Symmetrie von $\widetilde{T}$ analog definiert

(ähnlich wie alternierend in Definition 2.14, nur ohne Vorzeichen)

Definition 13.28 (Höhere Ableitungen)

V, W Vektorräume und $D \subset V$ offen

Höhere Ableitungen von $f : D \to W$ sind iterativ definiert durch:

f ist k-mal auf D differenzierbar mit k-ten Ableitungen gegeben durch k-lineare Abbildungen $f^{(k)} : D \to \mathcal{B}(V, \ldots, V; W) = \mathcal{B}(V^{\times k}; W)$

$\iff$ f $(k-1)$-mal differenzierbar und $f^{(k-1)} : D \to \mathcal{B}(V^{\times k-1}; W)$ ist differenzierbar mit Ableitung $\left(f^{(k-1)}\right)' = f^{(k)}$

Falls $f^{(k)} : D \to \mathcal{B}(V^{\times k}; W)$ stetig, heißt f k-mal stetig differenzierbar

Die Menge dieser Funktionen wird mit $C^k(D, W)$ bezeichnet

Korollar 13.29

(i) $f \in C^k(D, W) \iff$ *alle k-ten partiellen Ableitungen sind stetig*

(ii) *Die k-multilineare Abbildung $f^{(k)}(x)$ ist symmetrisch,*
d.h. für jede Permutation $\sigma \in S_k$ und $v_1, \dots, v_k \in V$ gilt

$$f^{(k)}(x)(v_{\sigma(1)}, \dots, v_{\sigma(k)}) \;=\; f^{(k)}(x)(v_1, \dots, v_k)$$

Beweis: (i) und (ii) folgen aus der iterativen Anwendung von Satz 13.18 und Satz 13.26 respektive. □

Satz 13.30 (Satz von Taylor)

V, W normierte Vektorräume und $D \subset V$ offen

Zudem $x \in D$ und f n-mal differenzierbar auf D

Dann gilt für $v \in V$ mit $x + v \in D$ die Taylor Formel

$$f(x+v) \;=\; \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x) v^k \;+\; o(\|v\|^n) \qquad \textit{für } \|v\| \to 0$$

Erläuterung: Hierbei ist $f^{(k)}(x)v^k = f^{(k)}(x)(v, \dots, v)$

Erinnerung: $g(v) = o(\|v\|^n)$ für $\|v\| \to 0$ heißt $\lim_{v\to 0} \frac{\|g(v)\|}{\|v\|^n} = 0$

Beweis: Durch Induktion über n

Für $n = 1$ ist die Aussage genau die Definition der Ableitung

Für den Schritt von $n-1$ nach n betrachte den Rest

$$g(v) \;=\; f(x+v) \;-\; \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(x) v^k$$

Um die Ableitung zu berechnen, verwende

Behauptung: Zu $T \in \mathcal{B}(V^{\times k}; W)$ k-multilinear und symmetrisch, sei

$$h : V \to W \qquad h(v) \;=\; Tv^k$$

$$\Longrightarrow \quad h'(v) = kTv^{k-1} \in \mathcal{B}(V, W)$$

Begründung: Für $\varepsilon \in V$ gilt nach Multilinearität und Symmetrie

$$\frac{1}{\|\varepsilon\|}\|h(v+\varepsilon) - h(v) - h'(v)\varepsilon\| \;=\; \frac{1}{\|\varepsilon\|}\|T(v+\varepsilon)^k - Tv^k - kT(v^{k-1}, \varepsilon)\|$$

$$\leqslant \; \frac{1}{\|\varepsilon\|}\|T\| \sum_{\ell=2}^{k} \binom{k}{\ell} \|\varepsilon\|^\ell \;\xrightarrow{\varepsilon\to 0}\; 0$$

Dies zeigt die Behauptung

Also:

$$g'(v) = f'(x+v) - \sum_{k=1}^{n} \frac{1}{(k-1)!} f^{(k)}(x) v^{k-1}$$

Jetzt gilt nach dem Schrankensatz

$$\begin{aligned} \|g(v)\| &= \|g(v) - g(0)\| \\ &\leqslant \|v\| \sup_{0 \leqslant t \leqslant 1} \|g'(tv)\| \\ &\leqslant \|v\| \, o(\|v\|^{n-1}) \end{aligned}$$

Letzteres nach Induktionsannahme angewandt auf die $(n-1)$-mal differenzierbare Funktion $t \in [0,1] \mapsto g'(tv) \in W$

Da $\|v\| o(\|v\|^{n-1}) = o(\|v\|^n)$, folgt der Satz $\square$

Definition 13.31

V normierter Vektorraum und D offen

$f : D \subset V \to \mathbb{R}$ differenzierbare Abbildung

(i) $x \in D$ kritischer Punkt von $f \iff f'(x) = 0 = 0$-Abbildung

(ii) $w \in \mathbb{R}$ kritischer Wert von $f \iff w = f(x)$ für kritischen Punkt x

Beispiel 13.32

Für $f(x, y) = \cos(x) + \cos(y)$ ist $f'(x, y) = (-\sin(x), -\sin(y))$

$(n\pi, m\pi)$ mit $n, m \in \mathbb{Z}$ kritische Punkte und $-2, 0, 2$ kritische Werte

Satz 13.33

Sei $f : D \subset V \to \mathbb{R}$ differenzierbar und $x \in D$ lokales Extremum von f, d.h. z.B für ein lokales Minimum $\exists\, \varepsilon > 0$ mit $f(y) \geqslant f(x)\ \forall\ \|y - x\| < \varepsilon$

$\Longrightarrow x$ kritischer Punkt von f

Beweis: Für jedes $v \in V$ hat eindimensionale Abbildung $t \mapsto f(x+tv)$ lokales Extremum bei $t = 0$. Somit $\partial_v f(x) = 0 \;\; \forall \; v \in V$

Da f differenzierbar, gilt nach Kettenregel $\partial_v f(x) = f'(x)v = 0 \; \forall \; v \in V$

Also ist $f'(x)$ die 0-Abbildung $\square$

Sei nun $f : D \subset \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar

Zudem $x \in D$ kritischer Punkt von f

Nach Satz von Taylor gilt dann

$$f(x+v) \;=\; f(x) + \frac{1}{2} f''(x) v^2 \;+\; o(\|v\|^2)$$

Um zu überprüfen, ob x lokales Extremum ist, berechne $f''(x)v^2$

Zunächst

$$f'(x)v \;=\; \langle \nabla f(x) \mid v \rangle_{\mathbb{R}^n} = \sum_{j=1}^{n} \partial_{x_j} f(x) v_j \qquad v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Weiter

$$(f''(x)v)w = (f'(x)v)'w \qquad w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

$$= \langle \nabla \left(\sum_{j=1}^{n} \partial_{x_j} f(x) v_j \right) | w \rangle$$

$$= \sum_{i,j=1}^{n} (\partial_{x_i} \partial_{x_j} f)(x) v_j w_i$$

$$= \sum_{i,j=1}^{n} (\partial_{x_i} \partial_{x_j} f)(x) v_i w_j \quad \text{(Satz von Schwarz)}$$

Definition 13.34

$f : D \subset \mathbb{R}^n \to \mathbb{R}$ zweimal partiell differenzierbar in $x \in D$

Dann ist $\partial^2 f(x) = (\partial_{x_i} \partial_{x_j} f(x))_{i,j=1,\ldots,n}$ Hess'sche Matrix von f in x

Alternative Schreibweisen: $\partial^2 f(x) = \text{Hess}\,(f)(x) = H_f(x) = \nabla^2 f(x)$

Für $f \in C^2(D, \mathbb{R})$ gilt nach Satz von Schwarz:

$$(\partial^2 f(x))^T = \partial^2 f(x)$$

Also ist Hess'sche reell symmetrisch $\Longrightarrow$ orthogonal diagonalisierbar

Nach obiger Rechnung:

$$(f''(x)v)w = \langle v | \partial^2 f(x) w \rangle = \langle w | \partial^2 f(x) v \rangle$$

Erinnerung (Kapitel 7): für $A = A^T \in \mathrm{Mat}(n \times n, \mathbb{R})$ gilt

(i) $A > 0$ positiv $\Longleftrightarrow \langle v | Av \rangle > 0 \quad \forall\, v \in \mathbb{R}^n,\ v \neq 0$

(ii) $A > 0 \Longleftrightarrow$ alle Eigenwerte von A positiv

(iii) $A > 0 \Longleftrightarrow \det_k((T_{i,j})_{i,j=1,\ldots,k}) > 0$ für $k = 1, \ldots, n$ (Sylvester)

(iv) $A \geqslant 0$ nicht-negativ $\Longleftrightarrow \langle v | Av \rangle \geqslant 0 \quad \forall\, v \in \mathbb{R}^n,\ v \neq 0$

(v) Analog sind $A < 0$ und $A \leqslant 0$ definiert

(vi) A indefinit $\Longleftrightarrow$ es gibt positive und negative Eigenwerte

Satz 13.35

$f : D \subset \mathbb{R}^n \to \mathbb{R}$ zweimal stetig differenzierbar auf D offen

(i) *$x \in D$ lokales Maximum (Minimum) $\Longrightarrow \partial^2 f(x) \leqslant 0 \quad (\partial^2 f(x) \geqslant 0)$*

(ii) *$x \in D$ kritischer Punkt von f und $\partial^2 f(x) < 0$ (bzw. $\partial^2 f(x) > 0$)*

$\Longrightarrow x$ strenges lokales Maximum (bzw. Minimum) von f, d.h.

$\exists\, \delta > 0$ mit $f(x) > f(y)$ (bzw. $f(x) < f(y)$) $\quad \forall\, \|y - x\| < \delta,\ y \neq x$

(iii) *$x \in D$ kritischer Punkt von f und $\partial^2 f(x)$ indefinit*

$\Longrightarrow x$ kein lokales Extremum (Sattelpunkt)

Beweis: Alle Aussagen folgen direkt aus Satz von Taylor

$$f(y) \;=\; f(x) \;+\; \langle (x-y) | \partial^2 f(x)(x-y) \rangle \;+\; o(\|x-y\|^2)$$

In (iii) gibt es positive und negative Eigenwerte von $\partial^2 f(x)$ $\square$

Bemerkung 13.36

Wenn $\partial^2 f(x)$ semi-definit, keine Aussage möglich

Höhere Ordnungen notwendig

Beispiel 13.37

$f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x,y) = x^2 + xy + y^2 + x + y + 1$. Dann

$$\nabla f(x,y) = \begin{pmatrix} 2x + y + 1 \\ 2y + x + 1 \end{pmatrix}$$

Einziger kritischer Punkt:

$$\nabla f(x,y) = 0 \iff 2x + y = -1\,,\; x + 2y = -1 \iff x = y = -\frac{1}{3}$$

Hess'sche nach Sylvester positiv:

$$\partial^2 f(x,y) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} > 0$$

$\Longrightarrow (-\frac{1}{3}, -\frac{1}{3})$ lokales Minumum

Da $x^2 + xy + y^2 = \frac{1}{2}(x^2 + y^2 + (x+y)^2) \geqslant \frac{1}{2}(x^2 + y^2)$, wächst f für $|x|, |y| \to \infty \Longrightarrow$ globales Minimum

Beispiel 13.38

$f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x,y) = \cos(x) + \cos(y)$

$$\nabla f(x,y) = \begin{pmatrix} -\sin(x) \\ -\sin(y) \end{pmatrix} = 0 \quad \Longleftrightarrow \quad x \in \pi\mathbb{Z}\,, \quad y \in \pi\mathbb{Z}$$

Hess'sche:

$$\partial^2 f(x,y) \; = \; \begin{pmatrix} -\cos(x) & 0 \\ 0 & -\cos(y) \end{pmatrix}$$

Also: lokale Maxima bei $x \in 2\pi\mathbb{Z},\; y \in 2\pi\mathbb{Z}$

lokale Minima bei $x \in \pi + 2\pi\mathbb{Z},\; y \in \pi + 2\pi\mathbb{Z}$

Sattelpunkte bei $x \in \pi + 2\pi\mathbb{Z}\,,\; y \in 2\pi\mathbb{Z}$ und $x \in 2\pi\mathbb{Z}\,,\; y \in \pi + 2\pi\mathbb{Z}$

Bemerkung 13.39 (Gradienten-Fluss und Morse Theorie)

Für $f : D \subset \mathbb{R}^n \to \mathbb{R}$ ist Morse Vektorfeld $X : D \to \mathbb{R}^n$ definiert als

$$X(x) \;=\; \frac{\nabla f(x)}{\|\nabla f(x)\|^2}$$

Für jedes Vektorfeld X existiert der Fluss $\Theta_t : D \to D$, $t \in (-T, T) \subset \mathbb{R}$, definiert als Lösung der gewöhnlichen Differentialgleichung

$$\partial_t \Theta_t(x) \;=\; X(\Theta_t(x)) \qquad x \in D$$

Satz von Picard-Lindelöf: $f \in C^2(D, \mathbb{R}) \Longrightarrow \Theta_t$ existiert und eindeutig

Dann

$$\partial_t f(\Theta_t(x)) \;=\; \langle \nabla f(\Theta_t(x)) | \partial_t \Theta_t(x) \rangle \;=\; \langle \nabla f | \frac{\nabla f}{\|\nabla f\|^2} \rangle (\Theta_t(x)) \;=\; 1$$

Also $\Theta_t(\Sigma_E) = \Sigma_{E+t}$ für Niveau-Flächen $\Sigma_E \;=\; \{x \in D \mid f(x) = E\}$

Verhalten an kritischen Werten untersucht die Morse-Theorie

14 Nichtlineare Analysis

Erstes Hauptziel:

Satz 14.1 (Lokale inverse Funktion bzw. lokale Umkehrbarkeit)

V, W Banachräume und $D \subset V$ offen

$f : D \to W$ stetig differenzierbar

Bei $x_0 \in D$ sei $f'(x_0) \in \mathcal{B}(V, W)$ invertierbar mit $f'(x_0)^{-1} \in \mathcal{B}(W, V)$

$\Longrightarrow\ \exists$ offene Kugel $B_\delta(x_0) = \{y \in D \mid \|y - x_0\| < \delta\}$, so dass

$$f : B_\delta(x_0) \to f(B_\delta(x_0))$$

invertierbar ist mit inverser Abbildung $f^{-1} : f(B_{x_0}(x_0)) \to B_\delta(x_0)$, deren Ableitung stetig ist und gegeben durch

$$(f^{-1})'(y) \ = \ [f'(f^{-1}(y))]^{-1}$$

Falls $f \in C^k(D, W)$, so ist auch f^{-1} k-mal stetig differenzierbar

Bemerkung 14.2

1. Voraussetzung an nur einen Punkt, Aussage lokal
2. Seien $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$

   $f'(x_0)$ invertierbar $\implies n = m$
3. $\dim(V) = \infty$ und $f'(x_0)$ invertierbar $\implies \dim(W) = \infty$
4. $\dim(W) = \infty$ und $f'(x_0)$ invertierbar $\implies \dim(V) = \infty$
5. Im Fall $V = W = \mathbb{R}$ ist sogar globale Aussage möglich (Ana I):

   $f'(x) > 0 \ \forall \ x \in I \subset \mathbb{R} \Longrightarrow f$ invertierbar auf I

   Dies ist im Höherdimensionalen nicht möglich
6. Wesentliches Beweiselement: Banachscher Fixpunktsatz

Beispiel 14.3 (Notwendigkeit der C^1-Voraussetzung)

Voraussetzung stetiger Differenzierbarkeit kann nicht abgeschwächt werden zur Differenzierbarkeit: Die Funktion

$$f(x) = \begin{cases} x + x^2 \sin\left(\frac{1}{x}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

ist differenzierbar und

$$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x} = \lim_{x\to 0} \left(1 + x \sin \frac{1}{x}\right) = 1 > 0$$

Aber $f'(x) = 1 + 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$, so dass f' nicht stetig

In der Tat hat f' positive und negative Werte in jeder Umgebung von 0 da $f''(x) = 2\sin\left(\frac{1}{x}\right) - \frac{2}{x}\cos\left(\frac{1}{x}\right) + \frac{1}{x^2}\sin\left(\frac{1}{x}\right) = -\frac{2}{x}$ für $x = \frac{1}{2\pi k}$, $k \in \mathbb{N}$ und $f'(x) = 0$ für diese x, so dass Vorzeichen von f' hier wechselt

Also f nicht lokal monoton und somit nicht lokal umkehrbar bei 0

Beweis von Satz 14.1: Es reicht zu zeigen, dass Funktion

$$g \;:\; B_\delta(0) \subset V \to V \qquad g(x) \;=\; f'(x_0)^{-1}\big(f(x_0 + x) - f(x_0)\big)$$

für δ ausreichend klein invertierbar (dann auch f invertierbar). Nun:

$$g(0) \;=\; 0 \qquad g'(0) \;=\; \mathbf{1}_V$$

Somit nur Fall:

$$W = V\,, \quad x_0 = 0\,, \quad f(0) = 0\,, \quad f'(0) = \mathbf{1}_V$$

Da f stetig differenzierbar, $\exists\; \varepsilon > 0$ mit

$$\|f'(x) - \mathbf{1}_V\| \;\leqslant\; \tfrac{1}{2} \qquad \forall\; x \in \overline{B_\varepsilon(0)} \;=\; \{y \in V \mid \|y\| \leqslant \varepsilon\}$$

Für $y \in V$ betrachte die Funktion

$$F_y(x) \;=\; x + y - f(x) \qquad x \in \overline{B_\varepsilon(0)}$$

Idee hierbei: eindeutiger Fixpunkt x von $F_y(x) = x$ löst $f(x) = y$

Schrankensatz für $F_y(x) = x + y - f(x)$ und $x, x' \in \overline{B_\varepsilon(0)}$:

$$\begin{aligned}\|F_y(x) - F_y(x')\| &= \|x - f(x) - (x' - f(x'))\| \\ &\leqslant \|x - x'\| \sup_{t\in[0,1]} \|\mathbf{1}_V - f'(tx' + (1-t)x)\| \\ &\leqslant \tfrac{1}{2}\|x - x'\|\end{aligned}$$

Also F_y Lipshitz-stetig mit Konstante $L \leqslant \frac{1}{2} < 1$. Spezialfall $x' = 0$:

$$\|F_y(x) - y\| = \|F_y(x) - F_y(0)\| \leqslant \tfrac{1}{2}\|x\| \leqslant \tfrac{1}{2}\varepsilon$$

d.h. $F_y : \overline{B_\varepsilon(0)} \to \overline{B_{\frac{\varepsilon}{2}}(y)}$

Zudem $\overline{B_{\frac{\varepsilon}{2}}(y)} \subset \overline{B_\varepsilon(0)}$ für $y \in \overline{B_{\frac{\varepsilon}{2}}(0)}$

$\implies F_y : \overline{B_\varepsilon(0)} \to \overline{B_\varepsilon(0)}$ Lipshitz-stetig

$\overline{B_\varepsilon(0)}$ vollständig, weil abgeschlossen in vollständigem Raum V

$\implies$ Banachscher Fixpunktsatz kann auf F_y angewandt werden

$\Longrightarrow \ \forall\, y \in \overline{B_{\frac{\varepsilon}{2}}(0)}\ \ \exists$ eindeutiger Fixpunkt $x \in \overline{B_\varepsilon(0)}$ von F_y:

$$x \ = \ F_y(x) \ = \ x + y - f(x) \quad \Longleftrightarrow \quad y = f(x)$$

Für $y \in \overline{B_{\frac{\varepsilon}{2}}(0)}$ hat Gleichung $y = f(x)$ also genau eine Lösung und

$$f \ : \ \{x \in B_\varepsilon(0) \mid f(x) \in B_{\frac{\varepsilon}{2}}(0)\} \ = \ f^{-1}(B_{\frac{\varepsilon}{2}}(0)) \ \longrightarrow \ B_{\frac{\varepsilon}{2}}(0)$$

ist also invertierbar

Da f stetig, ist $f^{-1}\left(B_{\frac{\varepsilon}{2}}(0)\right)$ offen,

enthält somit eine offene Kugel $B_\delta(0)$, wie im Satz 14.1 behauptet

Verbleibt: f^{-1} auf $f(B_\delta(0)) \subset B_{\frac{\varepsilon}{2}}(0)$ stetig differenzierbar

Hierfür folgender Zusatz zu Satz 12.12:

Satz 14.4 (vergleiche Satz 12.12)

Seien V, W Banachräume und $T \in \mathcal{B}(V, W)$ invertierbar
Dann sind alle Elemente der offenen Kugel um T mit Radius $r = \frac{1}{\|T^{-1}\|}$

$$B_r(T) = \{S \in \mathcal{B}(V, W) \mid \|T - S\| < r\}$$

invertierbar
Definiere $\varphi : B_r(T) \subset \mathcal{B}(V, W) \to \mathcal{B}(W, V)$ durch $\varphi(S) = S^{-1}$
$\Longrightarrow$ φ differenzierbar und $\varphi'(S)R = -S^{-1}RS^{-1}$ für $R \in \mathcal{B}(V, W)$

Beweis: Erster Teil schon in Satz 12.12. Für letzte Aussage:

$$\begin{aligned}
&\|\varphi(S + R) - \varphi(S) - \varphi'(S)R\| \\
&= \|(S + R)^{-1} - S^{-1} + S^{-1}RS^{-1}\| \\
&\leqslant \|(S + R)^{-1}\| \|S - (S + R) + (S + R)S^{-1}R\| \|S^{-1}\| \\
&\leqslant \|(S + R)^{-1}\| \|S^{-1}\|^2 \|R\|^2 = o(\|R\|)
\end{aligned}$$

$\square$

Weiter im Beweis von Satz 14.1:

Zu zeigen: f^{-1} auf $f(B_\delta(0)) \subset B_{\frac{\varepsilon}{2}}(0)$ stetig differenzierbar

Zunächst: $\|f'(x) - \mathbf{1}\| < \frac{1}{2}$ für $x \in \overline{B_{\frac{\varepsilon}{2}}(0)}$

Also $f'(x)$ invertierbar (Neumann Reihe)

Da $x \mapsto f'(x)$ stetig, ist $x \in B_{\frac{\varepsilon}{2}}(0) \mapsto f'(x)^{-1} \in \mathcal{B}(V)$ stetig (Satz 14.4)

Außerdem: f^{-1} Lipshitz-stetig mit Lipshitz-Konstanten 2 auf $B_{\frac{\varepsilon}{2}}(0)$,

da für $x, x' \in \overline{B_\varepsilon(0)}$ gilt, wegen Lipschitz-Konstante $\frac{1}{2}$ von F_0,

$$\begin{aligned}
\|x - x'\| &= \|f(x) - f(x') - F_0(x) + F_0(x')\| \\
&\leqslant \|f(x) - f(x')\| + \tfrac{1}{2}\|x - x'\| \\
\iff \|x - x'\| &\leqslant 2\,\|f(x) - f(x')\| \\
\iff \|f^{-1}(y) - f^{-1}(y')\| &\leqslant 2\,\|y - y'\|
\end{aligned}$$

Somit auch Abbildung $y \in B_{\frac{\varepsilon}{2}}(0) \mapsto (f'(f^{-1}(y)))^{-1}$ stetig

Nun zeigen wir, dass $(f^{-1})'(y) = (f'(f^{-1}(y)))^{-1}$ Ableitung ist:

$$\begin{aligned}
&\|f^{-1}(y') - f^{-1}(y) - (f^{-1})'(y)(y'-y)\| \\
&= \|x' - x - (f'(x))^{-1}(f(x') - f(x))\| \\
&\leqslant \|f'(x)^{-1}\| \|f'(x)(x-x) - f(x') - f(x)\| \\
&\leqslant \|f'(x)^{-1}\|\, \eta\, \|x' - x\| \qquad (\forall\, \eta > 0 \text{ nach Def. von } f'(x)) \\
&\leqslant \|f'(x)^{-1}\|\, \eta\, 2\, \|y' - y\| \qquad (\text{nach Lipshitz-Stetigkeit})
\end{aligned}$$

Zuletzt verbleibt Aussage über k-fache Differenzierbarkeit. Hierzu

$$(f^{-1})'(y) \;=\; (f'(x))^{-1} \qquad x \;=\; f^{-1}(y)$$

iterativ unter Verwendung von Satz 14.4 ableiten, z.B.

$$(f^{-1})''(y) \;=\; -(f'(x))^{-1} f''(x)(f^{-1})'(y)(f'(x))^{-1}$$

etc. $\square$

Beispiel 14.5

$f : \mathbb{R}^2 \to \mathbb{R}^2$ gegeben durch $f(x,y) = (x^2 + xy + y^2 + x + y + 1, x + y)^T$

$$\partial f(x,y) = \begin{pmatrix} 2x + y + 1 & 2y + x + 1 \\ 1 & 1 \end{pmatrix}$$

Dies ist stetig und

$$\det(\partial f(x,y)) = (2x + y + 1) - (2y + x + 1) = x - y$$

Also f lokal invertierbar in (x,y) falls $x \neq y$

Satz 14.1 ist eine lokale Aussage. Folgende Definition ist global:

Definition 14.6

V, W Banachräume, $D \subset V$ offen und $f : D \to W$

f ist C^k-Diffeomorphismus $\iff f : D \to f(D)$ bijektiv

und f sowie $f^{-1} : f(D) \to D$ sind k-mal stetig differenzierbar

Satz 14.7 (Satz über die implizite Funktion)

Seien V, W Banachräume und $D \subset V \times W$ offen
$F : D \to W$ k-mal stetig differenzierbar für $k \geqslant 1$
Sei $(x_0, y_0) \in D$, so dass $\partial_y F(x_0, y_0) \in \mathcal{B}(W)$ invertierbar
$\implies \exists\, \varepsilon > 0$ *und eindeutiges* $G : B_\varepsilon(x_0) \times B_\varepsilon(F(x_0, y_0)) \subset V \times W \to W$ *mit*

$$F(x, G(x, y)) \;=\; y$$

Insbesondere gilt für $g(x) = G(x, F(x_0, y_0))$, $g : B_\varepsilon(x_0) \to W$,

$$F(x, g(x)) \;=\; F(x_0, y_0) \qquad \forall\, x \in B_\varepsilon(x_0)$$

Dann heißt g implizite Funktion zu F durch $(x_0, y_0) = (x_0, g(x_0))$
(Oft wird F so normiert, dass $F(x_0, y_0) = 0$)
Zudem sind G und g k-mal stetig differenzierbar und

$$g'(x) \;=\; -\left(\partial_y F(x, g(x))\right)^{-1} \partial_x F(x, g(x))$$

Beweis Betrachte $f : D \subset V \times W \to V \times W$ definiert durch

$$f(x, y) = (x, F(x, y))^T$$

Dann

$$\partial f(x, y) = \begin{pmatrix} \mathbf{1}_V & 0 \\ \partial_x F(x, y) & \partial_y F(x, y) \end{pmatrix}$$

Nun ist $\partial_y F(x_0, y_0)$ invertierbar. Also (wie für 2×2-Matrizen)

$$(\partial f(x_0, y_0))^{-1} = \begin{pmatrix} \mathbf{1}_V & 0 \\ -\partial_x F(x_0, y_0)(\partial_y F(x_0, y_0))^{-1} & (\partial_y F(x_0, y_0))^{-1} \end{pmatrix}$$

Nach Satz 14.1 existiert somit $\delta > 0$, so dass f ein lokales Inverses hat

$$f^{-1} : f(B_\delta(x_0, y_0)) \to B_\delta(x_0, y_0) \subset V \times W$$

welches zudem k-mal stetig differenzierbar ist

Dieses Inverse muss von folgender Gestalt sein (ohne Transponieren):

$$f^{-1}(x,y) \;=\; (x, G(x,y)) \qquad (x,y) \in f(B_\delta(x_0,y_0))$$

Somit

$$(x,y) \;=\; f(f^{-1}(x,y)) \;=\; (x, F(x, G(x,y)))$$

d.h.

$$y \;=\; F(x, G(x,y)) \qquad \forall\, (x,y) \in f(B_\delta(x_0,y_0))$$

Nun $f(B_\delta(x_0,y_0)) = (f^{-1})^{-1}(B_\delta(x_0,y_0))$ offen (da f^{-1} stetig)

Zudem $(x_0, F(x_0,y_0)) \in f(B_\delta(x_0,y_0))$

Somit existiert $\varepsilon > 0$ mit $B_\varepsilon(x_0) \times B_\varepsilon(F(x_0,y_0)) \subset f(B_\delta(x_0,y_0))$

und f^{-1} sowie G sind dann auf $B_\varepsilon(x_0) \times B_\varepsilon(F(x_0,y_0))$ definiert

Die erste Aussage folgt. Zur letzten:

$$0 \;=\; \partial_x F(x, g(x)) \;=\; \partial_x F(x, g(x)) + \partial_y F(x, g(x)) g'(x) \qquad \square$$

Beispiel 14.8

Sei $F : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $F(x, y) = y^2 - 2y - x^2$

$$F(x, y) = 0 \iff (y-1)^2 = 1 + x^2 \iff y = g(x) = 1 \pm \sqrt{1 + x^2}$$

Zwei glatte Lösungen. In der Tat $\partial_y F(x, y) = 2(y - 1) \neq 0 \ \forall \ y \neq 1$

Beispiel 14.9

Sei $F : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $F(x, y) = y^2 - 2xy - x^4$

$$F(x, y) = 0 \iff (y-x)^2 = x^4 + x^2 \iff y = x \pm x\sqrt{1 + x^2}$$

wieder zwei Lösungen, aber nicht mehr getrennt

Bei $(0, 0)$ gilt $\partial_y F(0, 0) = 0$, also keine Eindeutigkeit nach Satz 14.7

Beispiel 14.10

Nun $F(x, y) = y^2 + x^2 = 0 \iff x = y = 0$

Keine Lösungsfunktion, sondern Punkt. In der Tat wieder $\partial_y F(0, 0) = 0$

Definition 14.11 (Untermannigfaltigkeiten im Euklidischen Raum)

Seien $k \in \mathbb{N}$ und $\emptyset \neq \mathcal{M} \subset \mathbb{R}^{n+m}$

$\mathcal{M}$ C^k-Untermannigfaltigkeit des $\mathbb{R}^{n+m}$ der Dimension m

$\iff$ $\forall$ Punkte $p \in \mathcal{M}$ $\exists$ offene Umgebung $U = B_\varepsilon(p) \subset \mathbb{R}^{n+m}$ und

k-mal stetig differenzierbare Funktion $F : U \to \mathbb{R}^n$,

so dass F' auf U Maximalrang n hat und

$$\mathcal{M} \cap U \ = \ \{x \in U \mid F(x) = 0\}$$

Bemerkung 14.12

Wenn $F'(p)$ Maximalrang, so heißt p ein regulärer Punkt von F, andernfalls ein singulärer oder kritischer Punkt

Dies verallgemeinert Definition 13.31 zu vektorwertigen Funktionen

In Definition 14.11 tauchen nur reguläre Punkte auf

Beispiel 14.13

Sei $\mathcal{M} = \mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid \|x\| = 1\}$ n-Sphäre (bez. euklidische Norm)

Sie ist Untermannigfaltigkeit des $\mathbb{R}^{n+1}$ die Dimension n

Hier globale Funktion $F : \mathbb{R}^{n+1} \to \mathbb{R}$ gegeben durch $F(x) = \|x\| - 1$

Dann $\mathbb{S}^n = \{x \in \mathbb{R}^{n+1} \mid F(x) = 0\}$

Bemerkung 14.14

Lokal ist Untermannigfaltigkeit also simultane Niveaufläche der Komponentenfunktion F_j von F, d.h.

$$\mathcal{M} \cap U \;=\; \bigcap_{j=1}^{n} \{x \in U \mid F_j(x) = 0\}$$

Jede Niveau-Fläche $\{x \in U \mid F_j(x) = 0\}$ von Kodimension 1 im $\mathbb{R}^{n+m}$

Rangbedingung besagt: Hyperflächen schneiden sich alle transvers

Deswegen ist $\mathcal{M}$ von Kodimension n im $\mathbb{R}^{n+m}$, also von Dimension m

Satz 14.15

Eine m-dimensionale C^k-Untermannigfaltigkeit $\mathcal{M}$ besitzt einen Atlas

$$\mathcal{A} = \{\varphi \mid \varphi : U_\varphi \to \mathbb{R}^m \text{ Homöomorphismus, } U_\varphi \subset \mathcal{M}\}$$

bestehend aus Karten φ, die Folgendes erfüllen:

(i) *$U_\varphi \subset \mathcal{M}$ offen in $\mathcal{M}$ ($\mathcal{M}$ versehen mit Unterraumtopologie)*

(ii) $\bigcup_{\varphi \in A} U_\varphi = \mathcal{M}$

(iii) *Für alle $\varphi, \psi \in \mathcal{A}$ mit $U_\varphi \cap U_\psi \neq \emptyset$ sind die Kartenwechsel*

$$\psi \circ \varphi^{-1} : \varphi(U_\varphi \cap U_\psi) \subset \mathbb{R}^m \to \psi(U_\varphi \cap U_\psi) \subset \mathbb{R}^m$$

C^k-Diffeomorphismen

Bemerkung 14.16

(i), (ii), (iii) sind Definition (abstrakter) C^k-Mannigfaltigkeit

Jede C^k-Untermannigfaltigkeit ist also eine C^k-Mannigfaltigkeit

Umkehrung gilt: Satz von Whitney

Beweis:

Sei $U \subset \mathbb{R}^{m+n}$ offene Umgebung von $p = (x_0, y_0) \in \mathcal{M}$ im $\mathbb{R}^m \times \mathbb{R}^n$ und $F : U \to \mathbb{R}^n$ von Maximalrang, so dass

$$\mathcal{M} \cap U \;=\; \{(x, y) \in U \subset \mathbb{R}^m \times \mathbb{R}^n \mid F(x, y) = 0\}$$

Sei $\partial_y F(x_0, y_0) \in \mathcal{B}(\mathbb{R}^n)$ invertierbar (sonst permutiere Argumente)
Nach Satz über implizite Funktionen $\exists$ Umgebung $V \subset \mathbb{R}^m$ von x_0 und $g : V \to \mathbb{R}^n$, so dass

$$F(x, g(x))) \;=\; 0 \qquad \forall\, x \in V$$

Nun setze

$$\varphi^{-1} \;:\; V \to \mathcal{M} \qquad \varphi^{-1}(x) \;=\; (x, g(x))$$

und

$$\varphi \;:\; \varphi^{-1}(V) \to \mathbb{R}^m \qquad \varphi(x, g(x)) \;=\; x$$

Nun wird φ zu Diffeomorphismus $\tilde{\varphi}$ erweitert:

$$\tilde{\varphi}^{-1}(x,t) \;=\; (x, g(x)+t) \qquad x \in V\,,\; t \in B_\varepsilon(0) \subset \mathbb{R}^n$$

Tatsächlich ist $\tilde{\varphi}$ invertierbar, da

$$(\tilde{\varphi}^{-1})'(x,t) \;=\; \begin{pmatrix} \mathbf{1}_m & g'(x) \\ 0 & \mathbf{1}_n \end{pmatrix}$$

Also $\tilde{\varphi}^{-1}$ lokal bei $(x_0, 0)$ invertierbar (Satz 14.1)

Gegeben zweite Karte ψ, ist also $\tilde{\psi} \circ \tilde{\varphi}^{-1}$ lokaler C^k-Diffeomorphismus (auf adäquatem Definitionsbereich)

Zudem: $\psi \circ \varphi^{-1}(x) = \tilde{\psi} \circ \tilde{\varphi}^{-1}(x, 0)$

Somit Kartenwechsel lokale C^k-Diffeomorphismen

Atlas durch die Karte zu jedem Punkt $p \in \mathcal{M}$ gegeben $\square$

Nächste Problemstellung:

Gegeben (Energie) Funktion $f : D \subset \mathbb{R}^{n+m} \to \mathbb{R}$

Bestimme ihre Extrema unter vorgegebenen Nebenbedingungen

Letztere z.B. durch Untermannigfaltigkeit $\mathcal{M} \subset \mathbb{R}^{n+m}$ gegeben

Somit gesucht: Extrema der Funktion $f : \mathcal{M} \to \mathbb{R}$

Alternativ: n Nebenbedingungen durch Funktion $F : D \to \mathbb{R}^n$ gegeben

Menge $\{x \in D \mid F(x) = 0\}$ muss keine Untermannigfaltigkeit sein,

sondern kann neben regulären Teilstücken auch Singularitäten haben

Suche dann lokale Extrema auf jedem regulären Teil separat
und analysiere Funktionswerte von f bei Singularität separat

Beispiel 14.17

Sei $f : \mathbb{R}^2 \to \mathbb{R}$ gegeben durch $f(x, y) = ax + by^2$ mit $a, b \in \mathbb{R}$

Nebenbedingung: $F(x, y) = x^2 + y^2 - r^2 = 0$ (Kreis)

Parametrisierung $t \in [0, 2\pi) \mapsto h(t) = (r\cos(t), r\sin(t)) \in \mathbb{R}^2$

Dann müssen lediglich Extrema gesucht werden von

$$\tilde{f}(t) \;=\; f \circ h(t) = ar\cos(t) + br^2(1 - \cos^2(t))$$

Da $\tilde{f}'(t) = -ar\sin(t) + 2br^2\sin(t)\cos(t)$, liegen sie bei

$$\sin(t) \;=\; 0 \quad \iff \quad t = 0, \pi \quad \text{und} \quad \cos(t) = \frac{a}{2rb}$$

Also sind die Extremwerte zwei der folgenden Zahlen:

$$+ar \quad , \quad -ar \quad , \quad \frac{a^2}{2b} + br^2\left(1 - \frac{a^2}{4r^2b^2}\right)$$

Beispiel 14.18 (mit singulärer Nebenbedingung)

Sei $(x, y) \in \mathbb{R}^2 \mapsto f(x, y) = x^2 - \frac{4}{5}xy$

Nebenbedingung: $F(x, y) = y^2 - x^3 = 0$

Also: $y^2 = x^3$, d.h. $x \geqslant 0$ und $y = \pm x^{\frac{3}{2}}$

Parametrisierung also möglich (mit zwei Zweigen):

$$\begin{aligned}\tilde{f}_\pm(x) &= f(x, \pm x^{\frac{3}{2}}) = x^2 \mp \frac{4}{5}x^{\frac{5}{2}} \\ \tilde{f}'_\pm(x) &= 2x \mp 2x^{\frac{3}{2}} = 2x(1 \mp x^{\frac{1}{2}})\end{aligned}$$

Also liegen folgende kritische Werte vor:

1. An der Singularität $x = y = 0$: $f(0, 0) = 0$
2. Für Zweig $+$ bei $x = 1,\ y = 1$: $f(1, 1) = \frac{1}{5}$

   lokales Maximum weil $\tilde{f}''_+(x) = 2 - 3x^{\frac{1}{2}}$, also $\tilde{f}''_+(1) = -1 < 0$

Für globale Extrema noch Asymptotiken: $\tilde{f}_\pm(x) \to \mp\infty$ für $x \to \infty$

Obige Lösungen basieren auf Methode der Parametrisierung

Gut, wenn Nebenbedingungen geometrisch oder analytisch "einfach"

Für kompliziertere Fälle gibt es folgende Methode:

Satz 14.19 (Lagrange Multiplikatoren)

$\varnothing \neq D \subset \mathbb{R}^{m+n}$ offen und $f : D \to \mathbb{R}$ differenzierbar

$\mathcal{M}$ m-dimensionale Untermannigfaltigkeit

$p \in \mathcal{M} \cap D$ lokales Extremum von f unter der Nebenbedingung $\mathcal{M}$

$\Longrightarrow \quad \exists$ Lagrange Multiplikatoren $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$, so dass

$$\nabla f(p) \; = \; \sum_{j=1}^{n} \lambda_j \, \nabla F_j(p) \tag{14.1}$$

wobei $F : D \to \mathbb{R}^n$ die Nebenbedingung lokal parametrisiert,

d.h. $\mathcal{M} \cap D = \{x \in D \mid F(x) = 0\}$ (vergleiche Definition 14.11)

Bemerkung 14.20

(14.1) sind $n + m$ Gleichungen, da $\nabla f(p) \in \mathbb{R}^{n+m}$

Zusammen mit n Gleichungen $F(p) = 0$ gibt es $2n + m$ Gleichungen

Anzahl der Unbekannten $p_1, \ldots, p_{n+m}, \lambda_1, \ldots, \lambda_n$ auch $2n + m$

Typischerweise ist also Lösungsmenge diskret

Schwierigere Situation bei entarteten lokalen Extrema (Beispiel)

Bemerkung 14.21

Geometrisch besagt (14.1): Gradient von f in p orthogonal auf $\mathcal{M}$

Vektoren $\nabla F_j(p),\ j = 1, \ldots, n$, spannen orthogonales Komplement des Tangentialraumes an $\mathcal{M}$ bei p auf

Beweis von Satz 14.19:

Sei $p = (x_0, y_0) \in \mathbb{R}^m \times \mathbb{R}^n$ und, ohne Einschränkung,

$$\det_n(\partial_y F(x_0, y_0)) \neq 0$$

Nach Satz 14.7 $\exists$ differenzierbares $g : B_\varepsilon(x_0) \subset \mathbb{R}^m \to \mathbb{R}^n$ mit

$$F(x, g(x)) = 0 \quad \Longleftrightarrow \quad (x, g(x)) \in \mathcal{M}$$

Nun betrachte f eingeschränkt auf $\mathcal{M}$:

$$\tilde{f} : B_\varepsilon(x_0) \to \mathbb{R} \qquad \tilde{f}(x) = f(x, g(x))$$

Bei lokalen Extremum auf $\mathcal{M}$ bei $p = (x_0, y_0)$ gilt

$$\tilde{f}'(x_0) = \nabla \tilde{f}(x_0)^T = 0$$

Also nach Kettenregel

$$\begin{aligned}
0 \;=\; \tilde{f}'(x_0) \;&==\; (\partial_x f(p), \partial_y f(p)) \begin{pmatrix} \mathbf{1} \\ g'(x_0) \end{pmatrix} \\
&=\; \partial_x f(p) + \partial_y f(p) g'(x_0) \\
&=\; \partial_x f(p) - \partial_y f(p) (\partial_y F(p))^{-1} \partial_x F(p) \qquad \text{(nach Satz 14.7)}
\end{aligned}$$

Nun: $(\partial_y F(p))^{-1} \; : \; \mathbb{R}^n \to \mathbb{R}^n$ und $\partial_y f(p) : \mathbb{R}^n \to \mathbb{R}$,
also $\lambda = \partial_y f(p)(\partial_y F(p))^{-1}$ eine lineare Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}$,
die durch den Vektor $\lambda \in \mathbb{R}^n$ gegeben ist. Es gilt:

$$\partial_x f(p) \;=\; \lambda \, \partial_x F(p) \qquad \partial_y f(p) \;=\; \lambda \, \partial_y F(p)$$

Zusammen ist dies gerade (14.1) $\square$

Beispiel 14.22

Sei $f : \mathbb{R}^{m+1} \to \mathbb{R}$ gegeben durch $f(x) = \prod_{j=1}^{m+1} x_j^2$

Nebenbedingung m-Sphäre: $\mathcal{M} = \left\{ x \in \mathbb{R}^{m+1} \mid \sum_{j=1}^{m+1} x_j^2 = 1 \right\}$

f hat als stetige Funktion auf einem Kompaktum globale Extrema

$\mathcal{M}$ Niveaufläche zu $F : \mathbb{R}^{m+1} \to \mathbb{R}, \quad F(x) = \sum_{j=1}^{m+1} x_j^2 - 1 = 0$

Es gibt einen Lagrange Multiplikator $\lambda \in \mathbb{R}$. Gleichungen sind

$$\nabla f(x) = \lambda \nabla F(x) \quad \text{und} \quad F(x) = 0$$

d.h.

$$2x_i \prod_{j \neq i} x_j^2 = \lambda 2 x_i \qquad i = 1, \ldots, m+1 \qquad \sum_{j=i}^{m+1} x_j^2 = 1$$

oder

$$\prod_{j \neq i} x_j^2 = \lambda \qquad i = 1, \ldots, m+1 \qquad \sum_{j=1}^{m+1} x_j^2 = 1$$

Beispiel (Fortsetzung)

Lösungen: Falls ein $x_i = 0$, z.B. $x_{m+1} = 0$, dann ist $\lambda = 0$

es gibt nur eine Gleichung für m Unbekannte $x_1, \ldots, x_m$, nämlich

$$\sum_{j=1}^{m} x_j^2 \; = \; 1$$

Bei allen Lösungen dieser Gleichungen gilt $f(x) = 0$

Somit ist das Minimum 0 stark entartet

Wert 0 auf $(m-1)$-dimensionalen Untermannigfaltigkeiten,

gegeben durch die $(m-1)$-Sphären,

die man durch Schnitt mit den Ebenen $x_i = 0$ erhält

Beispiel (Fortsetzung)

Falls alle $x_j \neq 0$, so folgt aus den beiden ersten Gleichungen

$$1 = \frac{\lambda}{\lambda} = \frac{\prod_{j\neq 1} x_j^2}{\prod_{j\neq 2} x_j^2} = \frac{x_2^2}{x_1^2} \implies x_1^2 = x_2^2$$

Analog folgt aus den anderen Gleichungen

$$x_1^2 = x_2^2 = \ldots = x_{m+1}^2 = \frac{1}{m+1}$$

Letzteres wegen $\sum_{j=i}^{m+1} x_j^2 = (m+1)x_1^2 = 1$

Also ist Maximalwert von f gegeben durch $\left(\frac{1}{m+1}\right)^{m+1}$

Er wird an 2^{m+1} Punkten angenommen, mit Komponenten

$$x_j = \sigma_j \frac{1}{\sqrt{m+1}} \qquad \sigma_j \in \{-1, 1\}, \quad j = 1, \ldots, m+1$$