

Mathematik II für Studierende der Geophysik/Ozeanographie, Meteorologie und Physik

Vorlesungsskript

Janko Latschev
Fachbereich Mathematik
Universität Hamburg
www.math.uni-hamburg.de/home/latschev

Hamburg, Sommersemester 2019

⁰Version vom 10. Juli 2019, basierend auf Versionen der Vorjahre. Ursprüngliche Version: Vicente Cortés, 2006/07

-1 / 862

Inhaltsverzeichnis I

- 1 Die Determinante - Fortsetzung
 - Charakterisierung der Determinante
 - Explizite Formel für die Determinante
 - Transponierte Matrix
 - Determinantenmultiplikationssatz
 - Matrixinversion und Determinanten
- 2 Äquivalenzrelationen
 - Definition und erste Beispiele
 - Quotientenraum
 - Orientierung
- 3 Diagonalisierung von Endomorphismen
 - Eigenwert, Eigenvektor und Eigenraum
 - Charakteristisches Polynom
 - Diagonalisierbarkeit
- 4 Euklidische und Hermiteische Vektorräume
 - Bilinearformen

0 / 862

Inhaltsverzeichnis II

- Euklidische Vektorräume
- Orthonormale Basen
- Die Orthogonale Gruppe
- Hermitesche Vektorräume
- Die unitäre Gruppe
- Normalformen unitärer Endomorphismen
- Selbstadjungierte Endomorphismen
- Hauptachsentransformation

5 Metrische Räume und Vollständigkeit

- Metrische Räume
- Normierte Vektorräume
- Die Parallelogrammgleichung
- Äquivalente Normen
- Vollständigkeit und Hilberträume
- Banachscher Fixpunktsatz
- Topologische Grundbegriffe in metrischen Räumen

1 / 862

Inhaltsverzeichnis III

- Kompaktheit
- Abbildungsräume

6 Beschränkte Operatoren, Hilbertbasen, Fourier-Reihen

- Beschränkte Operatoren, Darstellungssatz von Riesz
- Orthonormale Familien und Hilbertbasen
- Fourier-Reihen

7 Differentialrechnung mehrerer Veränderlicher

- Richtungsableitung
- Partielle Ableitungen
- Höhere partielle Ableitungen
- Der Gradient einer Funktion
- Divergenz eines Vektorfeldes
- Der Laplace-Operator
- Die Rotation eines Vektorfeldes

8 Differenzierbare Abbildungen

- Differenzierbarkeit und Differential

2 / 862

Inhaltsverzeichnis IV

- Differenzierbarkeit und Stetigkeit
- Rechenregeln und Kettenregel für das Differential
- Mittelwertsatz
- Niveaumengen und lokale Extrema
- Taylorentwicklung
- Hessematrix und lokale Extrema
- 9 Der Umkehrsatz und seine Anwendungen
 - Umkehrsatz
 - Satz über implizite Funktionen
 - Abbildungen von konstantem Rang
- 10 Mannigfaltigkeiten
 - Untermannigfaltigkeiten im Euklidischen Raum
 - Tangentialraum
 - Extrema mit Nebenbedingungen
- 11 Gewöhnliche Differentialgleichungen
 - Definition und Beispiele

3 / 862

Inhaltsverzeichnis V

- Elementare Lösungsmethoden für DG'en 1. Ordnung
- Existenz und Eindeigkeitssätze für gewöhnliche DG'en
- Abhängigkeit der Lösung von den AB'en
- Lineare Differentialgleichungen im \mathbb{R}^n
- Gewöhnliche Differentialgleichungen höherer Ordnung

4 / 862

Determinante einer quadratischen Matrix

Zur Erinnerung:

Definition (Determinante)

Eine Abbildung $\det: \text{Mat}(n, \mathbb{K}) \rightarrow \mathbb{K}$ heißt **Determinante**, wenn sie folgende Eigenschaften hat:

(D1) \det ist linear in **jeder Spalte**, d.h.

$$\begin{aligned} & \det(a_1 \cdots a_{i-1} \quad \lambda a_i + \mu b_i \quad a_{i+1} \cdots a_n) \\ &= \lambda \det(a_1 \cdots a_i \cdots a_n) + \mu \det(a_1 \cdots a_{i-1} \quad b_i \quad a_{i+1} \cdots a_n) \end{aligned}$$

für alle Spaltenvektoren $a_1, \dots, a_n, b_i \in \mathbb{K}^n$ und $\lambda, \mu \in \mathbb{K}$.

(D2) \det ist **alternierend**, d.h. $\det A = 0$, falls zwei Spalten von A übereinstimmen.

(D3) (**Normierung**) Für die Einheitsmatrix $\mathbf{1}_n := (e_1 \cdots e_n)$ gilt $\det(\mathbf{1}_n) = 1$.

Am Ende des letzten Semesters haben wir folgenden Satz bewiesen:

Satz (Explizite Formel für die Determinante)

Es gibt genau eine Abbildung $\det: \text{Mat}(n, \mathbb{K}) \rightarrow \mathbb{K}$ mit den Eigenschaften (D1)-(D3).

Die Determinante einer $(n \times n)$ -Matrix $A = (a_{ij})_{i,j}$ ist durch folgende Formel gegeben:

$$\det A := \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

Für große $n \in \mathbb{N}$ ist diese Formel für Rechnungen ungeeignet, da $n!$ Produkte gebildet und summiert werden müssten.

435 / 862

Regel von Sarrus:

Andererseits erhalten wir für $n = 3$ direkt aus der Formel:

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

dabei entsprechen die ersten drei Terme den zyklischen Permutationen [123], [231] und [312], die letzten drei den Transpositionen τ_{13} , τ_{12} und τ_{23} .

Praktisches Vorgehen: Wir schreiben die ersten beiden Spalten noch einmal hinter die Matrix, addieren dann die Produkte der blauen Diagonalen und subtrahieren davon die Produkte der roten Diagonalen:

$$\begin{array}{ccccc} + & + & + & & \\ a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \\ - & - & - & & \end{array}$$

436 / 862

Transponierte Matrix

Wir unterbrechen nun kurz die Diskussion von Determinanten, um ein anderes nützliches Konzept einzuführen.

Definition (Transponierte Matrix)

Sei $A = (a_{ij})_{i,j}$ eine $(m \times n)$ -Matrix.

Die $(n \times m)$ -Matrix A^T mit den Einträgen

$$(A^T)_{ij} := a_{ji}$$

heißt die zu A **transponierte Matrix**.

Bemerkung

Auf der Menge $\text{Mat}(n, \mathbb{K})$ der quadratische Matrizen ist die Abbildung $\alpha: \text{Mat}(n, \mathbb{K}) \ni A \mapsto A^T \in \text{Mat}(n, \mathbb{K})$ eine **Involution**, d.h. $\alpha \circ \alpha = \text{Id}$.

437 / 862

Transponierte Matrix

Beispiel

Ist $v \in \text{Mat}(n, 1, \mathbb{K})$ ein Spaltenvektor, so ist $v^T \in \text{Mat}(1, n, \mathbb{K})$ der entsprechende Zeilenvektor.

Satz

Für alle $A \in \text{Mat}(m, n, \mathbb{K})$ und $B \in \text{Mat}(n, r, \mathbb{K})$ gilt

$$(AB)^T = B^T A^T.$$

Beweis. Für $C = AB$ und $D = B^T A^T$ gilt

$$(C^T)_{ki} = c_{ik} = \sum_{j=1}^n a_{ij} b_{jk} = \sum_{j=1}^n (B^T)_{kj} (A^T)_{ji} = (D)_{ki}.$$

□

438 / 862

Satz

Für jede Matrix $A = (a_{ij})_{i,j} \in \text{Mat}(n, \mathbb{K})$ gilt

$$\det A = \det A^T.$$

Beweis. Wir haben

$$\begin{aligned} \det A &= \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} \quad \text{und} \\ \det A^T &= \sum_{\tau \in S_n} \varepsilon(\tau) (A^T)_{\tau(1)1} \cdots (A^T)_{\tau(n)n} \\ &= \sum_{\tau \in S_n} \varepsilon(\tau) a_{1\tau(1)} \cdots a_{n\tau(n)}. \end{aligned}$$

Die Übereinstimmung der beiden Summen folgt mit Hilfe der Substitution $\tau = \sigma^{-1}$ durch Umordnung der Faktoren und unter Beachtung von $\varepsilon(\tau) = \varepsilon(\sigma)$. □

439 / 862

Folgerung

(i) *Ebenso wie durch die Eigenschaften*

(D1) *det ist linear in jeder Spalte,*

(D2) *det ist alternierend bzgl. der Spalten, d.h. $\det A = 0$, falls zwei Spalten von A übereinstimmen,*

(D3) *det ist normiert: $\det(\mathbf{1}_n) = 1$*

lässt sich $\det: \text{Mat}(n, \mathbb{K}) \rightarrow \mathbb{K}$ auch durch die folgenden Eigenschaften charakterisieren:

(D1') *det ist linear in jeder Zeile,*

(D2') *det ist alternierend bzgl. der Zeilen, d.h. $\det A = 0$, falls zwei Zeilen von A übereinstimmen,*

(D3') *$\det(\mathbf{1}_n) = 1$.*

Folgerung

- (ii) Addition des λ -fachen der j -ten Zeile der Matrix A zur i -ten Zeile ($i \neq j$) ändert den Wert der Determinante nicht.
- (iii) Vertauschen der i -ten und j -ten Zeile ($i \neq j$) ändert die Determinante nur um ein Vorzeichen.
- (iv) Insbesondere lässt sich die Determinante durch Anwendung des Gaußalgorithmus auf die Zeilen der Matrix berechnen (bei Buchführung über die Zeilenvertauschungen!).

441 / 862

Determinantenmultiplikationssatz

Satz (Determinantenmultiplikationssatz)

Es gilt für alle $A, B \in \text{Mat}(n, \mathbb{K})$:

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

Beweis.

1. Fall: Gilt $\det(B) = 0$, so ist $\ker(B) \neq 0$ und somit $\ker(AB) \neq 0$. Dann folgt auch $\det(AB) = 0 = \det(A) \cdot 0 = \det(A) \det(B)$.

2. Fall: Wir können also annehmen, dass $\det(B) \neq 0$. Wir fixieren ein solches B und zeigen, dass für die Funktion

$$A \mapsto \widetilde{\det} A := \frac{\det(A \cdot B)}{\det B}$$

die Axiome (D1')-(D3') einer Determinantenfunktion erfüllt sind, woraus wegen der Eindeutigkeit einer solchen Funktion $\widetilde{\det} A = \det A$ und somit die Behauptung folgt.

442 / 862

(D1) (ii) Um die Additivität von $\widetilde{\det}$ zu zeigen, drücken wir die Matrix $B = (b_1, \dots, b_n)$ durch (Spalten-)Vektoren $b_j \in \mathbb{K}^n$ aus, und die Matrix A schreiben wir mit Vektoren $a_j \in \mathbb{K}^n$ als $A = \begin{pmatrix} a_1^T \\ \vdots \\ a_n^T \end{pmatrix}$.

Für die i -te Zeile von A gelte $a_i^T = a_i'^T + a_i''^T$. Dann ist

$$A \cdot B = \begin{pmatrix} a_1^T b_1 & \dots & a_1^T b_n \\ \vdots & & \vdots \\ (a_i'^T b_1 + a_i''^T b_1) & \dots & (a_i'^T b_n + a_i''^T b_n) \\ \vdots & & \vdots \\ a_n^T b_1 & \dots & a_n^T b_n \end{pmatrix}$$

Aus der Additivität von \det folgt $\det AB = \det A'B + \det A''B$, und daraus die Additivität von $\widetilde{\det}$:

$$\widetilde{\det} A = \frac{\det AB}{\det B} = \widetilde{\det} A' + \widetilde{\det} A''.$$

□

445 / 862

Folgerung

(i)

$$\det: \text{GL}(n, \mathbb{K}) \rightarrow (\mathbb{K} \setminus \{0\}, \cdot)$$

ist ein Gruppenhomomorphismus.

(ii) Insbesondere gilt für alle $A \in \text{GL}(n, \mathbb{K})$:

$$\det(A^{-1}) = (\det A)^{-1}$$

(iii)

$$\text{SL}(n, \mathbb{K}) := \{A \in \text{GL}(n, \mathbb{K}) \mid \det A = 1\} \subseteq \text{GL}(n, \mathbb{K}).$$

ist eine Untergruppe, die sogenannte **spezielle lineare Gruppe**.

Beweis. (i-ii) sind klar. (iii) folgt, da $\text{SL}(n, \mathbb{K}) = \det^{-1}(1)$ Kern eines Gruppenhomomorphismus ist. □

Achtung: Für $n \geq 2$ gilt **nicht** $\det(A + B) = \det(A) + \det(B)$!

Gegenbeispiel $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ und $B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$.

446 / 862

Definition (Determinante eines Endomorphismus)

Sei V ein \mathbb{K} -Vektorraum der Dimension n , $F \in \text{End}(V)$ ein Endomorphismus und $A = M_B(F) \in \text{Mat}(n, \mathbb{K})$ seine darstellende Matrix bezüglich einer beliebigen Basis B von V . Wir definieren die **Determinante von F** als

$$\det F := \det A.$$

Behauptung: Dies ist wohldefiniert, d.h. die Definition hängt nicht von der Wahl der Basis B ab.

Beweis. Sei B' eine weitere Basis von V , $A' := M_{B'}(F)$. Dann entspricht A einer Selbstabbildung $\phi_B^{-1} \circ F \circ \phi_B$ von \mathbb{R}^n und A' einer Abbildung $\phi_{B'}^{-1} \circ F \circ \phi_{B'}$.

Letztere schreibt sich auch als

$$\underbrace{\phi_{B'}^{-1} \circ \phi_B}_{=: S} \circ \underbrace{\phi_B^{-1} \circ F \circ \phi_B}_{= A} \circ \underbrace{\phi_{B'}^{-1} \circ \phi_{B'}}_{=: S^{-1}}$$

d.h. $A' = SAS^{-1}$ mit der invertierbaren $n \times n$ -Matrix S .

Also ist $\det A' = \det(SAS^{-1}) = \det S \det A (\det S)^{-1} = \det A$. □

447 / 862

Methoden zur Matrixinversion

Matrixinversion mittels Gaußalgorithmus

Wir wollen die Inverse A^{-1} zu $A \in \text{GL}(n, \mathbb{K})$ bestimmen. Dazu verknüpfen wir folgende Überlegungen:

- Die i -te Spalte der Matrix A^{-1} ist $A^{-1}e_i$, wobei e_i der i -te Vektor der kanonischen Basis ist.
- Gleichzeitig ist der Vektor $\mathbf{x} := A^{-1}e_i$ offenbar die eindeutige Lösung des inhomogenen Gleichungssystems $A\mathbf{x} = e_i$.
- Wir können also \mathbf{x} bestimmen, indem wir den Gaußschen Algorithmus auf $(A|e_i)$ anwenden bis wir $(\mathbf{1}_n|\mathbf{x})$ erhalten.

Auf diese Weise haben wir die i -te Spalte $\mathbf{x} = A^{-1}e_i$ von A^{-1} ermittelt. Das Verfahren auf alle Basisvektoren e_i gleichzeitig angewandt

$$(A \mid \mathbf{1}_n) \xrightarrow{\text{Gaußscher Alg.}} (\mathbf{1}_n \mid \mathbf{X})$$

liefert die inverse Matrix $A^{-1} = \mathbf{X}$.

Ein Zahlenbeispiel:

Wir wollen $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 1 & -1 \end{pmatrix}$ invertieren (das geht, da $\det A = -2$):

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{array} \right) \xrightarrow{III-I} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & -2 & -1 & 0 & 1 \end{array} \right) \xrightarrow{I-II, III \cdot (-1/2)}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 1 & -1 & 0 \\ 0 & 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & -\frac{1}{2} \end{array} \right) \xrightarrow{I+III, II-2 \cdot III} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{2} & -1 & -\frac{1}{2} \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & -\frac{1}{2} \end{array} \right),$$

$$\text{d.h. } A^{-1} = \begin{pmatrix} \frac{3}{2} & -1 & -\frac{1}{2} \\ -1 & 1 & 1 \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{pmatrix}.$$

ÜA: Als explizite Probe rechne man $AA^{-1} = \mathbf{1}_3$ nach.

Frage: Was bedeutet es für dieses Verfahren, wenn der Gaußalgorithmus in der linken Hälfte eine Nullzeile erzeugt?

449 / 862

Berechnung der Inversen mittels Determinanten

Für das nächste Verfahren benötigen wir den Begriff der Streichungsmatrix.

Definition (Streichungsmatrix)

Sei $A = (a_{kl})_{k,l=1\dots n} \in \text{Mat}(n, \mathbb{K})$ und $1 \leq i, j \leq n$ zwei Indizes. Die Matrix $A_{ij} \in \text{Mat}(n-1, \mathbb{K})$, die aus A durch Streichen der i -ten Zeile und j -ten Spalte entsteht, heißt (i, j) -te **Streichungsmatrix** von A .

Lemma

- (i) die (i, j) -te Streichungsmatrix von A^T stimmt mit der Matrix $(A_{ji})^T$ überein.
- (ii) $(-1)^{i+j} \det A_{ij} = \det(\mathbf{a}_1 \cdots \mathbf{a}_{j-1} \mathbf{e}_i \mathbf{a}_{j+1} \cdots \mathbf{a}_n)$.

Beweis. (i) ist klar.

Beweis von (ii). Die $(n \times n)$ -Matrix $(\mathbf{a}_1 \cdots \mathbf{a}_{j-1} \mathbf{e}_j \mathbf{a}_{j+1} \cdots \mathbf{a}_n)$ lässt sich durch Addition von Vielfachen der j -ten Spalte zu den anderen Spalten in folgende Form bringen:

$$B_{ij} := \begin{pmatrix} a_{11} & \cdots & a_{1j-1} & 0 & a_{1j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{i-11} & \cdots & a_{i-1j-1} & 0 & a_{i-1j+1} & \cdots & a_{i-1n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ a_{i+11} & \cdots & a_{i+1j-1} & 0 & a_{i+1j+1} & \cdots & a_{i+1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj-1} & 0 & a_{nj+1} & \cdots & a_{nn} \end{pmatrix}$$

Die Matrix B_{ij} lässt sich durch $i - 1$ Zeilen- und $j - 1$ Spaltenvertauschungen in die Blockdiagonalgestalt $\begin{pmatrix} 1 & 0 \\ 0 & A_{ij} \end{pmatrix}$ bringen.

Also folgt $\det(\mathbf{a}_1 \cdots \mathbf{a}_{j-1} \mathbf{e}_j \mathbf{a}_{j+1} \cdots \mathbf{a}_n) =$

$$\det B_{ij} = (-1)^{(i-1)+(j-1)} \det A_{ij} = (-1)^{i+j} \det A_{ij}. \quad \square$$

451 / 862

Satz

Sei $A = (a_{ij})_{i,j} \in \text{Mat}(n, \mathbb{K})$ und sei $\tilde{A} = (\tilde{a}_{ij})_{i,j}$ definiert durch

$$\tilde{a}_{ij} := (-1)^{i+j} \det A_{ji} \stackrel{\text{Lemma}}{=} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{e}_j \mathbf{a}_{i+1} \cdots \mathbf{a}_n).$$

Dann gilt $\tilde{A}A = A\tilde{A} = (\det A)\mathbf{1}_n$.

Beweis. Wir zeigen zuerst $\tilde{A}A = (\det A)\mathbf{1}_n$:

$$\begin{aligned} \sum_j \tilde{a}_{ij} a_{jk} &\stackrel{\text{Lemma}}{=} \sum_j a_{jk} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{e}_j \mathbf{a}_{i+1} \cdots \mathbf{a}_n) \\ &= \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \underbrace{\sum_j a_{jk} \mathbf{e}_j}_{=\mathbf{a}_k} \mathbf{a}_{i+1} \cdots \mathbf{a}_n) = \delta_{ik} \det A. \end{aligned}$$

Weiter im Beweis: Ist $\det(A) \neq 0$, so erhalten wir aus $\tilde{A}A = (\det A)\mathbf{1}_n$ auch $A\tilde{A} = \det(A)\mathbf{1}_n$, und zwar durch Multiplikation von $\tilde{A}A = (\det A)\mathbf{1}_n$ von links mit A und von rechts mit A^{-1} .
Ist aber $0 = \det(A) = \det(A^T)$, so gilt

$$0 = (\widetilde{A^T})A^T \stackrel{\text{Lemma}(i)}{=} (\tilde{A})^T A^T = (A\tilde{A})^T$$

und damit auch $A\tilde{A} = 0$. □

453 / 862

Folgerung (Entwicklungssatz von Laplace)

Für jede Matrix $A \in \text{Mat}(n, \mathbb{K})$ gilt:

(Z) (Entwicklung nach der i -ten Zeile)

$$\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det A_{ij}$$

(S) (Entwicklung nach der j -ten Spalte)

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}.$$

Beweis. Auf der rechten Seite von (Z) bzw (S) steht, siehe Lemma (ii), genau der i -te bzw. j -te Diagonaleintrag der Matrix $A\tilde{A} = (\det A)\mathbf{1}_n$ bzw. der Matrix $\tilde{A}A = (\det A)\mathbf{1}_n$. □

454 / 862

Ein Zahlenbeispiel für den Entwicklungssatz von Laplace:

Wir wollen die Determinante von $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ \pi & 1 & -1 \end{pmatrix}$ durch Entwicklung nach der ersten Spalte berechnen:

$$\begin{aligned} \det \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ \pi & 1 & -1 \end{pmatrix} &= 1 \cdot \det A_{11} - 1 \cdot \det A_{21} + \pi \cdot \det A_{31} \\ &= 1 \cdot \det \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix} - 1 \cdot \det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} + \pi \cdot \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \\ &= -3 + 2 + \pi = \pi - 1. \end{aligned}$$

455 / 862

Folgerung (Cramersche Regel)

Sei $A \in GL(n, \mathbb{K})$ und $\mathbf{b} \in \mathbb{K}^n$. Dann gilt:

- (i) $A^{-1} = \frac{1}{\det A} \tilde{A}$, wobei $\tilde{A} = (\tilde{a}_{ij})$ die zuvor definierte Matrix mit $\tilde{a}_{ij} = (-1)^{i+j} \det(A_{ji}) = \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{e}_j \mathbf{a}_{i+1} \cdots \mathbf{a}_n)$ ist.
- (ii) Die eindeutig bestimmte Lösung $\mathbf{x} = (x_i)_{i=1 \dots n}$ der Gleichung $A\mathbf{x} = \mathbf{b}$ berechnet sich nach der folgenden **Cramerschen Regel**

$$x_i = \frac{1}{\det A} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{b} \mathbf{a}_{i+1} \cdots \mathbf{a}_n).$$

Beweis.

- (i) folgt sofort aus $\det A \neq 0$ und $\tilde{A}A = (\det A)\mathbf{1}_n$.
- (ii) Mit $\tilde{a}_{ij} = \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{e}_j \mathbf{a}_{i+1} \cdots \mathbf{a}_n)$ und $\mathbf{x} = A^{-1}\mathbf{b}$ ist

$$\begin{aligned} x_i &= \frac{1}{\det A} \sum_j \tilde{a}_{ij} b_j = \frac{1}{\det A} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \sum_j b_j \mathbf{e}_j \mathbf{a}_{i+1} \cdots \mathbf{a}_n) \\ &= \frac{1}{\det A} \det(\mathbf{a}_1 \cdots \mathbf{a}_{i-1} \mathbf{b} \mathbf{a}_{i+1} \cdots \mathbf{a}_n) \quad \square \end{aligned}$$

456 / 862

Cramersche Regel für $n = 2$:

Wie bereits im vergangenen Semester diskutiert, hat in Dimension 2 die Lösung eines Gleichungssystems $A\mathbf{x} = \mathbf{b}$ die Koordinaten

$$x_1 = \frac{b_1 a_{22} - a_{12} b_2}{a_{11} a_{22} - a_{12} a_{21}} \quad \text{und}$$
$$x_2 = \frac{a_{11} b_2 - b_1 a_{21}}{a_{11} a_{22} - a_{12} a_{21}}.$$

Außerdem gilt (vgl. Spezialfälle: $\mathbf{b} = \mathbf{e}_1$ bzw. $\mathbf{b} = \mathbf{e}_2$):

$$A^{-1} = \frac{1}{a_{11} a_{22} - a_{12} a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

457 / 862

Kapitel 14

Äquivalenzrelationen

458 / 862

Äquivalenzrelationen

Im vergangenen Semester haben wir bereits Teilmengen $R \subseteq X \times X$ als *Relationen auf der Menge X* aufgefasst (z.B. die Ordnung \leq auf den reellen Zahlen). Dabei haben wir auch $x R y$ geschrieben, falls $(x, y) \in R$.

Definition

Eine Relation \sim auf X heißt **Äquivalenzrelation**, wenn sie folgende Eigenschaften hat:

- (i) $x \sim x$ für alle $x \in X$, (reflexiv)
- (ii) $x \sim y \implies y \sim x$ (symmetrisch)
- (iii) $x \sim y$ und $y \sim z \implies x \sim z$ (transitiv)

Jedes Element $x \in X$ definiert eine **Äquivalenzklasse**

$$[x] := \{y \in X \mid y \sim x\}.$$

Die Menge der Äquivalenzklassen wird oft mit X / \sim bezeichnet.

459 / 862

Beispiele

- X sei die Menge der Schüler einer Schule, \sim die Äquivalenzrelation "in die gleiche Klasse gehen". Dann ist X / \sim die Menge der Klassen.
- \mathcal{C} sei die Menge der Cauchyfolgen $\mathbf{x} = (x_n)_{n \geq 0}$ in \mathbb{Q} . Wir sagen $\mathbf{x} \sim \mathbf{y}$, falls die Folge $\mathbf{x} \# \mathbf{y}$, definiert durch

$$(\mathbf{x} \# \mathbf{y})_n = \begin{cases} x_n & \text{falls } n \text{ ungerade} \\ y_n & \text{falls } n \text{ gerade,} \end{cases}$$

ebenfalls eine Cauchyfolge ist. Man kann zeigen, dass dies eine Äquivalenzrelation auf \mathcal{C} definiert. Die Menge der Äquivalenzklassen kann als ein Modell für \mathbb{R} interpretiert werden.

- Wir nennen zwei quadratische Matrizen $A, B \in \text{Mat}(n, \mathbb{K})$ *zueinander konjugiert*, wenn es eine Matrix $S \in \text{GL}(n, \mathbb{K})$ gibt, so dass $A = SBS^{-1}$. Dies ist eine Äquivalenzrelation auf $\text{Mat}(n, \mathbb{K})$.

460 / 862

Beispiele (Fortsetzung)

- Sei $X = \mathbb{Z}$ die Menge der ganzen Zahlen, und sei $n > 1$ eine natürliche Zahl. Dann wird durch

$$x \sim y \iff n \text{ teilt } (x - y)$$

eine Äquivalenzrelation auf \mathbb{Z} definiert.

Zwei ganze Zahlen sind hier genau dann äquivalent, wenn sie bei Division durch n denselben Rest lassen. Die Menge dieser *Restklassen* bezeichnet man mit $\mathbb{Z}/n\mathbb{Z}$ oder auch einfach \mathbb{Z}_n .

Durch

$$[x] + [y] := [x + y] \quad \text{und} \quad [x] \cdot [y] := [x \cdot y]$$

werden auf \mathbb{Z}_n Rechenoperationen eingeführt.

$(\mathbb{Z}_n, +)$ ist eine kommutative Gruppe mit neutralem Element $[0]$.

Ist p prim, so ist $(\mathbb{Z}_p, +, \cdot)$ sogar ein Körper, wobei $[1]$ das neutrale Element der Multiplikation ist.

461 / 862

Definition (Quotientenraum)

Sei \mathbb{K} ein Körper und sei V ein Vektorraum über \mathbb{K} , und sei $U \subseteq V$ ein Unterraum. Dann definieren wir eine Äquivalenzrelation \sim auf V wie folgt:

$$x \sim y \iff x - y \in U.$$

Also $R = \{(x, y) \mid x, y \in V \text{ und } x - y \in U\} \subseteq V \times V$.

$V/U := \{[x] \mid x \in V\}$ ist mit den Operationen

$$[v_1] + [v_2] := [v_1 + v_2] \quad \text{und} \quad \lambda \cdot [v] := [\lambda v]$$

(in kanonischer Weise) ein Vektorraum über \mathbb{K} .

V/U heißt **Quotientenvektorraum** von V nach U .

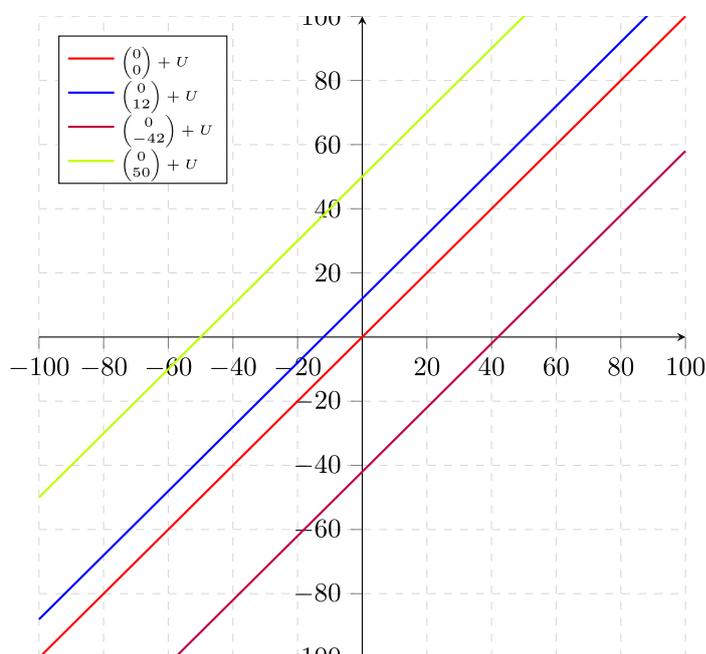
Die natürliche Abbildung $\pi : V \rightarrow V/U$, welche $v \in V$ auf $[v] \in V/U$ abbildet, ist linear und surjektiv mit Kern $\ker \pi = U$. Insbesondere erhalten wir aus der Dimensionsformel

$$\dim V/U = \dim V - \dim U.$$

462 / 862

Ein Beispiel

Hier sehen Sie einige Äquivalenzklassen von \mathbb{R}^2 nach dem in hellem Rot gezeichneten ein-dimensionalen Unterraum:



463 / 862

Orientierung

Definition

Sei V ein endlichdimensionaler **reeller** Vektorraum. Zwei (geordnete) Basen $B = (b_1, \dots, b_n)$ und $B' = (b'_1, \dots, b'_n)$ heißen **gleich orientiert**, wenn der **Basiswechsel** $F = \phi_{B'} \circ \phi_B^{-1} : V \rightarrow V$ positive Determinante hat.

Bemerkungen

- Beachte, dass $F(b_i) = \phi_{B'} \phi_B^{-1}(b_i) = \phi_{B'}(e_i) = b'_i$.
- Die darstellende Matrix von F bezüglich der Basis B ist die Matrix der Abbildung $\phi_B^{-1} \circ (\phi_{B'} \circ \phi_B^{-1}) \circ \phi_B = \phi_B^{-1} \circ \phi_{B'} : \mathbb{R}^n \rightarrow \mathbb{R}^n$. Insbesondere:

$$\det F = \det(\phi_B^{-1} \circ \phi_{B'}).$$

464 / 862

Orientierung

Satz

Sei V ein reeller Vektorraum der Dimension $n \in \mathbb{N}$. Wir schreiben $B \sim B'$, falls B und B' gleich orientierte (geordnete) Basen von V sind.

- Die Relation " \sim " ist eine Äquivalenzrelation auf der Menge der (geordneten) Basen von V .
- Jede (geordnete) Basis B definiert eine Äquivalenzklasse (ihre "Händigkeit")

$$[B] := \{B' \text{ Basis von } V \mid B' \sim B\}$$

und es gibt genau zwei Äquivalenzklassen.

Beweis.

- ist eine einfache Übungsaufgabe.
- Sei $B = (b_1, \dots, b_n)$ eine Basis. Jede Basis ist entweder gleich orientiert zu (b_1, \dots, b_n) oder zu $(-b_1, b_2, \dots, b_n)$. Daher gibt es genau zwei Äquivalenzklassen.

465 / 862



Orientierung

Definition

Sei V ein reeller Vektorraum der Dimension $n \in \mathbb{N}$. Eine **Orientierung** von V ist eine Äquivalenzklasse gleich orientierter Basen von V .

Bemerkungen

- Jede Basis $B = (b_1, \dots, b_n)$ definiert eine Orientierung $[B]$. Wie wir gesehen haben, gibt es genau zwei Orientierungen $[B]$ und $[B]^{op} = [(-b_1, b_2, \dots, b_n)]$. Letztere heißt die zu $[B]$ **entgegengesetzte** (oder **umgekehrte**) Orientierung.
- Automorphismen $F \in \text{GL}(V)$ mit $\det F > 0$ heißen **orientierungserhaltend**, denn $FB = (Fb_1, \dots, Fb_n) \in [B]$ für jede Basis $B = (b_1, \dots, b_n)$.
Automorphismen $F \in \text{GL}(V)$ mit $\det F < 0$ heißen **orientierungsumkehrend**, denn $FB = (Fb_1, \dots, Fb_n) \in [B]^{op}$ für jede Basis B .

Orientierung

Beispiele I

- Die kanonische Orientierung des \mathbb{R}^n ist die durch die kanonische Basis definierte Orientierung $[(e_1, \dots, e_n)]$. Die Basen (e_2, e_3, e_1) und $(-e_2, e_1, e_3)$ definieren auch die **kanonische Orientierung** des \mathbb{R}^3 .
- Die durch die Basis (e_2, e_1, e_3) definierte Orientierung ist zur durch die kanonische Basis definierten Orientierung $[(e_1, e_2, e_3)]$ entgegengesetzt. Es gilt $[(e_2, e_1, e_3)] = [(-e_1, e_2, e_3)]$.
- Die Drehung (um die z-Achse, entgegen dem Uhrzeigersinn um den Winkel φ)

$$D = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ist wegen $\det D = 1 > 0$ orientierungserhaltend.

467 / 862

Orientierung

Beispiele II

- Die Spiegelung (an der (x, y) -Ebene)

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

ist wegen $\det S = -1 < 0$ orientierungsumkehrend.

- Die Punktspiegelung $-\text{Id} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist eine Verknüpfung von n Spiegelungen, und somit
 - ▶ orientierungsumkehrend falls n ungerade und
 - ▶ orientierungserhaltend falls n gerade ist.

468 / 862

Kapitel 15

Diagonalisierung von Endomorphismen

469 / 862

Eigenwerte, Eigenvektor und Eigenraum

Definition

Sei F ein Endomorphismus eines \mathbb{K} -Vektorraums V , d.h. die Selbstabbildung $F: V \rightarrow V$ sei \mathbb{K} -linear.

- $\lambda \in \mathbb{K}$ heißt **Eigenwert** von F , wenn es einen **vom Nullvektor verschiedenen** Vektor $0 \neq v \in V$ gibt, so dass

$$F(v) = \lambda v.$$

- Jeder solche Vektor heißt **Eigenvektor** von F zum Eigenwert λ .
- Der Unterraum von V definiert durch

$$V_\lambda := \{v \in V \mid F(v) = \lambda v\}$$

heißt **Eigenraum** von F zum Eigenwert λ .

470 / 862

Beispiele

- Für die Identität $\text{Id} : V \rightarrow V$ ist jeder von 0 verschiedene Vektor ein Eigenvektor zum Eigenwert 1, d.h. der gesamte Raum V ist Eigenraum zum Eigenwert 1.
- Sei $V = \mathbb{R}^2$ und

$$F = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \text{End}(V).$$

$e_1 + e_2$ ist ein Eigenvektor zum Eigenwert 1,

$e_1 - e_2$ ist ein Eigenvektor zum Eigenwert -1 ,

Die Eigenräume sind $V_1 = \mathbb{R}(e_1 + e_2)$ und $V_{-1} = \mathbb{R}(e_1 - e_2)$,

Da $V = V_1 \oplus V_{-1}$, hat F keine weiteren Eigenwerte.

471 / 862

Beispiele (Fortsetzung)

Die lineare Abbildung von \mathbb{R}^2 nach \mathbb{R}^2 , gegeben durch die Matrix

$$D_\varphi := \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix},$$

ist eine Drehung um den Winkel φ und hat daher **keine reellen Eigenwerte**, falls φ kein ganzzahliges Vielfaches von π ist.

Zum Beispiel ist $D_{\pi/2} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ und die Eigenwertgleichung liefert

$$D_{\pi/2} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} \stackrel{!}{=} \lambda \begin{pmatrix} x \\ y \end{pmatrix},$$

d.h. $-y = \lambda x$ und $x = \lambda y$. Dies impliziert aber $x = -\lambda^2 x$ und $y = -\lambda^2 y$. Wegen $\lambda^2 \geq 0$, ist die einzige Lösung dafür $x = y = 0$.

472 / 862

Charakteristisches Polynom

Das charakteristische Polynom dient zur Bestimmung der Eigenwerte.

Verabredung. Wir nehmen im Folgenden an, dass $\mathbb{K} \supseteq \mathbb{Q}$ ein unendlicher Körper ist. Wir unterscheiden nicht zwischen Polynomen $P(t) \in \mathbb{K}[t]$ und polynomialen Funktionen $P: \mathbb{K} \rightarrow \mathbb{K}$.

Definition (Charakteristisches Polynom)

Sei V ein \mathbb{K} -Vektorraum der Dimension $n < \infty$ und $F \in \text{End}(V)$.

Das **charakteristische Polynom** $P_F: \mathbb{K} \rightarrow \mathbb{K}$, $t \mapsto P_F(t)$, von F ist definiert durch

$$\begin{aligned} P_F(t) &:= \det(F - t\text{Id}) = \det(A - t\mathbf{1}_n) \\ &= \sum_{\sigma \in S_n} \varepsilon(\sigma) (a_{1\sigma(1)} - t\delta_{1\sigma(1)}) \cdots (a_{n\sigma(n)} - t\delta_{n\sigma(n)}), \end{aligned}$$

wobei $A = (a_{ij})_{i,j}$ die darstellende Matrix von F bezüglich (irgend)einer Basis von V ist.

473 / 862

Wie sieht das charakteristische Polynom aus?

Satz

Das charakteristische Polynom von $F \in \text{End}(V)$ hat Grad $n = \dim V$.
Schreibt man

$$P_F(t) = \sum_{k=0}^n \alpha_k t^k$$

mit Koeffizienten $\alpha_j \in \mathbb{K}$, die von F abhängen, dann gilt

$$\begin{aligned} \alpha_0 &= \det F = \det A, \\ \alpha_{n-1} &= (-1)^{n-1} (a_{11} + \cdots + a_{nn}) = (-1)^{n-1} \text{Tr } A, \\ \alpha_n &= (-1)^n, \end{aligned}$$

wobei $A = (a_{ij})$ eine darstellende Matrix von F ist.

474 / 862

Bemerkung

Wie alle Koeffizienten des charakteristischen Polynoms ist auch die Spur $\text{Tr } F$ des Endomorphismus F (definiert als $\text{Tr } F := \text{Tr } A$ für eine beliebige darstellende Matrix A) unabhängig von der Wahl der Basis.

Beweis des Satzes.

Das Polynom $(a_{1\sigma(1)} - t\delta_{1\sigma(1)}) \cdot \dots \cdot (a_{n\sigma(n)} - t\delta_{n\sigma(n)})$ hat für jede Permutation $\sigma \neq \text{Id}$ einen Grad $\leq n - 2$, da dann mindestens zwei Ausdrücke $\delta_{i\sigma(i)}$ gleich 0 sind. Daher gilt

$$P_F(t) = (a_{11} - t) \cdot \dots \cdot (a_{nn} - t) + Q(t)$$

mit einem Polynom $Q(t)$ vom Grad $\leq n - 2$.

Ausmultiplizieren liefert nun:

$$P_F(t) = (-t)^n + (-t)^{n-1}(a_{11} + \dots + a_{nn}) + \dots + P_F(0)$$

mit $P_F(0) = \det(F - 0 \cdot \text{Id}) = \det F$.

□

475 / 862

Charakteristisches Polynom und Eigenwerte

Satz

Sei V ein endlichdimensionaler Vektorraum und $F \in \text{End}(V)$. Dann gilt:
Die Eigenwerte von F sind genau die Nullstellen des charakteristischen Polynoms von F .

Beweis.

$\lambda \in \mathbb{K}$ ist genau dann ein Eigenwert von F , wenn es einen Vektor $0 \neq v \in V$ gibt mit $F(v) = \lambda v$, d.h. wenn $(F - \lambda \text{Id})(v) = 0$ gilt, d.h. wenn $\ker(F - \lambda \text{Id}) \neq 0$.

Dies ist gleichbedeutend mit $\det(F - \lambda \text{Id}) = P_F(\lambda) = 0$.

□

Diagonalisierbarkeit von Endomorphismen

Definition

Sei V ein \mathbb{K} -Vektorraum der Dimension n .

$F \in \text{End}(V)$ heißt **diagonalisierbar**, wenn es eine Basis $B = (b_1, \dots, b_n)$ von V gibt, so dass die darstellende Matrix von F Diagonalgestalt hat:

$$M_B(F) = \text{diag}(\lambda_1, \dots, \lambda_n) := \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}, \quad \text{mit } \lambda_1, \dots, \lambda_n \in \mathbb{K}.$$

Bemerkung

Die Basisvektoren b_i sind dann Eigenvektoren von F und die Diagonaleinträge λ_i sind die zugehörigen Eigenwerte λ_i .

F ist also genau dann diagonalisierbar, wenn es eine Basis aus Eigenvektoren gibt.

Diese Basis, wenn sie existiert, ist nicht eindeutig.

477 / 862

Charakteristisches Polynom und Diagonalisierbarkeit

Satz

Das charakteristische Polynom eines **diagonalisierbaren** Endomorphismus F zerfällt in Linearfaktoren.

Beweis. Aus $M_B(F) = \text{diag}(\lambda_1, \dots, \lambda_n)$ folgt

$$\begin{aligned} P_F(t) &= \det(\text{diag}(\lambda_1 - t, \dots, \lambda_n - t)) \\ &= (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t) \\ &= (-1)^n (t - \lambda_1) \cdot \dots \cdot (t - \lambda_n). \end{aligned}$$

□

Bemerkung

Die Umkehrung des Satzes gilt nicht, wie wir gleich sehen werden.

478 / 862

Beispiel 1

Sei

$$F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \text{End}(\mathbb{K}^2).$$

$P_F(t) = (1 - t)^2$ zerfällt in Linearfaktoren, $t = 1$ ist der einzige Eigenwert. Eigenvektoren zum Eigenwert 1 erfüllen die Gleichung

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 0.$$

Also ist der Eigenraum zum Eigenwert 1 gleich $V_1 = \mathbb{K}e_1$.

Es gibt also keine Basis aus Eigenvektoren und F ist nicht diagonalisierbar.

479 / 862

Beispiel 2

Sei

$$F = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \text{End}(\mathbb{K}^2).$$

Dann gilt $P_F(t) = t^2 + 1$.

Über dem Körper $\mathbb{K} = \mathbb{R}$ hat $P_F(t)$ keine Nullstellen und zerfällt also nicht in Linearfaktoren.

Insbesondere ist F nicht diagonalisierbar. Wir haben schon gesehen, dass F die 90° -Drehung ist und deshalb keine Eigenvektoren hat.

Über dem Körper $\mathbb{K} = \mathbb{C}$ zerfällt $P_F(t) = t^2 + 1 = (t + i)(t - i)$.

Darüber hinaus ist F diagonalisierbar (mit komplexen Eigenwerten $\pm i$):

$$\begin{aligned} F(e_1 - ie_2) &= e_2 + ie_1 = i(e_1 - ie_2) \\ F(e_1 + ie_2) &= e_2 - ie_1 = -i(e_1 + ie_2). \end{aligned}$$

480 / 862

Multiplizität der Nullstellen des char. Polynoms

Ist λ eine Nullstelle eines Polynoms $P(t)$, so kann man P schreiben als

$$P(t) = (t - \lambda)^m Q(t)$$

mit einem Polynom Q mit $Q(\lambda) \neq 0$. Die nicht-negative Zahl m heißt **Vielfachheit** der Nullstelle λ .

Definition (Multiplizität der Nullstellen des char. Polynoms)

Sei F ein Endomorphismus eines endlichdimensionalen \mathbb{K} -Vektorraums V mit Eigenwert λ .

Die **algebraische Vielfachheit** m_λ des Eigenwerts λ ist definiert durch

$$P_F(t) = (t - \lambda)^{m_\lambda} Q(t), \text{ wobei } Q \in \mathbb{K}[t] \text{ mit } Q(\lambda) \neq 0.$$

Die **geometrische Vielfachheit** n_λ des Eigenwerts λ ist definiert als die Dimension $\dim V_\lambda$ des zu λ gehörenden Eigenraums $V_\lambda \subseteq V$.

481 / 862

Multiplizität und Dimension des Eigenraumes

Satz

Es sei F ein Endomorphismus eines endlichdimensionalen \mathbb{K} -Vektorraums V mit Eigenwert λ . Dann ist die geometrische Vielfachheit von λ durch die algebraische Vielfachheit beschränkt:

$$n_\lambda \leq m_\lambda.$$

Beweis. Sei $B_\lambda = (b_1, \dots, b_{n_\lambda})$ eine Basis des Eigenraums V_λ .

Nach dem Basisergänzungssatz können wir B_λ zu einer Basis B von V ergänzen.

Dann gilt

$$M_B(F) = \begin{pmatrix} \lambda \mathbf{1}_{n_\lambda} & * \\ 0 & D \end{pmatrix},$$

wobei D eine quadratische Matrix ist.

Daraus folgt $P_F(t) = (-1)^{n_\lambda} (t - \lambda)^{n_\lambda} P_D(t)$ und somit $n_\lambda \leq m_\lambda$. □

482 / 862

Lineare Unabhängigkeit von Eigenvektoren

Satz

Sei V ein Vektorraum und $F \in \text{End}(V)$. Die Vektoren v_i , $i = 1, \dots, k$ seien Eigenvektoren zu **paarweise verschiedenen** Eigenwerten λ_i .

Dann ist die Familie (v_1, \dots, v_k) linear unabhängig.

Beweis. Wir führen den Beweis durch Induktion nach k .

IA: Für $k = 1$ ist nichts zu zeigen, da ein Eigenvektor per definitionem ungleich dem Nullvektor ist.

IV: Sei $k \geq 2$, und es gelte, dass die Familie (v_1, \dots, v_{k-1}) von $k - 1$ Eigenvektoren v_1, \dots, v_{k-1} zu verschiedenen Eigenwerten λ_i linear unabhängig ist.

IS: Seien v_1, \dots, v_k Eigenvektoren zu verschiedenen Eigenwerten λ_i . Wir müssen zeigen, dass die Familie (v_i) linear unabhängig ist.

483 / 862

Weiter im Beweis:

Angenommen, es gelte $0 = \sum_{i=1}^k c_i v_i$, so folgt einerseits

$$0 = \sum_{i=1}^k \lambda_k c_i v_i$$

und andererseits durch Anwendung von F

$$0 = \sum_{i=1}^k \lambda_i c_i v_i.$$

Subtraktion der beiden Gleichungen liefert

$$0 = \sum_{i=1}^{k-1} \underbrace{(\lambda_k - \lambda_i)}_{\neq 0} c_i v_i$$

und somit (wegen der Induktionsvoraussetzung) $c_i = 0$ für $i = 1, \dots, k - 1$.

Es folgt $\sum_{i=1}^k c_i v_i = c_k v_k$ und somit aus der Annahme auch $c_k = 0$. \square

484 / 862

Erinnerung und Erweiterung:

Definition (Innere Summe von mehr als zwei Unterräumen)

Sei V ein Vektorraum und $V_1, V_2, \dots, V_k \subseteq V$ Unterräume. Die **innere Summe** der Unterräume V_i ist dann der Unterraum

$$V_1 + V_2 + \dots + V_k := \{v_1 + v_2 + \dots + v_k \mid v_i \in V_i, i = 1, 2, \dots, k\}.$$

Haben die Unterräume die Eigenschaft

$$V_i \cap \operatorname{span}\left\{\bigcup_{j \neq i} V_j\right\} = \{0\} \quad \forall i$$

so heißt die Summe **direkt** und wir schreiben $V_1 \oplus V_2 \oplus \dots \oplus V_k$.

485 / 862

Es gilt:

- $V_1 \oplus \dots \oplus V_k = (\dots ((V_1 \oplus V_2) \oplus V_3) \oplus \dots \oplus V_{k-1}) \oplus V_k$.
- Jeder Vektor $v \in V_1 \oplus \dots \oplus V_k$ hat eine **eindeutige** Darstellung $v = v_1 + v_2 + \dots + v_k$ mit $v_i \in V_i$.
- In einer direkten Summe $V_1 \oplus \dots \oplus V_k$ ist jedes k -Tupel (v_1, v_2, \dots, v_k) von Vektoren $v_i \in V_i \setminus 0$ linear unabhängig.

Bemerkung

Allgemeiner:

Sei $(V_i)_{i \in I}$ eine beliebige (nicht notwendiger Weise endliche) Familie von Unterräumen $V_i \subseteq V$.

Wir definieren $\sum_{i \in I} V_i := \operatorname{span}\{\cup_{i \in I} V_i\}$ und schreiben $V = \bigoplus_{i \in I} V_i$, falls (i) $V = \sum_{i \in I} V_i$ und (ii) jede Familie $(v_i)_{i \in I}$ von Vektoren $v_i \in V_i \setminus 0$ linear unabhängig ist.

486 / 862

Diagonalisierbarkeit und Aufspaltung in Eigenräume

Satz

Sei V ein endlichdimensionaler Vektorraum und $F \in \text{End}(V)$. Dann sind äquivalent:

- (i) Der Endomorphismus F ist diagonalisierbar.
- (ii) Das charakteristische Polynom P_F zerfällt in Linearfaktoren, und für alle Eigenwerte λ stimmen geometrische und algebraische Vielfachheit überein, $m_\lambda = n_\lambda$.
- (iii) V ist die direkte Summe der Eigenräume von F , $V = \bigoplus_{\lambda \text{ EW}} V_\lambda$.

Beweis. Wir zeigen $(iii) \Rightarrow (i) \Rightarrow (ii) \Rightarrow (iii)$.

$(iii) \Rightarrow (i)$ Es ist klar, dass jeder Eigenraum unter F invariant ist (warum?), und dass die Einschränkung von F auf jeden Eigenraum die Form $\lambda \cdot \mathbf{1}_{n_\lambda}$ hat.

Eine Blockdiagonalmatrix, deren Blöcke Diagonalmatrizen sind, ist aber eine Diagonalmatrix.

487 / 862

Weiter im Beweis:

$(i) \Rightarrow (ii)$ Ist klar, denn das charakteristische Polynom eines diagonalisierbaren Endomorphismus F zerfällt in Linearfaktoren.

$(ii) \Rightarrow (iii)$ Seien $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von F .

Behauptung: Die innere Summe $U = \bigoplus_{i=1}^k V_{\lambda_i} \subseteq V$ der Eigenräume zu den verschiedenen Eigenwerten ist eine direkte Summe.

Diese Behauptung beweisen wir indirekt. Wir nehmen also an, für ein gewisses $1 \leq i \leq k$ existiert ein von Null verschiedener Vektor

$$0 \neq v \in V_{\lambda_i} \cap \text{span}\left\{\bigcup_{j \neq i} V_{\lambda_j}\right\}.$$

Dann gilt also $v = \sum_{j \neq i} v_j$ mit $v_j \in V_{\lambda_j}$.

Die von Null verschiedenen Vektoren der Menge $\{v\} \cup \{v_j \mid j \neq i\}$ erfüllen dann die Voraussetzung des letzten Satzes, bilden also eine linear unabhängige Menge von Vektoren. Dies widerspricht aber gerade der Wahl von v . Die Behauptung ist also bewiesen.

488 / 862

Weiter im Beweis:

Wir wissen dann

$$\dim U = \sum_{\lambda \in EW} n_{\lambda}.$$

Andererseits gilt, da P_F vollständig in Linearfaktoren zerfällt, dass

$$\dim V = \deg P_F(t) = \sum_{\lambda \in EW} m_{\lambda}.$$

Da aber für alle Eigenwerte die Vielfachheiten m_{λ} und n_{λ} übereinstimmen, folgt $\dim U = \dim V$, und somit wegen $U \subseteq V$ auch $U = V$. \square

Beispiele

- Wir haben gesehen, dass der Endomorphismus

$$F = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \text{End}(\mathbb{K}^2)$$

das charakteristische Polynom $P_F(t) = (1 - t)^2$ hat, welches in Linearfaktoren zerfällt. F ist aber nicht diagonalisierbar. Der einzige Eigenwert $\lambda = 1$ hat $n_1 = 1 < 2 = m_1$.

- Der Endomorphismus

$$F = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \in \text{End}(\mathbb{K}^2)$$

hat das charakteristische Polynom $P_F(t) = (1 - t)(2 - t)$, welches ebenfalls in Linearfaktoren zerfällt. Da die geometrische Vielfachheit jedes Eigenwerts mindestens 1 ist, stimmen hier geometrische und algebraische Vielfachheiten überein. Die Abbildung ist also diagonalisierbar. ÜA: Finden Sie eine Basis aus Eigenvektoren!

Kapitel 15

Euklidische und Hermitesche Vektorräume

491 / 862

Wir betrachten nun Vektorräume, die mit einer zusätzlichen Struktur, einem **Skalarprodukt**, ausgestattet sind. Mit dessen Hilfe können wir geometrische Größen wie Abstand, Länge und Winkel definieren.

Erinnerung: Linearformen

Sei V ein \mathbb{K} -Vektorraum.

- Der Vektorraum

$$V^* := L(V, \mathbb{K}) = \{L: V \rightarrow \mathbb{K} \mid L \text{ ist } \mathbb{K}\text{-linear}\}$$

heißt **Dualraum** zu V oder auch Raum der **Linearformen**.

- Hat V endliche Dimension n , so auch V^* . Ist (b_1, \dots, b_n) eine Basis von V , so ist b_1^*, \dots, b_n^* definiert durch

$$b_i^*(b_j) = \delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases},$$

eine Basis von V^* , die wir die zu (b_1, \dots, b_n) **duale Basis** nennen.

492 / 862

Bi- und Multilinearformen

Definition

Sei V ein \mathbb{K} -Vektorraum.

- Eine Abbildung $\beta: V \times V \rightarrow \mathbb{K}$ heißt **Bilinearform**, wenn sie linear in beiden Argumenten ist, d.h. wenn für alle $x, y, z \in V$ und $a, b \in \mathbb{K}$ gilt:

$$\begin{aligned}\beta(ax + by, z) &= a\beta(x, z) + b\beta(y, z) \text{ und} \\ \beta(z, ax + by) &= a\beta(z, x) + b\beta(z, y)\end{aligned}$$

- Eine Abbildung

$$\begin{aligned}\mu: V^k = V \times \dots \times V &\rightarrow \mathbb{K} \\ (v_1, \dots, v_k) &\mapsto \mu(v_1, \dots, v_k)\end{aligned}$$

heißt **Multilinearform** (genauer **k -Linearform**), wenn μ in jedem der k Argumente linear ist.

493 / 862

Beispiele

- Linearformen sind 1-Linearformen
Beispielsweise ist $f \mapsto \int_a^b f(x) dx$ eine Linearform auf dem \mathbb{R} -Vektorraum der integrierbaren Funktionen auf dem Intervall $[a, b]$.
- Sei $V = C^0([a, b])$ der \mathbb{R} -Vektorraum der stetigen Funktionen auf dem Intervall $[a, b]$. Dann definiert $(f, g) \mapsto \int_a^b f(x) \cdot g(x) dx$ eine Bilinearform auf V .
- Die Determinante $\det: \text{Mat}(n, \mathbb{K}) \cong (\mathbb{K}^n)^n \rightarrow \mathbb{K}$ ist n -linear.
- Sei $A \in \text{Mat}(n, \mathbb{K})$. Die Abbildung

$$\begin{aligned}\alpha: \mathbb{K}^n \times \mathbb{K}^n &\rightarrow \mathbb{K} \\ (v_1, v_2) &\mapsto \alpha(v_1, v_2) := v_1^T \cdot A \cdot v_2\end{aligned}$$

ist eine Bilinearform auf \mathbb{K}^n .

Wir werden sehen, dass sich jede Bilinearform auf \mathbb{K}^n in dieser Form schreiben lässt.

494 / 862

Symmetrische und schiefsymmetrische Bilinearformen

Definition

Eine Bilinearform $\beta: V \times V \rightarrow \mathbb{K}$ heißt

- **symmetrisch**, wenn

$$\beta(v, w) = \beta(w, v) \quad \text{für alle } v, w \in V.$$

- **schiefsymmetrisch**, wenn

$$\beta(v, w) = -\beta(w, v) \quad \text{für alle } v, w \in V.$$

- **nicht entartet**, wenn die lineare Abbildung

$$\begin{aligned} V &\rightarrow V^* \\ v &\mapsto \beta(v, \cdot) \end{aligned}$$

injektiv ist, d.h. wenn $v = 0$ der einzige Vektor ist, für den $\beta(v, w) = 0$ für alle $w \in V$ gilt.

495 / 862

Definition

Sei $\dim V = n$, $\alpha: V \times V \rightarrow \mathbb{K}$ eine Bilinearform und $B = (b_1, \dots, b_n)$ eine Basis von V . Als **darstellende Matrix** von α bezüglich B bezeichnen wir die Matrix $A = (\alpha_{ij})_{1 \leq i, j \leq n}$ mit den Einträgen

$$\alpha_{ij} = \alpha(b_i, b_j).$$

Bemerkungen/ÜA

- Für Vektoren $v, w \in V$ mit $v = \sum_i v_i b_i$ und $w = \sum_j w_j b_j$ gilt

$$\alpha(v, w) = \alpha \left(\sum_i v_i b_i, \sum_j w_j b_j \right) = \sum_{i,j} v_i \alpha(b_i, b_j) w_j.$$

Fassen wir $\mathbf{v} = (v_i)$ und $\mathbf{w} = (w_j)$ als Spaltenvektoren auf, so erhalten wir $\alpha(v, w) = \mathbf{v}^T \cdot A \cdot \mathbf{w}$, wobei A die darstellende Matrix von α ist.

- Die Bilinearform α ist genau dann nicht entartet, wenn ihre darstellende Matrix invertierbar ist.

496 / 862

Bemerkungen/ÜA

- Die Bilinearform α ist symmetrisch (bzw. schief-symmetrisch) genau dann, wenn ihre darstellende Matrix A symmetrisch (bzw. schief-symmetrisch) ist, d.h. wenn $A^T = A$ (bzw. $A^T = -A$). Der Raum der symmetrischen (bzw. schief-symmetrischen) Bilinearformen bildet einen Vektorraum über \mathbb{K} der Dimension $\frac{n(n+1)}{2}$ (bzw. $\frac{n(n-1)}{2}$).

- Die darstellende Matrix einer Bilinearform hat folgende Verallgemeinerung:

Ist V ein \mathbb{K} -Vektorraum mit $\dim V = n$ und (b_1, \dots, b_n) eine Basis, so ist eine k -Linearform $\mu : V \times \dots \times V \rightarrow \mathbb{K}$ auf V durch die n^k Werte

$$\mu_{i_1, \dots, i_k} := \mu(b_{i_1}, \dots, b_{i_k}) \quad \text{mit } 1 \leq i_1, \dots, i_k \leq n$$

eindeutig bestimmt.

497 / 862

Beispiel

Auf dem Vektorraum $V = \mathbb{R}^2$ betrachten wir die (schief-symmetrische und nicht entartete) Bilinearform ω mit der darstellenden Matrix

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

Für Vektoren $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ und $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$ gilt dann

$$\begin{aligned} \omega(\mathbf{v}, \mathbf{w}) &= (v_1 \quad v_2) \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \\ &= (v_1 \quad v_2) \cdot \begin{pmatrix} w_2 \\ -w_1 \end{pmatrix} \\ &= v_1 w_2 - v_2 w_1. \end{aligned}$$

498 / 862

Orthogonale Unterräume

Satz

Sei β eine **nicht entartete** Bilinearform auf einem endlichdimensionalen Vektorraum V und $U \subseteq V$ ein Unterraum.

- (i) Der Unterraum $U^{\perp\beta} := \{v \in V \mid \beta(u, v) = 0 \ \forall u \in U\}$ hat die Dimension $\dim U^{\perp\beta} = \dim V - \dim U$.
- (ii) Die Unterräume U und $U^{\perp\beta}$ sind genau dann komplementär, d.h. $V = U \oplus U^{\perp\beta}$, wenn $U \cap U^{\perp\beta} = 0$.

Wir nennen $U^{\perp\beta}$ den β -orthogonalen Unterraum zu U .

Beispiel

Für die im letzten Beispiel betrachtete Bilinearform ω auf \mathbb{R}^2 und jeden Vektor $\mathbf{v} \neq 0$ gilt $\omega(\mathbf{v}, \mathbf{v}) = 0$ und somit $(\mathbb{R}\mathbf{v})^{\perp\omega} = \mathbb{R}\mathbf{v}$. Der ω -orthogonale Unterraum ist also für *keinen* 1-dimensionalen Unterraum in \mathbb{R}^2 ein lineares Komplement.

499 / 862

Beweis des Satzes:

Sei (u_1, \dots, u_r) eine Basis von U , ausgedehnt zu einer Basis (u_1, \dots, u_n) von V . Der Unterraum $U^{\perp\beta}$ ist genau der Lösungsraum des linearen Gleichungssystems

$$\beta(u_i, \sum_{j=1}^n x_j u_j) = \sum_{j=1}^n \beta(u_i, u_j) x_j = \sum_{j=1}^n \beta_{ij} x_j = 0, \quad \text{für alle } i = 1, \dots, r.$$

Da β nicht entartet ist, hat die darstellende Matrix $(\beta_{ij})_{i,j}$ vollen Rang n . Die Matrix des Gleichungssystems besteht aus den ersten r Zeilen dieser darstellenden Matrix, hat also Rang $r = \dim U$. Damit ist die Dimension des Lösungsraumes $U^{\perp\beta}$ gleich $\dim V - r = \dim V - \dim U$.

(ii) folgt aus (i) und der Dimensionsformel:

$$\begin{aligned} \dim(U + U^{\perp\beta}) &= \dim U + \dim U^{\perp\beta} - \dim(U \cap U^{\perp\beta}) \\ &\stackrel{(i)}{=} \dim V - \dim(U \cap U^{\perp\beta}). \end{aligned}$$



500 / 862

Euklidische Vektorräume

Definition (Euklidisches Skalarprodukt/Euklidischer Vektorraum)

Sei V ein **reeller** Vektorraum. Eine **symmetrische Bilinearform** $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ heißt **positiv definit**, wenn für alle $v \in V$, $v \neq 0$ gilt

$$\langle v, v \rangle > 0.$$

Ein **(Euklidisches) Skalarprodukt** auf V ist eine positiv definite symmetrische Bilinearform $V \times V \rightarrow \mathbb{R}$.

Ein **Euklidischer Vektorraum** ist ein reeller Vektorraum, zusammen mit einem (Euklidischen) Skalarprodukt.

Bemerkung

Skalarprodukte sind nicht entartet, denn aus $\langle v, v \rangle = 0$ folgt $v = 0$.

501 / 862

Beispiel

Das **kanonische Skalarprodukt** auf \mathbb{R}^n ist definiert durch

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^n x_i y_i = \mathbf{x}^T \cdot \mathbf{y},$$

wobei $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ und $\mathbf{y} = \sum_{i=1}^n y_i \mathbf{e}_i$. Die darstellende Matrix des kanonischen Skalarprodukts von \mathbb{R}^n bezüglich der kanonischen Basis ist die Einheitsmatrix $\mathbf{1}_n$, denn

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} := \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{sonst.} \end{cases}$$

502 / 862

Geometrische Größen 1: Länge und Abstand

Definition

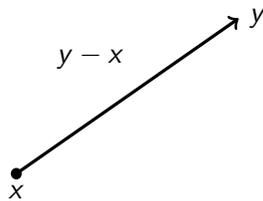
Sei V ein Euklidischer Vektorraum.

- Die **Länge** (auch **Norm**) eines Vektors $v \in V$ ist die nicht-negative Zahl

$$\|v\| := \sqrt{\langle v, v \rangle}.$$

- Der **Abstand** $d(x, y)$ zweier Punkte $x, y \in V$ ist die Länge des Vektors $y - x$:

$$d(x, y) := \|y - x\|.$$



503 / 862

Die Cauchy-Schwarzsche Ungleichung

Die folgende Ungleichung ist fundamental.

Satz (Cauchy-Schwarzsche Ungleichung)

Sei V ein Euklidischer Vektorraum. Dann gilt für alle $v, w \in V$:

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|,$$

mit Gleichheit genau dann, wenn v und w linear abhängig sind.

504 / 862

Beweis. Wir können annehmen, dass $v \neq 0$ und $w \neq 0$ gilt, denn sonst ist nichts zu zeigen.

Für alle $\lambda, \mu \in \mathbb{R}$ gilt wegen der Bilinearität

$$0 \leq \langle \lambda v + \mu w, \lambda v + \mu w \rangle = \lambda^2 \|v\|^2 + \mu^2 \|w\|^2 + 2\lambda\mu \langle v, w \rangle.$$

Einsetzen von $\lambda = \|w\|$ und $\mu = \pm \|v\|$ liefert

$$0 \leq 2\|v\|^2 \|w\|^2 \pm 2\|v\| \|w\| \langle v, w \rangle,$$

d.h. $\mp \langle v, w \rangle \leq \|v\| \|w\|$, und damit die Cauchy-Schwarzsche Ungleichung. Gleichheit gilt genau dann, wenn für eine dieser beiden Wahlen von λ und μ die Beziehung $\lambda v + \mu w = 0$ gilt, d.h. wenn v und w linear abhängig sind. \square

505 / 862

Geometrische Größen 2: Winkel und Orthogonalität

Definition

Sei V ein Euklidischer Vektorraum und $v, w \in V \setminus \{0\}$.

- Der **Winkel** $\angle(v, w)$ zwischen v und w ist definiert als

$$\angle(v, w) := \arccos \underbrace{\frac{\langle v, w \rangle}{\|v\| \|w\|}}_{\in [-1, 1]} \in [0, \pi].$$

- Man sagt, dass $v, w \in V$ **senkrecht** aufeinander stehen oder **orthogonal** sind, wenn

$$\langle v, w \rangle = 0$$

und schreibt dafür $v \perp w$.

- In Euklidischen Vektorräumen schreibt man für den zu $U \subseteq V$ orthogonale Unterraum einfach U^\perp .

506 / 862

Lemma

Sei $U \subseteq V$ ein Unterraum eines endlichdimensionalen Euklidischen Vektorraums. Dann ist U^\perp ein lineares Komplement zu U in V , d.h. $U \cap U^\perp = \{0\}$ und $V = U \oplus U^\perp$.

Beweis. Nach einem bereits bewiesenen Satz sind die beiden Aussagen der Behauptung äquivalent.

Ist aber $U \subseteq V$ ein Unterraum und $v \in U \cap U^\perp$, so gilt $\langle v, v \rangle = 0$.

Da das Skalarprodukt positiv definit ist, folgt hieraus $v = 0$, also

$U \cap U^\perp = \{0\}$. □

Folgerung

Ist $U \subseteq V$ ein Unterraum eines endlichdimensionalen Euklidischen Vektorraums, so bildet die kanonische Projektion $\pi : V \rightarrow V/U$ das orthogonale Komplement U^\perp von U bijektiv auf V/U ab.

Beweis. Die Aussage gilt für jedes lineare Komplement von U . □

507 / 862

Orthonormale Basen

In Euklidischen Vektorräumen gibt es eine ausgezeichnete Klasse von Basen.

Definition

Sei $(V, \langle \cdot, \cdot \rangle)$ ein n -dimensionaler Euklidischer Vektorraum.

- (i) Eine Vektor $v \in V$ heißt **Einheitsvektor**, wenn $\|v\| = 1$.
- (ii) Eine Basis (w_1, \dots, w_n) von V aus n Einheitsvektoren $w_i \in V$, die paarweise orthogonal zueinander sind, heißt **Orthonormalbasis** (kurz: **ONB**).

Für eine Orthonormalbasis (w_1, \dots, w_n) gilt also $\langle w_i, w_j \rangle = \delta_{ij}$.

Beispiel

Die kanonische Basis von \mathbb{R}^n ist orthonormal bzgl. des kanonischen Skalarproduktes.

508 / 862

Lemma

Die Komponenten eines Vektors $v \in V$ in einer Orthonormalbasis (w_1, \dots, w_n) sind gegeben durch $\langle v, w_i \rangle \in \mathbb{R}$, d.h.

$$v = \sum_{i=1}^n \langle v, w_i \rangle w_i.$$

Beweis. Da (w_1, \dots, w_n) eine Basis ist, lässt sich v (eindeutig) als Linearkombination schreiben,

$$v = \sum_{k=1}^n \lambda_k w_k \quad \text{mit } \lambda_k \in \mathbb{R}.$$

Wir nehmen das Skalarprodukt mit w_i :

$$\langle v, w_i \rangle = \sum_{k=1}^n \lambda_k \langle w_k, w_i \rangle = \sum_{k=1}^n \lambda_k \delta_{ki} = \lambda_i$$

□

509 / 862

Satz

Sei V ein Euklidischer Vektorraum und $(v_i)_{i \in I}$ eine Familie von orthogonalen Vektoren $v_i \in V$ mit $v_i \neq 0$, d.h.

$$v_i \perp v_j \quad \text{für alle } i \neq j.$$

Dann ist die Familie $(v_i)_{i \in I}$ linear unabhängig.

Insbesondere ist in einem n -dimensionalen Euklidischen Vektorraum jede Familie aus n orthogonalen Einheitsvektoren eine Orthonormalbasis.

Beweis.

Sei $0 = \sum_{j \in J} \lambda_j v_j$, wobei $J \subseteq I$ eine endliche Teilmenge ist und $\lambda_j \in \mathbb{R}$. Daraus folgt für alle $i \in J$

$$0 = \left\langle \sum_{j \in J} \lambda_j v_j, v_i \right\rangle = \sum_{j \in J} \lambda_j \langle v_j, v_i \rangle = \lambda_i \|v_i\|^2$$

und somit $\lambda_i = 0$.

□

510 / 862

Orthonormalisierungsverfahren

Satz (Gram-Schmidt)

Jeder endlich-dimensionale Euklidische Vektorraum besitzt eine Orthonormalbasis. (Diese ist natürlich nicht eindeutig.)

Beweis. Wir beweisen den Satz durch Induktion nach $n = \dim V \in \mathbb{N}$.

- Falls $\dim V = 1$, so existiert $v \neq 0$ und $w_1 := \frac{v}{\|v\|}$ ist eine ONB.
- Wir nehmen an, dass jeder Euklidische Vektorraum der Dimension $< n$ eine ONB hat und zeigen, dass dann auch jeder n -dimensionale Euklidische Vektorraum eine ONB hat.
- Sei $v \in V$, $v \neq 0$. Nach Induktion besitzt der $(n-1)$ -dimensionale Unterraum $v^\perp := (\mathbb{R}v)^\perp$ eine ONB (w_2, \dots, w_n) . Durch $w_1 := \frac{v}{\|v\|}$ wird diese zu einer ONB von V ergänzt.

□

511 / 862

Gram-Schmidt Verfahren

Eine beliebige Basis der Form (v_1, \dots, v_n) kann man algorithmisch zu einer ONB umformen:

Wir setzen $w_1 = \frac{v_1}{\|v_1\|}$. Iterativ nehmen wir an, dass wir schon Vektoren w_1, \dots, w_k konstruiert haben mit $\langle w_i, w_j \rangle = \delta_{ij}$. Ist (w_1, \dots, w_k) schon eine Basis, dann sind wir fertig. Andernfalls definieren wir die **senkrechte Projektion von v_{k+1} auf den von den Vektoren w_1, \dots, w_k aufgespannten Raum**:

$$\tilde{v}_{k+1} = \langle v_{k+1}, w_1 \rangle w_1 + \dots + \langle v_{k+1}, w_k \rangle w_k$$

und bilden: $\tilde{w}_{k+1} = v_{k+1} - \tilde{v}_{k+1}$. Es gilt

$$\langle \tilde{w}_{k+1}, w_i \rangle = \langle v_{k+1}, w_i \rangle - \langle v_{k+1}, w_i \rangle \langle w_i, w_i \rangle = 0 \text{ für } 1 \leq i \leq k.$$

Der Vektor $w_{k+1} = \frac{\tilde{w}_{k+1}}{\|\tilde{w}_{k+1}\|}$ erfüllt dann $\langle w_{k+1}, w_j \rangle = \delta_{k+1,j}$. Nach n Schritten erhalten wir eine ONB (w_1, \dots, w_n) .

512 / 862

Beispiel

Seien $v_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = e_1$ in \mathbb{R}^3 . Bezüglich des Standardskalarprodukts $\langle \cdot, \cdot \rangle$ erhalten wir:

$$\|v_1\|^2 = 9, \quad \text{also} \quad w_1 = \frac{1}{\sqrt{9}}v_1 = \frac{1}{3}v_1,$$

$$\tilde{w}_2 = e_1 - \langle e_1, w_1 \rangle w_1 = e_1 - \frac{1}{9}v_1 = \frac{1}{9}(8e_1 - 2e_2 - 2e_3),$$

$$w_2 = \frac{1}{\sqrt{\langle \tilde{w}_2, \tilde{w}_2 \rangle}} \tilde{w}_2 = \frac{1}{6\sqrt{2}}(8e_1 - 2e_2 - 2e_3).$$

(w_1, w_2) ist also eine Orthonormalbasis für $\text{span}(v_1, v_2) \subseteq \mathbb{R}^3$.

Frage: Welche Orthonormalbasis erhält man, wenn man statt mit (v_1, v_2) mit (v_2, v_1) startet?

513 / 862

Die orthogonale Gruppe

Definition

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer Vektorraum.

Eine surjektive lineare Abbildung $F: V \rightarrow V$ heißt **orthogonal**, wenn

$$\langle F(v), F(w) \rangle = \langle v, w \rangle \quad \text{für alle } v, w \in V.$$

Bemerkung

- Aus der Gleichung folgt, dass F injektiv ist: aus $F(v) = 0$ folgt $0 = \langle F(v), F(v) \rangle = \langle v, v \rangle$, also $v = 0$. Falls $\dim V < \infty$, so folgt aus Injektivität die Surjektivität. Diese muss also bei endlich dimensionalen Vektorräumen nicht gefordert werden.
- Die orthogonalen Abbildungen bilden eine Untergruppe $O(V) \subseteq \text{Aut}(V)$, die sogenannte **orthogonale Gruppe**.
- Orthogonale Abbildungen erhalten Winkel, Längen und Abstände. Sie setzen sich aus Drehungen und Spiegelungen zusammen.

514 / 862

Die orthogonale Gruppe des \mathbb{R}^n

Beispiel

Wir betrachten den \mathbb{R}^n mit dem kanonischen Skalarprodukt $\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{i=1}^n x_i y_i$. Dann definiert man die **orthogonale Gruppe**

$$O(n) := O(\mathbb{R}^n, \langle \cdot, \cdot \rangle) \subseteq GL(n, \mathbb{R})$$

als Untergruppe der Gruppe der invertierbaren Matrizen. Sei $A = (a_{ij}) \in \text{Mat}(n, \mathbb{R})$ und e_i die kanonische Basis. Dann ist $A \in O(n) \iff$

$$\delta_{ij} = \langle e_i, e_j \rangle \stackrel{!}{=} \langle Ae_i, Ae_j \rangle = \sum_{k,l} a_{ki} a_{lj} \langle e_k, e_l \rangle = \sum_{k,l} a_{ki} a_{lj} \delta_{kl} = \sum_k a_{ki} a_{kj}$$

für alle $i, j = 1, \dots, n$. Dies ist aber äquivalent zu $\mathbf{1}_n = A^T A$. Es gilt also für $A \in \text{Mat}(n, \mathbb{R})$

$$A \in O(n) \iff A^T A = \mathbf{1}_n \iff A^T = A^{-1}.$$

515 / 862

Hermitesche Vektorräume

Um auf **komplexen** Vektorräumen geometrische Größen einführen zu können, benötigt man den Begriff eines Hermiteschen Skalarprodukts.

Definition

Sei V ein **komplexer** Vektorraum. Eine **Hermitesche Form** auf V ist eine Abbildung $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{C}$, die \mathbb{C} -linear im ersten Argument ist, und die für alle $v, w \in V$ folgende Bedingung erfüllt:

$$\overline{\langle v, w \rangle} = \langle w, v \rangle.$$

Bemerkung

Daraus folgt, dass $\langle \cdot, \cdot \rangle$ **konjugiert-linear** im zweiten Argument ist, d.h.

$$\langle u, \lambda v + \mu w \rangle = \bar{\lambda} \langle u, v \rangle + \bar{\mu} \langle u, w \rangle \quad \text{für alle } u, v, w \in V, \lambda, \mu \in \mathbb{C},$$

denn $\langle u, \lambda v \rangle = \overline{\langle \lambda v, u \rangle} = \bar{\lambda} \overline{\langle v, u \rangle} = \bar{\lambda} \langle u, v \rangle$.

Außerdem ist $\langle v, v \rangle$ für alle $v \in V$ reell (warum?).

516 / 862

Bemerkung

Die Menge der Hermiteschen Formen eines komplexen Vektorraumes ist ein **reeller** (kein komplexer!) Vektorraum, denn: Ist β eine Hermitesche Form und $\lambda \in \mathbb{C}$, so gilt

$$(\lambda\beta)(u, v) = \lambda\beta(u, v) = \lambda\overline{\beta(v, u)},$$

d.h. $(\lambda\beta)(u, v) = \overline{(\lambda\beta)(v, u)} = \bar{\lambda} \cdot \overline{\beta(v, u)}$, falls $\lambda = \bar{\lambda}$ reell ist.

Definition

Eine Hermitesche Form $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{C}$ heißt **positiv definit**, wenn

$$\langle v, v \rangle > 0 \quad \text{für alle } v \in V \setminus 0.$$

Ein **(Hermitesches) Skalarprodukt** auf V ist eine positiv definite Hermitesche Form $V \times V \rightarrow \mathbb{C}$. Ein **Hermitescher Vektorraum** ist ein komplexer Vektorraum zusammen mit einem Hermiteschen Skalarprodukt. (Hermitesche Vektorräume heißen auch **unitäre** Vektorräume.)

517 / 862

Beispiel

- Das **kanonische (Hermitesche) Skalarprodukt** auf \mathbb{C}^n ist definiert durch

$$\langle z, w \rangle := \sum_{i=1}^n z_i \bar{w}_i,$$

wobei $z = \sum_{i=1}^n z_i e_i$ und $w = \sum_{i=1}^n w_i e_i$.

- Sei $V = \mathbb{R}^n$. Dann definiert jedes Euklidische Skalarprodukt $\langle \cdot, \cdot \rangle$ ein Hermitesches Skalarprodukt $\langle \cdot, \cdot \rangle^H$ auf \mathbb{C}^n mittels komplex linearer und konjugiert linearer Erweiterung. D.h. wir definieren $\langle \cdot, \cdot \rangle^H$ durch

$$\langle x + iy, u + iv \rangle^H := \langle x, u \rangle + \langle y, v \rangle + i(\langle y, u \rangle - \langle x, v \rangle)$$

für alle $x, y, u, v \in \mathbb{R}^n$.

ÜA: $\langle \cdot, \cdot \rangle^H$ ist in der Tat ein Hermitesches Skalarprodukt.

518 / 862

Bemerkung

Sei V ein Hermitescher Vektorraum.

- Auf V definiert man wieder die Länge eines Vektors durch

$$\|v\| := \sqrt{\langle v, v \rangle}, \quad v \in V.$$

- Ist $\langle \cdot, \cdot \rangle$ ein Hermitesches Skalarprodukt, so definiert sein Realteil $g(u, v) := \operatorname{Re}\langle u, v \rangle$ ein Euklidisches Skalarprodukt auf V , aufgefasst als reeller Vektorraum (ÜA).

Satz (Cauchy-Schwarzsche Ungleichung)

Sei $h = \langle \cdot, \cdot \rangle$ ein Hermitesches Skalarprodukt auf einem \mathbb{C} -Vektorraum V .

- (i) Dann gilt für alle $v, w \in V$:

$$(*) \quad |\langle v, w \rangle| \leq \|v\| \cdot \|w\|.$$

- (ii) Gleichheit in $(*)$ gilt genau dann, wenn v und w linear abhängig sind.

519 / 862

Beweis.

- (i) Wir können annehmen, dass $\langle v, w \rangle \neq 0$ gilt, und setzen

$$\lambda := \frac{\overline{\langle v, w \rangle}}{|\langle v, w \rangle|} \in \mathbb{C}.$$

Wegen $|\langle v, w \rangle|^2 = \langle v, w \rangle \overline{\langle v, w \rangle}$ gilt dann

$$(1) \quad \langle \lambda v, w \rangle = \lambda \langle v, w \rangle = \frac{\overline{\langle v, w \rangle} \langle v, w \rangle}{|\langle v, w \rangle|} = |\langle v, w \rangle|.$$

Die Cauchy-Schwarzsche Ungleichung für das *Euklidische* Skalarprodukt $g := \operatorname{Re}(h)$ liefert weiter

$$(2) \quad \langle \lambda v, w \rangle = g(\lambda v, w) \leq \sqrt{g(\lambda v, \lambda v)} \sqrt{g(w, w)} = \|v\| \cdot \|w\|,$$

woraus mit (1) die Behauptung (i) folgt.

- (ii) Gleichheit in (2) und somit in $(*)$ gilt genau dann, wenn λv und w linear abhängig über \mathbb{R} und somit v und w linear abhängig über \mathbb{C} sind.



520 / 862

Unitäre Basen

Wieder gibt es eine ausgezeichnete Klasse von Basen:

Definition (Unitäre Basen)

Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler Hermitescher Vektorraum. Eine **unitäre** Basis von V ist eine Basis (w_1, \dots, w_n) mit

$$\langle w_i, w_j \rangle = \delta_{ij}.$$

(Ist der Bezug auf das Hermitesche Skalarprodukt klar, so kann man auch wieder von einer ONB sprechen.)

Die Komponenten eines Vektors $v \in V$ in einer unitären Basis (w_1, \dots, w_n) sind wieder gegeben durch $\langle v, w_i \rangle$, d.h. $v = \sum_{i=1}^n \langle v, w_i \rangle w_i$.

Satz

Jeder endlichdimensionale Hermitesche Vektorraum besitzt eine unitäre Basis.

Beweis. Übungsaufgabe (wie für den Satz von Gram-Schmidt)

521 / 862

Die unitäre Gruppe

Definition (Unitäre Gruppe)

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Hermitescher Vektorraum.

Ein surjektiver Endomorphismus $F: V \rightarrow V$ heißt **unitär**, wenn

$$\langle F(v), F(w) \rangle = \langle v, w \rangle \quad \text{für alle } v, w \in V.$$

(Falls $\dim V < \infty$, braucht die Surjektivität nicht gefordert zu werden.)
Die unitären Endomorphismen bilden eine Untergruppe $U(V) \subseteq \text{Aut}(V)$, die sogenannte **unitäre Gruppe** (ÜA).

Beispiel

Für $U(n) := U(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$, $\langle z, w \rangle := \sum_{i=1}^n z_i \bar{w}_i$:

$A \in U(n) \iff \bar{A}^T A = \mathbf{1}_n$. Man schreibt auch $A^\dagger := \bar{A}^T$.

ÜA: Betrachtet man \mathbb{C}^n als reellen Vektorraum \mathbb{R}^{2n} , so kann $U(n)$ als Untergruppe von $O(2n)$ aufgefasst werden. Andererseits gilt $O(n) \subseteq U(n)$.

522 / 862

Normalformen von Endomorphismen

Für bestimmte Klassen von Endomorphismen $F: V \rightarrow V$ existiert eine **Normalform**. Damit ist gemeint, dass man stets eine Basis von V finden kann, in der die darstellende Matrix Normalform – z.B. Diagonalgestalt – hat (und dass “echt” verschiedene Endomorphismen in Normalform auch verschieden aussehen).

Sei z. B. $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Abbildung mit der darstellenden Matrix

$$A = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}.$$

bezüglich der kanonischen Basis (e_1, e_2) . Dann ist $b_1 = e_1 - e_2$ ein Eigenvektor zum Eigenwert $\frac{1}{2}$, und $b_2 = e_1 + e_2$ zum Eigenwert $\frac{3}{2}$. D.h. in der Basis (b_1, b_2) hat F die einfachere Diagonalgestalt

$$B = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{3}{2} \end{pmatrix}$$

Wir werden insbesondere Normalformen für orthogonale und unitäre Endomorphismen finden.

523 / 862

Satz (Normalform unitärer Endomorphismen)

Sei V ein endlichdimensionaler Hermitescher Vektorraum und $F \in \mathcal{U}(V)$ ein **unitärer** Endomorphismus.

- (i) Die Eigenwerte von F sind komplexe Zahlen vom Betrag 1.
- (ii) Eigenvektoren zu verschiedenen Eigenwerten stehen senkrecht aufeinander.
- (iii) Es gibt eine **unitäre** Basis von V bestehend aus Eigenvektoren von F . D.h. F ist **in einer unitären Basis** diagonalisierbar.

Beweis.

- (i) Sei v ein Eigenvektor von F und $\lambda \in \mathbb{C}$ der zugehörige Eigenwert. Aus

$$0 \neq \langle v, v \rangle = \langle Fv, Fv \rangle = \langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle$$

ergibt sich $|\lambda| = 1$.

- (ii) Seien $v, w \in V$ Eigenvektoren von F zu den Eigenwerten λ und μ .

Wegen (i) ist $\lambda^{-1} = \bar{\lambda}$. Daher impliziert $\mu \neq \lambda$, dass $\lambda \bar{\mu} \neq 1$.

Aus $\langle v, w \rangle = \langle Fv, Fw \rangle = \lambda \bar{\mu} \langle v, w \rangle$ folgt daher $\langle v, w \rangle = 0$.

524 / 862

(iii) Beweis durch Induktion nach $n = \dim V$. Der Fall $n = 1$ ist klar.

Sei $\dim V \geq 2$. Nach dem Fundamentalsatz der Algebra hat jedes nicht-konstante komplexe Polynom eine Nullstelle. Insbesondere hat das charakteristische Polynom von F eine Nullstelle und F mithin einen Eigenwert $\lambda_1 \neq 0$.

Wir wählen einen zugehörigen Eigenvektor $b_1 \in V$ der Länge 1 und betrachten den Unterraum $W := b_1^\perp$.

Es gilt $\dim W = n - 1$ und $FW \subseteq W$, denn aus $w \in W$ folgt

$$0 = \langle b_1, w \rangle = \langle Fb_1, Fw \rangle = \lambda_1 \langle b_1, Fw \rangle$$

und somit $Fw \in W$, da $\lambda_1 \neq 0$.

Nach Induktionsvoraussetzung gibt es eine unitäre Basis (b_2, \dots, b_n) von W bestehend aus Eigenvektoren von F , und (b_1, b_2, \dots, b_n) ist die gesuchte unitäre Basis von V . \square

525 / 862

Normalform orthogonaler Endomorphismen

Folgerung (Normalform orthogonaler Endomorphismen)

Sei V ein endlichdimensionaler Euklidischer Vektorraum und $F \in O(V)$ ein orthogonaler Endomorphismus.

- (i) Eigenwerte von F sind reelle Zahlen vom Betrag 1, dh. $= \pm 1$.
- (ii) Eigenvektoren zu verschiedenen Eigenwerten stehen senkrecht aufeinander.
- (iii) Es gibt eine Orthonormalbasis, bezüglich derer F durch eine Blockdiagonalmatrix folgender Art dargestellt wird

$$\text{diag}(\lambda_1, \dots, \lambda_p, D_{\varphi_1}, \dots, D_{\varphi_q}),$$

wobei $\lambda_j \in \{\pm 1\}$, $D_{\varphi_j} = \begin{pmatrix} \cos \varphi_j & -\sin \varphi_j \\ \sin \varphi_j & \cos \varphi_j \end{pmatrix}$ Drehungen um den Winkel $\varphi_j \in \mathbb{R}$ sind, und $p + 2q = \dim V$ gilt.

526 / 862

Beweis.

(i-ii) Gleicher Beweis wie für unitäre Endomorphismen.

(iii) Analog erhalten wir auch eine aus Eigenvektoren bestehende ONB (b_1, \dots, b_p) für die Summe $U := V_1 \oplus V_{-1}$ der Eigenräume $V_{\pm 1}$.

Es sei $U \neq V$. Wieder bildet F das orthogonale Komplement $W := U^\perp$ in sich ab. Außerdem enthält W keine Eigenvektoren von F .

Insbesondere hat W gerade Dimension $m = 2q$, denn sonst hätte das charakteristische Polynom der Einschränkung $F|_W$ als reelles Polynom ungeraden Grades eine reelle Nullstelle.

Es genügt zu zeigen, dass W eine ONB besitzt, bezüglich der die darstellende Matrix von $F|_W$ die Gestalt $\text{diag}(D_{\varphi_1}, \dots, D_{\varphi_q})$ hat.

527 / 862

Weiter im Beweis:

Durch Wahl einer ONB von W können wir annehmen, dass $W = \mathbb{R}^m$ der Euklidische Raum ist und $F|_W$ gegeben ist durch $A \in O(m)$.

Wir können A als komplexe Matrix, d.h. als Endomorphismus des \mathbb{C}^m , auffassen. Dann ist A unitär bezüglich des Hermiteschen Skalarproduktes auf \mathbb{C}^m .

Sei nun μ ein komplexer Eigenwert von A mit Eigenvektor $v = \mathbf{x} + i\mathbf{y}$ mit $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Da A reell ist, gilt

$$Av = \mu v \implies A\bar{v} = \overline{Av} = \overline{\mu v} = \bar{\mu} \cdot \bar{v}.$$

Also ist $\bar{\mu}$ auch ein Eigenwert von A mit Eigenvektor $\bar{v} := \mathbf{x} - i\mathbf{y}$. Die echt komplexen Eigenwerte $\mu_j = e^{i\varphi_j}$ ($\varphi_j \in \mathbb{R}$) treten also in komplex konjugierten Paaren **gleicher Vielfachheit** auf, denn die komplexe Konjugation der Komponenten bildet den Eigenraum zum Eigenwert μ isomorph auf den Eigenraum zum Eigenwert $\bar{\mu}$ ab.

528 / 862

Weiter im Beweis: Nach dem Satz gibt es eine unitäre Basis

$$(\beta_1, \dots, \beta_q, \bar{\beta}_1, \dots, \bar{\beta}_q)$$

von $\mathbb{C}^m = \mathbb{C}^{2q}$ bestehend aus Eigenvektoren von $A \in U(m)$ mit zugehörigen Eigenwerten $(\mu_1, \dots, \mu_q, \bar{\mu}_1, \dots, \bar{\mu}_q)$.

Hierbei ist $\mu_j = e^{i\varphi_j}$ mit $\varphi_j \in \mathbb{R}$, und $\bar{\mu}_j = e^{-i\varphi_j}$. Wir definieren

$$\beta'_j = \frac{1}{\sqrt{2}}(\beta_j + \bar{\beta}_j) \quad \text{und} \quad \beta''_j = \frac{i}{\sqrt{2}}(\beta_j - \bar{\beta}_j) \quad \text{für } j = 1, \dots, q.$$

Diese Vektoren liegen in $\mathbb{R}^m \subset \mathbb{C}^m$ und bilden dort eine ONB

$(\beta'_1, \beta''_1, \dots, \beta'_q, \beta''_q)$. Die darstellende Matrix von A bezüglich dieser Basis ist von der gewünschten Form, d.h. es gilt

$$\begin{aligned} A\beta'_j &= \frac{1}{\sqrt{2}}(e^{i\varphi_j}\beta_j + e^{-i\varphi_j}\bar{\beta}_j) = \cos(\varphi_j)\frac{1}{\sqrt{2}}(\beta_j + \bar{\beta}_j) + \sin(\varphi_j)\frac{i}{\sqrt{2}}(\beta_j - \bar{\beta}_j) \\ &= \cos(\varphi_j)\beta'_j + \sin(\varphi_j)\beta''_j \quad \text{und} \\ A\beta''_j &= -\sin(\varphi_j)\beta'_j + \cos(\varphi_j)\beta''_j \end{aligned}$$



529 / 862

Selbstadjungierte Endomorphismen

Definition (Selbstadjungierte Endomorphismen)

Sei F ein Endomorphismus eines Euklidischen oder Hermiteschen Vektorraumes $(V, \langle \cdot, \cdot \rangle)$.

- Ein Endomorphismus $F^{ad} \in \text{End}(V)$ heißt zu F **adjungiert**, falls
$$\langle F(v), w \rangle = \langle v, F^{ad}(w) \rangle \quad \text{für alle } v, w \in V.$$
- Existiert F^{ad} , so nennt man F **normal**, wenn $F^{ad} \circ F = F \circ F^{ad}$.
Speziell nennt man F **selbstadjungiert**, wenn $F = F^{ad}$.

(Wir untersuchen jetzt die Diagonalisierbarkeit im Fall $F = F^{ad}$, könnten allerdings auch allgemeiner normale Endomorphismen betrachten.)

Satz

Sei F ein Endomorphismus eines Euklidischen oder Hermiteschen Vektorraums V . Dann gilt:

- Es existiert höchstens ein zu F adjungierter Endomorphismus F^{ad} .
- Ist V endlich-dimensional, so existiert F^{ad} .

530 / 862

Beweis. Die Eindeutigkeit folgt daraus, dass $\langle \cdot, \cdot \rangle$ nicht entartet ist:

Seien F_1 und F_2 zu F adjungiert. Dann gilt für alle $u, v \in V$, dass

$$\langle u, F_1(v) - F_2(v) \rangle = \langle F(u), v \rangle - \langle F(u), v \rangle = 0.$$

Damit muss aber gelten, dass $F_1(v) = F_2(v)$ für alle $v \in V$.

Sei nun $\dim(V) = n$, (b_1, \dots, b_n) eine Orthornomal- (bzw. unitäre) Basis und sei $A = (a_{ij})_{i,j=1,\dots,n}$ die darstellende Matrix von F bezüglich dieser Basis, d.h.

$$F(b_j) = \sum_{i=1}^n a_{ij} b_i, \quad \text{oder äquivalent dazu } \langle F(b_j), b_i \rangle = a_{ij}$$

Dann ist F^{ad} gegeben durch die darstellende Matrix $A^\dagger := \overline{A}^T$, denn

$$\langle b_j, F^{ad}(b_i) \rangle = \langle b_j, \sum_{k=1}^n \overline{a}_{ik} b_k \rangle = \sum_{k=1}^n a_{ik} \langle b_j, b_k \rangle = a_{ij},$$

und damit $\langle F(u), v \rangle = \langle u, F^{ad}(v) \rangle$ für alle $u, v \in V$. □

531 / 862

Symmetrische und Hermitesche Matrizen

Folgerung (aus dem Beweis des Satzes)

Sei V ein endlichdimensionaler Euklidischer oder Hermitescher Vektorraum und A die darstellende Matrix eines Endomorphismus F **bezüglich einer ONB**. Dann ist die darstellende Matrix von F^{ad} gegeben durch $A^\dagger := \overline{A}^T$ und es gilt

$$F \text{ selbstadjungiert} \iff A = A^\dagger.$$

Beispiel

Insbesondere gilt für $V = \mathbb{R}^n$ mit dem kanonischen Euklidischen Skalarprodukt $A^{ad} = A^T$, und für $V = \mathbb{C}^n$ mit dem kanonischen Hermiteschen Skalarprodukt gilt $A^{ad} = A^\dagger = \overline{A}^T$.

Matrizen mit $A = A^T$ (und ebenso die zugehörigen Endomorphismen) heißen **symmetrisch**. Analog verwendet man (über \mathbb{C}) den Begriff **Hermitesch**, wenn $A = A^\dagger$.

Kern und Bild von adjungierten Endomorphismen

Erinnerung: Für $F \in \text{Aut}(V)$ gilt

$$\ker(F) = \{v \in V \mid F(v) = 0\}$$

$$\text{im}(F) = \{F(w) \mid w \in V\}$$

Satz

Sei F ein Endomorphismus eines Euklidischen oder Hermiteschen Vektorraumes V . Dann gilt

$$\begin{aligned}\ker(F) &= (\text{im}(F^{ad}))^\perp \\ \ker(F^{ad}) &= (\text{im}(F))^\perp\end{aligned}$$

Beweis. Das folgt unmittelbar aus der definierenden Gleichung

$$\langle F(v), w \rangle = \langle v, F^{ad}(w) \rangle, \quad \text{für alle } v, w \in V.$$



533 / 862

Normalform selbstadjungierter Endomorphismen

Satz (Normalform selbstadjungierter Endomorphismen)

Sei V ein Euklidischer oder Hermitescher Vektorraum und $F \in \text{End}(V)$ selbstadjungiert.

- (i) Die Eigenwerte von F sind reell.
- (ii) Eigenvektoren zu verschiedenen Eigenwerten von F stehen senkrecht aufeinander.
- (iii) Falls $\dim V < \infty$, so besitzt V eine **orthonormale** bzw. **unitäre** Basis bestehend aus Eigenvektoren von F .

Beweis.

(i) Aus $Fv = \lambda v$ und $\|v\| = 1$, folgt $\lambda = \langle Fv, v \rangle = \langle v, Fv \rangle = \bar{\lambda}$.

(ii) v, w seien Eigenvektoren zu verschiedenen Eigenwerten $\lambda, \mu \in \mathbb{R}$.

Dann folgt aus

$$\lambda \langle v, w \rangle = \langle Fv, w \rangle = \langle v, Fw \rangle = \mu \langle v, w \rangle$$

mit $\lambda \neq \mu$, dass $\langle v, w \rangle = 0$.

534 / 862

Weiter im Beweis:

- (iii) Es genügt zu zeigen, dass F einen Eigenvektor v besitzt. F erhält dann das orthogonale Komplement $W := v^\perp$, denn

$$\langle v, w \rangle = 0 \implies \langle v, Fw \rangle = \langle Fv, w \rangle = 0.$$

Somit ist wiederum ein Induktionsbeweis nach der Dimension von V möglich.

Nach dem Fundamentalsatz der Algebra hat das charakteristische Polynom von F eine Nullstelle λ . Im Fall V komplex und hermitesch liefert das die Existenz eines Eigenwertes und eines Eigenvektors.

Im Fall V reell und Euklidisch fixieren wir eine ONB und identifizieren V mit \mathbb{R}^n . Wir können dann F auffassen als selbstadjungierten Endomorphismus von \mathbb{R}^n , aber auch von \mathbb{C}^n . Beide haben dasselbe charakteristische Polynom, welches nach (i) reelle Nullstellen hat und daher in reelle Linearfaktoren zerfällt.

Also hat F auch im Fall eines reellen Vektorraumes V einen Eigenvektor $v \in V$.

535 852

Dualisierung

Satz

Sei $(V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Euklidischer (oder Hermitescher) Vektorraum. Der \mathbb{R} -Vektorraum der symmetrischen Bilinearformen (bzw. der Hermiteschen Formen) auf V ist isomorph zum \mathbb{R} -Vektorraum der selbstadjungierten Endomorphismen von V .

Beweis. Der Isomorphismus ist gegeben durch die Beziehung

$$\beta(u, v) := \langle F(u), v \rangle, \quad \forall u, v \in V.$$

zwischen einem selbstadjungierten Endomorphismus F und einer symmetrischen/Hermiteschen Form β (ÜA). □

Folgerung

Sei V ein \mathbb{C} -Vektorraum mit $\dim(V) = n$ und h eine **Hermitesche Form** auf V .

- (i) Dann existiert eine Basis von V , bezüglich derer die darstellende Matrix von h folgende Gestalt hat:

$$\text{diag}(\mathbf{1}_{r_+}, -\mathbf{1}_{r_-}, 0_{r_0})$$

mit $r_+ + r_- + r_0 = n$ und $0_{r_0} \in \text{Mat}_{r_0}(\mathbb{C})$ die $r_0 \times r_0$ -Nullmatrix.

- (ii) Die Zahlen r_+ und r_- hängen nicht von der Wahl der Basis ab.

Beweis. (i) Wähle ein Hermitesches **Skalarprodukt** $\langle \cdot, \cdot \rangle$ auf V .

Dann entspricht der Hermiteschen Form h ein selbstadjungierter Endomorphismus F . Dieser kann in Normalform gebracht werden mit reellen Eigenwerten λ_i und einer unitären Basis (u_1, \dots, u_n) aus Eigenvektoren von F .

Es gebe r_+ positive und r_- negative Eigenwerte. Dann ist $\dim(\ker(F)) = n - r_+ - r_-$.

537 / 862

Für $\lambda_i \neq 0$ setze $\tilde{u}_i := \frac{1}{\sqrt{|\lambda_i|}} u_i$. Dann ist für $\lambda_i \neq 0$

$$h(\tilde{u}_i, \tilde{u}_j) = \langle F \tilde{u}_i, \tilde{u}_j \rangle = \frac{\lambda_i}{|\lambda_i|} \langle u_i, u_j \rangle = \pm 1 \delta_{i,j}$$

und die Basis hat die gewünschte Eigenschaft.

Beweis von (ii): Aus

$$V_0 = \ker F = \{v \mid h(v, w) = 0 \text{ für alle } w \in V\}$$

folgt die Unabhängigkeit von $r_0 = \dim V_0$ von allen Wahlen.

Um zu zeigen, dass etwa r_+ unabhängig von allen Wahlen ist, zeigen wir

$$r_+ = \max\{\dim W \mid W \subseteq V, h \text{ ist auf } W \text{ positiv-definit}\}.$$

Dazu sei V_+ bzw. V_- die Summe der Eigenräume zu den positiven bzw. negativen Eigenwerten. Die Form h ist auf V_+ positiv definit und auf dem Komplement $V_- \oplus V_0$ negativ semi-definit, d.h. $h(v, v) \leq 0$ für alle $v \in V_- \oplus V_0$.

Ist also $W \subseteq V$ ein Unterraum, auf dem h positiv definit ist, dann gilt $W \cap (V_- \oplus V_0) = 0$ und somit $\dim W \leq r_+$. □

538 / 862

Hauptachsentransformation

Folgerung (Hauptachsentransformation)

Sei V ein endlichdimensionaler **Euklidischer Vektorraum** und β eine symmetrische Bilinearform auf V .

- (i) Dann existiert eine **orthogonale** Basis (b_1, \dots, b_n) von V , bezüglich derer die darstellende Matrix von β folgende Gestalt hat:

$$\text{diag}(\mathbf{1}_{r_+}, -\mathbf{1}_{r_-}, 0_{r_0}).$$

(Die Ursprungsgeraden $\mathbb{R}b_i$ heißen **Hauptachsen** von β .)

- (ii) Die Zahlen r_+ , r_- und r_0 hängen nicht von der Wahl der Basis ab.

Beweis. Analog zum Beweis des vorhergehenden Satzes. □

Aussage (i) ist bekannt als Satz von der **Hauptachsentransformation**, (ii) als **Trägheitssatz von Sylvester**.

539 / 862

Beispiele

1) Sei $V = \mathbb{R}^2$ mit kanonischen Skalarprodukt $\langle \cdot, \cdot \rangle$.

Dann ist $\{v \in \mathbb{R}^2 \mid \langle v, v \rangle = 1\}$ gleich dem Kreis mit Radius 1.

2) Nun betrachten wir die symmetrische Bilinearform

$$\beta \left(\begin{pmatrix} x \\ y \end{pmatrix}, \begin{pmatrix} x' \\ y' \end{pmatrix} \right) = 5xx' + 5yy' + 3xy' + 3yx'.$$

Bezüglich des Standardskalarprodukts $\langle \cdot, \cdot \rangle$ wird die Bilinearform durch die Matrix

$$F := \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

beschrieben. Deren charakteristisches Polynom $P_F(T) = t^2 - 10t + 16$ hat die beiden verschiedenen Eigenwerte 8 und 2 mit normierten orthogonalen Eigenvektoren $u_1 := \frac{1}{\sqrt{2}}(e_1 + e_2)$ und $u_2 := \frac{1}{\sqrt{2}}(e_1 - e_2)$. Für die orthogonale Basis $b_1 = \frac{1}{\sqrt{\lambda_1}}u_1 = \frac{1}{4}(e_1 + e_2)$ und $b_2 = \frac{1}{\sqrt{\lambda_2}}u_2 = \frac{1}{2}(e_1 - e_2)$ gilt $\beta(b_i, b_j) = \delta_{ij}$.

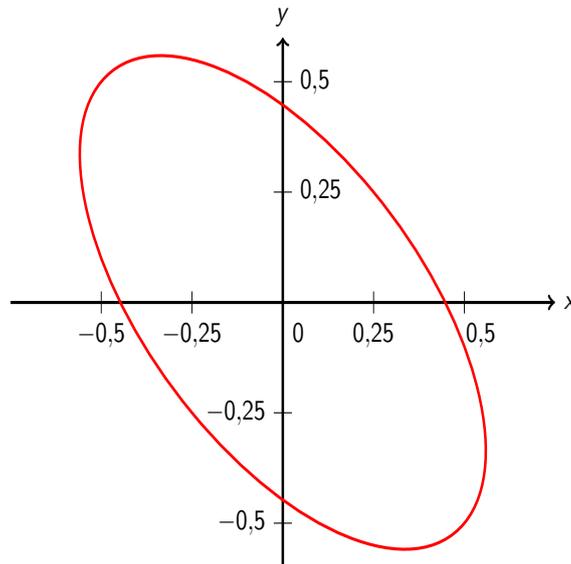
540 / 862

Ellipse

Insbesondere sind die Hauptachsen von β die Diagonalen $y = x$ und $y = -x$. Das sind genau die Hauptachsen der Ellipse

$$E = \left\{ v = \begin{pmatrix} x \\ y \end{pmatrix} \mid \beta(v, v) = 1 \right\} \text{ mit}$$

$$\beta(v, v) = 5x^2 + 5y^2 + 6xy = \left(\frac{x+y}{\frac{1}{2}}\right)^2 + \left(\frac{x-y}{1}\right)^2.$$



541 / 862

Kapitel 16

Metrische Räume und Vollständigkeit

542 / 862

Metrische Räume

Erinnerung:

Metrische Räume sind Mengen, auf denen eine reellwertige Abstandsfunktion definiert ist.

Definition

Sei X eine Menge.

Eine **Metrik** (eine Abstandsfunktion) auf X ist eine Abbildung

$$d: X \times X \rightarrow [0, \infty)$$

mit folgenden Eigenschaften:

- M1) Für $x, y \in X$ gilt genau dann $d(x, y) = 0$, wenn $x = y$.
- M2) $d(x, y) = d(y, x)$ für alle $x, y \in X$.
- M3) $d(x, z) \leq d(x, y) + d(y, z)$ für alle $x, y, z \in X$ (**Dreiecksungleichung**).

543 / 862

Wichtige Beispielfamilie: Normierte Vektorräume

Definition

Sei V ein \mathbb{K} -Vektorraum, wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Eine **Norm** auf V ist eine Abbildung

$$\|\cdot\|: V \rightarrow [0, \infty), \quad v \mapsto \|v\|,$$

mit folgenden Eigenschaften:

- N1) Für $v \in V$ gilt genau dann $\|v\| = 0$, wenn $v = 0$.
- N2) $\|\lambda v\| = |\lambda| \cdot \|v\|$ für alle $v \in V$, $\lambda \in \mathbb{K}$.
- N3) $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$ (**Dreiecksungleichung**).

Ein **normierter Vektorraum** ist ein reeller oder komplexer Vektorraum zusammen mit einer Norm.

Beispiel(Übungsaufgabe)

Auf $V = \mathbb{R}^n$ wird durch $\|\mathbf{x}\|_{\max} := \max(|x_1|, \dots, |x_n|)$ eine Norm definiert.

544 / 862

Satz

Sei $(V, \langle \cdot, \cdot \rangle)$ ein Euklidischer oder Hermitescher Vektorraum. Dann wird durch

$$\|v\| := \sqrt{\langle v, v \rangle}, \quad v \in V,$$

eine Norm $\|\cdot\|$ auf V definiert. Jeder Euklidische oder Hermitesche Vektorraum ist also insbesondere ein normierter Vektorraum.

Beweis.

N1) folgt daraus, dass $\langle \cdot, \cdot \rangle$ positiv definit ist.

N2) folgt aus den Eigenschaften von $\langle \cdot, \cdot \rangle$:

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle = |\lambda|^2 \|v\|^2.$$

N3) folgt aus der Cauchy-Schwarzschen Ungleichung:

$$\begin{aligned} \|v + w\|^2 &= \|v\|^2 + \|w\|^2 + \langle v, w \rangle + \langle w, v \rangle \\ &= \|v\|^2 + \|w\|^2 + 2 \operatorname{Re} \langle v, w \rangle \\ &\leq \|v\|^2 + \|w\|^2 + 2 |\langle v, w \rangle| \\ &\leq \|v\|^2 + \|w\|^2 + 2 \|v\| \|w\| = (\|v\| + \|w\|)^2. \quad \square \end{aligned}$$

545 / 862

Eine Norm auf einem Vektorraum definiert einen Abstand**Bemerkungen**

- Die **Euklidische Metrik** des \mathbb{R}^n ist gegeben durch:

$$d(x, y) := \sqrt{\sum_{i=1}^n |x_i - y_i|^2}$$

- Allgemeiner wird durch

$$d(x, y) := \|x - y\|, \quad \text{für } x, y \in V$$

auf jedem normierten Vektorraum $(V, \|\cdot\|)$ eine Metrik $d: V \times V \rightarrow \mathbb{R}_{\geq 0}$ definiert. Dabei folgt die Dreiecksungleichung für die Metrik aus der für die Norm:

$$\begin{aligned} d(x, z) &= \|x - z\| = \|x - y + y - z\| \\ &\leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z) \end{aligned}$$

546 / 862

Konvergenz und Stetigkeit in metrischen Räumen

Definition (Erinnerung)

- (i) Sei (X, d) ein metrischer Raum. Eine Folge $x_n \in X$, $n \in \mathbb{N}$, heißt **konvergent** mit **Grenzwert** $x \in X$, in Zeichen $\lim_{n \rightarrow \infty} x_n = x$, wenn die Folge $d(x_n, x) \in \mathbb{R}$ eine Nullfolge ist, d.h. $\lim_{n \rightarrow \infty} d(x_n, x) = 0$.
- (ii) Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f: X \rightarrow Y$ heißt **im Punkt $x \in X$ stetig**, wenn $\lim_{n \rightarrow \infty} f(x_n) = f(x) \in Y$ für jede gegen $x \in X$ konvergente Folge $x_n \in X$.
 f heißt **stetig**, wenn f in allen Punkten $x \in X$ stetig ist.

Beispiel

Sei $(V, \|\cdot\|)$ ein normierter Vektorraum mit zugehöriger Metrik d . Dann ist die Norm $\|\cdot\|: V \rightarrow \mathbb{R}$ eine stetige Funktion, denn aus der Dreiecksungleichung für die Norm folgt $|\|x\| - \|y\|| \leq \|x - y\|$.

547 / 862

Die Parallelogrammgleichung

Die Frage, wann eine Norm auf einem reellen Vektorraum von einem Skalarprodukt kommt, kann mittels der Parallelogrammgleichung entschieden werden.

Satz (Parallelogrammgleichung)

Eine Norm $\|\cdot\|$ auf einem reellen Vektorraum V wird genau dann durch ein Skalarprodukt auf V induziert, wenn für alle $v, w \in V$ die **Parallelogrammgleichung** gilt:

$$(*) \quad \|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2.$$

Beweis.

- (\Rightarrow) Ist die Norm durch ein Skalarprodukt definiert, d.h. ist $\|u\|^2 = \langle u, u \rangle$ für alle $u \in V$, so ergibt sich (*) durch Berechnung der linken Seite mittels der Bilinearität des Skalarproduktes.

548 / 862

Weiter im Beweis:

(\Leftarrow) Wir definieren

$$\langle v, w \rangle := \frac{1}{4} (\|v + w\|^2 - \|v - w\|^2).$$

Dann gilt offenbar $\langle v, w \rangle = \langle w, v \rangle$ und $\langle v, v \rangle = \|v\|^2$.

Um zu beweisen, dass $\langle \cdot, \cdot \rangle$ ein Skalarprodukt ist, genügt es daher, die Linearität im ersten Argument nachzuweisen. Wir beginnen mit einer Hilfsrechnung: Für alle $u, u', w \in V$ gilt

$$\begin{aligned} & \langle u + u', w \rangle + \langle u - u', w \rangle \\ \stackrel{(Def)}{=} & \frac{1}{4} (\|u + u' + w\|^2 - \|u + u' - w\|^2) \\ & + \frac{1}{4} (\|u - u' + w\|^2 - \|u - u' - w\|^2) \\ \stackrel{(*)}{=} & \frac{1}{2} (\|u + w\|^2 + \|u'\|^2) - \frac{1}{2} (\|u - w\|^2 + \|u'\|^2) \\ = & \frac{1}{2} (\|u + w\|^2 - \|u - w\|^2) \stackrel{(Def)}{=} 2\langle u, w \rangle. \end{aligned}$$

549 / 862

Weiter im Beweis:

Wir haben gezeigt, dass für alle $u, u', w \in V$ gilt

$$(1) \quad \langle u + u', w \rangle + \langle u - u', w \rangle = 2\langle u, w \rangle.$$

Für $u = u'$ erhält man $\langle 2u, w \rangle = 2\langle u, w \rangle$.

Nun zeigen wir die Additivität: Seien $v, v' \in V$ beliebig. Definiert man $u := \frac{1}{2}(v + v')$ und $u' := \frac{1}{2}(v - v')$, so ist $v = u + u'$ und $v' = u - u'$. Aus (1) erhält man dann

$$\langle v, w \rangle + \langle v', w \rangle \stackrel{(1)}{=} 2\langle u, w \rangle = \langle 2u, w \rangle = \langle v + v', w \rangle,$$

d.h. die Additivität im ersten Argument.

Daraus ergibt sich per Induktion $\langle nv, w \rangle = n\langle v, w \rangle$ für alle $n \in \mathbb{N}$.

Desweiteren folgt aus der Definition und der Norm-Eigenschaft (N2), dass

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

und damit $\langle nv, w \rangle = n\langle v, w \rangle$ für alle $n \in \mathbb{Z}$.

550 / 862

Weiter im Beweis: Daraus ergibt sich wiederum, dass

$$\frac{1}{n}\langle u, w \rangle = \frac{1}{n}\langle n\frac{1}{n}u, w \rangle = \langle \frac{1}{n}u, w \rangle$$

für alle $u \in V, n \in \mathbb{Z}^*$. Somit

$$(2) \quad \langle \lambda u, w \rangle = \lambda \langle u, w \rangle$$

für alle $\lambda \in \mathbb{Q}$.

Da es zu jeder reellen Zahl λ eine Folge λ_n rationaler Zahlen gibt, die gegen λ konvergiert, liefert die Stetigkeit der Norm die Gleichung (2) für alle $\lambda \in \mathbb{R}$. \square

551 / 862

Beispiele

(i) Die Euklidische Norm auf \mathbb{R}^n erfüllt natürlich die Parallelogrammgleichung.

(ii) Durch

$$\|x\|_1 := \sum_{i=1}^n |x_i|, \quad x = \sum_{i=1}^n x_i e_i \in \mathbb{R}^n,$$

wird eine Norm auf \mathbb{R}^n definiert.

Diese erfüllt nicht die Parallelogrammgleichung, denn z.B. ist

$$8 = \|e_1 + e_2\|_1^2 + \|e_1 - e_2\|_1^2 \neq 2\|e_1\|_1^2 + 2\|e_2\|_1^2 = 4.$$

(iii) Die **Maximumsnorm**

$$\|x\|_{\max} := \max\{|x_1|, |x_2|, \dots, |x_n|\}$$

auf \mathbb{R}^n erfüllt nicht die Parallelogrammgleichung.

552 / 862

Äquivalente Normen

Definition

Zwei Normen $\|\cdot\|$ und $\|\cdot\|'$ auf einem Vektorraum V heißen **äquivalent**, wenn es positive Konstanten c und C gibt, so dass

$$c\|v\| \leq \|v\|' \leq C\|v\| \quad \text{für alle } v \in V.$$

Bemerkung

- Dies definiert eine Äquivalenzrelation zwischen Normen (ÜA). Insbesondere sind zwei Normen, die beide zu einer dritten äquivalent sind, auch zueinander äquivalent (Transitivität).
- Seien d_1 und d_2 die durch zwei äquivalente Normen auf V definierten Metriken. Dann konvergiert eine Folge $x_n \in V$ gegen x bzgl. d_1 genau dann, wenn sie auch bzgl. d_2 gegen x konvergiert: die metrischen Räume (V, d_1) und (V, d_2) haben also denselben Konvergenzbegriff.

553 / 862

Auf endlich-dimensionalen Vektorräumen sind alle Normen äquivalent

Satz

Je zwei Normen $\|\cdot\|$ und $\|\cdot\|'$ auf einem **endlichdimensionalen** reellen oder komplexen Vektorraum V sind äquivalent.

Beweis. Sei $B := (b_1, \dots, b_n)$ eine Basis von V .

Für $x = \sum_{i=1}^n x_i b_i \in V$ betrachten wir wieder die Maximumsnorm

$$\|x\|_{\max} := \max\{|x_1|, \dots, |x_n|\}.$$

Wir zeigen, dass jede Norm auf V äquivalent zu $\|\cdot\|_{\max}$ ist.

Für jede Norm $\|\cdot\|$ auf V gilt nach der Dreiecksungleichung

$$\|x\| = \left\| \sum_{i=1}^n x_i b_i \right\| \leq \sum_{i=1}^n |x_i| \cdot \|b_i\| \leq C \|x\|_{\max},$$

mit $C := \sum_{i=1}^n \|b_i\|$. Dies beweist die eine Ungleichung.

554 / 862

Weiter im Beweis:

Wir nehmen an, es gäbe kein $c > 0$ mit $c\|\cdot\|_{\max} \leq \|\cdot\|$.

D.h. es gibt für jedes $c > 0$ ein $y \in V$ mit $c\|y\|_{\max} > \|y\|$. Damit findet man für jedes c einen Vektor $x := \frac{y}{\|y\|_{\max}}$ mit $\|x\|_{\max} = 1$ und $c > \|x\|$.

Für alle $k \in \mathbb{N}$ gibt es also einen Vektor $x^{(k)} = \sum_{i=1}^n x_i^{(k)} b_i \in V$ mit $|x_i^{(k)}| \leq 1$ und $\frac{1}{k} > \|x^{(k)}\|$. Nach Bolzano-Weierstraß hat jede der beschränkten Zahlenfolgen $(x_i^{(k)})_{k \in \mathbb{N}}$, $i = 1, \dots, n$, eine konvergente Teilfolge.

Wir können daher annehmen, dass $(x^{(k)})$ bezüglich der Norm $\|\cdot\|_{\max}$ gegen $x = \sum_{i=1}^n x_i b_i \in V$ konvergiert.

Es gilt dann

$$\|x\| \leq \|x - x^{(k)}\| + \|x^{(k)}\| \leq C\|x - x^{(k)}\|_{\max} + \frac{1}{k} \xrightarrow{k \rightarrow \infty} 0.$$

Also $x = 0$ und somit $0 = \|x\|_{\max} = \lim_{k \rightarrow \infty} \|x^{(k)}\|_{\max}$, im Widerspruch zu $\|x^{(k)}\|_{\max} = 1$. □

555 / 862

Vollständigkeit, Banachräume, Hilberträume

Definition (Vollständigkeit metrischer Räume)

Erinnerung: Sei (X, d) ein metrischer Raum.

- Eine Folge $x_n \in X$ heißt **Cauchy-Folge**, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass

$$d(x_m, x_n) < \varepsilon \quad \text{für alle } m, n \geq N.$$

- Der metrische Raum (X, d) heißt **vollständig**, wenn jede Cauchy-Folge in X gegen einen Grenzwert $x \in X$ konvergiert.

Definition

- Ein **Banachraum** ist ein normierter Vektorraum, der (bezüglich der zur Norm gehörenden Metrik) vollständig ist.
- Ein reeller bzw. komplexer **Hilbertraum** ist ein Euklidischer bzw. Hermitescher Vektorraum, der (bezüglich der vom Skalarprodukt induzierten Metrik) vollständig ist.

556 / 862

Bemerkungen/Beispiele

- ① Wieder gilt: Seien d_1 und d_2 die durch zwei äquivalente Normen auf V definierten Metriken. Dann ist $x_n \in V$ eine Cauchyfolge bzgl. d_1 genau dann, wenn sie eine Cauchyfolge bzgl. d_2 ist.
- ② Aus der Vollständigkeit von $\mathbb{K} = \mathbb{R}, \mathbb{C}$ folgt, dass $(\mathbb{K}^n, \|\cdot\|_{\max})$ ein Banachraum ist: Ist $x^{(k)} \in \mathbb{K}^n$ eine Cauchyfolge, so auch alle ihre Komponenten.
- ③ Mittels Wahl einer Basis, folgt daraus: Jeder endlichdimensionale normierte reelle oder komplexe Vektorraum ist ein Banachraum.
- ④ Damit sind auch endlichdimensionale Euklidische und Hermitesche Vektorräume Hilberträume.
Insbesondere ist \mathbb{R}^n mit dem Euklidischen und \mathbb{C}^n mit dem Hermiteschen Skalarprodukt ein reeller bzw. komplexer Hilbertraum.
- ⑤ Unendlich-dimensionale Hilberträume spielen eine wichtige Rolle in der Quantenmechanik.

557 / 862

Beispiel für einen ∞ -dimensionalen Hilbertraum: Quadratisch summierbare Folgen

Satz

Der Vektorraum der quadratisch summierbaren komplexen Zahlenfolgen

$$\ell^2 := \ell^2(\mathbb{N}, \mathbb{C}) := \left\{ (x_n)_{n \in \mathbb{N}} \mid x_n \in \mathbb{C}, \sum_{n=0}^{\infty} |x_n|^2 < \infty \right\}$$

mit dem Hermiteschen Skalarprodukt (ÜA!)

$$\langle x, y \rangle = \sum_{n=0}^{\infty} x_n \bar{y}_n$$

ist ein Hilbertraum.

Beweis. $(\ell^2, \langle \cdot, \cdot \rangle)$ ist ein Hermitescher Vektorraum (siehe ÜA).

558 / 862

Wir beweisen die Vollständigkeit von ℓ^2 :

Sei $(x^{(k)})_{k \in \mathbb{N}} = (x_0^{(k)}, x_1^{(k)}, \dots)$ eine Cauchy-Folge in ℓ^2 .

Dann ist jede der Komponentenfolgen $(x_n^{(k)})_{k \in \mathbb{N}}$ eine Cauchy-Folge und konvergiert daher gegen einen komplexen Grenzwert $x_n := \lim_{k \rightarrow \infty} x_n^{(k)}$.

Behauptung: Es gilt $x := (x_n)_{n \in \mathbb{N}} \in \ell^2$ und $x = \lim_{k \rightarrow \infty} x^{(k)}$.

Da $(x^{(k)})$ eine Cauchy-Folge ist, gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit

$$\sum_{n=0}^m |x_n^{(p)} - x_n^{(q)}|^2 \leq \sum_{n=0}^{\infty} |x_n^{(p)} - x_n^{(q)}|^2 < \varepsilon^2$$

für alle $p, q \geq N$ und alle $m \in \mathbb{N}$.

Grenzübergang $p \rightarrow \infty$ in der *endlichen* Summe liefert

$$\sum_{n=0}^m |x_n - x_n^{(q)}|^2 \leq \varepsilon^2$$

für alle $q \geq N$ und jedes $m \in \mathbb{N}$, d.h. $\|x - x^{(q)}\| \leq \varepsilon$ und somit $x - x^{(q)} \in \ell^2$. Dann folgt aber auch $x = x^{(q)} + (x - x^{(q)}) \in \ell^2$. □

559 / 862

Banachscher Fixpunktsatz

Den Banachschen Fixpunktsatz werden wir in der Theorie der gewöhnlichen Differentialgleichungen benutzen. Wir werden ihn dann auf bestimmte Banachräume von Funktionen anwenden, um die Existenz der Lösung einer Differentialgleichung zu zeigen.

Satz (Banachscher Fixpunktsatz)

Sei (X, d) ein vollständiger metrischer Raum und $F: X \rightarrow X$ eine **kontrahierende Abbildung**, d.h. es existiert $0 \leq \theta < 1$, so dass

$$d(F(x), F(y)) \leq \theta d(x, y)$$

für alle $x, y \in X$.

Dann existiert ein **eindeutiger Fixpunkt von F** , d.h. es gibt genau ein $x \in X$ mit $F(x) = x$.

Des Weiteren konvergiert für alle Anfangswerte $x_0 \in X$ die durch $x_{k+1} = F(x_k)$ rekursiv definierte Folge (x_k) gegen diesen Fixpunkt.

Beweis. 1) Existiert ein Fixpunkt, so ist er eindeutig bestimmt:

Sind nämlich $x, y \in X$ Fixpunkte, $F(x) = x$ und $F(y) = y$, dann gilt

$$0 \leq d(x, y) = d(F(x), F(y)) \leq \theta d(x, y).$$

Wegen $\theta < 1$ folgt $d(x, y) = 0$, also $x = y$.

2) Die Folge (x_k) ist für jeden Startpunkt x_0 eine Cauchyfolge, denn es gilt, mit $C := d(x_1, x_0)$:

$$\begin{aligned} d(x_{n+k}, x_n) &= d(F(x_{n+k-1}), F(x_{n-1})) \leq \theta d(x_{n+k-1}, x_{n-1}) \leq \dots \\ &\leq \theta^n d(x_k, x_0) \leq \theta^n \sum_{l=0}^{k-1} d(x_{l+1}, x_l) \leq \theta^n \sum_{l=0}^{k-1} \theta^l \cdot C \leq \frac{C\theta^n}{1-\theta} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Da (X, d) vollständig ist, existiert ein Grenzwert $x \in X$.

Dieser ist aber auch der Fixpunkt von F , denn wegen der Dreiecksungleichung gilt

$$\begin{aligned} d(F(x), x) &\leq d(F(x), x_k) + d(x_k, x) \\ &\leq \theta d(x, x_{k-1}) + d(x_k, x) \xrightarrow{k \rightarrow \infty} 0. \end{aligned}$$

Somit ist $d(F(x), x) = 0$, d.h. $F(x) = x$. □

561 / 862

Banachscher Fixpunktsatz: Anwendungsbeispiel

Beispiel aus dem 1. Semester: Berechnung von $\sqrt{2}$

Sei $X := [1, \infty)$ mit der Euklidischen Metrik versehen. Dies ist ein vollständiger metrischer Raum.

Wir betrachten die Abbildung

$$F: X \ni x \mapsto \frac{1}{2} \left(x + \frac{2}{x} \right) \in X.$$

Diese Abbildung ist kontrahierend, denn

$$F(x) - F(y) = \frac{1}{2}(x - y) + \frac{1}{x} - \frac{1}{y} = \left(\frac{1}{2} - \frac{1}{xy} \right) (x - y)$$

impliziert wegen $xy \geq 1$ dass $\|F(x) - F(y)\| \leq \frac{1}{2}\|x - y\|$.

Somit hat F einen Fixpunkt, nämlich (wie man leicht nachrechnet) $\sqrt{2}$.

Definition (Offene und abgeschlossene Mengen)

Sei (X, d) ein metrischer Raum.

- Die Teilmenge $B_r(x) := \{y \in X \mid d(y, x) < r\}$ heißt (offene) **Kugel** (oder **Ball**) vom Radius r und Mittelpunkt x .
- Eine Teilmenge $U \subseteq X$ heißt **offen**, wenn es zu jedem $x \in U$ ein $r > 0$ gibt, so dass $B_r(x) \subseteq U$.
- Eine Teilmenge $A \subseteq X$ heißt **abgeschlossen**, wenn ihr Komplement $X \setminus A$ offen ist.

Bemerkung/ÜA

- Beliebige Vereinigungen offener Mengen sind wieder offen.
- Beliebige Durchschnitte abg. Mengen sind wieder abgeschlossen.
- **Endliche** Durchschnitte offener Mengen sind wieder offen.
- Endliche Vereinigungen abg. Mengen sind abgeschlossen.
- X und \emptyset sind offen und abgeschlossen.

563 / 862

Inneres, Abschluss und Rand

Definition

Sei X ein metrischer Raum und $A \subseteq X$.

(i) Das **Innere** von A ist definiert als die offene Teilmenge

$$\overset{\circ}{A} := \bigcup_{B \subseteq A \text{ offen}} B$$

(ii) Der **Abschluss** von A ist definiert als die abgeschlossene Teilmenge

$$\bar{A} := \bigcap_{B \supseteq A \text{ abgeschl.}} B$$

(iii) $\partial A := \bar{A} \setminus \overset{\circ}{A}$ heißt (topologischer) **Rand** von A .

Das Innere $\overset{\circ}{A}$ ist also die größte offene Teilmenge, der Abschluss \bar{A} die kleinste abgeschlossene Obermenge.

564 / 862

Beispiele/ ÜA:

- Der Abschluss der offenen Kugel $B_r(x) \subseteq \mathbb{R}^n$ im Euklidischen Raum ist die abgeschlossene Kugel

$$\bar{B}_r(x) = \{y \in \mathbb{R}^n \mid d(y, x) \leq r\}.$$

Der (topologische) Rand der Kugel ist die **Sphäre** vom Radius r :

$$\partial B_r(x) = S_r(x) := \{y \in \mathbb{R}^n \mid d(y, x) = r\}.$$

- Für $\mathbb{Q} \subset \mathbb{R}$ gilt $\bar{\mathbb{Q}} = \mathbb{R}$ und $\overset{\circ}{\mathbb{Q}} = \emptyset$, also auch $\partial\mathbb{Q} = \mathbb{R}$.

565 / 862

Stetige Abbildungen

Satz

Für eine Abbildung $F: X \rightarrow Y$ zwischen metrischen Räumen sind äquivalent:

- (i) F ist stetig
- (ii) Urbilder $F^{-1}(U)$ aller offenen Mengen $U \subseteq Y$ sind offen.
- (iii) Urbilder $F^{-1}(A)$ aller abgeschlossenen Mengen A sind abgeschlossen.

Beweis. Wir beweisen nur die Äquivalenz von (i) und (ii):

(ii) \Rightarrow (i):

Sei $\epsilon > 0$. Betrachte die offene Kugel $B_\epsilon(F(x)) \subseteq Y$. Ihr Urbild ist nach Voraussetzung offen, d.h. wir finden ein $\delta > 0$, so dass $B_\delta(x) \subseteq F^{-1}(B_\epsilon(F(x)))$, d.h. $F(B_\delta(x)) \subseteq B_\epsilon(F(x))$.

Wir haben also gezeigt:

$$\forall \epsilon > 0 \exists \delta > 0 : d(x, y) < \delta \implies d(F(x), F(y)) < \epsilon,$$

und dies ist gerade die ϵ - δ -Definition von Stetigkeit.

566 / 862

Weiter im Beweis: (i) \Rightarrow (ii):

Angenommen, es existiert eine offene Menge $V \subseteq Y$, so dass die Menge $F^{-1}(V) \subseteq X$ nicht offen ist.

Es gibt dann $x \in F^{-1}(V)$, so dass $B_{\frac{1}{n}}(x) \not\subseteq F^{-1}(V)$ für alle $n \in \mathbb{N}$. Sei also x_n eine Folge mit $x_n \in B_{\frac{1}{n}}(x) \setminus F^{-1}(V)$.

Insbesondere gilt dann $x_n \rightarrow x$ und wegen der Stetigkeit folgt daraus auch $F(x_n) \rightarrow F(x)$.

Da $V \subseteq Y$ offen ist, gibt es ein $\delta > 0$ mit $B_\delta(F(x)) \subseteq V$. Es existiert aber auch ein $N \in \mathbb{N}$ mit

$$F(x_n) \in B_\delta(F(x)) \subseteq V \quad \text{für alle } n > N.$$

Das steht im Widerspruch zu $x_n \notin F^{-1}(V)$ für alle $n \in \mathbb{N}$. □

Frage: Welche Beweismethode haben wir hier benutzt?

567 / 862

Satz (Abgeschlossene Teilmengen metrischer Räume)

Für eine Teilmenge $A \subseteq X$ eines metrischen Raumes X sind äquivalent:

- (i) A ist abgeschlossen
- (ii) Wenn ein $x \in X$ Grenzwert einer konvergenten Folge von Elementen von A ist, so gilt $x \in A$.

Beweis.

(\Rightarrow) Sei $x_k \in A$ eine Folge mit Grenzwert $x \in X$.

Wenn $x \notin A$, so läge x in der offenen Teilmenge $U = X \setminus A$.

Insbesondere existiert ein $r > 0$ mit $B_r(x) \subseteq U$, d.h. $B_r(x) \cap A = \emptyset$.

Das widerspricht aber der Konvergenz der Folge $x_k \in A$ gegen x .

(\Leftarrow) Wir zeigen, dass es zu jedem $x \in X \setminus A$ ein $r > 0$ gibt mit

$$B_r(x) \subseteq X \setminus A.$$

Sei also $x \in X \setminus A$. Existiert kein solches $r > 0$, dann gäbe es eine Folge $x_k \in B_{1/k}(x) \cap A$. Da diese Folge gegen x konvergiert, folgt nach Voraussetzung $x \in A$.

Dies widerspricht aber $x \in X \setminus A$. □

568 / 862

Folgerung

Sei X metrischer Raum. Sei A eine Teilmenge (versehen mit der Metrik, die durch die Einschränkung der Metrik von X gegeben ist). Dann:

- (i) Ist A vollständig, so ist A abgeschlossen in X .
- (ii) Ist X vollständig und A abgeschlossen, so ist A auch vollständig.

Beweis.

- (i) Sei $x_k \in A$ konvergent mit Grenzwert $x \in X$.

Dann ist (x_k) eine Cauchy-Folge und somit $x \in A$, da A als vollständiger metrischer Raum vorausgesetzt wurde.

- (ii) Sei $x_k \in A$ eine Cauchy-Folge. Da X als vollständig vorausgesetzt wurde, konvergiert (x_k) gegen $x \in X$. Da A abgeschlossen sein soll, ist aber nach dem vorgehenden Satz $x \in A$.

□

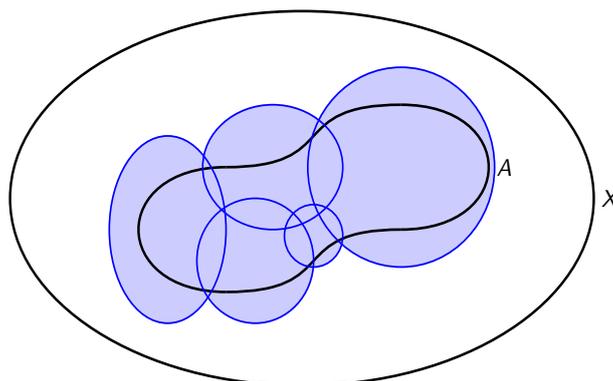
569 / 862

Kompakte Mengen

Definition (Kompaktheit)

Sei X ein metrischer Raum.

- Eine **(offene) Überdeckung** einer Teilmenge $A \subseteq X$ ist eine Familie $(U_i)_{i \in I}$ (offener) Teilmengen $U_i \subseteq X$ mit $A \subseteq \bigcup_{i \in I} U_i$.
- Eine Teilmenge $A \subseteq X$ heißt **kompakt**, wenn es zu **jeder** offenen Überdeckung $(U_i)_{i \in I}$ von A eine endliche Teilüberdeckung $(U_{i_1}, U_{i_2}, \dots, U_{i_k})$, $i_1, \dots, i_k \in I$, gibt.



570 / 862

Kompaktheit ist eine extrem nützliche Eigenschaft.

Satz

Sei X ein metrischer Raum und $A \subseteq X$ eine kompakte Teilmenge. Dann hat jede Folge in A eine konvergente Teilfolge mit Grenzwert in A .

Beweis.

Sei $x_k \in A$, $k \in \mathbb{N}$, eine Folge in A .

Wenn die Folge $(x_k)_{k \in \mathbb{N}}$ keine konvergente Teilfolge mit Grenzwert in A hat, dann gibt es zu jedem $a \in A$ eine offene Umgebung U_a , die nur endlich viele Glieder der Folge x_k enthält.

$(U_a)_{a \in A}$ ist eine offene Überdeckung der kompakten Menge A und besitzt daher eine endliche Teilüberdeckung $U_{a_1}, U_{a_2}, \dots, U_{a_r}$, d.h.

$$A \subseteq U_{a_1} \cup U_{a_2} \cup \dots \cup U_{a_r}.$$

Nach Konstruktion enthält diese Vereinigung aber nur endlich viele Glieder der Folge, im Widerspruch zur Annahme $x_k \in A$ für alle $k \in \mathbb{N}$. \square

571 / 862

Wir nennen eine Teilmenge $A \subset X$ eines metrischen Raumes *beschränkt*, falls sie in einer Kugel $B_r(x)$ enthalten ist.

Folgerung

Jede kompakte Teilmenge $A \subseteq X$ eines metrischen Raumes X ist abgeschlossen, vollständig und beschränkt.

Beweis. Cauchyfolgen konvergieren genau dann, wenn sie eine konvergente Teilfolge besitzen. Nach dem vorhergehenden Satz ist dies aber der Fall. Also ist A vollständig. Da der Grenzwert in A liegt, ist A abgeschlossen. Wähle ein $x \in X$. Dann ist $(B_k(x))_{k \in \mathbb{N}}$ eine offene Überdeckung von X und somit von A .

Da A als kompakt vorausgesetzt wurde, existiert eine endliche Teilüberdeckung $B_{k_1}(x), \dots, B_{k_r}(x)$ und mit $R = \max\{k_1, \dots, k_r\}$ erhalten wir $A \subseteq B_R(x)$. \square

572 / 862

Satz

X kompakt, $A \subseteq X$ abgeschlossen $\implies A$ kompakt.

Beweis. Sei $(U_i)_{i \in I}$ eine offene Überdeckung von A .

Dann ist $(X \setminus A, (U_i)_{i \in I})$ eine offene Überdeckung des Kompaktums X .

Also gibt es eine endliche Teilüberdeckung $(X \setminus A, U_{i_1}, \dots, U_{i_k})$ von X .

Insbesondere ist $(U_{i_1}, \dots, U_{i_k})$ eine endliche Teilüberdeckung von A . \square

Folgerung

Sei X ein metrischer Raum, in dem die abgeschlossenen Kugeln $\overline{B}_r(x)$ kompakt sind.

Die kompakten Teilmengen von X sind dann genau die abgeschlossenen beschränkten Teilmengen.

573 / 862

Satz (Satz von Heine-Borel)

Im \mathbb{R}^n , versehen mit der Euklidischen Metrik, (und in jedem **endlich-dimensionalen** normierten Vektorraum) gilt: Eine Teilmenge A ist genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis. Aufgrund der vorangegangenen Folgerung genügt es zu zeigen, dass die abgeschlossenen Kugeln $\overline{B}_r(x) \subseteq \mathbb{R}^n$ kompakt sind.

Da alle Normen auf \mathbb{R}^n äquivalent sind, können wir z.B. die Maximumsnorm zu Grunde legen.

Außerdem können wir $x = 0$ (Translationsinvarianz des Abstands) und $r = 1$ (Homogenität) annehmen, d.h. es reicht aus, zu zeigen, dass $\overline{B}_r(x) = \overline{B}_1(0) = [-1, 1]^n =: W$ kompakt ist.

Sei also $(U_i)_{i \in I}$ eine offene Überdeckung des Würfels $W =: W_0$.

Wir nehmen an, es gäbe keine endliche Teilüberdeckung und führen diese Annahme zum Widerspruch.

574 / 862

Weiter im Beweis:

Wir konstruieren durch sukzessive Halbierung der Kantenlängen eine Folge von Würfeln W_k mit Kantenlänge 2^{1-k} und mit der Eigenschaft, dass W_k nicht durch endlich viele der U_i überdeckt wird. Dabei ist $W_0 = [-1, 1]^n$ und $W_k \subset W_{k-1}$ für alle $k \geq 1$.

Halbierung der Kanten führt zu einer Zerlegung des Würfels W_k in 2^n Teilwürfel, von denen mindestens einer nicht durch endlich viele U_i überdeckt werden kann (sonst könnte man W_k durch endlich viele U_i überdecken); diesen Teilwürfel wählen wir als W_{k+1} .

Sei (x_k) eine beliebige Folge mit $x_k \in W_k$. Dann ist (x_k) eine Cauchy-Folge.

Wegen der Abgeschlossenheit von W_i gilt $x = \lim_{k \rightarrow \infty} x_k \in W_i$, denn $x_k \in W_i$ für alle $k \geq i$. Daraus folgt $x \in \bigcap_i W_i \subseteq W$.

Da W von den offenen Mengen U_i überdeckt wird, existiert ein i_0 mit $x \in U_{i_0}$ und somit $B_\varepsilon(x) \subseteq U_{i_0}$, wenn $\varepsilon > 0$ klein genug gewählt ist.

Daraus ergibt sich $W_k \subseteq B_\varepsilon(x) \subseteq U_{i_0}$, wenn $2^{1-k} < \varepsilon$. Ein Widerspruch! (Hierbei haben wir benutzt, dass $\|x - y\|_{\max} \leq 2^{1-k}$ für alle $y \in W_k$.) \square

575 / 862

Satz

Sei $F: X \rightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen und $A \subseteq X$ sei kompakt. Dann ist das Bild $F(A) \subseteq Y$ kompakt.

Beweis. Sei $(U_i)_{i \in I}$ eine offene Überdeckung von $F(A)$. Dann bilden die Urbilder $V_i := F^{-1}(U_i)$ eine offene Überdeckung des Kompaktums A .

Also gibt es eine endliche Teilüberdeckung $(V_{i_1}, \dots, V_{i_k})$, und deren Bilder $(U_{i_1}, \dots, U_{i_k})$ überdecken dann das Bild $F(A)$. \square

Folgerung

Sei $f: X \rightarrow \mathbb{R}$ eine stetige Funktion und X kompakt.

Dann ist f beschränkt und nimmt ihr Minimum und Maximum in X an.

Beweis. Nach dem vorhergehenden Satz ist $f(X) \subseteq \mathbb{R}$ kompakt und somit beschränkt und abgeschlossen. \square

Gleichmäßige Stetigkeit

Satz

Jede stetige Abbildung $F: X \rightarrow Y$ von einem kompakten metrischen Raum X in einen metrischen Raum Y ist **gleichmäßig stetig**, d.h. für alle $\varepsilon > 0$ existiert $\delta > 0$, so dass **für alle** $x, x' \in X$ gilt

$$d(F(x), F(x')) < \varepsilon \quad \text{falls} \quad d(x, x') < \delta.$$

Beweis. Sei $\varepsilon > 0$. Wegen der Stetigkeit von F existiert für jedes $x \in X$ ein $\delta(x) > 0$, so dass

$$d(F(x), F(x')) < \frac{\varepsilon}{2},$$

für alle $x' \in X$ mit $d(x, x') < \delta(x)$.

577 / 862

Weiter im Beweis:

Da die offenen Kugeln $U_x := B_{\delta(x)/2}(x)$ das Kompaktum X überdecken, gibt es endlich viele $x_1, \dots, x_k \in X$ mit $X = \bigcup_{i=1}^k U_{x_i}$.

Seien nun $x, x' \in X$ mit

$$d(x, x') < \delta := \min_{i=1, \dots, k} \delta(x_i)/2.$$

Dann gibt es x_i mit $x \in U_{x_i}$, d.h. $d(x, x_i) < \delta(x_i)/2$, und somit $d(x', x_i) \leq d(x', x) + d(x, x_i) < \delta(x_i)/2 + \delta(x_i)/2 = \delta(x_i)$.

Daraus ergibt sich

$$d(F(x), F(x_i)) < \varepsilon/2 \quad \text{und} \quad d(F(x'), F(x_i)) < \varepsilon/2$$

und somit nach der Dreiecksungleichung

$$d(F(x), F(x')) \leq d(F(x), F(x_i)) + d(F(x_i), F(x')) < \varepsilon.$$



578 / 862

Abbildungsräume als metrische Räume

Seien X und Y zwei metrische Räume und bezeichne

$$B(X, Y) := \{f: X \rightarrow Y \mid f \text{ beschränkt}\}$$

die Menge der Abbildungen von X nach Y mit *beschränktem Bild*, d.h. mit $f(X) \subseteq Y$ beschränkt.

Übung

Durch

$$d_\infty(f, g) := \sup_{x \in X} d_Y(f(x), g(x))$$

wird auf dem Abbildungsraum $B(X, Y)$ eine Metrik definiert, die sogenannte **Supremumsmetrik**.

Eine Folge (f_n) von Abbildungen konvergiert in $(B(X, Y), d_\infty)$ gegen f , wenn die Abbildungen *gleichmäßig* gegen f konvergieren, d.h. falls

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N \forall x \in X : d_Y(f_n(x), f(x)) < \varepsilon.$$

579 / 862

Ist nun X **kompakt**, so ist das Bild jeder *stetigen* Abbildung $f: X \rightarrow Y$ beschränkt, d.h. in diesem Fall gilt

$$C(X, Y) \subseteq B(X, Y).$$

Wie wir im letzten Semester gesehen haben, ist $C(X, Y)$ abgeschlossen unter gleichmäßiger Konvergenz.

Wir halten also fest:

Für einen kompakten metrischen Raum X und jeden metrischen Raum Y ist $C(X, Y) \subseteq (B(X, Y), d_\infty)$ eine abgeschlossene Teilmenge.

Satz

Sei X ein **kompakter** metrischer Raum und Y ein **vollständiger** metrischer Raum. Dann ist $(C(X, Y), d_\infty)$ ein **vollständiger** metrischer Raum.

580 / 862

Beweis. Wir müssen zeigen, dass jede Cauchyfolge in $C(X, Y)$ gleichmäßig gegen eine geeignete Abbildung $f : X \rightarrow Y$ konvergiert. Deren Stetigkeit folgt dann aus dem bereits erwähnten Resultat des letzten Semesters. Sei $\{f_n\}$ eine Cauchyfolge in $C(X, Y)$ und sei $\varepsilon > 0$. Dann gibt es ein $N_\varepsilon > 0$, so dass für alle $x \in X$ gilt

$$d_Y(f_n(x), f_m(x)) \leq d_\infty(f_n, f_m) < \varepsilon \quad \text{für alle } n, m \geq N_\varepsilon.$$

D.h. aber, dass für jedes $x \in X$ die Folge $f_n(x)$ eine Cauchyfolge im **vollständigen** metrischen Raum Y ist und daher für jedes $x \in X$ einen Grenzwert $f(x) := \lim_{n \rightarrow \infty} f_n(x)$ hat.

Da die Metrik d_Y stetig ist, erhalten wir

$$\begin{aligned} d_Y(f_n(x), f(x)) &= d_Y(f_n(x), \lim_{m \rightarrow \infty} f_m(x)) \\ &= \lim_{m \rightarrow \infty} d_Y(f_n(x), f_m(x)) < \varepsilon \end{aligned}$$

für jedes $x \in X$ und jedes $n \geq N_\varepsilon$.

Die Folge f_n konvergiert also **gleichmäßig** gegen f , und hieraus folgt die Behauptung. □

581 / 862

Bemerkung: Vertauschung von Limes und Integral

- Das bestimmte Integral definiert eine stetige Abbildung $(C([a, b], \mathbb{R}), d_\infty) \rightarrow \mathbb{R}$.

Beweis: Konvergiert eine Folge stetiger Funktionen $f_n : [a, b] \rightarrow \mathbb{R}$ **gleichmäßig** gegen eine (stetige!) Funktion $f : [a, b] \rightarrow \mathbb{R}$, so gilt

$$\left| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right| \leq (b-a) \|f - f_n\|_\infty \xrightarrow{n \rightarrow \infty} 0.$$

Die Integrale konvergieren also gegen das Integral der Grenzfunktion:

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

- Ist $f_n : [a, b] \rightarrow \mathbb{R}$ eine Folge stetiger Funktionen, die nur punktweise gegen eine stetige Funktion f konvergiert, dann gilt die obige Gleichheit im allgemeinen nicht.

Beispiel / Übung: Man betrachte, für $n \geq 2$, $f_n : [0, 1] \rightarrow \mathbb{R}$, $f_n(x) := \max\{n - n^2|x - \frac{1}{n}|, 0\}$.

582 / 862

Kapitel 17

Beschränkte Operatoren, Hilbertbasen, Fourier-Reihen

583 / 862

Beschränkte Operatoren, Hilbertbasen, Fourier-Reihen

Beschränkte Operatoren, Darstellungssatz von Riesz

Beschränkte Operatoren

Definition

Seien $(V_1, \|\cdot\|_1)$ und $(V_2, \|\cdot\|_2)$ zwei normierte Vektorräume.
Eine lineare Abbildung $F: V_1 \rightarrow V_2$ heißt **beschränkt**, falls es eine Konstante $C \geq 0$ gibt, so dass

$$\|F(u)\|_2 \leq C\|u\|_1 \quad \text{für alle } u \in V_1.$$

Dies ist äquivalent zu

$$\sup_{0 \neq u \in V_1} \frac{\|Fu\|_2}{\|u\|_1} < \infty.$$

Die Zahl

$$\|F\| := \sup_{0 \neq u \in V_1} \frac{\|Fu\|_2}{\|u\|_1} = \sup_{x \in V_1 \text{ mit } \|x\|_1=1} \|Fx\|_2$$

heißt **Operatornorm** der linearen Abbildung F .

584 / 862

Bemerkungen/ÜA

- Durch die Operatornorm wird der Vektorraum

$$L_b(V_1, V_2) := \{F \in L(V_1, V_2) \mid \|F\| < \infty\}$$

der *beschränkten* Operatoren zu einem normierten Vektorraum.

- Insbesondere ist für jeden normierten \mathbb{K} -Vektorraum V der Raum $V' := L_b(V, \mathbb{K})$ der beschränkten Linearformen ebenfalls ein normierter Vektorraum. Hierbei ist \mathbb{K} ($= \mathbb{R}$ oder \mathbb{C}) durch den üblichen Betrag normiert.
- Lineare Abbildungen zwischen **endlich-dimensionalen** reellen oder komplexen Vektorräumen sind beschränkt.
- Ist V_1 unendlich-dimensional, so gilt dies nicht: Sei $V_1 = V_2 = C^\infty([0, 1])$ versehen mit der Norm $\|f\|_\infty = \sup_{x \in [0, 1]} f(x)$, und sei $F = (\cdot)'$ der Ableitungsoperator. Dann gilt für $f_n(x) := x^n$ dass $\|f_n\|_\infty = 1$ und $\|F(f_n)\|_\infty = \|nx^{n-1}\|_\infty = n$.

585 / 862

Satz

Eine lineare Abbildung $F: V_1 \rightarrow V_2$ zwischen normierten Vektorräumen $(V_1, \|\cdot\|_1)$ und $(V_2, \|\cdot\|_2)$ ist genau dann **stetig**, wenn sie **beschränkt** ist.

Beweis. Ist F stetig, so ist F auch in $0 \in V_1$ stetig. Insbesondere zu $\varepsilon = 1$ gibt es ein $\delta > 0$ mit

$$\|Fv\|_2 < 1, \text{ falls } \|v\|_1 < \delta.$$

Sei nun $u \in V_1 \setminus \{0\}$. Für $v := \frac{\delta}{2\|u\|_1} u$ gilt dann, dass $\|v\|_1 < \delta$, und somit

$$\|F(u)\|_2 = \frac{2\|u\|_1}{\delta} \|F(v)\|_2 < \frac{2\|u\|_1}{\delta}.$$

Damit ist die stetige lineare Abbildung F beschränkt.

Ist andererseits F beschränkt, d.h. $\|F(u)\|_2 \leq C\|u\|_1$ für alle $u \in V_1$, so gilt

$$\|F(u) - F(u_n)\|_2 = \|F(u - u_n)\|_2 \leq C\|u - u_n\|_1.$$

Somit konvergiert $F(u_n) \rightarrow F(u)$ in V_2 , falls $u_n \rightarrow u$ in V_1 . □

586 / 862

Orthogonalprojektion auf vollständige Unterräume

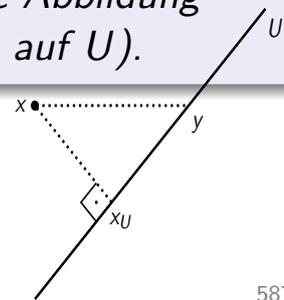
Satz

Sei V ein Euklidischer oder Hermitescher Vektorraum und $U \subseteq V$ ein **vollständiger** Unterraum.

- (i) Dann existiert zu jedem $x \in V$ genau ein $x_U \in U$ mit kleinstem Abstand zu x , d.h.

$$\|x - x_U\| \leq \|x - y\| \quad \text{für alle } y \in U.$$

- (ii) Das Element $x_U \in U$ ist durch $x - x_U \in U^\perp$ eindeutig charakterisiert.
 (iii) Es gilt $V = U \oplus U^\perp$.
 (iv) Die Zuordnung $x \mapsto x_U$ definiert eine stetige lineare Abbildung $P: V \rightarrow U$ (die sogenannte **orthogonale Projektion** auf U).



587 / 862

Beweis. (i) Sei $y_n \in U$ eine Folge mit

$$\lim_{n \rightarrow \infty} \|x - y_n\| = d := \inf_{y \in U} \|x - y\|.$$

(y_n) ist eine Cauchy-Folge, denn wegen der **Parallelogrammgleichung** gilt

$$\begin{aligned} \|y_n - y_m\|^2 &= \|(y_n - x) - (y_m - x)\|^2 \\ &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - \|(y_n - x) + (y_m - x)\|^2 \\ &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\left\|\frac{y_n + y_m}{2} - x\right\|^2 \\ &\leq 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4d^2 \xrightarrow{(m,n \rightarrow \infty)} 0 \end{aligned}$$

Aufgrund der **Vollständigkeit** des Unterraums U konvergiert (y_n) in U , $x_U := \lim_{n \rightarrow \infty} y_n \in U$, und $\|x - x_U\| = d$, wegen der Stetigkeit der Norm.

Aus $x_U, x'_U \in U$ mit $\|x_U - x\| = \|x'_U - x\| = d$ folgt wie oben

$$\begin{aligned} \|x_U - x'_U\|^2 &= 2\|x_U - x\|^2 + 2\|x'_U - x\|^2 - \|x_U + x'_U - 2x\|^2 \\ &= 4d^2 - \|x_U + x'_U - 2x\|^2 \leq 4d^2 - 4d^2 = 0. \end{aligned}$$

Das beweist die Eindeutigkeit von x_U mit $\|x_U - x\| = d$.

588 / 862

Weiter im Beweis:

(ii) Für alle $y \in U$, $t \in \mathbb{R}$ gilt mit $z := x - x_U$:

$$d^2 = \|z\|^2 \leq \|z + ty\|^2 = d^2 + 2t \operatorname{Re} \langle z, y \rangle + t^2 \|y\|^2.$$

Das ist nur möglich, wenn $\operatorname{Re} \langle z, y \rangle = 0$ für alle $y \in U$.

Durch Substitution $t \rightsquigarrow it$ zeigt man $\operatorname{Im} \langle z, y \rangle = 0$ für alle $y \in U$. Es folgt $z = x - x_U \perp U$.

Umgekehrt folgt für $u \in U$ mit $z := x - u \perp U$ für alle $y \in U$:

$$\|z + y\|^2 = \|z\|^2 + \|y\|^2 \geq \|z\|^2,$$

d.h. $u = x_U$.

(iii) In (i-ii) haben wir gezeigt, dass jeder Vektor $x \in V$ eine eindeutige Darstellung $x = x_U + (x - x_U)$ besitzt mit $x_U \in U$ und $x - x_U \in U^\perp$. Das beweist $V = U \oplus U^\perp$.

(iv) Die Linearität der Abbildung $P: x \mapsto x_U$ folgt aus der Charakterisierung (ii). Aus der Orthogonalität der Zerlegung $V = U \oplus U^\perp$ folgt $\|x_U\| \leq \|x\|$ und somit die Stetigkeit von P . \square

589 / 862

Satz (Darstellungssatz von Riesz)

Sei V ein Hilbertraum über \mathbb{K} ($= \mathbb{R}$ oder \mathbb{C}) und V' der Vektorraum aller **stetigen** linearen Abbildungen $V \rightarrow \mathbb{K}$.

Dann ist durch

$$\phi(x) := \langle \cdot, x \rangle, \quad x \in V,$$

ein konjugiert-linearer Isomorphismus normierter Vektorräume $\phi: V \rightarrow V'$ gegeben.

Beweis.

Es ist klar, dass ϕ (konjugiert-)linear ist, denn das (Hermitesche) Skalarprodukt ist (konjugiert-)linear im zweiten Argument.

Die Stetigkeit der Linearform $\phi(x): V \rightarrow \mathbb{K}$ ergibt sich aus der Cauchy-Schwarzschen Ungleichung:

$$|\phi(x)(y)| = |\langle y, x \rangle| \leq \|x\| \cdot \|y\| \quad \text{für alle } y \in V.$$

Daraus folgt $\|\phi(x)\| \leq \|x\|$.

590 / 862

Weiter im Beweis:

$$|\phi(x)(y)| = |\langle y, x \rangle| \leq \|x\| \cdot \|y\| \quad \text{für alle } y \in V.$$

Daraus folgt $\|\phi(x)\| \leq \|x\|$. Für $y = x$ erhalten wir Gleichheit:

$\|\phi(x)\| = \|x\|$. Da ϕ die Norm erhält, ist ϕ injektiv.

Es bleibt noch die Surjektivität von ϕ zu zeigen.

Sei dazu $0 \neq \alpha \in V'$. Wähle $v \in V$ mit $\alpha(v) = 1$. Aufgrund der Stetigkeit von α ist $U := \ker \alpha \subseteq V$ abgeschlossen und daher vollständig im Hilbertraum V .

Wir zerlegen $V = U \oplus U^\perp$ und betrachten die Orthogonalprojektion $v_U \in U$ von v auf U . Wir setzen $x_0 := v - v_U \in U^\perp$.

Dann gilt $\alpha(x_0) = \alpha(v) - \alpha(v_U) = \alpha(v) = 1$ und für alle $y \in V$ ist daher

$$y = \underbrace{(y - \alpha(y)x_0)}_{\in \ker \alpha = U} + \alpha(y)x_0 \in U \oplus U^\perp.$$

Somit $\langle y, x_0 \rangle = \alpha(y)\|x_0\|^2$ und daher $\alpha = \phi\left(\frac{x_0}{\|x_0\|^2}\right)$. □

591 / 862

Separabilität

Definition

Sei (X, d) ein metrischer Raum.

- Eine Teilmenge $Y \subseteq X$ heißt **dicht in X** , falls ihr Abschluss gleich X ist, d.h. $\overline{Y} = X$.
Äquivalent dazu ist: für jedes $x \in X$ und jedes $\varepsilon > 0$ gibt es $y \in Y$ mit $d(x, y) < \varepsilon$.
- (X, d) heißt **separabel**, falls es eine **abzählbare** dichte Teilmenge in X gibt.

Beispiel

Für jedes n ist \mathbb{Q}^n dicht in \mathbb{R}^n ;

jeder endlich-dimensionale reelle oder komplexe Vektorraum ist daher separabel.

Bemerkung

Der Hilbertraum ℓ^2 ist separabel, denn:

Die Folgen $x \in \ell^2$ mit $x_n \in \mathbb{Q} + i\mathbb{Q} \subseteq \mathbb{C}$ für alle $n \in \mathbb{N}$ und $x_n = 0$ für fast alle $n \in \mathbb{N}$ bilden eine abzählbare dichte Teilmenge S , d.h. jedes Element $x \in \ell^2$ ist Grenzwert einer Folge $x^{(k)} \in S$.

Weiterhin gilt:

Jeder unendlich-dimensionale **separable** komplexe Hilbertraum $(V, \langle \cdot, \cdot \rangle_V)$ ist isomorph zu ℓ^2 .

Genauer: Es gibt einen Isomorphismus von Vektorräumen $\phi: \ell^2 \rightarrow V$, der

$$\langle \phi(x), \phi(y) \rangle_V = \langle x, y \rangle_{\ell^2}$$

für alle $x, y \in \ell^2$ erfüllt.

Für den Beweis dieser Isomorphie benutzt man Hilbertbasen (s.u.: Jede Hilbertbasis $B = (b_1, b_2, \dots)$ von V definiert einen solchen Isomorphismus $\phi_B: \sum_k x_k e_k \mapsto \sum_k x_k b_k$).

593 / 862

Sei $\langle \cdot, \cdot \rangle$ eine symmetrische Bilinearform oder Hermitesche Form auf einem \mathbb{K} -Vektorraum V . Eine Teilmenge $U \subseteq V$ heißt **orthonormale Familie**, wenn $\langle u, u \rangle = 1$ und $\langle u, v \rangle = 0$ für alle $u, v \in U$ mit $u \neq v$.

Satz

Sei V ein **separabler** Euklidischer oder Hermitescher Vektorraum.

Dann ist jede orthonormale Familie $(v_i)_{i \in I}$ **abzählbar**.

Beweis.

Sei $(v_i)_{i \in I}$ eine orthonormale Familie. Für alle $i, j \in I$ mit $i \neq j$ gilt $\|v_i - v_j\| = \sqrt{2}$.

Sei $A \subseteq V$ eine abzählbare dichte Teilmenge.

Zu jedem $i \in I$ existiert ein Element $a(i) \in A$ mit $\|v_i - a(i)\| < \frac{\sqrt{2}}{2}$.

Wir zeigen, dass die Zuordnung $I \rightarrow A, i \mapsto a(i)$, injektiv ist:

$$\begin{aligned} \|a(i) - a(j)\| &= \|v_i - v_j + a(i) - v_i + v_j - a(j)\| \\ &\geq \|v_i - v_j\| - \|a(i) - v_i\| - \|v_j - a(j)\| > 0 \end{aligned}$$

für alle $i, j \in I$ mit $i \neq j$.

□

594 / 862

Satz (Besselsche Ungleichung)

Sei V ein Euklidischer oder Hermitescher Vektorraum und (v_1, v_2, \dots) eine (endliche oder abzählbar unendliche) orthonormale Familie.

Dann gilt für alle $x \in V$:

$$\sum_k |\langle x, v_k \rangle|^2 \leq \|x\|^2.$$

Ist V **vollständig**, so besitzt die Reihe $\sum_k \langle x, v_k \rangle v_k$ in V einen Grenzwert x_∞ und

$$x_\infty = x \iff \sum_k |\langle x, v_k \rangle|^2 = \|x\|^2.$$

595 / 862

Beweis. Wir betrachten für $N \in \mathbb{N}$ den endlich-dimensionalen (und somit vollständigen) Unterraum

$$U_N := \text{span}\{v_1, v_2, \dots, v_N\} \subseteq V.$$

Sei $x_N \in U_N$ die Orthogonalprojektion von x auf den Unterraum U_N .

Es gilt $x_N = \sum_{k=1}^N \langle x, v_k \rangle v_k$, denn $x - \sum_{k=1}^N \langle x, v_k \rangle v_k \perp U_N$.

Aus der orthogonalen Zerlegung $x = x_N + (x - x_N)$ folgt dann

$$\sum_{k=1}^N |\langle x, v_k \rangle|^2 = \|x_N\|^2 \leq \|x_N\|^2 + \|x - x_N\|^2 = \|x\|^2.$$

Daraus folgt die Besselsche Ungleichung und die absolute Konvergenz der Reihe $x_\infty := \sum_k \langle x, v_k \rangle v_k$ mit Werten in V .

Für x_∞ gilt $x_\infty \perp x - x_\infty$ und wegen $\|x\|^2 = \|x_\infty\|^2 + \|x - x_\infty\|^2$

$$\|x_\infty\|^2 = \|x\|^2 \iff \|x - x_\infty\|^2 = 0 \iff x = x_\infty.$$



596 / 862

Definition

Sei V ein Euklidischer oder Hermitescher Vektorraum.

Eine abzählbare orthonormale Familie (v_1, v_2, \dots) heißt **Hilbertbasis**, wenn jeder Vektor $x \in V$ eine Darstellung als Reihe $x = \sum_k \langle x, v_k \rangle v_k$ besitzt.

Beispiele

(i) Falls $\dim V < \infty$, so sind Hilbertbasen nichts anderes als unitäre bzw. orthonormale Basen.

(ii) Sei $V = \ell^2$ der Hilbertraum der quadratisch summierbaren Folgen $x: \mathbb{N} \rightarrow \mathbb{C}$.

Die Folgen $e_j \in \ell^2$ mit $e_j(k) := \delta_{jk}$ ($j, k \in \mathbb{N}$) bilden eine Hilbertbasis $(e_j)_{j \in \mathbb{N}}$, die sogenannte **kanonische Hilbertbasis** von ℓ^2 .

$(e_j)_{j \in \mathbb{N}}$ ist **keine Vektorraumbasis**, denn die lineare Hülle der e_j ist

$$\{x: \mathbb{N} \rightarrow \mathbb{C} \mid x_k := x(k) = 0 \text{ für fast alle } k \in \mathbb{N}\} \subsetneq \ell^2.$$

(Bei der linearen Hülle sind nur **endliche** Linearkombinationen zugelassen!)

597 / 862

Satz

Für eine abzählbare orthonormale Familie $B = (b_1, b_2, \dots)$ von Vektoren eines Euklidischen oder Hermiteschen Vektorraum V sind folgende Aussagen äquivalent:

- (i) $\text{span}\{b_1, b_2, \dots\}$ ist dicht in V .
- (ii) B ist eine Hilbertbasis.
- (iii) Für alle $x, y \in V$ gilt

$$\langle x, y \rangle = \sum_k \langle x, b_k \rangle \overline{\langle y, b_k \rangle}.$$

- (iv) Für alle $x \in V$ gilt die **Parsevalsche Gleichung**

$$\|x\|^2 = \sum_k |\langle x, b_k \rangle|^2.$$

Beweis.(i) \Rightarrow (ii):Zu $x \in V$ existiert eine Folge $x_i \in \text{span}\{b_1, b_2, \dots\}$ mit

$$x = \lim_{i \rightarrow \infty} x_i.$$

Wir wählen N_i so, dass

$$x_i \in U_i := \text{span}\{b_1, \dots, b_{N_i}\},$$

wobei wir N_i jeweils so groß wählen können, dass $N_i \in \mathbb{N}$ sogar eine monoton wachsende Folge ist.Für die orthogonale Projektion $x'_i = \sum_{k=1}^{N_i} \langle x, b_k \rangle b_k$ von x in U_i gilt

$$\|x - x'_i\| \leq \|x - x_i\| \quad \text{da } x_i \in U_i$$

und somit

$$x = \lim_{i \rightarrow \infty} x'_i = \sum_k \langle x, b_k \rangle b_k.$$

599 / 862

Weiter im Beweis:(ii) \Rightarrow (iii):Wenn $B = (b_1, b_2, \dots)$ eine Hilbertbasis ist, so können wir $x, y \in V$ schreiben als Reihen

$$x = \sum_i x_i b_i, \quad y = \sum_j y_j b_j,$$

wobei $x_i := \langle x, b_i \rangle$, $y_j := \langle y, b_j \rangle$.

Daraus folgt wegen der Stetigkeit des Skalarprodukts

$$\begin{aligned} \langle x, y \rangle &= \left\langle \sum_i x_i b_i, y \right\rangle = \sum_i x_i \langle b_i, y \rangle \\ \langle b_i, y \rangle &= \left\langle b_i, \sum_j y_j b_j \right\rangle = \sum_j \bar{y}_j \langle b_i, b_j \rangle = \bar{y}_i \end{aligned}$$

und somit

$$\langle x, y \rangle = \sum_i x_i \bar{y}_i.$$

600 / 862

Weiter im Beweis:

(iii) \Rightarrow (iv): Setze $x = y$.

(iv) \Rightarrow (i): Aus der Parsevalschen Gleichung folgt (nach dem Satz über die Besselsche Ungleichung), dass $x = \sum_k \langle x, b_k \rangle b_k$ für alle $x \in V$.

Damit gilt (i).

□

601 / 862

Satz (Hilbertbasen)

Jeder unendlichdimensionale **separable** Hilbertraum V besitzt eine **abzählbare** unendliche Hilbertbasis $B = (b_1, b_2, \dots)$.

Ist eine höchstens abzählbare orthonormale Familie F_0 in V vorgegeben, so kann B so gewählt werden, dass $F_0 \subseteq B$.

Zum Beweis des Satzes.

Sei F_0 die gegebene orthonormale Familie in einem separablen Hilbertraum. Sei X die durch \subseteq geordnete Menge aller orthonormalen Familien, die die Familie F_0 enthalten. Es gibt in X eine **maximale** Familie $B = (b_1, b_2, \dots)$, d.h. zu B gibt es keine Familie B' in X mit $B' \supsetneq B$.

Dieses B ist eine Hilbertbasis: Sei $y := x - \sum_k \langle x, b_k \rangle b_k$. Dann ist $\langle y, b_i \rangle = 0$ für alle $i \in \mathbb{N}$ und aufgrund der Maximalität von B muss $y = 0$ sein. □

602 / 862

[Für den Beweis der Existenz der maximalen Familie B wird das Zornsche Lemma benötigt, das aus dem Auswahlaxiom der Mengenlehre folgt. Mit dem Zornschen Lemma beweist man auch, dass jeder Vektorraum eine Basis besitzt. Zunächst:

Definition

Eine Relation \leq auf einer Menge X heißt **Ordnungsrelation**, wenn folgenden Eigenschaften für alle $x, y, z \in X$ erfüllt sind:

$$x \leq x,$$

aus $x \leq y$ und $y \leq x$ folgt $x = y$ und

aus $x \leq y$ und $y \leq z$ folgt $x \leq z$.

Die Ordnungsrelation heißt **total**, wenn für alle $x, y \in X$ eine der beiden Ungleichungen $x \leq y$ oder $y \leq x$ gilt.

$x \in X$ heißt **maximal**, wenn es kein $y \in X$ mit $x < y$ gibt (d.h. mit $x \leq y$ und $x \neq y$).

$x \in X$ heißt eine **obere Schranke** einer Teilmenge $Y \subseteq X$, falls $y \leq x$ für alle $y \in Y$.

603 / 862

Beispiel

Betrachte auf $X := \mathbb{N} \setminus \{1\}$ die Relation $n \leq m$, wenn die natürliche Zahl m die Zahl n teilt.

Dies ist eine Ordnungsrelation, die nicht total ist.

$p \in X$ ist maximal, wenn es keine natürliche Zahl $m \neq 1$ gibt, die p teilt und von p verschieden ist, also wenn p prim ist. Alle Primzahlen sind also maximal bezüglich dieser nicht-totalen Ordnung.

604 / 862

Lemma (Zornsches Lemma)

Auf der Menge X sei eine Ordnungsrelation \leq gegeben, derart dass jede total geordnete Teilmenge $Y \subseteq X$ eine obere Schranke besitzt. Dann hat X ein maximales Element. (Dieses ist i.a. nicht eindeutig.) \square

Erläuterung zum Beweis des Satzes (Hilbertbasen):

Man beachte, dass für jede bezüglich \subseteq total geordnete Teilmenge $Y \subseteq X$ die Vereinigung $\cup_{F \in Y} F$ eine obere Schranke bildet, und daher das Zornsche Lemma anwendbar ist.

]

605 / 862

Fourier-Reihen

Definition

Eine auf ganz \mathbb{R} definierte reell- oder komplexwertige Funktion f heißt **periodisch** mit Periode $L > 0$, wenn

$$f(x + L) = f(x) \quad \text{für alle } x \in \mathbb{R}.$$

- Es ist dann auch $f(x + nL) = f(x)$ für alle $n \in \mathbb{Z}$.
- Hat die Funktion f die Periode L , so hat

$$F(x) := f\left(\frac{L}{2\pi}x\right)$$

die Periode 2π .

Wir beschränken uns im Folgenden auf solche Funktionen, und bezeichnen mit V den Vektorraum der 2π -periodischen Funktionen $f: \mathbb{R} \rightarrow \mathbb{C}$, für die $f|_{[0,2\pi]}$ (Riemann-)integrierbar ist.

606 / 862

Auf dem Vektorraum V der 2π -periodischen auf $[0, 2\pi]$ integrierbaren Funktionen betrachten wir die Hermitesche Form

$$(*) \quad \langle f, g \rangle := \frac{1}{2\pi} \int_0^{2\pi} f(x) \overline{g(x)} dx, \quad f, g \in V.$$

(Vgl. auch MfP I: Integrierbarkeit von Produkten).

Die hermitesche Form $(*)$ ist positiv semi-definit, d.h. $\langle f, f \rangle \geq 0$ für alle $f \in V$. Auf dem Unterraum der stetigen Funktionen ist sie positiv definit und somit ein Skalarprodukt. Wir schreiben $\|f\|_2 := \sqrt{\langle f, f \rangle}$ (auch im semi-definiten Fall).

Satz

Die Funktionen $e_k(x) := \exp(ikx)$, $k \in \mathbb{Z}$, bilden ein orthonormales System bezüglich der Hermiteschen Form $(*)$.

(Beweis: MfP I)

607 / 862

Definition

Ein **trigonometrisches Polynom** vom Grad $\leq n$ ist eine Linearkombination

$$p(x) = \sum_{k=-n}^n \gamma_k e^{ikx} \quad \text{mit} \quad \gamma_k \in \mathbb{C}.$$

Es gilt $p \in V$ und $\gamma_k = \langle p, e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} p(x) e^{-ikx} dx$.

Definition

Es sei $f: \mathbb{R} \rightarrow \mathbb{C}$ eine 2π -periodische Funktion, so dass $f|_{[0, 2\pi]}$ integrierbar ist (kurz: $f \in V$). Die komplexe Zahl

$$c_k := \langle f, e_k \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx$$

heißt **k -ter Fourier-Koeffizient** und das trigonometrische Polynom

$F_n(f) := \sum_{k=-n}^n c_k e_k$ heißt **n -tes Fourier-Polynom** von f .

Die Reihe $(F_n(f))_{n=0,1,\dots}$, also die Folge der Partialsummen, heißt **Fourier-Reihe** von f .

608 / 862

Bemerkung.

Es gilt

$$F_n(f) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx),$$

wobei

$$a_k = c_k + c_{-k} = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx$$

$$b_k = i(c_k - c_{-k}) = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx.$$

Wenn f reellwertig ist, dann ist $c_{-k} = \bar{c}_k$ und auch die Fourierreihe $F_n(f)$ ist reellwertig.

609 / 862

Ein Beispiel

Wir bestimmen die reelle Fourierreihe der (periodischen Fortsetzung der) Funktion $f : [-\pi, \pi] \rightarrow \mathbb{R}$, $f(x) = \cos(ax)$ mit $a \notin \mathbb{Z}$.

- (ÜA) f ist gerade, also verschwinden die Fourierkoeffizienten b_k .
- Es gilt

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(ax) \cos(kx) dx.$$

Insbesondere erhalten wir

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(ax) dx = \frac{1}{a\pi} \sin(ax) \Big|_{-\pi}^{\pi} = \frac{2 \sin(a\pi)}{a\pi}.$$

- (ÜA) Aus den Additionstheoremen folgt

$$\cos(ax) \cos(kx) = \frac{1}{2} (\cos((a+k)x) + \cos((a-k)x)).$$

610 / 862

(Fortsetzung)

- Für das Integral ergibt sich also

$$\begin{aligned}
 a_k &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \cos((a+k)x) + \cos((a-k)x) dx \\
 &= \frac{1}{2\pi} \left(\frac{1}{a+k} \sin((a+k)x) + \frac{1}{a-k} \sin(a-k)x \right) \Big|_{-\pi}^{\pi} \\
 &= \frac{1}{2\pi} (-1)^k \left(\frac{2}{a+k} + \frac{2}{a-k} \right) \sin(a\pi) \\
 &= (-1)^k \frac{2a}{\pi(a^2 - k^2)} \sin(a\pi).
 \end{aligned}$$

- Damit ergibt sich die Fourierreihe

$$F(f) = \frac{\sin(a\pi)}{a\pi} + \sum_{k=1}^{\infty} \frac{(-1)^k 2a \sin(a\pi)}{\pi(a^2 - k^2)} \cos(kx).$$

611 / 862

Ähnlich wie bei einer Taylorreihe ist nicht garantiert, dass die Fourierreihe einer Funktion f konvergiert und dass sie im Fall der Konvergenz gegen f konvergiert.

Wenn aber $f \in V$ sich in der Form

$$f(x) = \sum_{k=-\infty}^{\infty} \gamma_k e_k$$

mit einer **gleichmäßig** konvergenten Reihe darstellen lässt, muss diese Reihe die Fourierreihe sein, denn man kann Integration und Limesbildung bei gleichmäßiger Konvergenz vertauschen und findet für den Fourierkoeffizienten:

$$\begin{aligned}
 c_k &= \frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{m=-\infty}^{\infty} \gamma_m e_m \right) e_{-k} dx \\
 &= \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \int_0^{2\pi} \gamma_m e_{m-k} dx = \gamma_k.
 \end{aligned}$$

Im Allgemeinen konvergiert aber die Fourierreihe von f weder gleichmäßig noch punktweise gegen f . Ein anderer Konvergenzbegriff auf V ist angemessen.

612 / 862

Aus der allgemeinen Theorie erhalten wir folgende Aussagen:

Satz (Fourier-Polynome und Besselsche Ungleichung)

- (i) Das Fourierpolynom $F_n(f)$ ist die beste Approximation **im quadratischen Mittel** an $f \in V$ durch ein trigonometrisches Polynom vom Grad $\leq n$, d.h. $\|f - F_n(f)\|_2 = \inf_{g \in V_n} \|f - g\|_2$, wobei $V_n := \text{span}\{e_k \mid -n \leq k \leq n\}$.
- (ii) Für die Fourier-Koeffizienten c_k von $f \in V$ gilt die Besselsche Ungleichung

$$\sum_{k=-\infty}^{\infty} |c_k|^2 \leq \|f\|_2^2$$

und Gleichheit gilt genau dann, wenn

$$\lim_{n \rightarrow \infty} \|f - F_n(f)\|_2 = 0.$$

613 / 862

Beweis. Teil (i) folgt aus der Tatsache, dass $F_n(f)$ gerade die orthogonale Projektion von f auf den Unterraum $V_n = \text{span}\{e_k \mid -n \leq k \leq n\}$ ist, denn für $-n \leq k \leq n$ gilt

$$\langle f - F_n(f), e_k \rangle = \langle f, e_k \rangle - \sum_{\ell=-n}^n \langle f, e_\ell \rangle \underbrace{\langle e_\ell, e_k \rangle}_{=\delta_{\ell k}} = 0.$$

Teil (ii) ist nur noch einmal eine Wiederholung der Aussage des Satzes zur Besselschen Ungleichung in unserer konkreten Situation, d.h. mit der orthonormalen Familie $\{\exp(ikx)\}_{k \in \mathbb{Z}}$. \square

614 / 862

Wir wollen nun zeigen, dass für *alle* $f \in V$ in der Besselschen Ungleichung Gleichheit gilt (Parsevalsche Gleichung).

Satz (Konvergenz der Fourier-Reihe im quadratischen Mittel)

Sei V der Vektorraum der 2π -periodische Funktionen $f: \mathbb{R} \rightarrow \mathbb{C}$, für die $f|_{[0,2\pi]}$ (Riemann-)integrierbar ist.

(i) Für alle $f \in V$ gilt

$$\|f\|_2^2 = \sum_{k=-\infty}^{\infty} |c_k|^2.$$

(ii) Die Fourier-Reihe von $f \in V$ konvergiert im quadratischen Mittel gegen f , d.h. $\lim_{n \rightarrow \infty} \|f - F_n(f)\|_2 = 0$.

Zum Beweis: Aus dem vorhergehenden Satz wissen wir, dass (i) und (ii) äquivalent sind. Es genügt also, (ii) zu zeigen.

Für den Beweis benötigen wir eine Reihe von Vorbereitungen.

615 / 862

Lemma

Für jede stetig differenzierbare Funktion $f: [a, b] \rightarrow \mathbb{R}$ gilt mit $r \in \mathbb{R}$:

$$\lim_{|r| \rightarrow \infty} \int_a^b f(t) \sin rt \, dt = 0.$$

Beweis. partielle Integration. □

Lemma

Für alle $x \in (0, 2\pi)$ gilt

$$(*) \quad \sum_{k=1}^{\infty} \frac{\sin kx}{k} = \frac{\pi - x}{2},$$

und die Konvergenz der Funktionenreihe ist auf jedem abgeschlossenen Teilintervall $[\delta, 2\pi - \delta]$ gleichmäßig.

616 / 862

Beweis. Wir zeigen zunächst die gleichmäßige Konvergenz auf abgeschlossenen Teilintervallen. Dazu fixieren wir $\delta > 0$, und betrachten

$$s_n(x) := \sum_{k=1}^n \sin kx = \operatorname{Im} \left(\sum_{k=1}^n e^{ikx} \right).$$

Dies können wir für $x \in [\delta, 2\pi - \delta]$ abschätzen als

$$\begin{aligned} |s_n(x)| &\leq \left| \sum_{k=1}^n e^{ikx} \right| = \left| e^{ix} \frac{1 - e^{inx}}{1 - e^{ix}} \right| \\ &= \left| e^{i\frac{x}{2}} \frac{1 - e^{inx}}{e^{-ix/2} - e^{ix/2}} \right| \leq \frac{2}{|e^{ix/2} - e^{-ix/2}|} \leq \frac{1}{\sin \delta/2}. \end{aligned}$$

617 / 862

Aus $|s_n(x)| \leq \frac{1}{\sin \delta/2} =: \sigma$ erhalten wir weiter

$$\begin{aligned} \left| \sum_{k=n}^m \frac{\sin kx}{k} \right| &= \left| \sum_{k=n}^m \frac{s_k(x) - s_{k-1}(x)}{k} \right| \\ &= \left| \sum_{k=n}^m s_k(x) \left(\frac{1}{k} - \frac{1}{k+1} \right) + \frac{s_m(x)}{m+1} - \frac{s_{n-1}(x)}{n} \right| \\ &\leq \sum_{k=n}^m |s_k(x)| \left(\frac{1}{k} - \frac{1}{k+1} \right) + \frac{|s_m(x)|}{m+1} + \frac{|s_{n-1}(x)|}{n} \\ &\leq \sigma \left(\frac{1}{n} - \frac{1}{m+1} + \frac{1}{m+1} + \frac{1}{n} \right) = \frac{2\sigma}{n}. \end{aligned}$$

Für $m \rightarrow \infty$ erhalten wir daraus $\left| \sum_{k=n}^{\infty} \frac{\sin kx}{k} \right| \leq \frac{2\sigma}{n}$ für $x \in [\delta, 2\pi - \delta]$, was die gleichmäßige Konvergenz der Reihe auf jedem solchen Intervall zeigt.

618 / 862

Es bleibt noch, den Wert der Reihe für festes $x \in (0, 2\pi)$ zu bestimmen.

Wir wissen $\frac{\sin kx}{k} = \int_{\pi}^x \cos kt \, dt$ und

$$\begin{aligned} \sum_{k=1}^n \cos kt &= \frac{1}{2} \sum_{k=1}^n (e^{ikt} + e^{-ikt}) = \frac{1}{2} \sum_{k=-n}^n e^{ikt} - \frac{1}{2} \\ &= \frac{e^{-int}(1 - e^{i(2n+1)t})}{2(1 - e^{it})} - \frac{1}{2} = \frac{\sin(n + \frac{1}{2})t}{2 \sin \frac{t}{2}} - \frac{1}{2}. \end{aligned}$$

Wir wenden nun das vorherige Lemma mit $a = \pi$, $b = x$ und $f(t) = \frac{1}{2 \sin \frac{t}{2}}$ an und erhalten

$$\sum_{k=1}^{\infty} \frac{\sin kx}{k} = \lim_{n \rightarrow \infty} \int_{\pi}^x \frac{\sin(n + \frac{1}{2})t}{2 \sin \frac{t}{2}} dt - \frac{x - \pi}{2} = \frac{\pi - x}{2}.$$

□

619 / 862

Folgerung

Für alle $x \in [0, 2\pi]$ gilt

$$(**) \quad F(x) := \sum_{k=1}^{\infty} \frac{\cos kx}{k^2} = \left(\frac{x - \pi}{2} \right)^2 - \frac{\pi^2}{12}.$$

Insbesondere gilt

$$F(0) = \sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Beweis. Wegen $|\cos(kx)| \leq 1$ hat die Reihe die Majorante $\sum_{k=1}^{\infty} \frac{1}{k^2}$, konvergiert also auf ganz \mathbb{R} gleichmäßig. Insbesondere ist die Grenzfunktion F überall stetig.

Wir erinnern uns nun an einen Satz aus dem letzten Semester: konvergiert eine Funktionenreihe in mindestens einem Punkt, und konvergiert die Reihe der Ableitungen gleichmäßig, so konvergiert die Reihe gegen eine differenzierbare Grenzfunktion, und die Reihe der Ableitungen konvergiert gegen die Ableitung der Grenzfunktion.

620 / 862

In unserem Fall gilt

$$\left(\frac{\cos(kx)}{k^2}\right)' = -\frac{\sin(kx)}{k}.$$

Nach dem letzten Lemma konvergiert die Reihe der Ableitungen somit auf jedem Teilintervall $[-\delta, 2\pi - \delta] \subset (0, 2\pi)$ gleichmäßig. F ist also auf dem Intervall $(0, 2\pi)$ differenzierbar mit

$$F'(x) = -\sum_{k=1}^{\infty} \frac{\sin(kx)}{k} \stackrel{(*)}{=} \frac{x - \pi}{2}.$$

Integration liefert nun für $x \in (0, 2\pi)$ die Identität

$$F(x) = \sum_{k=1}^{\infty} \frac{\cos kx}{k^2} = \frac{(x - \pi)^2}{4} + c$$

mit einer geeigneten Konstante $c \in \mathbb{R}$. Da F auf ganz \mathbb{R} stetig ist, setzt sich diese Formel auf das abgeschlossene Intervall $[0, 2\pi]$ fort.

621 / 862

Mit nochmaliger gliedweiser Integration erhalten wir nun

$$0 = \int_0^{2\pi} F(x) dx = \left. \frac{(x - \pi)^3}{12} \right|_0^{2\pi} + 2\pi c = \frac{\pi^3}{6} + 2\pi c,$$

also $c = -\frac{\pi^2}{12}$, und damit gilt die Behauptung (**).

Setzen wir schließlich $x = 0$ ein, so vereinfacht sich die Behauptung zu

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{4} - \frac{\pi^2}{12} = \frac{\pi^2}{6}.$$

□

Nach diesen Vorbereitungen kommen wir nun zum eigentlichen Beweis des Satzes. Diesen werden wir aus folgendem Spezialfall gewinnen:

Lemma über Treppenfunktionen

Sei $f \in V$ so, dass $f|_{[0, 2\pi]}$ eine **Treppenfunktion** ist.

Dann konvergiert die Fourier-Reihe $F_n(f)$ im quadratischen Mittel gegen f .

622 / 862

Beweis des Satzes aus diesem Lemma:

$f|_{[0,2\pi]}$ ist als Riemann-integrierbare Funktion beschränkt. Wir können o.B.d.A. annehmen, dass f reellwertig ist und $|f| \leq 1$.

Dann existieren zu $\varepsilon > 0$ Treppenfunktionen φ und ψ mit $-1 \leq \varphi \leq f \leq \psi \leq 1$, so dass

$$\frac{1}{2\pi} \int_0^{2\pi} (\psi(x) - \varphi(x)) dx \leq \frac{\varepsilon^2}{8}.$$

Für $0 \leq y \leq 2$ gilt $y^2 \leq 2y$. Wir wenden diese Formel auf die Differenz der Treppenfunktionen $y = \psi - \varphi$ an und erhalten

$$\|f - \varphi\|_2^2 \leq \|\psi - \varphi\|_2^2 \leq \frac{1}{2\pi} \int_0^{2\pi} 2(\psi - \varphi) \leq \frac{\varepsilon^2}{4},$$

und somit $\|f - \varphi\|_2 \leq \frac{\varepsilon}{2}$.

Wegen des Lemmas über Treppenfunktionen existiert für die Treppenfunktion φ ein $N \in \mathbb{N}$, so dass $\|\varphi - F_n(\varphi)\|_2 < \frac{\varepsilon}{2}$ für alle $n \geq N$.

Mit $g := f - \varphi$ gilt $\|g - F_n(g)\|_2 \leq \|g\|_2 \leq \frac{\varepsilon}{2}$ und somit

$$\|f - F_n(f)\|_2 \leq \|g - F_n(g)\|_2 + \|\varphi - F_n(\varphi)\|_2 < \varepsilon.$$

□

623 / 862

Beweis des Lemmas über Treppenfunktionen:

Wir können annehmen, dass $f|_{[0,2\pi]}$ eine Treppenfunktion mit $f|_{(0,a)} = 1$ und $f|_{(a,2\pi)} = 0$ ist, denn jede Treppenfunktion lässt sich als Linearkombination solcher Funktionen darstellen.

Als Fourier-Koeffizienten erhalten wir $c_0 = \frac{a}{2\pi}$ und

$$c_k = \frac{1}{2\pi} \int_0^a e^{-ikx} dx = \frac{i}{2\pi k} e^{-ikx} \Big|_0^a = \frac{i}{2\pi k} (e^{-ika} - 1), \quad k \in \mathbb{N}, k \neq 0.$$

Mit $|c_k|^2 = \frac{(e^{-ika} - 1)(e^{ika} - 1)}{4\pi^2 k^2} = \frac{1 - \cos ka}{2\pi^2 k^2}$ (für $k \neq 0$) erhalten wir

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

Weiter im Beweis:

Mit Hilfe der Formel

$$\sum_{k=1}^{\infty} \frac{\cos kx}{k^2} = \left(\frac{x - \pi}{2} \right)^2 - \frac{\pi^2}{12}$$

folgt daraus

$$\sum_{k=-\infty}^{\infty} |c_k|^2 = \frac{a^2}{4\pi^2} + \frac{1}{6} - \frac{1}{\pi^2} \left(\left(\frac{a - \pi}{2} \right)^2 - \frac{\pi^2}{12} \right) = \frac{a}{2\pi} = \|f\|_2^2.$$

Also konvergiert $F_n(f)$ im quadratischen Mittel gegen die Treppenfunktion f . □

Damit ist der Beweis des Satzes abgeschlossen.

625 / 862

Bemerkungen

- Aus der Besselschen Ungleichung folgt, dass die Folgen c_k, c_{-k} quadratisch summierbare komplexe Zahlenfolgen sind, d.h. Elemente von $\ell^2(\mathbb{C})$.
- Aus der Parsevalschen Gleichung folgt die Vollständigkeit des Orthonormalsystems $e_k(x), k \in \mathbb{Z}$.
- Die $e_k(x), k \in \mathbb{Z}$ bilden eine Hilbertbasis des **Hilbertraums** der L^2 -Funktionen auf $[0, 2\pi]$, den wir in Teil III der Vorlesung einführen werden.

Im Verlauf des Beweises haben wir auch schon die Fourierreihen der (periodisch fortgesetzten) Funktionen $\frac{\pi-x}{2}$ und $\frac{(x-\pi)^2}{2}$ auf dem Intervall $[0, 2\pi]$ bestimmt.

626 / 862

Anders als bei der Taylorreihe können auch periodische Funktionen dargestellt werden, die nur **stückweise stetig differenzierbar** sind.

Satz (Fourierentwicklung)

Sei $f \in V$ stetig und stückweise C^1 , d.h. es gibt eine Unterteilung $0 = t_0 < t_1 < \dots < t_r = 2\pi$, so dass $f_k := f|_{[t_{k-1}, t_k]}$ von der Klasse C^1 ist, $k = 1, \dots, r$.

Dann konvergiert die Fourierreihe $F_n(f)$ **gleichmäßig** gegen f .

Beweis.

Sei φ die periodische Funktion, die durch $\varphi|_{[t_{k-1}, t_k)} = f'_k$ definiert wird.

Die Fourier-Koeffizienten c_n von f ergeben sich durch partielle Integration aus den Fourier-Koeffizienten γ_n von φ :

$$\int_{t_{k-1}}^{t_k} f(x) e^{-inx} dx = \frac{i}{n} f(x) e^{-inx} \Big|_{t_{k-1}}^{t_k} - \frac{i}{n} \int_{t_{k-1}}^{t_k} \varphi(x) e^{-inx} dx.$$

Wegen der Periodizität von f erhält man daraus durch Summation $c_n = -\frac{i}{n} \gamma_n$ ($n \neq 0$).

627 / 862

Weiter im Beweis:

Wegen der allgemeinen Ungleichung $|ab| \leq \frac{1}{2}(|a|^2 + |b|^2)$ folgt nun

$$|c_n| \leq \frac{1}{2} \left(\frac{1}{n^2} + |\gamma_n|^2 \right).$$

Da $\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$ und $\sum_{n=-\infty}^{\infty} |\gamma_n|^2 < \infty$ (Besselsche Ungleichung) ist $\sum |c_n|$ (von x unabhängige) absolut konvergente Majorante der Fourier-Reihe, und daraus folgt die absolute und gleichmäßige Konvergenz der Fourier-Reihe $F_n(f)$ gegen eine stetige Grenzfunktion g .

Andererseits konvergiert $F_n(f)$ im quadratischen Mittel gegen f .

Das ist nur möglich, wenn $f = g$, denn beide Funktionen sind stetig. \square

Kapitel 18

Differentialrechnung mehrerer Veränderlicher

629 / 862

Differentialrechnung mehrerer Veränderlicher Richtungsableitung

Zur Erinnerung:

Eine Funktion $\tilde{f}: \mathbb{R} \rightarrow \mathbb{R}$ heißt **differenzierbar** in 0, wenn der Grenzwert $\lim_{t \rightarrow 0} \frac{\tilde{f}(t) - \tilde{f}(0)}{t}$ existiert.

Die **Ableitung** $\tilde{f}'(0)$ an der Stelle 0 ist dann gegeben durch

$$\tilde{f}'(0) = \lim_{t \rightarrow 0} \frac{\tilde{f}(t) - \tilde{f}(0)}{t} .$$

630 / 862

Definition

Sei $U \subseteq \mathbb{R}^n$ eine offene Teilmenge und $v \in \mathbb{R}^n \setminus \{0\}$ ein Vektor.

- (i) Eine Funktion $f: U \rightarrow \mathbb{R}$ heißt an der Stelle $x \in U$ **in Richtung v differenzierbar**, wenn die Funktion

$$\tilde{f}: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}, \quad t \mapsto f(x + tv),$$

an der Stelle $t = 0$ differenzierbar ist.

Hierbei ist $\varepsilon > 0$ so klein gewählt, dass $x + tv \in U$ für alle $t \in (-\varepsilon, \varepsilon)$.

- (ii) Die Zahl

$$(\partial_v f)(x) := \left. \frac{d}{dt} \right|_{t=0} f(x + tv) = \tilde{f}'(0)$$

heißt **Ableitung von f an der Stelle x in Richtung v** oder auch **Richtungsableitung**.

631 / 862

Rechenregeln für die Richtungsableitung

Satz

Seien $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$ zwei Funktionen, die beide in $x \in \mathbb{R}^n$ in Richtung $v \in \mathbb{R}^n$ differenzierbar sind. Dann gilt:

$$\partial_v(f + g)(x) = \partial_v(f)(x) + \partial_v(g)(x)$$

$$\partial_v(\lambda f)(x) = \lambda \partial_v(f)(x) \quad \text{für } \lambda \in \mathbb{R},$$

$$\partial_v(f \cdot g)(x) = \partial_v(f)(x)g(x) + f(x)\partial_v(g)(x).$$

Ist $g(x) \neq 0$, so ist $\frac{f}{g}$ in x in Richtung v differenzierbar, und es gilt:

$$\partial_v \left(\frac{f}{g} \right) (x) = \frac{\partial_v(f)(x)g(x) - f(x)\partial_v(g)(x)}{g(x)^2}.$$

Der **Beweis** folgt aus den Rechenregeln für die differenzierbaren reellen Funktionen in **einer** Variablen \tilde{f} und $\tilde{g}: \mathbb{R} \rightarrow \mathbb{R}$ definiert durch $\tilde{f}(t) := f(x + tv)$, $\tilde{g}(t) = g(x + tv)$. □

632 / 862

Partielle Ableitungen

Definition

Sei (e_1, \dots, e_n) die kanonische Basis des \mathbb{R}^n , $U \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$.

- Die Funktion f heißt an der Stelle $x \in U$ **partiell differenzierbar**, wenn f an der Stelle x in alle Koordinatenrichtungen e_i differenzierbar ist.
- Die Zahl

$$(\partial_i f)(x) := (\partial_{e_i} f)(x), \quad i = 1, 2, \dots, n,$$

heißt i -te **partielle Ableitung** von f an der Stelle x .

- f heißt **partiell differenzierbar**, wenn f in allen Punkten $x \in U$ partiell differenzierbar ist.
- Die Funktion $\partial_i f: U \rightarrow \mathbb{R}$ heißt i -te **partielle Ableitung** von f .
Man schreibt dafür auch

$$\frac{\partial f}{\partial x_i}.$$

633 / 862

Bemerkung

Die i -te partielle Ableitung $(\partial_i f)(x)$ an der Stelle $x = \sum_{j=1}^n x_j e_j$ ist genau die Ableitung der Funktion $h(t) := f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n)$ an der Stelle $t = x_i$:

$$\begin{aligned} (\partial_i f)(x) &= \left. \frac{d}{dt} \right|_{t=x_i} f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_n) \\ &= \left. \frac{d}{dt} \right|_{t=0} f(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n) \\ &= h'(x_i). \end{aligned}$$

Beim Berechnen der i -ten partiellen Ableitung sind also die Variablen x_j für $j \neq i$ als Konstanten zu behandeln. Differenziert wird nach der Variablen x_i .

634 / 862

Beispiele

- Die Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = \|x\|^2 = \sum_{j=1}^n x_j^2$, ist partiell differenzierbar und

$$\frac{\partial f}{\partial x_i}(\|x\|^2) = \frac{\partial}{\partial x_i} \sum_{j=1}^n x_j^2 = 2x_i.$$

- Es gibt partiell differenzierbare Funktionen, für die trotzdem nicht alle Richtungsableitungen existieren: Wir betrachten dafür die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, gegeben als

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad \text{für } (x, y) \neq (0, 0), \quad f(0, 0) = 0.$$

Dann gilt $\frac{\partial f}{\partial x}(0) = \frac{\partial f}{\partial y}(0) = 0$, aber

$$\partial_{e_1+e_2} f(0) = \lim_{t \rightarrow 0} \frac{f(t, t)}{t} = \lim_{t \rightarrow 0} \frac{t^2}{t \cdot 2t^2} = \lim_{t \rightarrow 0} \frac{1}{2t} \quad \text{existiert nicht!}$$

635 / 862

Beispiel

Partielle Differenzierbarkeit impliziert im Allgemeinen nicht die **Stetigkeit!**

Gegenbeispiel: Wir betrachten

$$f(x, y) := \begin{cases} \frac{xy^2}{x^2+y^4}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

f ist nicht stetig in $(0, 0)$: Betrachten wir etwa $x = ay^2$, so ist $\lim_{y \rightarrow 0} f(ay^2, y) = a/(a^2 + 1)$.

Aber alle Richtungsableitungen existieren in $p_0 = (0, 0)$:

Sei $v = (v_1, v_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. In $p_0 = (0, 0)$ gilt dann für $t \neq 0$:

$$\begin{aligned} \frac{f(p_0 + tv) - f(p_0)}{t} &= \frac{f(tv)}{t} = \frac{t^3 v_1 v_2^2}{t(t^2 v_1^2 + t^4 v_2^4)} = \frac{v_1 v_2^2}{v_1^2 + t^2 v_2^4} \\ &\xrightarrow{t \rightarrow 0} \begin{cases} \frac{v_2^2}{v_1}, & v_1 \neq 0 \\ 0, & v_1 = 0 \end{cases} \end{aligned}$$

636 / 862

Höhere partielle Ableitungen

Definition

Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$.

(i) Für $k \geq 2$ definiert man rekursiv:

f heißt **einmal partiell differenzierbar**, wenn f partiell differenzierbar ist.

f heißt **k -mal partiell differenzierbar**, wenn f partiell differenzierbar ist und alle partiellen Ableitungen $\partial_i f$ $(k-1)$ -mal partiell differenzierbar sind.

(ii) f heißt **k -mal stetig differenzierbar**, wenn f k -mal partiell differenzierbar ist und alle k -ten partiellen Ableitungen

$$\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_k}} := \partial_{i_1} \partial_{i_2} \cdots \partial_{i_k} f$$

stetig sind.

637 / 862

Lemma von Schwarz

Satz (Hermann Amandus Schwarz)

Sei $f: U \rightarrow \mathbb{R}$ ($U \subseteq \mathbb{R}^n$ offen) **zweimal stetig** differenzierbar.

Dann gilt für alle $i, j \in \{1, 2, \dots, n\}$:

$$\partial_i \partial_j f = \partial_j \partial_i f.$$

Beweis. Da es nur um zwei Koordinatenrichtungen e_i und e_j geht, können wir annehmen, dass $n = 2$.

Wir berechnen die zweiten partiellen Ableitungen an der Stelle $(x_0, y_0) \in U$.

Wir können, gegebenenfalls nach Verschiebung, annehmen, dass $(x_0, y_0) = (0, 0)$ und dass die offene Menge U ein Quadrat $Q = \{(x, y) \mid |x| < \varepsilon \text{ und } |y| < \varepsilon\}$ enthält für ein geeignetes $\varepsilon > 0$.

638 / 862

Weiter im Beweis:

Betrachte für festes $y \in (-\varepsilon, \varepsilon)$ die Funktion $g_y(x) := f(x, y) - f(x, 0)$.
Nach dem **Mittelwertsatz** gibt es zu jedem x ein ξ zwischen 0 und x mit

$$(1) \quad g_y(x) - g_y(0) = xg'_y(\xi) = x(\partial_1 f(\xi, y) - \partial_1 f(\xi, 0)).$$

Halte dieses ξ fest. Die erneute Anwendung des Mittelwertsatzes auf die Funktion $y \mapsto \partial_1 f(\xi, y)$ liefert ein η zwischen 0 und y mit

$$(2) \quad \partial_1 f(\xi, y) - \partial_1 f(\xi, 0) = y\partial_2 \partial_1 f(\xi, \eta).$$

Aus (1) und (2) ergibt sich also die Existenz von ξ, η mit $|\xi| < |x|$,
 $|\eta| < |y|$ und

$$(3) \quad g_y(x) - g_y(0) = xy\partial_2 \partial_1 f(\xi, \eta).$$

Analog finden wir für die Funktion $h_x(y) := f(x, y) - f(0, y)$ Werte $\tilde{\xi}, \tilde{\eta}$
mit

$$(3') \quad h_x(y) - h_x(0) = xy\partial_1 \partial_2 f(\tilde{\xi}, \tilde{\eta}),$$

und $|\tilde{\xi}| < |x|$ und $|\tilde{\eta}| < |y|$.

639 / 862

Weiter im Beweis:

Aus (3) und (3') erhalten wir wegen

$$g_y(x) - g_y(0) = f(x, y) - f(x, 0) - f(0, y) + f(0, 0) = h_x(y) - h_x(0):$$

$$(4) \quad \partial_2 \partial_1 f(\xi, \eta) = \partial_1 \partial_2 f(\tilde{\xi}, \tilde{\eta}).$$

Die **Stetigkeit** der zweiten partiellen Ableitungen erlaubt nun den
Grenzübergang $(x, y) \rightarrow (0, 0)$ in (4), woraus sich
 $\partial_2 \partial_1 f(0, 0) = \partial_1 \partial_2 f(0, 0)$ ergibt. □

640 / 862

Der Gradient einer Funktion

Definition (Gradient)

Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ an der Stelle $x \in U$ partiell

differenzierbar. Der Spaltenvektor $\text{grad } f(x) := \begin{pmatrix} \partial_1 f(x) \\ \partial_2 f(x) \\ \vdots \\ \partial_n f(x) \end{pmatrix}$ heißt

Gradient von f an der Stelle x .

Bemerkungen

Die Rechenregeln für die Richtungsableitungen implizieren Rechenregeln für den Gradienten. Z.B. ist grad additiv und für zwei in x partiell differenzierbare Funktionen gilt als Folge der Leibnizregel

$$\text{grad}(f \cdot g)(x) = f(x) \cdot (\text{grad } g)(x) + g(x) \cdot (\text{grad } f)(x).$$

641 / 862

Beispiel

Sei $r = r(x) := \|x\|$ die Euklidische Norm von $x \in \mathbb{R}^n$ und $f: \mathbb{R}_+ \rightarrow \mathbb{R}$ eine differenzierbare Funktion.

Dann ist die rotationssymmetrische Funktion $F(x) = f(r) = f(r(x))$ auf $U = \mathbb{R}^n \setminus \{0\}$ partiell differenzierbar.

Die partiellen Ableitungen erhält man mittels der Kettenregel:

$$\begin{aligned} \partial_i F(x) &= f'(r) \partial_i r \\ \partial_i r &= \partial_i \sqrt{r^2} = \frac{1}{2r} \partial_i \|x\|^2 = \frac{x_i}{\|x\|}. \end{aligned}$$

Der Gradient der Funktion r ist daher $\text{grad } r(x) = \frac{x}{\|x\|}$ und somit

$$\text{grad } F(x) = f'(r) \text{grad } r(x) = f'(\|x\|) \frac{x}{\|x\|}.$$

642 / 862

Divergenz eines Vektorfeldes

Definition (Divergenz)

Sei $U \subseteq \mathbb{R}^n$ offen.

- (i) Ein **Vektorfeld** auf $U \subseteq \mathbb{R}^n$ ist eine Abbildung $v: U \rightarrow \mathbb{R}^n$. Man denkt sich an jeden Punkt $x \in U$ den Vektor $v(x)$ angeheftet.
- (ii) Eine Abbildung $F = \sum_{i=1}^m F_i e_i: U \rightarrow \mathbb{R}^m$ heißt im Punkt $x \in U$ **partiell differenzierbar**, wenn ihre Komponentenfunktionen F_i , $i = 1, 2, \dots, m$, in x partiell differenzierbar sind.
- (iii) Für ein in $x \in U$ partiell differenzierbares Vektorfeld $v = \sum_{i=1}^n v_i e_i: U \rightarrow \mathbb{R}^n$ ist die **Divergenz** an der Stelle x gegeben durch

$$\operatorname{div} v(x) := \sum_{i=1}^n \partial_i v_i(x).$$

Der Gradient einer partiell differenzierbaren Funktion f auf U ist ein Vektorfeld auf U .

643 / 862

Leibnizregel für die Divergenz

Satz

Sei v ein in x partiell differenzierbares Vektorfeld und f eine in x partiell differenzierbare reellwertige Funktion. Dann ist $f \cdot v$ ein in x partiell differenzierbares Vektorfeld und es gilt, mit $\langle \cdot, \cdot \rangle$ dem kanonischen Skalarprodukt,

$$\operatorname{div}(fv)(x) = \langle \operatorname{grad} f(x), v(x) \rangle + f(x) \cdot \operatorname{div} v(x)$$

Beweis. Der Beweis ergibt sich aus der Leibnizregel:

$$\begin{aligned} \operatorname{div}(f \cdot v)(x) &= \sum_{i=1}^n \partial_i (f \cdot v_i)(x) \\ &= \sum_{i=1}^n (\partial_i f(x) \cdot v_i(x) + f(x) \cdot \partial_i v_i(x)) \\ &= \langle \operatorname{grad} f(x), v(x) \rangle + f(x) \cdot \operatorname{div} v(x). \end{aligned}$$



644 / 862

Der Laplace-Operator

Definition (Laplace-Operator)

Der Gradient einer zweimal partiell differenzierbaren Funktion $f: U \rightarrow \mathbb{R}$ ist ein partiell differenzierbares Vektorfeld $\text{grad } f: U \rightarrow \mathbb{R}^n$.

Somit definiert

$$\Delta f := \text{div grad } f = \sum_{i=1}^n \partial_i^2 f$$

eine Funktion $U \rightarrow \mathbb{R}$.

- Der Differentialoperator Δ heißt **Laplace-Operator**.
- Lösungen f der (Laplace- oder auch Potential-)Gleichung

$$\Delta f = 0$$

heißen **harmonische Funktionen**.

645 / 862

Leibnizregel für den Laplace Operator

Der Laplace-Operator ist linear:

$$\Delta(\lambda f + \mu g) = \lambda \Delta(f) + \mu \Delta(g)$$

für $\lambda, \mu \in \mathbb{C}$ und f, g zweimal partiell differenzierbare Funktionen. Aus der Leibnizregel für div und grad erhalten wir eine Leibnizregel für den Laplace-Operator.

Satz

Seien f und g zwei zweimal partiell differenzierbare Funktionen. Dann gilt

$$\Delta(f \cdot g) = f \cdot \Delta g + g \cdot \Delta f + 2\langle \text{grad } f, \text{grad } g \rangle.$$

646 / 862

Beispiel: Rotationssymmetrische harmonische Funktionen auf $\mathbb{R}^n \setminus \{0\}$

Sei $f: \mathbb{R}_+ \rightarrow \mathbb{R}$ eine zweimal differenzierbare Funktion, $r: \mathbb{R}^n \rightarrow \mathbb{R}^+$ die Norm von x , d.h. $r(x) := \|x\| = \sqrt{\sum_{i=1}^n x_i^2}$ und $F(x) = f(r(x))$ die zugehörige rotationssymmetrische zweimal partiell differenzierbare Funktion auf $\mathbb{R}^n \setminus \{0\}$.

Wir haben gesehen, dass gilt $\partial_i F(x) = f'(r) \frac{x_i}{r}$. Daraus folgt

$$\partial_j \partial_i F(x) = f''(r) \frac{x_j x_i}{r^2} + f'(r) \frac{\delta_{ij} r - \frac{x_i x_j}{r}}{r^2} = f''(r) \frac{x_j x_i}{r^2} + f'(r) \left(\frac{\delta_{ij}}{r} - \frac{x_i x_j}{r^3} \right)$$

Daraus folgt

$$\Delta F(x) = \sum_{i=1}^n \partial_i^2 F(x) = f''(r) + \frac{n-1}{r} f'(r).$$

Jede Lösung $f: \mathbb{R}^+ \rightarrow \mathbb{R}$ der Differentialgleichung $f''(r) + \frac{n-1}{r} f'(r) = 0$ liefert also eine rotationssymmetrische harmonische Funktion auf dem Raum $\mathbb{R}^n \setminus \{0\}$.

647 / 862

Beispiel (Fortsetzung)

Für $n = 1$ vereinfacht sich die Gleichung zu $f''(r) = 0$, und wir erhalten $f(r) = ar + b$ mit $a, b \in \mathbb{R}$ als allgemeine Lösung.

Für $n \geq 2$ setzen wir $h(r) := f'(r)$ und erhalten aus $f''(r) + \frac{n-1}{r} f'(r) = 0$ die Gleichung

$$h'(r) = -\frac{n-1}{r} h(r).$$

Unter der Voraussetzung, dass $h(r)$ keine Nullstellen hat, können wir die Differentialgleichung für $n \geq 2$ umformen:

$$\frac{n-1}{r} = -\frac{h'(r)}{h(r)} = -(\ln |h(r)|)'$$

Dies impliziert $\ln |h(r)| = -(n-1) \ln r + c$. Anwenden von \exp ergibt als Lösung

$$|h(r)| = \frac{e^c}{r^{n-1}} \quad \text{mit } c \in \mathbb{R}.$$

648 / 862

Beispiel (Fortsetzung)

Somit lautet die allgemeine Lösung $f(r) = \frac{a}{r^{n-2}} + b$ falls $n > 2$ und $f(r) = a \ln r + b$ für $n = 2$, jeweils mit $a, b \in \mathbb{R}$.

Für $n = 3$ erhalten wir mit $b = 0$ das **Newtonsche Gravitationspotential** $-G \frac{M}{r}$ einer im Nullpunkt konzentrierten Masse M . (Hierbei ist $G > 0$ die Gravitationskonstante.)

Ganz analog ist das **elektrische Potential** einer Punktladung Q im Ursprung gegeben durch $\gamma \frac{Q}{r}$. (Dabei ist wieder $\gamma > 0$ eine Konstante.)

649 / 862

Das Newtonsche Gravitationsfeld

Jede partiell differenzierbare Funktion $V: U \rightarrow \mathbb{R}$ auf einer offenen Teilmenge $U \subseteq \mathbb{R}^n$ definiert ein Vektorfeld $-\text{grad} V$ auf U .

Für das Newtonpotential $V(x) = -G \frac{M}{\|x\|}$ erhalten wir als Gradientenfeld

$$-\text{grad} V = -GM \frac{x}{\|x\|^3}, \quad \text{für } x \in \mathbb{R}^3 \setminus \{0\},$$

das **Newtonsche Gravitationsfeld**.

Die auf eine Probemasse m im Punkt x wirkende Kraft wird beschrieben durch das Vektorfeld

$$F = -m \text{grad} V = -GmM \frac{x}{\|x\|^3}.$$

Beschreibt $t \mapsto x(t) \in \mathbb{R}^3$ die Bewegung der Probemasse im Gravitationsfeld, so genügt diese nach dem Newtonschen Gesetz $F = m\ddot{x}$ der (von $m \neq 0$ unabhängigen) Differentialgleichung

$$\ddot{x} = -GM \frac{x}{\|x\|^3}.$$

650 / 862

Das elektrostatische Feld

Aus dem elektrischen Potential $V(x) = \gamma \frac{Q}{\|x\|}$ ergibt sich entsprechend

$$E = -\text{grad}V = \gamma Q \frac{x}{\|x\|^3}, \quad x \in \mathbb{R}^3 \setminus \{0\},$$

als Vektorfeld für das **elektrische Feld**.

Die auf eine Probeladung q im Punkt x wirkende Kraft ist

$$F = qE = \gamma qQ \frac{x}{\|x\|^3}.$$

(Die elektrische Kraft ist abstoßend, wenn $qQ > 0$ und anziehend wenn $qQ < 0$.)

Die Bewegung einer Probeladung mit Masse $m \neq 0$ genügt der Differentialgleichung

$$\ddot{x} = \gamma \frac{qQ}{m} \frac{x}{\|x\|^3}.$$

651 / 862

Das Vektorprodukt von Vektoren im \mathbb{R}^3 (vgl. Übung)

Definition (Vektorprodukt/Kreuzprodukt)

Das **Vektorprodukt** (oder **Kreuzprodukt**) von $x = \sum_{i=1}^3 x_i e_i \in \mathbb{R}^3$ und $y = \sum_{i=1}^3 y_i e_i \in \mathbb{R}^3$ ist der Vektor

$$\begin{aligned} x \times y &:= \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix} \\ &= (x_2 y_3 - x_3 y_2) e_1 + (x_3 y_1 - x_1 y_3) e_2 + (x_1 y_2 - x_2 y_1) e_3. \end{aligned}$$

Man schreibt manchmal auch $x \wedge y$.

Die durch das Kreuzprodukt definierte Abbildung $\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $(x, y) \mapsto x \times y$ ist bilinear.

652 / 862

Satz (Eigenschaften des Vektorprodukts, vgl. ÜA)

Sei der \mathbb{R}^3 versehen mit der durch die geordnete Basis (e_1, e_2, e_3) gegebenen **Orientierung**. Für alle $x, y \in \mathbb{R}^3$ gilt dann:

- (i) $x \times y = -y \times x$, d.h. die Abbildung $(x, y) \mapsto x \times y$ ist **schief-symmetrisch**.
- (ii) $x \times y$ ist **senkrecht** zu x und y .
- (iii) $\|x \times y\|^2 = \|x\|^2 \|y\|^2 - \langle x, y \rangle^2$.
- (iv) $\|x \times y\| = \|x\| \cdot \|y\| \sin \varphi$, wenn $x, y \in \mathbb{R}^3 \setminus \{0\}$ und $\varphi = \angle(x, y) \in [0, \pi]$.
- (v) $x \times y = 0 \iff (x, y)$ **linear abhängig**.
- (vi) Sind (x, y) **linear unabhängig**, so ist $(x, y, x \times y)$ eine **positiv orientierte Basis**.
- (vii) Sind x und y **orthonormal**, so ist $(x, y, x \times y)$ eine **positiv orientierte Orthonormalbasis** von \mathbb{R}^3 . □

Die Zahl $\|x \times y\| \in \mathbb{R}$ ist die Fläche des von den Vektoren x und y aufgespannten Parallelogramms.

653 / 862

Das Spatprodukt

Satz

Für alle $x, y, z \in \mathbb{R}^3$ gilt: $\langle x \times y, z \rangle = \det(x \ y \ z)$.

Beweis. Wegen der Trilinearität von $\langle x \times y, z \rangle$ und \det genügt es, die Gleichung für $x, y, z \in \{e_1, e_2, e_3\}$ zu überprüfen.

Offenbar ist $\langle e_i \times e_j, e_k \rangle = 0 = \det(e_i \ e_j \ e_k)$, wenn mindestens zwei der drei Indizes übereinstimmen.

Sei (i, j, k) eine Permutation von $(1, 2, 3)$. Da (e_i, e_j) orthonormal ist, gilt (nach Teil (vii) des letzten Satzes) $e_i \times e_j = \varepsilon_{ijk} e_k$, wobei ε_{ijk} das Vorzeichen der Permutation ist.

Also $\langle e_i \times e_j, e_k \rangle = \varepsilon_{ijk} = \det(e_i \ e_j \ e_k)$ für jede Permutation. □

Bemerkung

Die alternierende Trilinearform $(x, y, z) \mapsto \langle x \times y, z \rangle$ heißt **Spatprodukt**.

$\langle x \times y, z \rangle$ ist das orientierte Volumen des durch x, y, z aufgespannten **Parallelepipeds** oder **Spats** $P = \{ax + by + cz \mid a, b, c \in [0, 1]\} \subseteq \mathbb{R}^3$.

654 / 862

Definition

Sei $v = \sum_{i=1}^3 v_i e_i$ ein partiell differenzierbares Vektorfeld auf einer offenen Teilmenge $U \subseteq \mathbb{R}^3$.

Das Vektorfeld

$$\operatorname{rot} v := \begin{pmatrix} \partial_2 v_3 - \partial_3 v_2 \\ \partial_3 v_1 - \partial_1 v_3 \\ \partial_1 v_2 - \partial_2 v_1 \end{pmatrix}$$

heißt **Rotation** von v .

655 / 862

Bemerkung

Wir können den Gradienten als vektorwertigen Differentialoperator auffassen. Für jede offene Menge $U \subseteq \mathbb{R}^3$ haben wir dann

$$\operatorname{grad} = \begin{pmatrix} \partial_1 \\ \partial_2 \\ \partial_3 \end{pmatrix} : \{f : U \rightarrow \mathbb{R} \mid f \text{ part. diffb.}\} \rightarrow \text{Vektorfelder auf } U$$

$$f \mapsto \operatorname{grad}(f) = \begin{pmatrix} \partial_1 f \\ \partial_2 f \\ \partial_3 f \end{pmatrix}.$$

Mit dieser Schreibweise ergibt sich formal:

$$\operatorname{div} v = \langle \operatorname{grad}, v \rangle, \quad \operatorname{rot} v = \operatorname{grad} \times v.$$

656 / 862

Satz

Sei $U \subseteq \mathbb{R}^3$ offen und $f: U \rightarrow \mathbb{R}$, $v: U \rightarrow \mathbb{R}^3$ zweimal stetig partiell differenzierbar. Dann gilt

$$\operatorname{rot} \operatorname{grad} f = 0 \quad \text{und} \quad \operatorname{div} \operatorname{rot} v = 0,$$

d.h. (formal): $\operatorname{grad} \times \operatorname{grad} f = 0$ und $\langle \operatorname{grad}, \operatorname{grad} \times v \rangle = 0$.

Beweis. Das rechnet man unter Benutzung des Lemmas von Schwarz direkt nach (ÜA). □

Warnung:

Die oben angegebenen Formeln für Gradient, Divergenz und Rotation gelten nur in den Standardkoordinaten.

657 / 862

Kapitel 19

Differenzierbare Abbildungen

658 / 862

Wir erinnern an den Zusammenhang von Differenzierbarkeit und linearer Approximation:

Sei $D \subseteq \mathbb{R}$ und $f: D \rightarrow \mathbb{R}$ eine Funktion, $a \in D$ ein Häufungspunkt von D . Wenn es ein $c \in \mathbb{R}$ gibt, so dass

$$\lim_{x \rightarrow a} \frac{f(x) - [f(a) + c(x - a)]}{x - a} = 0$$

gilt, so ist $c = f'(a)$. Wir schreiben $x = a + \xi$ und finden äquivalent:

$$\lim_{\xi \rightarrow 0} \frac{f(a + \xi) - f(a) - c\xi}{\xi} = 0$$

Man beachte, dass wir $\xi \mapsto c\xi$ als lineare Funktion $\mathbb{R} \rightarrow \mathbb{R}$ auffassen können.

659 / 862

Definition (Differenzierbarkeit)

Sei $U \subseteq \mathbb{R}^n$ offen und $F: U \rightarrow \mathbb{R}^m$.

- (i) Die Funktion F heißt im Punkt $x \in U$ **(total) differenzierbar**, wenn es eine **lineare Abbildung** $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ gibt, so dass

$$\lim_{\xi \in \mathbb{R}^n \setminus \{0\}, \xi \rightarrow 0} \frac{\|F(x + \xi) - F(x) - A(\xi)\|}{\|\xi\|} = 0$$

- (ii) Die Funktion $F: U \rightarrow \mathbb{R}^m$ heißt **differenzierbar**, wenn F in allen Punkten $x \in U$ differenzierbar ist.

Bemerkung

Wir könnten in der Definition ohne weiteres \mathbb{R}^n und \mathbb{R}^m durch beliebige Banachräume ersetzen. Dann fordert man zusätzlich, dass die lineare Abbildung A stetig (also beschränkt) ist.

660 / 862

Satz (und Definition des Differentials)

Sei $U \subseteq \mathbb{R}^n$ offen und sei $F: U \rightarrow \mathbb{R}^m$ differenzierbar im Punkt $x \in U$.
Dann wird durch

$$\lim_{\xi \rightarrow 0} (\|F(x + \xi) - F(x) - A(\xi)\| / \|\xi\|) = 0$$

die lineare Abbildung

$$dF_x := A: \mathbb{R}^n \rightarrow \mathbb{R}^m$$

eindeutig bestimmt. Sie heißt **Differential** (oder **Ableitung**) von F im Punkt x . Wir identifizieren dF_x oft mit der beschreibenden Matrix.

661 / 862

Beweis der Eindeutigkeit von A :

Seien A und B zwei lineare Abbildungen mit der geforderten Eigenschaft.
Mit Hilfe der Dreiecksungleichung für die Norm erhält man

$$\begin{aligned} 0 &\leq \lim_{\xi \rightarrow 0} \frac{\|B(\xi) - A(\xi)\|}{\|\xi\|} \\ &\leq \lim_{\xi \rightarrow 0} \frac{\|F(x + \xi) - F(x) - A(\xi)\|}{\|\xi\|} + \lim_{\xi \rightarrow 0} \frac{\| -F(x + \xi) + F(x) + B(\xi)\|}{\|\xi\|} \\ &= 0 \end{aligned}$$

Setzt man nun $\xi = t\eta$ für $\eta \in \mathbb{R}^n$ mit $\|\eta\| = 1$ so erhält man daraus wegen der Linearität von A und B , dass

$$0 = \lim_{t \rightarrow 0} \frac{|t|}{|t|} (B(\eta) - A(\eta)) = B(\eta) - A(\eta).$$

Hieraus folgt wegen der Linearität $A(\eta) = B(\eta)$ für alle η und somit $A = B$. □

662 / 862

Beispiel

Die Funktion $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $f(x) = \|x\|^2$, ist überall differenzierbar, denn wegen

$$f(x + \xi) = \|x + \xi\|^2 = \|x\|^2 + 2\langle x, \xi \rangle + \|\xi\|^2 = f(x) + 2\langle x, \xi \rangle + \|\xi\|^2$$

gilt

$$\frac{f(x + \xi) - f(x) - 2\langle x, \xi \rangle}{\|\xi\|} = \frac{\|\xi\|^2}{\|\xi\|} = \|\xi\| \xrightarrow{\xi \rightarrow 0} 0.$$

Das Differential $df_x \in (\mathbb{R}^n)^*$ ist also gegeben durch die Linearform:

$$df_x(\xi) = 2\langle x, \xi \rangle = 2x^T \xi, \quad \xi \in \mathbb{R}^n.$$

Also hat df_x die darstellende Matrix $(2x_1, 2x_2, \dots, 2x_n)$, d.h.

$$df_x = (\text{grad} f(x))^T.$$

663 / 862

Jacobi-Matrix (Matrix der Ableitungen)

Satz (Differenzierbarkeit impliziert partielle Differenzierbarkeit)

Sei $U \subseteq \mathbb{R}^n$ offen und $F: U \rightarrow \mathbb{R}^m$ in x differenzierbar.

Dann existiert für jede Komponente F_j mit $j = 1, \dots, m$ von F und für jede Richtung $v \in \mathbb{R}^n \setminus \{0\}$ die Richtungsableitung $\partial_v F_j(x)$, und es gilt

$$\partial_v F_j(x) = (dF_j)_x(v).$$

Insbesondere sind alle Komponentenfunktionen F_j in x partiell differenzierbar, d.h. F ist in x partiell differenzierbar. Weiterhin gilt:

$$dF_x = \begin{pmatrix} \partial_1 F_1(x) & \cdots & \partial_n F_1(x) \\ \vdots & & \vdots \\ \partial_1 F_m(x) & \cdots & \partial_n F_m(x) \end{pmatrix} = \begin{pmatrix} (\text{grad} F_1(x))^T \\ \vdots \\ (\text{grad} F_m(x))^T \end{pmatrix}.$$

Diese Matrix heißt auch **Jacobi-Matrix** von F .

664 / 862

Beweis: Sei $v \in \mathbb{R}^n$ mit $\|v\| = 1$ und $F = \sum_{j=1}^m F_j e_j$.

Da F differenzierbar ist, sind alle Komponentenfunktionen F_j als Verkettung differenzierbarer Funktionen differenzierbar. D.h. es gilt für jede Komponente F_j

$$\begin{aligned} 0 &= \lim_{t \rightarrow 0} \frac{|F_j(x + tv) - F_j(x) - (dF_j)_x(tv)|}{|t|\|v\|} \\ &= \lim_{t \rightarrow 0} \left| \frac{F_j(x + tv) - F_j(x)}{t} - (dF_j)_x(v) \right|, \end{aligned}$$

d.h. $(dF_j)_x(v) = \frac{d}{dt}|_{t=0} F_j(x + tv) = \partial_v F_j(x)$.

Damit erhält man für $v = e_i$ die i -te partielle Ableitung

$$\partial_i F_j(x) = (dF_j)_x(e_i).$$

Dies gilt für jede Komponente F_j und somit

$$dF_x(e_i) = \begin{pmatrix} \partial_i F_1(x) \\ \vdots \\ \partial_i F_m(x) \end{pmatrix} = i\text{-te Spalte von } dF_x.$$

Dies beweist die Formel für die Ableitungsmatrix. □

665 / 862

Folgerung

Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$ eine differenzierbare Funktion.

Dann existiert für jedes $v \in \mathbb{R}^n \setminus \{0\}$ die Richtungsableitung $\partial_v f(x)$ und es gilt

$$\partial_v f(x) = df_x v = \langle \text{grad } f(x), v \rangle.$$

Die Jacobi-Matrix von f ist gegeben durch die Linearform

$$df = (\partial_1 f, \dots, \partial_n f) = (\text{grad } f)^T \quad : \quad U \rightarrow (\mathbb{R}^n)^*, \\ x \mapsto df_x.$$

Bemerkung

Wir erhalten damit eine koordinatenunabhängige Definition des Gradienten:

Ist $U \subseteq \mathbb{R}^n$ offen und $f : U \rightarrow \mathbb{R}$ in x differenzierbar, so ist $\text{grad } f(x) \in \mathbb{R}^n$ der nach dem Satz von Riesz eindeutig bestimmte Vektor mit

$$df_x(v) = \langle \text{grad } f(x), v \rangle \quad \text{für alle } v \in \mathbb{R}^n.$$

666 / 862

Beispiel

Für die Koordinatenfunktionen $f(x) = x_i$ haben wir $\text{grad}(x_i) = e_i$ und somit die Linearformen $dx_i = \langle e_i, \cdot \rangle = e_i^*$. Man schreibt daher für jede differenzierbare Funktion f auf einer offenen Teilmenge $U \subseteq \mathbb{R}^n$ auch

$$df = \sum_{i=1}^n \partial_i f \, dx_i.$$

667 / 862

Differenzierbarkeit impliziert Stetigkeit

Satz

Sei $U \subseteq \mathbb{R}^n$ offen und $F = \sum_{j=1}^m F_j e_j: U \rightarrow \mathbb{R}^m$ in $x \in U$ total differenzierbar. Dann ist F in x stetig.

Beweis. Sei $x_n := x + \xi_n$ mit $\xi_n \rightarrow 0$ eine Folge, die gegen x konvergiert. Dann ist wegen der Dreiecksungleichung

$$\|F(x_n) - F(x)\| \leq \underbrace{\frac{\|F(x_n) - F(x) - dF_x(\xi_n)\|}{\|\xi_n\|}}_{\rightarrow 0, \text{ da } f \text{ diffb.}} \underbrace{\|\xi_n\|}_{\rightarrow 0} + \|dF_x(\xi_n)\|.$$

Nun ist aber $dF_x: \mathbb{R}^n \rightarrow \mathbb{R}^m$ linear und damit stetig, d.h. auch

$\|dF_x(\xi_n)\| \xrightarrow{\xi_n \rightarrow 0} 0$. Also $\|F(x_n) - F(x)\| \rightarrow 0$. □

668 / 862

Hinreichendes Kriterium für Differenzierbarkeit

Satz

Sei $U \subseteq \mathbb{R}^n$ offen, $F = \sum_{j=1}^m F_j e_j: U \rightarrow \mathbb{R}^m$ **partiell differenzierbar** und alle partiellen Ableitungen $\partial_i F_j$ seien an der Stelle $x \in U$ **stetig**.
Dann ist F in x total **differenzierbar**.

Beweis. Wir können o.B.d.A. annehmen, dass $m = 1$.

Wähle $\varepsilon > 0$ so klein, dass $B_\varepsilon(x) \subseteq U$ gilt. Wähle $\xi = \sum_{i=1}^n \xi_i e_i \in B_\varepsilon(0)$.
Für $i = 1, \dots, n$ betrachten wir die reellen Funktionen einer Variablen

$$\begin{aligned} g_1(t) &:= F(x_1 + t, x_2, \dots, x_n) \\ &\vdots \\ g_i(t) &:= F(x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i + t, x_{i+1}, \dots, x_n) \\ &\vdots \\ g_n(t) &:= F(x_1 + \xi_1, \dots, x_{n-1} + \xi_{n-1}, x_n + t) \end{aligned}$$

Dann gilt

$$g'_i(0) = \partial_i F(x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i, \dots, x_n)$$

669 / 862

Weiter im Beweis:

Auf jede dieser Funktionen wenden wir nun den **Mittelwertsatz** an und erhalten, für $i = 1 \dots n$, Zahlen $\theta_i \in (-|\xi_i|, |\xi_i|)$ mit

$$\begin{aligned} &F(x_1 + \xi_1, \dots, x_i + \xi_i, x_{i+1}, \dots, x_n) - F(x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i, \dots, x_n) \\ &= \xi_i \partial_i F(x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i + \theta_i, x_{i+1}, \dots, x_n) \end{aligned}$$

Setzt man nun

$$\begin{aligned} y^{(i)} &:= (x_1 + \xi_1, \dots, x_i + \xi_i, x_{i+1}, \dots, x_n), \quad i = 0, \dots, n, \\ z^{(i)} &:= (x_1 + \xi_1, \dots, x_{i-1} + \xi_{i-1}, x_i + \theta_i, x_{i+1}, \dots, x_n), \quad i \geq 1, \end{aligned}$$

so erhält man daraus

$$F(x + \xi) - F(x) = \sum_{k=1}^n F(y^{(k)}) - F(y^{(k-1)}) = \sum_{k=1}^n \xi_k \partial_k F(z^{(k)}).$$

670 / 862

Ende des Beweises: Mit Hilfe der Cauchy-Schwarzschen Ungleichung ergibt sich dann

$$\frac{|F(x + \xi) - F(x) - \sum_{k=1}^n \partial_k F(x) \xi_k|}{\|\xi\|} = \frac{|\sum_{k=1}^n (\partial_k F(z^{(k)}) - \partial_k F(x)) \xi_k|}{\|\xi\|} \leq \left(\sum_{k=1}^n (\partial_k F(z^{(k)}) - \partial_k F(x))^2 \right)^{\frac{1}{2}}$$

Für $\xi \rightarrow 0$ gilt auch $z^{(k)} \rightarrow x$.

Da nun alle partiellen Ableitungen stetig sein sollen, bekommen wir

$$\partial_k F(z^{(k)}) \xrightarrow{\xi \rightarrow 0} \partial_k F(x),$$

und somit

$$\lim_{\xi \rightarrow 0} \frac{|F(x + \xi) - F(x) - \sum_{k=1}^n \partial_k F(x) \xi_k|}{\|\xi\|} = 0.$$

Damit ist F in x differenzierbar. □

671 / 862

Satz

Sei $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine Abbildung, die in x differenzierbar ist. Dann gilt:

- Ist $G: \mathbb{R}^n \rightarrow \mathbb{R}^m$ in x differenzierbar, so auch $F + G$ und es gilt

$$d(F + G)_x = dF_x + dG_x$$

- Für $\lambda \in \mathbb{R}$ ist λF in x differenzierbar und es gilt

$$d(\lambda F)_x = \lambda dF_x$$

- Ist $h: \mathbb{R}^n \rightarrow \mathbb{R}$ in x differenzierbar, so auch $h \cdot F$ und es gilt

$$d(h \cdot F)_x = F(x) \cdot dh_x + h(x) dF_x.$$

- Ist $g: \mathbb{R}^n \rightarrow \mathbb{R}$ in x differenzierbar und $g(x) \neq 0$, so ist $\frac{1}{g} F$ in x differenzierbar, und es gilt:

$$d\left(\frac{1}{g} F\right)_x = \frac{1}{g(x)^2} (g(x) dF_x - F(x) \cdot dg_x).$$

672 / 862

Satz (Kettenregel)

Seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ offen, $F: U \rightarrow \mathbb{R}^m$, $F(U) \subseteq V$ und $G: V \rightarrow \mathbb{R}^k$.
Wenn F in x und G in $y = F(x)$ differenzierbar sind, dann ist die Verkettung $G \circ F: U \rightarrow \mathbb{R}^k$ in x differenzierbar und

$$d(G \circ F)_x = dG_{F(x)} \circ dF_x.$$

Beweis. Für $x \in U$, $y := F(x) \in V$ und $\xi \in \mathbb{R}^n$ und $\eta \in \mathbb{R}^m$ hinreichend klein definieren wir

$$\begin{aligned}\varphi(\xi) &:= F(x + \xi) - F(x) - A\xi, & \text{wobei } A = dF_x, \\ \psi(\eta) &:= G(y + \eta) - G(y) - B\eta, & \text{wobei } B = dG_y,\end{aligned}$$

Dann gilt wegen der Differenzierbarkeit von F und G , dass

$$\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|} = 0 \quad \text{und} \quad \lim_{\eta \rightarrow 0} \frac{\psi(\eta)}{\|\eta\|} = 0.$$

673 / 862

Weiter im Beweis:

Für die spezielle Wahl $\eta := F(x + \xi) - y = A\xi + \varphi(\xi)$ erhalten wir

$$\begin{aligned}(G \circ F)(x + \xi) - (G \circ F)(x) &= G(y + \eta) - G(y) \\ &= B\eta + \psi(\eta) \\ &= BA\xi + B\varphi(\xi) + \psi(\eta)\end{aligned}$$

D.h.

$$\lim_{\xi \rightarrow 0} \frac{\|(G \circ F)(x + \xi) - (G \circ F)(x) - BA\xi\|}{\|\xi\|} \leq \lim_{\xi \rightarrow 0} \frac{\|B\varphi(\xi)\|}{\|\xi\|} + \lim_{\xi \rightarrow 0} \frac{\|\psi(\eta)\|}{\|\xi\|},$$

Nun ist aber wegen der Linearität von B

$$\lim_{\xi \rightarrow 0} \frac{\|B\varphi(\xi)\|}{\|\xi\|} = \lim_{\xi \rightarrow 0} \left\| B \left(\frac{\varphi(\xi)}{\|\xi\|} \right) \right\| \leq \|B\|_{Op.-Norm} \lim_{\xi \rightarrow 0} \frac{\|\varphi(\xi)\|}{\|\xi\|} = 0.$$

Also bleibt nur noch zu zeigen, dass $\lim_{\xi \rightarrow 0} \frac{\psi(\eta)}{\|\xi\|} = 0$.

674 / 862

Ende des Beweises: Dazu bemerkt man, dass $\eta \rightarrow 0$ falls $\xi \rightarrow 0$, denn wegen der Dreiecksungleichung ist

$$\|\eta\| = \|A\xi + \varphi(\xi)\| \leq \|A\xi\| + \|\varphi(\xi)\| \xrightarrow{\xi \rightarrow 0} 0.$$

Somit gilt $\frac{\|\psi(\eta)\|}{\|\eta\|} \xrightarrow{\xi \rightarrow 0} 0$. Aus der Differenzierbarkeit von G und der Dreiecksungleichung folgt dann:

$$\begin{aligned} \frac{\|\psi(\eta)\|}{\|\xi\|} &= \frac{\|\psi(\eta)\|}{\|\eta\|} \frac{\|\eta\|}{\|\xi\|} = \frac{\|\psi(\eta)\|}{\|\eta\|} \frac{\|A\xi + \varphi(\xi)\|}{\|\xi\|} \\ &\leq \underbrace{\frac{\|\psi(\eta)\|}{\|\eta\|}}_{\xrightarrow{\xi \rightarrow 0} 0} \left(\frac{\|A\xi\|}{\|\xi\|} + \underbrace{\frac{\|\varphi(\xi)\|}{\|\xi\|}}_{\xrightarrow{\xi \rightarrow 0} 0} \right) \xrightarrow{\xi \rightarrow 0} 0, \end{aligned}$$

denn wegen der (stetigen) Linearität von A bleibt $\frac{\|A\xi\|}{\|\xi\|} \leq \|A\|_{\text{operator}}$ beschränkt. □

675 / 862

Folgerung (Kettenregel in Komponenten)

Seien $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ offen, $F: U \rightarrow \mathbb{R}^m$, $F(U) \subseteq V$ und $G: V \rightarrow \mathbb{R}^k$, $H := G \circ F: U \rightarrow \mathbb{R}^k$. Es sei F in x und G in $y = F(x)$ differenzierbar. Dann ist $H = G \circ F$ in x differenzierbar und für alle $i = 1, 2, \dots, n$ und $l = 1, \dots, k$ gilt:

$$\partial_i H_l(x) = \sum_{j=1}^m \partial_j G_l(F(x)) \partial_i F_j(x).$$

In anderer Notation, mit den Einträgen der Jacobi-Matrix:

$$\frac{\partial H_l}{\partial x_i}(x) = \sum_{j=1}^m \frac{\partial G_l}{\partial y_j}(F(x)) \frac{\partial F_j}{\partial x_i}(x).$$

Dies ist eine direkte Konsequenz des Satzes und der Formel für die Matrixmultiplikation.

676 / 862

Beispiele

In den folgenden Beispielen betrachten wir Abbildungen des Vektorraumes der $n \times n$ -Matrizen $Mat_n(\mathbb{R}) = \mathbb{R}^{n^2}$.

- (i) Sei $F: Mat_n(\mathbb{R}) \rightarrow \mathbb{R}$ gegeben durch $F(X) = \det(X)$. Dann gilt für das Differential in $\mathbf{1}_n$

$$dF_{\mathbf{1}_n}(A) = d \det_{\mathbf{1}_n}(A) = \partial_A \det(\mathbf{1}_n) = \text{spur}(A);$$

denn nach der Produktregel gilt

$$\begin{aligned} dF_{\mathbf{1}_n}(A) &= \left. \frac{d}{dt} \right|_{t=0} \det(\mathbf{1}_n + tA) \\ &= \left. \frac{d}{dt} \right|_{t=0} \sum_{\sigma \in S_n} \varepsilon(\sigma) (\delta_{1\sigma(1)} + ta_{1\sigma(1)}) \cdot \dots \cdot (\delta_{n\sigma(n)} + ta_{n\sigma(n)}) \\ &= a_{11} + \dots + a_{nn}. \end{aligned}$$

677 / 862

Beispiele

- (ii) Nun sei $F: Mat_n(\mathbb{R}) \rightarrow Mat_n(\mathbb{R})$ gegeben durch $F(X) = X^2$ und $G: Mat_n(\mathbb{R}) \rightarrow \mathbb{R}$ durch $G(Y) = \text{spur} Y$.
 F und G sind auf $Mat_n(\mathbb{R})$ differenzierbar mit

$$dF_X A = \left. \frac{d}{dt} \right|_{t=0} [(X + tA)^2] = XA + AX,$$

und

$$dG_Y B = \left. \frac{d}{dt} \right|_{t=0} [\text{spur}(Y + tB)] = \text{spur} B,$$

für $Y, B \in Mat_n(\mathbb{R})$.

Somit ist $G \circ F: X \mapsto \text{spur}(X^2)$ differenzierbar und

$$d(G \circ F)_X A = dG_{F(X)} \circ dF_X(A) = \text{spur}(XA + AX) = 2\text{spur}(AX).$$

678 / 862

Beispiel

Sei $c: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n$ ($\varepsilon > 0$) eine differenzierbare Abbildung, d.h. eine **differenzierbare Kurve**. Setze $x := c(0)$, $v := c'(0) = (c'_1(0), \dots, c'_n(0))^T$. Ferner sei $f: U \rightarrow \mathbb{R}$ eine in einer offenen Umgebung $U \subseteq \mathbb{R}^n$ von x definierte und in x differenzierbare Funktion.

Dann gilt für die Ableitung der Funktion $f \circ c: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$:

$$(f \circ c)'(0) = df_{c(0)}(c'(0)) = df_x(v) = \partial_v f(x).$$

D.h.

$$\left. \frac{d}{dt} \right|_{t=0} (f(c(t))) = \partial_v f(x) = \left. \frac{d}{dt} \right|_{t=0} (f(x + tv)).$$

Es genügt hier übrigens vorauszusetzen, dass c eine stetige Abbildung ist, die im Nullpunkt differenzierbar ist.

679 / 862

Mittelwertsatz

Satz (Mittelwertsatz für Funktionen auf dem \mathbb{R}^n)

Sei $U \subseteq \mathbb{R}^n$ offen und $F: U \rightarrow \mathbb{R}$ stetig differenzierbar.

Seien weiterhin $x \in U$ und $\xi \in \mathbb{R}^n$, so dass $x + t\xi \in U$, für alle $t \in [0, 1]$.

Dann existiert ein $x_0 = x + t_0\xi \in U$ mit $t_0 \in (0, 1)$, so dass

$$F(x + \xi) - F(x) = dF_{x_0}(\xi).$$

Beweis. Die Aussage ist eine direkte Folge des Mittelwertsatzes für Funktionen $f: [0, 1] \rightarrow \mathbb{R}$.

Sei $\varphi: [0, 1] \rightarrow \mathbb{R}^n$, $t \mapsto x + t\xi$ und $f = F \circ \varphi: [0, 1] \rightarrow \mathbb{R}$.

Nach dem Mittelwertsatz für f erhalten wir ein $t_0 \in (0, 1)$ mit

$$f(1) - f(0) = f'(t_0).$$

Aus der Kettenregel ergibt sich dann

$$f(1) - f(0) = f'(t_0) = (F \circ \varphi)'(t_0) = dF_{x+t_0\xi}(\varphi'(t_0))$$

Für $\varphi(t) = x + t\xi$ ist $\varphi' \equiv \xi$ und damit $F(x + \xi) - F(x) = dF_{x_0}(\xi)$. \square

680 / 862

Ist eine Abbildung $F: U \rightarrow \mathbb{R}^m$ konstant, so ist $dF_x = 0$ für alle $x \in U$. Die Umkehrung gilt im Allgemeinen nicht (man betrachte eine Vereinigung $U = U_1 \cup U_2$ zweier disjunkter offener Mengen). Allerdings gilt:

Folgerung

Sei $U \subseteq \mathbb{R}^n$ offen. Ist U **wegzusammenhängend**, d.h. gibt es für je zwei Punkte $x, y \in U$ eine stetige Kurve $c: [0, 1] \rightarrow U$ mit $c(0) = x$ und $c(1) = y$, so gilt:
Ist $F: U \rightarrow \mathbb{R}^m$ differenzierbar mit $dF_x = 0$ für alle $x \in U$, so ist F konstant.

Beweis. Für jedes $\xi \in U$ finde $\varepsilon(\xi) > 0$ so, dass $K_\xi := \overline{B_{\varepsilon(\xi)}(\xi)} \subseteq U$.

Nach dem Mittelwertsatz ist F auf K_ξ konstant, d.h. $F|_{K_\xi} \equiv d_\xi$.

Sei nun c eine stetige Kurve von x nach y . Deren Bild wird überdeckt durch die offenen Kugeln $\{B_{\varepsilon(c(t))}(c(t))\}_{t \in [0,1]}$.

Da das Intervall $[0, 1]$ kompakt ist und c stetig, ist auch $c([0, 1])$ kompakt.

Daher finden wir eine endliche Teilüberdeckung durch offene Kugeln

$B_{\varepsilon(c(t_i))}(c(t_i))$ für $i = 1, \dots, k$ und $t_1 = 0, t_k = 1$. Damit ist aber

$F|_{B_{\varepsilon(c(t_i))}} \equiv d_i$, und somit $F(y) = F(c(t_k)) = \dots = F(c(t_1)) = F(x)$. \square

681 / 862

Niveaumengen und Gradient

Satz

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ stetig differenzierbar.

Dann steht $\text{grad } f$ senkrecht auf den **Niveaumengen**

$$N_f(k) := \{x \in \mathbb{R}^n \mid f(x) = k\} \quad \text{mit } k \in \mathbb{R},$$

d.h. für jede innerhalb der Niveaumenge $N_f(k)$ verlaufende differenzierbare Kurve $c: (a, b) \rightarrow \mathbb{R}^n$ gilt

$$\text{grad } f|_{c(t)} \perp c'(t) \quad \text{für alle } t \in (a, b).$$

Beweis. Das folgt durch Ableiten der Gleichung $f \circ c = k$ nach der Kettenregel

$$0 = df_c \circ c' = \langle \text{grad } f, c' \rangle.$$

 \square

Beispiel

Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ gegeben durch die Norm, $f(x) = \|x\|$. Die Niveaumengen sind dann Sphären vom Radius k und der Gradient ist $\text{grad}f(x) = \frac{x}{\|x\|}$.

Bemerkung

Sei $U \subseteq \mathbb{R}^n$ offen. Wenn $f: U \rightarrow \mathbb{R}$ im Punkt $x \in U$ differenzierbar und $\text{grad}f(x) \neq 0$ ist, so gibt der **Gradient** die Richtung des **stärksten Anstiegs** der Funktion f im Punkt $x \in U$ an, denn für jeden Einheitsvektor $v \in \mathbb{R}^n$ gilt wegen der Cauchy-Schwarzschen Ungleichung

$$\partial_v f(x) = \langle \text{grad}f(x), v \rangle \leq \|\text{grad}f(x)\|$$

mit Gleichheit genau für den Einheitsvektor $v = \frac{\text{grad}f(x)}{\|\text{grad}f(x)\|}$.

683 / 862

Lokale Extrema

Die folgende Definition ist eine direkte Verallgemeinerung der entsprechenden Konzepte für Funktionen $f: \mathbb{R} \rightarrow \mathbb{R}$.

Definition

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ eine Funktion und $x \in U$.

Man sagt, dass f in x ein **lokales Maximum** (bzw. ein **lokales Minimum**) annimmt, falls es ein $\varepsilon > 0$ gibt, derart dass

$$f(x) \geq f(\xi) \quad \text{bzw.} \quad f(x) \leq f(\xi)$$

für alle $\xi \in U$ mit $\|x - \xi\| < \varepsilon$.

Lokale Minima und Maxima heißen auch **lokale Extrema**.

Man spricht von einem **isolierten** lokalen Extremum, falls zusätzlich $f(\xi) \neq f(x)$ für alle $\xi \in U \setminus \{x\}$ mit $\|x - \xi\| < \varepsilon$ für ein $\varepsilon > 0$.

684 / 862

Lokale Extrema und Ableitung

Satz

Sei $U \subseteq \mathbb{R}^n$ offen. Eine Funktion $f: U \rightarrow \mathbb{R}$ sei an der Stelle $x \in U$ **partiell differenzierbar**.

Wenn f an der Stelle $x \in U$ ein lokales Extremum annimmt, dann gilt $\text{grad}f(x) = 0$.

Beweis. Für alle $i = 1, \dots, n$ ist die Funktion $t \mapsto g(t) := f(x + te_i)$ im Nullpunkt differenzierbar und hat dort ein lokales Extremum.

Demnach gilt

$$0 = g'(0) = \partial_i f(x).$$

□

685 / 862

Wiederholung: Taylor-Entwicklung für Funktionen einer Variablen

Es sei $f: I \rightarrow \mathbb{R}$ eine $(m+1)$ -mal stetig differenzierbare Funktion auf einem Intervall $I \subseteq \mathbb{R}$.

Zu $x \in I$ ist das Taylor-Polynom $T_m(\xi) = T_m(f, x)(\xi) \in \mathbb{R}[\xi]$ gegeben durch

$$T_m(f, x)(\xi) = f(x) + f'(x)\xi + \frac{1}{2}f''(x)\xi^2 + \dots + \frac{1}{m!}f^{(m)}(x)\xi^m,$$

und zu jedem ξ mit $\{x + t\xi \mid t \in [0, 1]\} \subseteq I$ gibt es ein $\tau \in [0, 1]$ mit

$$f(x + \xi) = T_m(f, x)(\xi) + \frac{f^{(m+1)}(x + \tau\xi)}{(m+1)!}\xi^{m+1}.$$

Wir wollen die Taylor-Entwicklung nun für Funktionen mit Definitionsbereich $U \subseteq \mathbb{R}^n$ verallgemeinern.

686 / 862

Notation: Multi-Indizes

Um die Taylorentwicklung für Funktionen von n Veränderlichen kompakt schreiben zu können, führen wir die folgenden Abkürzungen ein.

Für $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{N}$ und $x = \sum_{i=1}^n x_i e_i \in \mathbb{R}^n$ setzen wir

$$\begin{aligned}\alpha &:= (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n \\ |\alpha| &:= \alpha_1 + \alpha_2 + \dots + \alpha_n \in \mathbb{N} \\ \alpha! &:= \alpha_1! \alpha_2! \dots \alpha_n! \in \mathbb{N} \\ x^\alpha &:= x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n} \\ \partial^\alpha &:= \partial_1^{\alpha_1} \partial_2^{\alpha_2} \dots \partial_n^{\alpha_n}\end{aligned}$$

Definition

Für $U \subseteq \mathbb{R}^n$ offen definieren wir

$$C^k(U) := \{f: U \rightarrow \mathbb{R} \mid f \text{ ist } k\text{-mal stetig differenzierbar}\}.$$

Dies ist ein reeller Vektorraum.

687 / 862

Taylorentwicklung für Funktionen von n Variablen

Satz (Taylorentwicklung)

Sei $U \subseteq \mathbb{R}^n$ offen, $f \in C^{m+1}(U)$, $x \in U$ und $\xi \in \mathbb{R}^n$, so dass $\{x + t\xi \mid t \in [0, 1]\} \subseteq U$.

Dann existiert $\tau \in [0, 1]$, so dass

$$f(x + \xi) = \sum_{|\alpha| \leq m} \frac{1}{\alpha!} \partial^\alpha f(x) \xi^\alpha + \sum_{|\alpha|=m+1} \frac{1}{\alpha!} \partial^\alpha f(x + \tau\xi) \xi^\alpha.$$

Definition

$T_m(\xi) = T_m(f, x)(\xi) := \sum_{|\alpha| \leq m} \frac{1}{\alpha!} \partial^\alpha f(x) \xi^\alpha$ ist ein Polynom vom (Gesamt-)Grad $\leq m$ in den n Variablen ξ_1, \dots, ξ_n und heißt *m-tes Taylorpolynom* von f .

688 / 862

Beweis. Für den Beweis betrachten wir zu gegebenem $x \in U$ und $\xi \in \mathbb{R}^n$ wie im Satz die Funktion $g : [0, 1] \rightarrow \mathbb{R}$, $g(t) := f(x + t\xi)$.

Nach Voraussetzung ist diese Funktion $m + 1$ mal stetig differenzierbar, und somit besagt die Taylorformel für Funktionen einer Variablen

$$(1) \quad f(x + \xi) = g(1) = \sum_{k=0}^m \frac{g^{(k)}(0)}{k!} \cdot 1^k + \frac{g^{(m+1)}(\tau)}{(m+1)!} \cdot 1^{m+1}$$

für ein $\tau \in [0, 1]$.

Nun gilt aber

$$g^{(k)}(t) = \frac{d^k g}{dt^k}(t) = \sum_{i_1, \dots, i_k=1}^n \frac{\partial}{\partial x^{i_k}} \left(\dots \left(\frac{\partial}{\partial x^{i_1}} f(x + t\xi) \right) \right) \cdot \xi_{i_1} \xi_{i_2} \cdots \xi_{i_k}.$$

Da $f \in C^{m+1}(U)$, sind jeweils alle Terme gleich, bei denen sich die i_j nur in der Reihenfolge unterscheiden (Lemma von Schwarz). Es gibt aber genau $\frac{k!}{\alpha_1! \cdots \alpha_n!}$ Terme, bei denen genau α_r der Indizes i_j gleich r sind.

689 / 862

Wir können also den letzten Ausdruck auch schreiben als

$$g^{(k)}(t) = \sum_{|\alpha|=k} \frac{k!}{\alpha!} \partial^\alpha f(x + t\xi) \xi^\alpha$$

Die Behauptung des Satzes folgt nun durch Einsetzen in (1). □

Folgerung (Approximation durch das Taylorpolynom)

Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ eine ***m-mal stetig differenzierbare Funktion*** und $x \in U$.

Dann existiert $\delta > 0$ und $\varphi : B_\delta(x) \rightarrow \mathbb{R}$, so dass $B_\delta(x) \subseteq U$ und

$$f(x + \xi) = \sum_{|\alpha| \leq m} \frac{1}{\alpha!} \partial^\alpha f(x) \xi^\alpha + \varphi(\xi)$$

für alle $\xi \in B_\delta(0)$, wobei

$$\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|^m} = 0.$$

690 / 862

Beweis (der Folgerung).

Da U offen ist, existiert $B_\delta(x) \subseteq U$. Aus der Taylorentwicklung der Ordnung $m - 1$ folgt für $\xi \in B_\delta(0)$:

$$f(x + \xi) = \sum_{|\alpha| \leq m-1} \frac{1}{\alpha!} \partial^\alpha f(x) \xi^\alpha + \sum_{|\alpha|=m} \frac{1}{\alpha!} \partial^\alpha f(x + \tau\xi) \xi^\alpha,$$

für ein $\tau \in [0, 1]$ (welches von ξ abhängt!).

Wir setzen

$$\varphi(\xi) := \sum_{|\alpha|=m} \frac{1}{\alpha!} (\partial^\alpha f(x + \tau\xi) - \partial^\alpha f(x)) \cdot \xi^\alpha.$$

Dann gilt $f(x + \xi) = \sum_{|\alpha| \leq m} \frac{1}{\alpha!} \partial^\alpha f(x) \xi^\alpha + \varphi(\xi)$ und, weil die m -ten partiellen Ableitungen von f nach Voraussetzung stetig in x sind,

$$\frac{|\varphi(\xi)|}{\|\xi\|^m} \leq \sum_{|\alpha|=m} \frac{1}{\alpha!} \underbrace{|\partial^\alpha f(x + \tau\xi) - \partial^\alpha f(x)|}_{\xrightarrow{\xi \rightarrow 0} 0} \cdot \underbrace{\frac{|\xi^\alpha|}{\|\xi\|^m}}_{\leq \frac{\|\xi\|_{\max}^m}{\|\xi\|^m} \leq C} \xrightarrow{\xi \rightarrow 0} 0.$$

□

691 / 862

Spezialfall:

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar und $x, \xi \in \mathbb{R}^n$ derart, dass $\{x + t\xi | t \in [0, 1]\} \subseteq U$.

Dann gilt

$$f(x + \xi) = f(x) + \underbrace{\sum_{i=1}^n \partial_i f(x) \xi_i + \frac{1}{2} \sum_{i,j=1}^n \partial_i \partial_j f(x) \xi_i \xi_j}_{\text{quadratisches Taylorpolynom}} + \varphi(\xi),$$

wobei $\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|^2} = 0$.

Hessematrix

Definition

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ **zweimal stetig differenzierbar**. Die nach dem Lemma von Schwarz symmetrische Matrix

$$\text{Hess } f(x) := (\partial_i \partial_j f(x))_{i,j=1,\dots,n}$$

heißt **Hessematrix** von f im Punkt x .

Nachtrag zum letzten Beispiel:

Mit der Hessematrix lässt sich die obige Gleichung unter Verwendung des Standardskalarprodukts auf \mathbb{R}^n nun indexfrei schreiben:

$$f(x + \xi) = f(x) + \langle \text{grad } f(x), \xi \rangle + \frac{1}{2} \langle \text{Hess } f(x) \xi, \xi \rangle + \varphi(\xi)$$

mit $\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|^2} = 0$.

693 / 862

Definition

(i) Eine symmetrische Bilinearform β auf einem reellen Vektorraum heißt **positiv semi-definit**, wenn

$$\beta(v, v) \geq 0 \quad \text{für alle } v \in V.$$

(ii) β heißt **negativ definit** (bzw. negativ semi-definit), wenn $-\beta$ positiv definit (bzw. positiv semi-definit) ist.

(iii) β heißt **indefinit**, wenn β weder positiv noch negativ semi-definit ist.

(iv) Ein symmetrischer Endomorphismus A eines Euklidischen Vektorraum heißt **positiv definit** (bzw. positiv semi-definit, indefinit etc.), wenn die zugehörige symmetrische Bilinearform

$$\beta(v, w) = \langle v, Aw \rangle, \quad v, w \in V,$$

positiv definit (bzw. positiv semi-definit, indefinit etc.) ist.

694 / 862

Hessematrix und lokale Extrema

Satz

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}$ zweimal stetig differenzierbar.

- (i) Wenn f an der Stelle $x \in U$ ein lokales Minimum (bzw. Maximum) hat, dann ist $\text{grad } f(x) = 0$ und $\text{Hess } f(x)$ ist positiv (bzw. negativ) **semi-definit**.
- (ii) Wenn der Gradient von f an der Stelle $x \in U$ verschwindet und
 - ▶ $\text{Hess } f(x)$ positiv definit ist, dann hat f in x ein **isoliertes** lokales Minimum,
 - ▶ $\text{Hess } f(x)$ negativ definit ist, dann hat f in x ein **isoliertes** lokales Maximum,
 - ▶ $\text{Hess } f(x)$ indefinit ist, dann hat f in x kein lokales Extremum.

[Animationen.]

695 / 862

Beweis.

- (i) f habe in x z.B. ein lokales Minimum. Wir wissen bereits, dass $\text{grad } f(x) = 0$ und mit $h(\xi) := \frac{1}{2} \langle \text{Hess } f(x) \xi, \xi \rangle$ gilt

$$f(x) \leq f(x + \xi) = f(x) + h(\xi) + \varphi(\xi).$$

Daraus folgt $0 \leq h(\xi) + \varphi(\xi)$, d.h. $h(\xi) \geq -\varphi(\xi)$ und somit mit $t \in \mathbb{R}^*$:

$$h\left(\frac{\xi}{\|\xi\|}\right) = \frac{h(\xi)}{\|\xi\|^2} = \frac{h(t\xi)}{\|t\xi\|^2} \geq -\frac{\varphi(t\xi)}{\|t\xi\|^2} \xrightarrow{t \rightarrow 0} 0.$$

Das impliziert $h(y) \geq 0$ für alle Einheitsvektoren y , d.h. $h \geq 0$.

696 / 862

- (ii) Sei nun umgekehrt $\text{grad } f(x) = 0$ und $h(\xi) > 0$ für alle $\xi \neq 0$.
Dann ist $m := \min_{\|y\|=1} h(y) > 0$.

Weiterhin findet man wegen $\lim_{\xi \rightarrow 0} \frac{\varphi(\xi)}{\|\xi\|^2} = 0$ zu jedem $0 < \varepsilon < m$ ein $\delta > 0$, so dass $|\varphi(\xi)| < \varepsilon \|\xi\|^2$ für alle $0 \neq \xi \in B_\delta(0)$.

Aus der Taylorentwicklung folgt dann für diese ξ :

$$f(x + \xi) = f(x) + h(\xi) + \varphi(\xi) > f(x) + m\|\xi\|^2 - \varepsilon\|\xi\|^2 > f(x),$$

d.h. x ist ein isoliertes lokales Minimum von f .

Ist h negativ definit, argumentiert man analog.

Ist h indefinit, so gibt es eine Zerlegung $\mathbb{R}^n = V^- \oplus V^0 \oplus V^+$ in Teilräume, so dass h auf V^0 verschwindet und auf V^\pm jeweils positiv bzw. negativ definit ist, mit V^\pm jeweils nicht trivial.

Dann ist x für die Einschränkung von f auf $U \cap (x + V^+)$ ein isoliertes lokales Minimum und für die Einschränkung auf $U \cap (x + V^-)$ ein isoliertes lokales Maximum. Es gibt also in jeder Umgebung von x Punkte y und z mit $f(y) < f(x) < f(z)$.

□
697 / 862

Kapitel 20

Der Umkehrsatz und seine Anwendungen

Umkehrsatz

Definition (Diffeomorphismen)

Seien U und V offene Mengen im \mathbb{R}^n . Eine bijektive, stetig differenzierbare Abbildung $f: U \rightarrow V$, deren Umkehrabbildung auch stetig differenzierbar ist, heißt **Diffeomorphismus** zwischen U und V .

Beispiele/ÜA

- Jede lineare Abbildung $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist stetig differenzierbar, mit $dA_x = A$ für alle $x \in \mathbb{R}^n$. Ist also A invertierbar, so ist diese Abbildung ein Diffeomorphismus.
- Die Abbildung $h: \mathbb{R} \times (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow (0, \infty) \times \mathbb{R}$, gegeben als

$$h(u, v) = (e^u \cos v, e^u \sin v)$$

ist ein Diffeomorphismus, mit Umkehrabbildung gegeben durch $h^{-1}(x, y) = (\frac{1}{2} \ln(x^2 + y^2), \arctan(\frac{y}{x}))$.

699 / 862

Satz

Sei $f: U \rightarrow V$ ein Diffeomorphismus. Dann ist für alle $x \in U$ die Ableitung df_x eine invertierbare lineare Abbildung und es gilt

$$(df^{-1})_{f(x)} = (df_x)^{-1}.$$

Beweis. Dies folgt aus der Kettenregel:

$$Id_{\mathbb{R}^n} = d Id_x = d(f^{-1} \circ f)_x = df_{f(x)}^{-1} \circ df_x$$

und damit $df_{f(x)}^{-1} = (df_x)^{-1}$. □

Beispiel/ÜA

Die Abbildung $h(u, v) = (e^u \cos v, e^u \sin v)$ aus dem letzten Beispiel hat die Jacobi-Matrix

$$dh_{(u,v)} = \begin{pmatrix} e^u \cos v & -e^u \sin v \\ e^u \sin v & e^u \cos v \end{pmatrix},$$

und diese ist in der Tat für alle $(u, v) \in \mathbb{R} \times (-\frac{\pi}{2}, \frac{\pi}{2})$ invertierbar.

700 / 862

Umkehrsatz

Der Umkehrsatz besagt, dass **lokal**, d.h. in einer geeignet gewählten Umgebung von x , eine Umkehrabbildung existiert, falls df_x invertierbar ist.

Satz (Umkehrsatz)

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}^n$ **stetig** differenzierbar und $p \in U$ so, dass die Ableitungsmatrix df_p invertierbar ist. Dann existieren offene Umgebungen $V \subseteq U$ von p und W von $q := f(p)$, so dass $f|_V: V \rightarrow W$ ein Diffeomorphismus ist.

Beispiele

- Sei $f: \mathbb{R} \rightarrow \mathbb{R}_+$, $f(x) = x^2$. Dann ist $f'(x) \neq 0$ für $x \in \mathbb{R} \setminus \{0\}$.
Für $x > 0$ betrachte $W = \mathbb{R}_+$ und $V = \mathbb{R}_+$ mit $f^{-1}(y) = \sqrt{y}$.
Für $x < 0$ betrachte $W = \mathbb{R}_+$ und $V = \mathbb{R}_-$ mit $f^{-1}(y) = -\sqrt{y}$.
- $f: \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^3$ besitzt eine stetige Umkehrabbildung (da streng monoton wachsend). Diese ist aber nicht differenzierbar in $y = 0 = f(0)$ da $f'(0) = 0$.

701 / 862

Der Beweis des Umkehrsatzes beruht auf dem **Banachschen Fixpunktsatz** und den folgenden beiden Sätzen.

Wir brauchen zunächst:

Definition

Seien V, W normierte Räume und $U \subseteq V$. Eine Funktion $f: U \rightarrow W$ heißt **Lipschitzstetig** auf U , wenn es eine Zahl $L \geq 0$ gibt, so dass

$$\|f(x) - f(y)\| \leq L\|x - y\|$$

für alle $x, y \in U$. Jede solche Zahl L heißt eine **Lipschitzkonstante** für f auf U .

Bemerkungen

- Lipschitzstetige Funktionen sind gleichmäßig stetig.
- Stetige lineare Funktionen sind Lipschitzstetig.

702 / 862

Satz (Schrankensatz)

Sei $f: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ stetig differenzierbar. Sei $K \subseteq U$ eine **kompakte** Menge, die **konvex** ist, d.h. für alle $x, y \in K$ liegt auch die Verbindungsstrecke $\{tx + (1-t)y \mid t \in [0, 1]\}$ in K . Dann ist f auf K **Lipschitzstetig** mit Lipschitzkonstante $L := \max_{x \in K} \|df_x\|_{\text{Operatornorm}}$. Es gilt also

$$\|f(x) - f(y)\| \leq L\|x - y\|$$

für alle $x, y \in K$.

Beweis. Da K kompakt ist und $x \mapsto df_x$ und die Operatornorm stetig sind, existiert

$$L = \max_{x \in K} \|df_x\|_{\text{Op.-norm}} = \max_{x \in K} \max_{\xi \in \mathbb{R}^n, \|\xi\|=1} \|df_x(\xi)\|$$

Nach dem **Mittelwertsatz** existiert dann für $x, y \in K$ ein $t_0 \in [0, 1]$, so dass

$$\|f(x) - f(y)\| = \|df_{t_0x+(1-t_0)y}(y-x)\| \leq L\|y-x\|. \quad \square$$

703 / 862

Satz (Diffeomorphie)

Seien U und V offene Mengen im \mathbb{R}^n und $f: U \rightarrow V$ stetig differenzierbar und bijektiv mit stetiger Umkehrabbildung $f^{-1}: V \rightarrow U$. Ist für jedes $x \in U$ das Differential df_x invertierbar, so ist f ein Diffeomorphismus, d.h. f^{-1} ist auch stetig differenzierbar mit $df_{f(x)}^{-1} = (df_x)^{-1}$.

Wir erinnern uns:

$f(x) = x^3$ hat eine stetige, aber in $x = 0$ nicht stetig differenzierbare Umkehrabbildung. Die Bedingung df_x invertierbar ist also wesentlich.

Beweis. Wir zeigen die stetige Differenzierbarkeit von f^{-1} in $y_0 = f(x_0)$. O.B.d.A. können wir annehmen, dass $x_0 = y_0 = 0$ und $df_{x_0} = \text{Id}$, denn: Sind $A, B: \mathbb{R}^n \rightarrow \mathbb{R}^n$ invertierbare affine Abbildungen, so gilt der Satz für f genau dann, wenn er für $A \circ f \circ B$ gilt. Mit $B(x) = x_0 + x$ und $A(y) = (df_{x_0})^{-1}(y - y_0)$ haben wir aber $A \circ f \circ B(0) = 0$ und

$$d(A \circ f \circ B)_0 = dA_{f(x_0)} \circ df_{x_0} \circ \text{Id} = df_{x_0}^{-1} \circ df_{x_0} = \text{Id}.$$

704 / 862

Weiter im Beweis des Satzes: Sei $y \in V$ und $x := f^{-1}(y) \in U$. Wir definieren

$$\begin{aligned}\varphi(x) &:= f(x) - f(0) - df_0(x) = f(x) - x \\ \psi(y) &:= f^{-1}(y) - y = x - f(x) = -\varphi(x) = -\varphi(f^{-1}(y)).\end{aligned}$$

Um die Differenzierbarkeit von f^{-1} zu zeigen, müssen zeigen, dass

$$\frac{\psi(y)}{\|y\|} \xrightarrow{y \rightarrow 0} 0.$$

Da f stetig differenzierbar ist, gilt $\frac{\varphi(x)}{\|x\|} \xrightarrow{x \rightarrow 0} 0$.

Daher können wir ein $\varepsilon > 0$ finden, so dass $\frac{\|\varphi(x)\|}{\|x\|} \leq \frac{1}{2}$ für alle $\|x\| < \varepsilon$.

Da f^{-1} stetig ist, finden wir zu diesem ε ein $\delta > 0$, so dass $\|f^{-1}(y)\| < \varepsilon$ für alle $\|y\| < \delta$. Somit gilt für alle y mit $\|y\| < \delta$, dass

$$\|\psi(y)\| = \|\varphi(f^{-1}(y))\| \leq \frac{1}{2}\|f^{-1}(y)\|, \quad (*)$$

und daher wegen der Dreiecksungleichung

$$\|f^{-1}(y)\| \leq \|f^{-1}(y) - y\| + \|y\| = \|\psi(y)\| + \|y\| \stackrel{(*)}{\leq} \frac{1}{2}\|f^{-1}(y)\| + \|y\|.$$

705 / 862

Ende des Beweises: Also haben wir für alle y mit $\|y\| < \delta$, dass

$$\|f^{-1}(y)\| \leq 2\|y\|$$

und somit

$$\frac{\|\psi(y)\|}{\|y\|} = \frac{\|\varphi(x)\|}{\|x\|} \frac{\|x\|}{\|y\|} = \frac{\|\varphi(x)\|}{\|x\|} \frac{\|f^{-1}(y)\|}{\|y\|} \leq 2 \cdot \frac{\|\varphi(x)\|}{\|x\|}.$$

Aus $y \rightarrow 0$ folgt nun, wegen der Stetigkeit von f^{-1} , dass auch $x = f^{-1}(y) \rightarrow 0$ und somit

$$\frac{\|\psi(y)\|}{\|y\|} \leq 2 \cdot \frac{\|\varphi(x)\|}{\|x\|} \xrightarrow{y \rightarrow 0} 0.$$

Damit ist f^{-1} differenzierbar in y und mit $d(f^{-1})_y = (df_x)^{-1}$ erhalten wir auch die Stetigkeit von df^{-1} . \square

706 / 862

Jetzt beweisen wir den Umkehrsatz:

Satz (Umkehrsatz)

Sei $U \subseteq \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}^n$ stetig differenzierbar und $p \in U$ so, dass die Ableitungsmatrix df_p invertierbar ist.

Dann existieren offene Umgebungen $V \subseteq U$ von p und W von $q := f(p)$, so dass $f|_V: V \rightarrow W$ ein Diffeomorphismus ist, d.h. $f|_V$ ist bijektiv und $(f|_V)^{-1}: W \rightarrow V$ ist stetig differenzierbar.

Beweis. Wir können wieder o.B.d.A. annehmen, dass $p = q = 0$ und $df_0 = \text{Id}$. Der Beweis des Umkehrsatzes erfolgt nun in mehreren Schritten:

1) Definition von W : Sei $\delta > 0$, so dass $\overline{B_{2\delta}(0)} \subseteq U$ und so, dass

$$\|Id - df_x\|_{Op.-Norm} \leq \frac{1}{2} \quad \text{für alle } x \in \overline{B_{2\delta}(0)}. \quad (*)$$

Wir setzen dann $W := B_\delta(0)$.

2) Definition von V : Wir setzen $V := f^{-1}(W) \cap B_{2\delta}(0) \subseteq U$.

Da W offen ist und f stetig, ist auch das Urbild $f^{-1}(W)$ offen.

Damit ist V als Durchschnitt zweier offener Mengen offen.

707 / 862

Weiter im Beweis: 3) Die Abbildung $f|_V: V \rightarrow W$ ist bijektiv: Zu $y \in W = B_\delta(0)$ definieren wir die differenzierbare Abbildung

$$\varphi_y: B_{2\delta}(0) \rightarrow \mathbb{R}^n, \quad \varphi_y(x) := y + x - f(x).$$

Ein Fixpunkt x von φ_y ist eine Lösung von $f(x) = y$.

Wegen $(d\varphi_y)_x = Id - df_x$ liefern die Wahl von $\delta > 0$ und der

Schranksatz

$$\|\varphi_y(x_1) - \varphi_y(x_2)\| \leq \frac{1}{2} \|x_1 - x_2\| \quad \text{für } x_1, x_2 \in B_{2\delta}(0). \quad (**)$$

Da $\|y\| < \delta$ und $\varphi_y(0) = y$, gilt für alle $x \in \overline{B_{2\delta}(0)}$ die Ungleichung

$$\|\varphi_y(x)\| \leq \|\varphi_y(x) - \varphi_y(0)\| + \|y\| < 2\delta,$$

D.h. $\varphi_y: \overline{B_{2\delta}(0)} \rightarrow B_{2\delta}(0)$. Wegen $(**)$ ist φ_y also eine kontrahierende Abbildung auf dem vollständigen metrischen Raum $\overline{B_{2\delta}(0)}$.

Nach dem **Banachschen Fixpunktsatz** finden wir also genau ein $x \in \overline{B_{2\delta}(0)}$ mit $\varphi_y(x) = x$.

Wegen $\|x\| = \|\varphi_y(x)\| < 2\delta$, gilt sogar $x \in B_{2\delta}(0)$. D.h. aber $f(x) = y$ und $x \in f^{-1}(W) \cap B_{2\delta}(0) = V$. Somit ist $f: V \rightarrow W$ bijektiv.

708 / 862

Weiter im Beweis:4) f^{-1} ist stetig:Wegen (**) und der Dreiecksungleichung erhalten wir für $x_1, x_2 \in V$

$$\begin{aligned} \|x_2 - x_1\| &\leq \|\varphi_0(x_2) - \varphi_0(x_1)\| + \|f(x_2) - f(x_1)\| \\ &\stackrel{(**)}{\leq} \frac{1}{2}\|x_2 - x_1\| + \|f(x_2) - f(x_1)\|. \end{aligned}$$

Damit ist für $x_1 = f^{-1}(y_1)$ und $x_2 = f^{-1}(y_2)$

$$\|f^{-1}(y_2) - f^{-1}(y_1)\| \leq 2\|y_2 - y_1\|,$$

und somit ist die Umkehrfunktion f^{-1} auf W (Lipschitz)stetig.5) Für alle $x \in V$ ist df_x ein Isomorphismus:Für $x \in V \subseteq B_{2\delta}(0)$ gilt wegen der Wahl von δ in (*) für $\xi \in \mathbb{R}^n$

$$\|(Id - df_x)(\xi)\| \leq \frac{1}{2}\|\xi\|.$$

Für $\xi \in \ker(df_x)$ ist dann $\|\xi\| \leq \frac{1}{2}\|\xi\|$, d.h. $\xi = 0$. Also ist df_x invertierbar.Aus dem vorhergehenden Satz folgt somit die Behauptung des Umkehrsatzes. 709 852**Bemerkung:**

f^{-1} ist sogar k -mal stetig differenzierbar, wenn f k -mal stetig differenzierbar ist ($k \in \mathbb{N}$). Die Gleichung $df_y^{-1} = (df_{g(y)})^{-1}$ zeigt nämlich, dass f^{-1} k -mal stetig differenzierbar ist, wenn f k -mal und $g = f^{-1}$ $(k-1)$ -mal stetig differenzierbar ist.

Folgerung (Offenheits-/Diffeomorphiesatz)

Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Abbildung mit df_x invertierbar für alle $x \in U$. Dann gilt

- $f(U) \subseteq \mathbb{R}^n$ ist offen.
- Ist f injektiv, so ist f ein Diffeomorphismus $U \rightarrow f(U)$.

Beweis. Dies folgt aus dem Umkehrsatz und dem Satz über Diffeomorphie bei stetiger Umkehrabbildung. (ÜA) □

Beispiel (ebene Polarkoordinaten)

Die Abbildung

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad \begin{pmatrix} r \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix},$$

ist unendlich oft differenzierbar. Ihr Differential ist

$$df = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{pmatrix} = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$

Da $\det df = r$, gibt es zu jedem Punkt $p = (r_0, \varphi_0) \in \mathbb{R}^* \times \mathbb{R}$ eine offene Umgebung $U \subseteq \mathbb{R}^* \times \mathbb{R}$, so dass f die Menge U diffeomorph auf eine offene Menge $V = f(U) \subseteq \mathbb{R}^2$ abbildet. Die Umkehrabbildung $(f|_U)^{-1}: V \rightarrow U$ ist ebenfalls unendlich oft differenzierbar.

Eine typische Wahl ist $U = \mathbb{R}_+ \times (\varphi_0 - \pi, \varphi_0 + \pi)$.

711 / 862

Satz über implizite Funktionen

Motivation

Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion. Wir betrachten die Gleichung für $x \in \mathbb{R}^n$

$$f(x_1, \dots, x_n) = 0,$$

bzw. die Lösungsmenge $N_f := \{x \in \mathbb{R}^n \mid f(x) = 0\}$.

- Ist f linear, d.h. $f(x) = \sum_i c_i x_i$, so ist die Lösungsmenge $N_f(0)$ ein $(n-1)$ -dimensionaler Unterraum.

Ist $c_i \neq 0$, so können wir auf der Menge $N_f(0)$ die Variable x_i als lineare Funktion der anderen Variablen darstellen.

Ist z.B. $c_n \neq 0$, so gilt

$$f(x_1, \dots, x_n) = 0 \iff x_n = -\frac{1}{c_n} \sum_{i=1}^{n-1} c_i x_i.$$

712 / 862

- Was passiert, wenn f nicht linear ist? Unter welchen Bedingungen kann man nun die Gleichung nach einer Koordinate auflösen, d.h. wann existiert eine Funktion $g: \mathbb{R}^{n-1} \rightarrow \mathbb{R}$, so dass

$$f(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1})) = 0, \quad \text{d.h. } N_f(0) \supset \text{graph}(g)?$$

Ist g differenzierbar, falls f differenzierbar ist?

Beispiel

- Sei $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := x^2 + y^2 - 1$. Es ist $N_f(0) = S^1 \subseteq \mathbb{R}^2$ der Kreis. Hier benötigen wir zwei Funktionen, um ganz $N_f(0)$ als Graphen darzustellen:

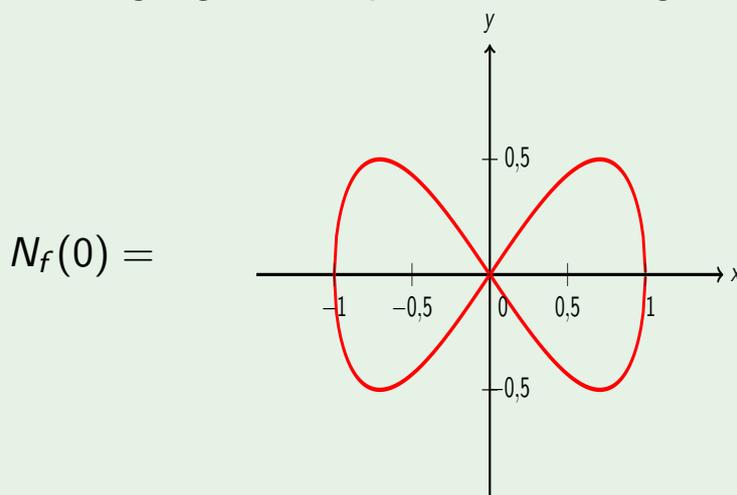
$$S^1 = \underbrace{\{(x, y) \in S^1 \mid y \geq 0\}}_{\text{graph}(g_+)} \cup \underbrace{\{(x, y) \in S^1 \mid y \leq 0\}}_{\text{graph}(g_-)}$$

mit $g_{\pm}: [-1, 1] \rightarrow \mathbb{R}_{\pm}$. Die Funktionen $g_{\pm}(x) := \pm\sqrt{1-x^2}$ sind differenzierbar nur auf dem offenen Intervall $(-1, 1)$.

713 / 862

Beispiele(Fortsetzung)

- $f(x, y, z) = x^2 + y^2$, d.h. $N_f = \{(0, 0, z) \mid z \in \mathbb{R}\}$. Hier gibt es kein $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ mit $\text{graph}(g) = N_f$.
- $f(x, y) = y^2 - x^2(1-x^2)$. $N_f(0) \setminus \{(-1, 0), (0, 0), (1, 0)\}$ ist Vereinigung von Graphen von 4 stetig differenzierbaren Funktionen.



- $f(x, y) := x - y^3$. Beim Versuch $N_f(0)$ als Graph einer Funktion von x zu schreiben, erhalten wir $g(x) = \sqrt[3]{x}$, was in 0 nicht differenzierbar ist.

714 / 862

Satz über implizite Funktionen

Satz (Satz über implizite Funktionen)

Sei $U \subseteq \mathbb{R}^m \times \mathbb{R}^n$ offen, $f: U \rightarrow \mathbb{R}^n$ k -mal stetig differenzierbar ($k \geq 1$) und $(p, q) \in U$, so dass $f(p, q) = 0$. Weiterhin sei das Differential der Abbildung $y \mapsto f(p, y)$ im Punkt $y = q$ **invertierbar**.

Dann gibt es offene Umgebungen $V \subseteq \mathbb{R}^m$ von p und $W \subseteq \mathbb{R}^n$ von q und eine k -mal stetig differenzierbare Abbildung $g: V \rightarrow W$, so dass für alle $(x, y) \in V \times W$ gilt: $f(x, y) = 0 \iff y = g(x)$.

Anders ausgedrückt gilt $N_f(0) \cap V \times W = \text{graph}(g)$.

Beweis. Die Hilfsfunktion $F: U \rightarrow \mathbb{R}^m \times \mathbb{R}^n$, $F(x, y) := (x, f(x, y))$ hat in (p, q) ein Differential, das sich wie folgt aus dem Differential der Abbildung $y \mapsto f(x, y) = \sum_{i=1}^n f_i(x, y)e_i$ berechnet:

$$dF_{(p,q)} = \begin{pmatrix} \mathbf{1}_m & 0 \\ * & \left(\frac{\partial f_i}{\partial y_j} \right)_{i,j} (p, q) \end{pmatrix}$$

715 / 862

Weiter im Beweis: Da $\left(\frac{\partial f_i}{\partial y_j} \right)_{i,j}$ im Punkt (p, q) invertierbar ist, ist auch $dF_{(p,q)}$ im Punkt (p, q) invertierbar.

Nach dem **Umkehrsatz** gibt es offene Umgebungen $V_1 \subseteq \mathbb{R}^m$ von p und $V_2 \subseteq \mathbb{R}^n$ von q , so dass $V_1 \times V_2 \subseteq U \subseteq \mathbb{R}^m \times \mathbb{R}^n$ durch F bijektiv auf eine offene Umgebung $\Omega \subseteq \mathbb{R}^m \times \mathbb{R}^n$ von $F(p, q) = (p, f(p, q)) = (p, 0)$ abgebildet wird, und zwar so, dass die Umkehrabbildung

$$G = (G_1, G_2): \Omega \rightarrow V_1 \times V_2$$

k -mal stetig differenzierbar ist.

Dann ist $V := \{x \in V_1 \mid (x, 0) \in \Omega\}$ eine offene Umgebung $V \subseteq V_1$ von p . Wir setzen nun $W := V_2$ und

$$g: V \rightarrow W, \quad g(x) := G_2(x, 0).$$

Da G k -mal stetig differenzierbar ist, gilt dies auch für g .

716 / 862

Weiter im Beweis: Aus

$$\Omega \ni (x, y) = F(G(x, y)) = (G_1(x, y), f(G_1(x, y), G_2(x, y))) \quad (*)$$

folgt dann $G_1(x, y) = x$ und $f(x, G_2(x, y)) = y$.

Wir überprüfen nun für alle $(x, y) \in V \times W$ die Äquivalenz

$$f(x, y) = 0 \iff y = g(x):$$

(\Leftarrow) Da $g(x) = G_2(x, 0)$ folgt aus (*), dass $f(x, y) = f(x, g(x)) = 0$.

(\Rightarrow) Umgekehrt folgt aus $f(x, y) = 0$, dass $F(x, y) = (x, f(x, y)) = (x, 0)$, und somit

$$(x, y) = G(F(x, y)) = G(x, 0) = (G_1(x, 0), G_2(x, 0)) = (x, g(x)),$$

d.h. $y = g(x)$.

Damit ist der Satz über implizite Funktionen bewiesen. \square

717 / 862

Beispiel (zweischaliges Hyperboloid)

Die Funktion $f: \mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$,

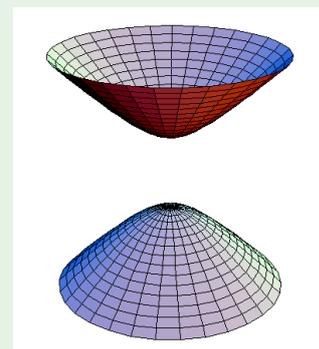
$$f(x, y, z) := x^2 + y^2 - z^2 + 1,$$

erfüllt $\frac{\partial f}{\partial z} = -2z \neq 0$ für alle $(x, y, z) \in H := f^{-1}(0)$.

Nach dem Satz über implizite Funktionen definiert die Gleichung $f(x, y, z) = 0$ also **lokal** eine unendlich oft differenzierbare Funktion $(x, y) \mapsto z = g(x, y)$. Diese kann man durch Auflösen der Gleichung $f(x, y, z) = 0$ nach z explizit angeben:

$$g_{\pm}(x, y) = \pm \sqrt{x^2 + y^2 + 1}.$$

Der Graph von $g_+: \mathbb{R}^2 \rightarrow \mathbb{R}^+$, bzw. $g_-: \mathbb{R}^2 \rightarrow \mathbb{R}_-$ ist die obere, bzw. untere, Schale des Hyperboloids H .



718 / 862

Differential einer implizit definierten Funktion

Unter den Voraussetzungen des Satzes über implizite Funktionen lässt sich das Differential der durch die Gleichung $f(x, y) = 0$ implizit definierten Abbildung $x \mapsto g(x)$ mit $N_f(0) = \text{graph } g$ wie folgt berechnen:

Bezeichnet $d_x f$ das Differential der Abbildung $x \rightarrow f(x, y)$ und $d_y f$ das Differential der Abbildung $y \rightarrow f(x, y)$, dann liefert Ableiten der Gleichung $f(x, g(x)) = 0$ nach der Kettenregel:

$$0 = d_x f + d_y f dg, \quad \text{also} \quad dg = -(d_y f)^{-1} d_x f.$$

Hierbei sind dg an der Stelle x und $d_x f, d_y f$ an der Stelle $(x, g(x))$ auszuwerten:

$$dg|_x = - (d_y f|_{(x, g(x))})^{-1} d_x f|_{(x, g(x))}.$$

In Komponenten, $i = 1, \dots, n$ und $j = 1, \dots, m$:

$$\frac{\partial f_i}{\partial x_j}(x, g(x)) + \sum_{k=1}^n \frac{\partial f_i}{\partial y_k}(x, g(x)) \cdot \frac{\partial g_k}{\partial x_j}(x) = 0.$$

719 / 862

Beispiel

Wir betrachten noch einmal die Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$,

$$f(x, y, z) = x^2 + y^2 - z^2 + 1.$$

Wie wir bereits gesehen haben, ist in allen Punkten $(x, y, z) \in N_f(0)$ die partielle Ableitung $\partial_z f \neq 0$, wir finden also in der Nähe eines solchen Punktes eine Funktion $g = g(x, y)$ mit

$$f(x, y, z) = 0 \iff z = g(x, y).$$

Durch partielles Ableiten der Funktion $F(x, y) = f(x, y, g(x, y))$ nach x und y erhalten wir die Gleichungen

$$\partial_x f + \partial_z f \partial_x g = 0 \quad \text{und} \quad \partial_y f + \partial_z f \partial_y g = 0.$$

Aus $\partial_x f = 2x$ und $\partial_z f = -2z$ folgt also z.B.

$$\partial_x g(x, y) = - \frac{\partial_x f(x, y, g(x, y))}{\partial_z f(x, y, g(x, y))} = \frac{x}{g(x, y)}.$$

720 / 862

Definition (Rang einer differenzierbaren Abbildung)

$U \subseteq \mathbb{R}^m$ sei offen und $f: U \rightarrow \mathbb{R}^n$ differenzierbar.

(i) Der **Rang** von f im Punkt $p \in U$ ist definiert als

$$\operatorname{rg}(f)_p := \operatorname{rg} df_p \leq \min\{m, n\}.$$

Das definiert eine Funktion $\operatorname{rg}(f): U \rightarrow \{0, 1, 2, \dots, \min\{m, n\}\}$.

- (ii) f heißt **Immersion**, wenn $\operatorname{rg}(f) \equiv m$, d.h. wenn das Differential überall injektiv ist.
- (iii) f heißt **Submersion**, wenn $\operatorname{rg}(f) \equiv n$, d.h. wenn das Differential überall surjektiv ist.
- (iv) k -mal stetig differenzierbare Abbildungen heißen auch **von der Klasse C^k oder C^k -Abbildungen**. Wir erlauben hier $k \in \mathbb{N} \cup \{\infty\}$.
- (v) $U \subseteq \mathbb{R}^m$, $V \subseteq \mathbb{R}^n$ seien offen. Eine C^k -Abbildung $f: U \rightarrow V$ heißt **C^k -Diffeomorphismus**, wenn f bijektiv ist und auch f^{-1} von der Klasse C^k ist.

721 / 862

Beispiele

(i) Sei $m \leq n$. Dann ist die Abbildung $\iota: \mathbb{R}^m \rightarrow \mathbb{R}^n$,

$$\iota(x_1, \dots, x_m) = (x_1, \dots, x_m, 0, \dots, 0) \in \mathbb{R}^m \times \mathbb{R}^{n-m} = \mathbb{R}^n$$

eine Immersion. Diese nennt man auch die kanonische Immersion.

Warnung: Immersionen sind nicht unbedingt injektive Abbildungen. Als Beispiel betrachte man die Abbildung $h: \mathbb{R} \rightarrow \mathbb{R}^2$,

$$h(t) = (\cos t, \sin t).$$

(ii) Sei $m \geq n$. Dann ist die Abbildung $\pi: \mathbb{R}^m \rightarrow \mathbb{R}^n$,

$$\pi(x_1, \dots, x_m) = (x_1, \dots, x_n)$$

eine Submersion. Man nennt sie auch die kanonische Submersion.

Warnung: Submersionen sind nicht unbedingt surjektive Abbildungen. Können Sie ein Beispiel angeben?

722 / 862

Beispiele

(iii) Sei $r \leq \min\{m, n\}$. Die Abbildung $\mathbb{R}^m \rightarrow \mathbb{R}^n$,

$$(x_1, \dots, x_m) \mapsto (x_1, \dots, x_r, 0, \dots, 0) \in \mathbb{R}^r \times \mathbb{R}^{n-r} = \mathbb{R}^n$$

hat konstanten Rang r .

(iv) Sei $f: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion. Dann ist $F: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^{n+1}$,
 $F(x) = (x, f(x))$ eine Immersion, denn $\text{rg}(dF_x) \equiv n$.

Das Bild von F ist der **Graph** von f .

(v) Der Rang eines Diffeomorphismus $f: U \rightarrow V$, mit $U \subseteq \mathbb{R}^m$, $V \subseteq \mathbb{R}^n$ offen, ist konstant gleich $m = n$.

Durch Ableiten der Gleichungen $f^{-1} \circ f = \text{Id}_U$ und $f \circ f^{-1} = \text{Id}_V$ folgt nämlich, dass df_p für alle $p \in U$ invertierbar ist, d.h.

$$m = n = \text{rg}(f).$$

Andererseits besagt der Umkehrsatz, dass jede C^k -Abbildung $f: U \rightarrow V$ vom konstantem maximalen Rang n zwischen offenen Mengen des \mathbb{R}^n **lokal** ein C^k -Diffeomorphismus ist.

723 / 862

Wir zeigen nun, dass jede Abbildung von konstantem Rang in geeigneten Koordinaten wie eine der oben angegebenen linearen Abbildungen aussieht.

Satz (Normalform für Abbildungen von konstantem Rang)

Sei $U \subseteq \mathbb{R}^m$ offen, $p \in U$ und $f: U \rightarrow \mathbb{R}^n$ eine C^k -Abbildung mit $\text{rg}(f) = r$ auf U .

Dann existieren offene Umgebungen $V \subseteq U$ von p und $W \subseteq \mathbb{R}^n$ von $f(p)$ und C^k -Diffeomorphismen $\varphi: V \rightarrow \varphi(V) \subseteq \mathbb{R}^m$, $\psi: W \rightarrow \psi(W) \subseteq \mathbb{R}^n$, so dass

$$(\psi \circ f \circ \varphi^{-1})(x_1, \dots, x_m) = (x_1, \dots, x_r, \underbrace{0, \dots, 0}_{n-r})$$

für alle $x \in \varphi(V)$.

Beweis. Nach eventueller Umnummerierung der Koordinaten im \mathbb{R}^m und \mathbb{R}^n können wir annehmen, dass die quadratische Matrix

$$A = \left(\frac{\partial f_i(p)}{\partial x_j} \right)_{i,j=1,\dots,r} \quad \text{invertierbar ist.}$$

724 / 862

Weiter im Beweis:

Wir betrachten die Hilfsabbildung $F : U \rightarrow \mathbb{R}^m$,
 $F(x) = (f_1(x), \dots, f_r(x), x_{r+1}, \dots, x_m)$.

Da

$$dF_p = \begin{pmatrix} A & * \\ 0 & \mathbf{1}_{m-r} \end{pmatrix}$$

invertierbar ist, existiert eine offene Umgebung $V \subseteq U$ von p , so dass $\varphi := F|_V : V \rightarrow \varphi(V)$ ein C^k -Diffeomorphismus ist.

Aus $F \circ \varphi^{-1}(x) = x$ folgt dann $f_i(\varphi^{-1}(x)) = x_i$ für alle $i \leq r$ und $x \in \varphi(V)$.

Für $\tilde{f} := f \circ \varphi^{-1}$ gilt dann also $\tilde{f}_i(x) = x_i$ für alle $i \leq r$. Außerdem können wir den Definitionsbereich von \tilde{f} zu einer Umgebung von p von der Form $\tilde{U} = U' \times U'' \subset \mathbb{R}^r \times \mathbb{R}^{m-r}$ verkleinern, mit U' und U'' konvex.

725 / 862

Weiter im Beweis:

Nun ist

$$d\tilde{f}_p = \begin{pmatrix} \mathbf{1}_r & 0 \\ * & B \end{pmatrix}, \quad B = \left(\frac{\partial \tilde{f}_i}{\partial x_j} \right)_{i,j \geq r+1},$$

und aus $\text{rg}(\tilde{f}) = \text{rg}(f) = r$ folgt dann $B = 0$, d.h. \tilde{f} hängt nur von (x_1, \dots, x_r) ab.

Also hat \tilde{f} die Form $\tilde{f}(x) = (x', f''(x'))$, wobei

$$x' = (x_1, \dots, x_r) \quad \text{und} \quad f'' = (\tilde{f}_{r+1}, \dots, \tilde{f}_n),$$

d.h. das Bild von \tilde{f} ist der Graph von $f'' : U' \rightarrow \mathbb{R}^{n-r}$. Nun ist

$W := U' \times \mathbb{R}^{n-r}$ eine offene Umgebung von $f(p) = \tilde{f}(\varphi(p)) = (p', f''(p'))$, wobei $p' \in U'$ die Projektion von $\varphi(p) = (p', p'')$ nach \mathbb{R}^r ist.

Wir definieren nun den C^k -Diffeomorphismus $\psi : W \rightarrow W \subseteq \mathbb{R}^n$,

$$\psi(x) := (x', x'' - f''(x')) \quad \text{für} \quad x = (x', x'') \in W = U' \times \mathbb{R}^{n-r}.$$

Dann gilt, wie gewünscht,

$$\psi(f(\varphi^{-1}(x))) = \psi\tilde{f}(x) = \psi(x', f''(x')) = (x', f''(x') - f''(x')) = (x', 0). \quad \square$$

726 / 862

Ist $r = n \leq m$, so können wir den zweiten Beweisschritt (die Konstruktion des Diffeomorphismus ψ) weglassen, und erhalten

Folgerung (Normalform für Submersionen)

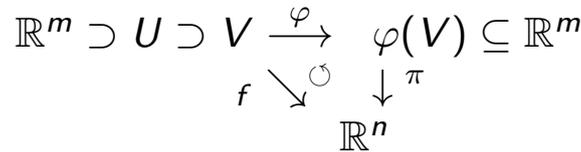
Sei $U \subseteq \mathbb{R}^m$ offen, $p \in U$ und $f: U \rightarrow \mathbb{R}^n$ eine C^k -Abbildung mit $\text{rg}(f)_p = n \leq m$.

Dann existiert eine offene Umgebung $V \subseteq U$ des Urbilds $p \in U$ und ein C^k -Diffeomorphismus $\varphi: V \rightarrow \varphi(V) \subseteq \mathbb{R}^m$, so dass

$$(f \circ \varphi^{-1})(x_1, \dots, x_m) = (x_1, \dots, x_n)$$

auf $\varphi(V)$. Insbesondere ist f lokal eine Submersion. □

Mit der kanonischen Projektion $\pi: \mathbb{R}^m \rightarrow \mathbb{R}^n$ bekommen wir folgendes Diagramm:



Das Symbol \circlearrowright bedeutet $f \circ \varphi^{-1} = \pi$, d.h. das Diagramm *kommutiert*. 727 / 862

Ist andererseits $r = m < n$, so können wir im ersten Beweisschritt des Satzes statt F die Hilfsabbildung $G: U \times \mathbb{R}^{n-m} \rightarrow \mathbb{R}^n$,

$$G(x, z) := f(x) + (0, z)$$

betrachten. G ist von der Klasse C^k und

$$dG_{(p,z)} = \left(df_p \mid \begin{array}{c} 0 \\ \mathbf{1}_{n-m} \end{array} \right) = \left(\begin{array}{c} A \\ * \end{array} \mid \begin{array}{c} 0 \\ \mathbf{1}_{n-m} \end{array} \right) \in \text{Mat}_n(\mathbb{R}).$$

Da wir o.B.d.A wieder annehmen, dass $A = \left(\frac{\partial f_i}{\partial x_j} \right)_{i,j=1,\dots,m}$ invertierbar ist, ist auch $dG_{(p,0)}$ invertierbar. Nach dem Umkehrsatz existiert dann eine offene Umgebung W von $(p, 0) \in \mathbb{R}^n$, so dass $G|_W: W \rightarrow G(W)$ ein C^k -Diffeomorphismus auf eine offene Umgebung $V := G(W)$ von $G(p, 0) = f(p)$ ist.

Der C^k -Diffeomorphismus $\psi := (G|_W)^{-1}: V \rightarrow W = \psi(V)$ erfüllt dann

$$\psi(f(x)) = \psi(G(x, 0)) = (x, 0).$$

Wir haben also gezeigt:

Satz (Normalform für Immersionen)

Sei $U \subseteq \mathbb{R}^m$ offen, $p \in U$ und $f: U \rightarrow \mathbb{R}^n$ eine C^k -Abbildung mit $\text{rg}(f)_p = m \leq n$.

Dann existiert eine offene Umgebung V des Bildes $f(p)$ und ein C^k -Diffeomorphismus $\psi: V \rightarrow \psi(V) \subseteq \mathbb{R}^n$, so dass

$$(\psi \circ f)(x_1, \dots, x_m) = (x_1, \dots, x_m, 0, \dots, 0)$$

in einer Umgebung von p . Insbesondere ist f lokal eine Immersion.

Schematisch sieht die Situation bei einer Immersion so aus:

$$\begin{array}{ccc} & & V \subseteq \mathbb{R}^n \\ & \nearrow f & \downarrow \psi \\ \mathbb{R}^m \supset U & \xrightarrow{\iota} & \psi(V) \subseteq \mathbb{R}^n \end{array}$$

wobei ι die kanonische Immersion ist, und $\psi \circ f = \iota$ gilt.

729 / 862

Beispiel einer Abbildung mit konstantem Rang

Die Abbildung $f: \text{Mat}(n, \mathbb{R}) \rightarrow \text{Mat}(n, \mathbb{R})$, $A \mapsto f(A) = A^T A$, hat auf der offenen Teilmenge $\text{GL}(n, \mathbb{R}) \subseteq \text{Mat}(n, \mathbb{R})$ konstanten Rang $r = \frac{n(n+1)}{2}$, wie im Folgenden gezeigt wird.

Wir können f als Abbildung $f: \text{Mat}(n, \mathbb{R}) \rightarrow \text{Sym}(n, \mathbb{R})$ in den Unterraum $\text{Sym}(n, \mathbb{R}) \subseteq \text{Mat}(n, \mathbb{R})$ der symmetrischen Matrizen auffassen.

Das Differential $df_A: B \mapsto B^T A + A^T B$, $\text{Mat}(n, \mathbb{R}) \rightarrow \text{Sym}(n, \mathbb{R})$, ist für $A \in \text{GL}(n, \mathbb{R})$ surjektiv:

Sei $C \in \text{Sym}(n, \mathbb{R})$. df_A bildet $B = \frac{1}{2}(A^{-1})^T C$ auf $\frac{1}{2}(C^T + C) = C$ ab.

Daraus folgt $\text{rg}(f)_A = \dim \text{Sym}(n, \mathbb{R}) = \frac{n(n+1)}{2}$.

Somit definiert f eine Submersion $\text{GL}(n, \mathbb{R}) \rightarrow \text{Sym}(n, \mathbb{R})$ und hat also konstanten Rang $r = \frac{n(n+1)}{2}$.

Frage: Ist f surjektiv?

730 / 862

Kapitel 21

Mannigfaltigkeiten

731 / 862

Untermannigfaltigkeiten im Euklidischen Raum

Wir wollen nun Teilmengen des Euklidischen Raumes betrachten, die lokal durch eine Immersion oder eine Submersion gegeben sind.

Bemerkung: Die auf Teilmengen induzierte Metrik

Sei $Y \subseteq (X, d)$ eine Teilmenge eines metrischen Raumes (X, d) .

Dann definiert die Einschränkung von d auf Y eine Metrik d_Y auf Y , so dass (Y, d_Y) ein metrischer Raum ist. Es heißt d_Y **induzierte Metrik**.

Beispiel

Die zweidimensionale Einheitssphäre

$S^2 = S_1(0) = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ ist eine abgeschlossene Teilmenge im Euklidischen Raum und bezüglich der induzierten Metrik ein vollständiger metrischer Raum.

Die obere Halbsphäre $\{(x, y, z) \in S^2 \mid z > 0\}$ ist eine offene Teilmenge der Sphäre S^2 und kein vollständiger topologischer Raum.

732 / 862

Definition

Eine bijektive stetige Abbildung $\Phi: X \rightarrow Y$ zwischen metrischen Räumen X und Y heißt **Homöomorphismus**, wenn auch die Umkehrfunktion $\Phi^{-1}: Y \rightarrow X$ stetig ist.

Bezeichnung: $\Phi: X \xrightarrow{\sim} Y$.

Beispiel

Die Abbildung

$$\Phi: [0, 2\pi) \rightarrow S^1 = \{z \in \mathbb{C} \mid |z| = 1\}, \quad \varphi \mapsto e^{i\varphi},$$

ist stetig und bijektiv, aber kein Homöomorphismus. Denn die Bildfolge $z_n = \Phi(2\pi - \frac{1}{n}) \in S^1$ konvergiert gegen 1, aber die Urbildfolge $\Phi^{-1}(z_n) = 2\pi - \frac{1}{n} \in [0, 2\pi)$ konvergiert nicht gegen $\Phi^{-1}(1) = 0$.

Die Einschränkung von Φ auf das offene Intervall $(0, 2\pi)$ definiert jedoch einen Homöomorphismus von $(0, 2\pi)$ auf die offene Teilmenge $S^1 \setminus \{1\}$ der Einheitskreislinie S^1 .

733 / 862

Untermannigfaltigkeiten

Definition (Untermannigfaltigkeiten im \mathbb{R}^n)

Eine Teilmenge $M \subseteq \mathbb{R}^n$ heißt **m -dimensionale C^k -Untermannigfaltigkeit**, wenn es zu jedem $p \in M$ eine offene Umgebung $V \subseteq \mathbb{R}^n$, eine offene Teilmenge $U \subseteq \mathbb{R}^m$ und eine C^k -Immersion $F: U \rightarrow \mathbb{R}^n$ gibt, die U homöomorph auf die offene Teilmenge $V \cap M$ von M abbildet.

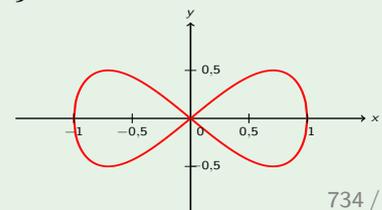
$F: U \xrightarrow{\sim} F(U)$ heißt **lokale Parametrisierung** von M oder **Karte** der Umgebung $F(U) \subseteq M$. **Zweidimensionale** Untermannigfaltigkeiten in einem beliebigen \mathbb{R}^n heißen **Flächen** und $(n-1)$ -dimensionale Untermannigfaltigkeiten von \mathbb{R}^n heißen **Hyperflächen**.

Beispiel

Die Menge $M := \{(x, y) \in \mathbb{R}^2 \mid x^2(1-x^2) - y^2 = 0\} \subseteq \mathbb{R}^2$

ist keine Untermannigfaltigkeit.

$M \setminus \{(0, 0)\}$ ist eine Untermannigfaltigkeit.



734 / 862

Mannigfaltigkeiten

Bemerkungen

- Sei $M \subseteq \mathbb{R}^n$ m -dimensionale C^k -Untermannigfaltigkeit und seien $F_i: U_i \xrightarrow{\sim} V_i$, $i = 1, 2$ zwei Karten mit $V := V_1 \cap V_2 \neq \emptyset$. Man zeigt: Die Urbildmengen $W_i := F_i^{-1}(V)$ sind offen in \mathbb{R}^m und $F_2^{-1} \circ F_1: W_1 \rightarrow W_2$ ist ein C^k -Diffeomorphismus.
- Man kann den Begriff der Untermannigfaltigkeit verallgemeinern: Gegeben seien ein metrischer Raum M , eine offene Überdeckung $(V_i)_{i \in I}$ von M mit offenen Mengen $U_i \subseteq \mathbb{R}^m$ und Homöomorphismen $F_i: U_i \xrightarrow{\sim} V_i$. Man spricht dann von einer **(abstrakten) m -dimensionalen C^k -Mannigfaltigkeit** wenn für je zwei offene Mengen $V_1, V_2 \subseteq M$ mit Abbildungen F_1 und F_2 die Abbildung

$$F_2^{-1} \circ F_1: F_1^{-1}(V_1 \cap V_2) \rightarrow F_2^{-1}(V_1 \cap V_2)$$

ein C^k -Diffeomorphismus ist.

735 / 862

Beispiel: Graph einer Abbildung als Untermannigfaltigkeit

Sei $U \subseteq \mathbb{R}^m$ offen und $f: U \rightarrow \mathbb{R}^r$ eine C^k -Abbildung.

Dann ist der **Graph** von f

$$\Gamma_f = \{(x, y) \in U \times \mathbb{R}^r \mid y = f(x)\}$$

eine m -dimensionale C^k -Untermannigfaltigkeit von \mathbb{R}^n , wobei $n = m + r$.

Denn die C^k -Immersion $F: U \rightarrow \mathbb{R}^n$, $F(x) = (x, f(x))$, bildet U homöomorph auf $F(U) = \Gamma_f$ ab.

Es gilt auch eine Umkehrung, jede Untermannigfaltigkeit des \mathbb{R}^n lässt sich **lokal** als Graph schreiben:

736 / 862

Satz

Eine Teilmenge $M \subseteq \mathbb{R}^n$ ist genau dann eine m -dimensionale C^k -**Untermannigfaltigkeit**, wenn es zu jedem Punkt $a \in M$, eventuell nach Umbenennung der Koordinaten des \mathbb{R}^n , offene Umgebungen $U' \subseteq \mathbb{R}^m$ von $a' := (a_1, \dots, a_m)$ und $U'' \subseteq \mathbb{R}^{n-m}$ von $a'' := (a_{m+1}, \dots, a_n)$ und eine C^k -Abbildung $g: U' \rightarrow U''$ gibt, so dass

$$M \cap (U' \times U'') = \{(x', x'') \in U' \times U'' \mid x'' = g(x')\}$$

gilt. **Lokal** lässt sich also M als **Graph** schreiben.

Beweis. Wir müssen nur noch zeigen, dass eine m -dimensionale Untermannigfaltigkeit M lokal der Graph einer Funktion $\mathbb{R}^m \rightarrow \mathbb{R}^{n-m}$ ist. Sei dazu $a \in M$, $U \subseteq \mathbb{R}^n$ eine Umgebung von a , so dass eine offene Menge $W \subset \mathbb{R}^m$ und eine lokale Parametrisierung $f: W \rightarrow \mathbb{R}^n$ von $M \cap U$ existiert. Da f konstanten Rang m hat, können wir wie im ersten Schritt des Beweises der Normalform für Abbildungen von konstanten Rang einen Diffeomorphismus $\varphi: W \rightarrow \varphi(W) \subseteq \mathbb{R}^m$ finden, so dass das Bild von $f \circ \varphi^{-1}$ Graph einer Funktion g ist. □

737 / 862

Beispiel

Die Sphäre S^2 ist eine C^∞ -Fläche in \mathbb{R}^3 .

Sie besitzt nämlich eine Überdeckung durch 6 Halbsphären

$$H_i^\pm := \{x = (x_1, x_2, x_3) \in S^2 \mid \pm x_i > 0\} \subseteq \mathbb{R}^3$$

und jede der Halbsphären ist ein Graph. Zum Beispiel ist

$$H_1^+ = \{(f(u), u) \mid u \in U\} \text{ mit } f: U = \{u \in \mathbb{R}^2 \mid \|u\| < 1\} \rightarrow \mathbb{R}, \\ f(u) = \sqrt{1 - \|u\|^2}.$$

Beispiel: Die stereographische Projektion der Sphäre

Wir betrachten die beiden Abbildungen $\varphi_{\pm}: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ definiert durch

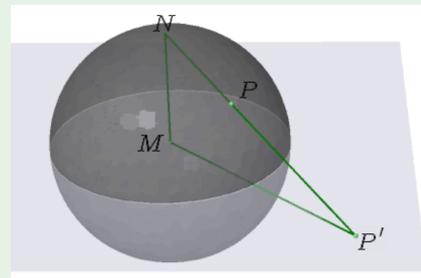
$$\varphi_{\pm}(x, y) := \frac{1}{1 + \|(x, y)\|^2} (2x, 2y, \pm(\|(x, y)\|^2 - 1)).$$

Beides sind Immersionen mit $\text{im}(\varphi_{\pm}) \subseteq S^2$ (ÜA).

Seien $N^{\pm} = (0, 0, \pm 1)$ der Nord- und Südpol der Sphäre. Dann sind $\varphi_{\pm}: \mathbb{R}^2 \rightarrow S^2 \setminus \{N^{\pm}\}$ Homöomorphismen mit der Umkehrabbildung (ÜA)

$$\varphi_{\pm}^{-1}: S^2 \setminus \{N^{\pm}\} \rightarrow \mathbb{R}^2, \quad \varphi_{\pm}^{-1}(x, y, z) := \left(\frac{x}{1 \mp z}, \frac{y}{1 \mp z} \right)$$

φ_{\pm}^{-1} heißt **stereographische Projektion aus dem Nord/Südpol**. Sie ordnet jedem Punkt $P \in S^2 \setminus \{N^{\pm}\}$ den Schnittpunkt P' der Gerade durch P und N^{\pm} mit dem $\mathbb{R}^2 = \{(x, y, 0) \in \mathbb{R}^3\}$ zu.



739 / 862

Untermannigfaltigkeiten als Urbilder unter Abbildungen von konstantem Rang

Satz

Sei $U \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}^m$ eine C^k -Abbildung von konstantem Rang r und $q \in f(U)$. Dann ist

$$M := f^{-1}(q) \subseteq U$$

eine C^k -Untermannigfaltigkeit des \mathbb{R}^n der Dimension $n - r$.

Beweis. Sei $p = (p_1, \dots, p_n) \in M$. O.B.d.A. können wir annehmen, dass

$$f: (x_1, \dots, x_n) \mapsto (x_1, \dots, x_r, 0, \dots, 0),$$

definiert auf einer in U enthaltenen offenen Umgebung $V_1 \times V_2 \subseteq \mathbb{R}^r \times \mathbb{R}^{n-r}$ von p , wobei $q = (p_1, \dots, p_r, 0, \dots, 0) \in f(V_1 \times V_2)$.

Dann ist $V_2 \ni (y_{r+1}, \dots, y_n) \mapsto (p_1, \dots, p_r, y_{r+1}, \dots, y_n) \in V_1 \times V_2$ eine Immersion, die V_2 homöomorph auf $f^{-1}(q)$ abbildet. \square

740 / 862

Beispiele

- (i) Die Abbildung $f: GL(n) \rightarrow \text{Sym}(n, \mathbb{R}), A \mapsto A^T A$, hat überall den Rang $n(n+1)/2$. Somit ist $O(n) = f^{-1}(\mathbf{1}_n) \subseteq \text{Mat}(n, \mathbb{R})$ eine (kompakte) C^∞ -Untermannigfaltigkeit der Dimension $n(n-1)/2$.
- (ii) Man zeigt auf ähnliche Weise, dass $U(n) \subseteq \text{Mat}(n, \mathbb{C})$ eine (kompakte) C^∞ -Untermannigfaltigkeit der (reellen) Dimension n^2 ist (ÜA).
- (iii) Sei $p \in \mathbb{R}^{n+1}$. Die Abbildung $f: \mathbb{R}^{n+1} \rightarrow \mathbb{R}, f(x) = \|x - p\|^2$, definiert eine Submersion von $\mathbb{R}^{n+1} \setminus \{p\}$ auf \mathbb{R}_+ . Die n -dimensionalen Sphären $S_r^n(p) = f^{-1}(r^2) \subseteq \mathbb{R}^{n+1}$ ($r > 0$) sind also C^∞ -Hyperflächen.
- (iv) Sei $U \subseteq \mathbb{R}^m$ offen und $F: U \rightarrow \mathbb{R}^r$ eine C^k -Abbildung und $\Gamma_F = \{(x, F(x)) \in U \times \mathbb{R}^r\}$ der Graph von F .
Dann gilt: die Abbildung $f: \mathbb{R}^{m+r} \rightarrow \mathbb{R}^r$ definiert durch $f(x, y) := y - F(x)$ ist eine Submersion und

$$\Gamma_F = f^{-1}(0).$$

741 / 862

Untermannigfaltigkeiten und Submersionen

Satz

Eine Teilmenge $M \subseteq \mathbb{R}^n$ ist genau dann eine m -dimensionale C^k -Untermannigfaltigkeit, wenn es zu jedem $p \in M$ eine Umgebung $V \subseteq \mathbb{R}^n$ gibt und eine C^k -Submersion $f: V \rightarrow \mathbb{R}^{n-m}$, so dass $M \cap V = f^{-1}(0)$.

Eine Untermannigfaltigkeit M ist also **lokal** durch $n - m$ Gleichungen gegeben.

Beweis. " \implies " Gegeben sei eine Untermannigfaltigkeit. Sei $F: U \rightarrow V \subseteq \mathbb{R}^n$ eine lokale Parametrisierung, wobei $U \subseteq \mathbb{R}^m$ und $p \in F(U) = M \cap V$.

Wegen der Normalform von Immersionen können wir annehmen, dass $F: (x_1, \dots, x_m) \mapsto (x_1, \dots, x_m, 0, \dots, 0)$.

Dann ist $f: V \rightarrow \mathbb{R}^{n-m}, f(x_1, \dots, x_n) = (x_{m+1}, \dots, x_n)$ die gesuchte Submersion.

Die Umkehrung folgt aus dem vorherigen Satz. □

742 / 862

Tangentialraum

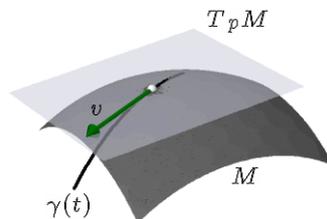
Definition (Tangentialvektoren)

Sei $M \subseteq \mathbb{R}^n$ eine C^1 -Untermannigfaltigkeit und $p \in M$.

Ein Vektor $v \in \mathbb{R}^n$ heißt **Tangentialvektor** an M in $p \in M$, wenn es ein $\varepsilon > 0$ und eine C^1 -Kurve $\gamma: (-\varepsilon, \varepsilon) \rightarrow M \subseteq \mathbb{R}^n$ gibt mit

$$\gamma(0) = p \quad \text{und} \quad v = \gamma'(0).$$

Die Menge aller Tangentialvektoren an M in p heißt **Tangentialraum** in p und wird mit $T_p M$ bezeichnet.



743 / 862

Satz (Eigenschaften des Tangentialraumes)

Sei $M \subseteq \mathbb{R}^n$ eine m -dimensionale C^1 -Untermannigfaltigkeit und $T_p M$ der Tangentialraum an M in $p \in M$. Dann gilt:

- (i) $T_p M$ ist ein **Vektorraum** der Dimension $m = \dim M$.
- (ii) Sei $F: U \subseteq \mathbb{R}^m \rightarrow V$ eine lokale Parametrisierung von M , $u \in U$ mit $p = F(u)$.
Dann bilden die m Vektoren $\partial_1 F(u), \dots, \partial_m F(u) \in \mathbb{R}^n$ eine Basis von $T_p M$.
- (iii) Sei $V \subseteq \mathbb{R}^n$ eine offene Umgebung von p und $f = (f_1, \dots, f_{n-m}): V \rightarrow \mathbb{R}^{n-m}$ eine C^1 -Submersion, so dass $M \cap V = f^{-1}(q)$, wobei $q = f(p)$. Dann ist
 $T_p M = \ker(df_p) = \bigcap_{j=1}^{n-m} (\text{grad } f_j(p))^\perp$.
- (iv) Insbesondere gilt $T_p M^\perp = \text{span}\{\text{grad } f_1(p), \dots, \text{grad } f_{n-m}(p)\} \subseteq \mathbb{R}^n$.

744 / 862

Beweis. Es genügt zu zeigen, dass

$$\text{span}\{\partial_1 F(u), \dots, \partial_m F(u)\} \subseteq T_p M \quad (*)$$

$$\text{und } \text{grad } f_j(p) \perp T_p M \text{ für alle } j = 1, \dots, n - m. \quad (**)$$

Wegen $\text{Rang}(F) = m$ und (*) folgt dann $\dim(T_p M) \geq m$. Wegen $\text{Rang}(f) = n - m$ und (**) folgt ebenso $\dim(T_p M) \leq n - (n - m) = m$.

Sei nun $v = \sum_{i=1}^m v_i e_i \in \mathbb{R}^m$ beliebig. Dann definiert $c(t) := F(u + tv)$ eine C^1 -Kurve $(-\varepsilon, \varepsilon) \rightarrow F(U) \subseteq M$ mit

$$c'(0) = dF_p v = \sum_{i=1}^m v_i \partial_i F(u) \in T_p M.$$

Das beweist (*).

Für jede C^1 -Kurve $c: (-\varepsilon, \varepsilon) \rightarrow F(U) \subseteq M$ mit $c(0) = p$ gilt $f(c(t)) \equiv q$ für alle $t \in (-\varepsilon, \varepsilon)$ und somit $0 = \left. \frac{d}{dt} \right|_{t=0} f(c(t)) = df_p c'(0)$.

Das zeigt $T_p M \subseteq \ker(df_p) = \bigcap_{j=1}^{n-m} (\text{grad } f_j(p))^\perp$ und damit (**). \square

745 / 862

Beispiele:

- Tangentialraum an die Sphäre:

Es ist $S^n = S^n(1) = f^{-1}(1)$, wobei $f: \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ durch $f(x) = \|x\|^2$ gegeben ist. Außerhalb von $0 \in \mathbb{R}^{n+1}$ ist f eine Submersion. Damit ist für $p \in S^2$ der Tangentialraum gerade $T_p M = (\text{grad } f_p)^\perp = 2p^\perp = p^\perp$.

- Tangentialraum an ein Hyperboloid:

Wir betrachten $H^n = H^n(1) = g^{-1}(1)$, wobei $g: \mathbb{R}^{n+1} \cong \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ durch $g(x, z) = \|x\|^2 - z^2$ gegeben ist. Wieder ist g außerhalb von $0 \in \mathbb{R}^{n+1}$ eine Submersion. Für $p = (x_1, \dots, x_n, z) \in H_1^n$ ist $T_p H_1^n = (\text{grad } f_p)^\perp = (2x, -2z)^\perp$.

Beispiele:

- Tangentialraum an einen Graphen:

Sei $\varphi: U \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ eine C^k -Abb. und $\Gamma_\varphi = \{(x, \varphi(x)) \in U \times \mathbb{R}\}$ der Graph von φ .

Dann ist $F: U \rightarrow \mathbb{R}^{m+1}$, $F(x) = (x, \varphi(x))$ eine Parametrisierung und

$$T_p \Gamma_\varphi = \text{span} \left((e_i, \partial_i \varphi(p))^T \right)_{i=1}^m \subseteq \mathbb{R}^{m+1}$$

Ist $f: \mathbb{R}^{m+1} \rightarrow \mathbb{R}$, $f(x, y) = \varphi(x) - y$ die durch den Graphen definierte Submersion, dann gilt auch

$$T_p \Gamma_\varphi = (\text{grad } f(p, \varphi(p)))^\perp = (\text{grad } \varphi(p), -1)^\perp.$$

Beachte, dass $\langle (\text{grad } \varphi(p), -1), (e_i, \partial_i \varphi(p)) \rangle = 0$.

Analoge Aussagen gelten natürlich auch für $\varphi: U \rightarrow \mathbb{R}^r$ mit $r > 1$.

747 / 862

Extrema mit Nebenbedingungen

Satz (Extrema mit Nebenbedingungen)

Sei $M \subseteq \mathbb{R}^n$ eine m -dimensionale Untermannigfaltigkeit, $U \subseteq \mathbb{R}^n$ offen und $f: U \rightarrow \mathbb{R}$ im Punkt $p \in M \cap U$ differenzierbar

- (i) Wenn die Einschränkung $F := f|_{U \cap M}$ im Punkt p ein lokales Extremum annimmt, so ist

$$(*) \quad T_p M \subseteq \ker df_p = (\text{grad } f(p))^\perp.$$

- (ii) Sei $r := n - m$ und $h = (h_1, \dots, h_r): U \rightarrow \mathbb{R}^r$ eine C^1 -Submersion, so dass $M \cap U = h^{-1}(0)$. Dann gilt (*) genau dann, wenn es Konstanten $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ (sogenannte **Lagrangemultiplikatoren**) gibt, so dass

$$\text{grad } f(p) = \sum_{j=1}^r \lambda_j \text{grad } h_j(p).$$

748 / 862

Beweis. Sei $c: (-\varepsilon, \varepsilon) \rightarrow M \cap U \subseteq \mathbb{R}^n$ eine C^1 -Kurve in der Untermannigfaltigkeit mit $c(0) = p$. Dann hat die Funktion $f \circ c: (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ ein lokales Extremum in 0 und somit

$$0 = \left. \frac{d}{dt} \right|_{t=0} f(c(t)) = df_p c'(0) = \langle \text{grad } f(p), c'(0) \rangle.$$

Das beweist (i), d.h. $T_p M \subseteq (\text{grad } f)^\perp$. Das impliziert aber $\text{grad } f \in (T_p M)^\perp$. Damit folgt (ii) aus

$$T_p M^\perp = \text{span}\{\text{grad } h_1(p), \dots, \text{grad } h_r(p)\}.$$

□

Bemerkung

Die Bedingung (*) ist notwendig aber nicht hinreichend. Ob tatsächlich ein Extremum vorliegt (und von welcher Art es ist) muss separat untersucht werden.

749 / 862

Beispiele

- Wir betrachten die Einschränkung der Funktion $f(x, y, z) = x$ auf die Sphäre $S^2 = \{p \in \mathbb{R}^3 : \|p\|^2 = 1\}$. Wir wissen, dass $T_p S^2 = p^\perp$, so dass die Bedingung (*) hier die Form

$$\text{grad } f(p) = \lambda \cdot p$$

annimmt. In $p = (x, y, z) \in \mathbb{R}^3$ gilt aber $\text{grad } f(p) = (1, 0, 0)$, und somit erhalten wir also Lösungen von (*) die beiden Punkte $p_\pm = (\pm 1, 0, 0)$. Offenbar ist p_- das absolute Minimum und p_+ das absolute Maximum von $f|_{S^2}$.

- Wir betrachten nun die Einschränkung derselben Funktion auf das Hyperboloid $H := g^{-1}(1)$ mit $g(x, y, z) := x^2 + y^2 - z^2$. Hier haben wir $T_p H = \text{grad } g(p)^\perp = (2x, 2y, -2z)^\perp$, und wir erhalten wiederum die beiden Punkte $p_\pm = (\pm 1, 0, 0)$ als Lösungen von (*). Diesmal sind diese Punkte aber keine lokalen Extrema von $f|_H$.

750 / 862

Beispiel

Mit dem Satz können wir auch (erneut) zeigen, dass jeder symmetrische Endomorphismus A eines **endlichdimensionalen** Euklidischen Vektorraums V (mindestens) einen Eigenvektor hat:

Da die Einheitskugel $S = S_1^n(0) \subseteq V$ **kompakt** ist, nimmt die quadratische Form

$$\begin{aligned} S &\rightarrow \mathbb{R} \\ x &\mapsto f(x) := \frac{1}{2} \langle x, Ax \rangle \end{aligned}$$

als stetige Funktion in einem Punkt $p \in S$ ihr Minimum an.

Nach dem vorherigen Satz gilt also

$$\text{grad } f(p) = Ap \in (T_p S)^\perp = \mathbb{R}p,$$

d.h. p ist ein Eigenvektor von A . Der Lagrangemultiplikator ist der Eigenwert.

751 / 862

Kapitel 22

Gewöhnliche Differentialgleichungen

752 / 862

Gewöhnliche Differentialgleichungen

Eine **gewöhnliche** Differentialgleichung (DG) besteht aus einer oder mehreren Gleichungen für eine oder mehrere Funktionen **einer Variablen** und deren Ableitungen.

Wir kennen schon einige Differentialgleichungen:

- Die differenzierbare Funktion $x: \mathbb{R} \rightarrow \mathbb{R}$, $t \mapsto x(t)$ erfülle die DG $x'(t) \equiv 0$. Jede Lösung ist dann von der Form $x(t) \equiv c$, wobei $c \in \mathbb{R}$ eine Konstante ist. Wir müssen noch eine **Anfangsbedingung** (AB) stellen: $x(t_0) = c$, um die Lösung festzulegen.
- Die DG **zweiter Ordnung** $x''(t) \equiv 0$ hat die Lösungen $x(t) = at + b$ wobei $a, b \in \mathbb{R}$ Konstanten sind. Diesmal müssen wir zwei Anfangsbedingungen stellen: $x(t_0) = b_0$ und $x'(t_0) = a_0$. Dann ist $x(t) = a_0(t - t_0) + b_0$ eine Lösung.
- Die DG $x'(t) = t^2$ mit der Anfangsbedingung $x(0) = c$ hat die Lösung $x(t) = \frac{1}{3}t^3 + c$.
- Die DG $x'(t) = x(t)$ mit der AB $x(0) = c$ hat die Lösung $x(t) = ce^t$.

753 / 862

Warum interessieren wir uns für Differentialgleichungen:

Die Bewegung eines Punktes im \mathbb{R}^3 wird beschrieben durch eine Kurve im \mathbb{R}^3

$$\begin{aligned} x: \mathbb{R} &\longrightarrow \mathbb{R}^3 \\ t &\mapsto x(t) = (x_1(t), x_2(t), x_3(t)) \end{aligned}$$

$x'(t)$ gibt dann die Geschwindigkeit und $x''(t)$ die Beschleunigung zum Zeitpunkt t an. Auf den Punkt wirke eine Kraft F , die vom Ort x , der Zeit t und der Geschwindigkeit $x'(t)$ des Punktes abhängt, d.h. $F = F(x, x', t)$. Das Newtonsche Bewegungsgesetz der Mechanik hat dann folgende Form

$$m \cdot x''(t) = F(x(t), x'(t), t).$$

Unter der Annahme, dass F und die Anfangswerte $x(t_0)$ und $x'(t_0)$ bekannt sind, versucht man, die Bewegungskurve des Punktes zu berechnen.

Dies ist ein Anfangswertproblem der Form

$$x''(t) = \frac{1}{m} F(x, x', t), \quad x(t_0) = x_0, \quad x'(t_0) = v_0.$$

754 / 862

Gewöhnliche Differentialgleichungen erster Ordnung

Definition (Gewöhnliche Differentialgleichung erster Ordnung)

Sei $\Omega \subseteq \mathbb{R}^2$ und $f: \Omega \rightarrow \mathbb{R}$ eine stetige Funktion.

- Dann heißt die Gleichung

$$x'(t) = f(x, t) \quad (1)$$

Differentialgleichung erster Ordnung in x . Meist schreiben wir auch nur $x' = f(x, t)$.

- Unter einer **Lösung** der DG (1) versteht man eine auf einem Intervall $I \subseteq \mathbb{R}$ differenzierbare Funktion $\varphi: I \rightarrow \mathbb{R}$, so dass
 - $(\varphi(t), t) \in \Omega$ für alle $t \in I$,
 - $\varphi'(t) = f(\varphi(t), t)$ für alle $t \in I$.
- Sei $t_0 \in I$ und $c \in \mathbb{R}$. Eine Lösung φ der DG (1) mit $\varphi(t_0) = c$ heißt Lösung des Anfangswertproblems (AWP) zu (1) mit der **Anfangsbedingung**

$$x(t_0) = c. \quad (2)$$

755 / 862

Gewöhnliche Differentialgleichungen höherer Ordnung

Definition (Gewöhnliche Differentialgleichung höherer Ordnung)

Sei $\Omega \subseteq \mathbb{R}^k \times \mathbb{R}$ und $f: \Omega \rightarrow \mathbb{R}$ stetig.

$$x^{(k)} = f(x, x', \dots, x^{(k-1)}, t) \quad (3)$$

heißt **gewöhnliche Differentialgleichung k -ter Ordnung**.

- Unter einer **Lösung** versteht man eine auf einem Intervall $I \subseteq \mathbb{R}$ definierte k -mal differenzierbare Funktion $\varphi: I \rightarrow \mathbb{R}$, so dass
 - $(\Phi(t), t) \in \Omega$ für alle t , wobei $\Phi := (\varphi, \varphi', \dots, \varphi^{(k-1)}): I \rightarrow \mathbb{R}^k$ und
 - $\varphi^{(k)}(t) = f(\varphi(t), \varphi'(t), \dots, \varphi^{(k-1)}(t), t)$ für alle $t \in I$.
- Sei $t_0 \in I$ und $(c_1, \dots, c_k) \in \mathbb{R}^k$. Das **Anfangswertproblem** ist gegeben durch (3) und

$$\left. \begin{aligned} x(t_0) &= c_1 \\ x'(t_0) &= c_2 \\ &\vdots \\ x^{(k-1)}(t_0) &= c_k \end{aligned} \right\} \quad (4)$$

756 / 862

Systeme gewöhnlicher Differentialgleichungen

Definition (Systeme gewöhnlicher Differentialgleichungen)

Sei $\Omega \subseteq (\mathbb{R}^n)^k \times \mathbb{R}$ und $F: \Omega \rightarrow \mathbb{R}^n$ stetig.

$$x^{(k)} = F(x, x', \dots, x^{(k-1)}, t) \quad (5)$$

heißt **System gewöhnlicher DG'en k-ter Ordnung** an $x = (x_1, \dots, x_n)$.

- Unter einer **Lösung** versteht man eine differenzierbare Kurve $\varphi: I \rightarrow \mathbb{R}^n$ mit
 - (i) $(\varphi(t), t) \in \Omega$ für alle t , wobei $\Phi = (\varphi, \varphi', \dots, \varphi^{(k-1)}): I \rightarrow \mathbb{R}^{nk}$ und
 - (ii) $\varphi^{(k)}(t) = F(\varphi(t), \varphi'(t), \dots, \varphi^{(k-1)}(t), t)$ für alle $t \in I$.
- Sei $t_0 \in I$ und c_1, \dots, c_k Vektoren in \mathbb{R}^n . Das **Anfangswertproblem** ist gegeben durch (5) und

$$\left. \begin{array}{l} x(t_0) = c_1 \\ x'(t_0) = c_2 \\ \vdots \\ x^{(k-1)}(t_0) = c_k \end{array} \right\} \quad (6)$$

757 / 862

Reduktion von Systemen höherer Ordnung

Satz

Sei $F: \Omega \subseteq (\mathbb{R}^n)^k \times \mathbb{R} \rightarrow \mathbb{R}^n$ stetig und $F^*: \Omega \rightarrow (\mathbb{R}^n)^k$ definiert durch $F^*(y_0, \dots, y_{k-1}, t) := (y_1, \dots, y_{k-1}, F(y_0, \dots, y_{k-1}, t))$, wobei $y_j \in \mathbb{R}^n$ für alle $j \in \{0, \dots, k-1\}$. Dann gilt:

- (1) Ist $x: I \rightarrow \mathbb{R}^n$ eine Lösung des AWP's **k-ter Ordnung**

$$x^{(k)} = F(x, \dots, x^{(k-1)}, t) \text{ mit AB'en } \left\{ \begin{array}{l} x(t_0) = a_1 \\ \vdots \\ x^{(k-1)}(t_0) = a_k \end{array} \right\} \quad (7)$$

so ist $y := (x, x', \dots, x^{(k-1)}): I \rightarrow \mathbb{R}^{n \cdot k}$ eine Lösung des AWP's

$$y' = F^*(y, t), \quad y(t_0) = (a_1, \dots, a_k) \quad (8)$$

- (2) Ist $y = (y_0, \dots, y_{k-1}): I \rightarrow (\mathbb{R}^n)^k$ eine Lsg. des AWP's (8) **erster Ordnung**, so ist $y_0: I \rightarrow \mathbb{R}^n$ eine Lsg. des AWP's (7) **k-ter Ordnung**.

758 / 862

Beweis: Nach Definition von F^* ist $y'(t) = F^*(y(t), t)$ äquivalent zu

$$\begin{aligned} y_0'(t) &= y_1(t), \\ y_1'(t) &= y_2(t), \\ &\dots \\ y_{k-2}'(t) &= y_{k-1}(t), \\ y_{k-1}'(t) &= F(y_0, \dots, y_{k-1}, t). \end{aligned}$$

Die Behauptung folgt dann sofort durch Einsetzen:

$$\begin{aligned} x(t) \text{ löst (7)} &\implies y(t) = (x(t), \dots, x^{(k-1)}(t)) \text{ löst (8)} \\ y(t) \text{ löst (8)} &\implies x(t) = y_0(t) \text{ löst (7)}. \end{aligned} \quad \square$$

759 / 862

Beispiel: Schwingungsgleichung ohne Reibung

Die Bewegung einer Masse m , die reibungsfrei an einer Feder auf der x -Achse um 0 gleitet, wird beschrieben durch die DG **zweiter Ordnung**

$$x''(t) = -\frac{k}{m}x = F(x, x', t) \quad \text{mit } F: \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}, \quad F(x_0, x_1, t) = -\frac{k}{m}x_0. \quad (9)$$

Wir führen nun (9) zurück auf das System **erster Ordnung**

$$(y_0'(t), y_1'(t)) = (y_1(t), -\frac{k}{m}y_0(t)) = F^*(y_0, y_1, t),$$

wobei $F^*: \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}^2$, $F^*(y_0, y_1, t) = (y_1, F(y_0, y_1, t)) = (y_1, -\frac{k}{m}y_0)$. Das heißt, $x(t)$ löst genau dann (9), wenn $(y_0(t), y_1(t)) := (x(t), x'(t))$ folgendes Differentialgleichungssystem erster Ordnung löst

$$\begin{pmatrix} y_0'(t) \\ y_1'(t) \end{pmatrix} = \begin{pmatrix} y_1(t) \\ -\frac{k}{m}y_0(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{k}{m} & 0 \end{pmatrix} \begin{pmatrix} y_0(t) \\ y_1(t) \end{pmatrix},$$

d.h. $y'(t) = A \cdot y(t)$, mit einer konstanten Matrix A : dies ist ein lineares Differentialgleichungs-System mit konstanten Koeffizienten.

760 / 862

Autonome Differentialgleichungen und Vektorfelder

Definition

- Ein DG-System der Form $x^{(k)} = F(x, \dots, x^{(k-1)})$, bzw. $x' = F(x)$ falls $k = 1$, heißt **autonom**. Entscheidend ist, dass F nicht von t selbst abhängt.
- Ein **stetiges Vektorfeld** (ab jetzt, kurz: Vektorfeld) auf einer offenen Menge $U \subseteq \mathbb{R}^n$ ist eine stetige Abbildung $V: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$.
- Eine C^1 -Kurve $\gamma_{x_0}: (-\varepsilon, \varepsilon) \subseteq \mathbb{R} \rightarrow U$ heißt **Integralkurve** des Vektorfeldes V durch $x_0 \in U$, falls gilt

$$\gamma'_{x_0}(t) = V(\gamma_{x_0}(t)) \quad \text{für alle } t \in (-\varepsilon, \varepsilon) \quad \text{und} \quad \gamma_{x_0}(0) = x_0.$$

Der Vektor $V(\gamma_{x_0}(t))$ des Vektorfelds ist gleich dem Tangentialvektor der Kurve γ_{x_0} in t . Die Integralkurven eines Vektorfeldes $V: U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ sind also Lösungen der autonomen Differentialgleichung

$$\gamma'(t) = V(\gamma(t)) \quad \text{mit der Anfangsbedingung } \gamma(0) = x_0.$$

761 / 862

Beispiel: Lineare Vektorfelder auf dem \mathbb{R}^2

- Sei $\lambda \in \mathbb{R}$ und $U: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ein Vektorfeld definiert durch $U(x, y) = \lambda \cdot (x, y)$. Eine Integralkurve von U durch $p := (x_0, y_0) \in \mathbb{R}^2$ ist dann gegeben durch

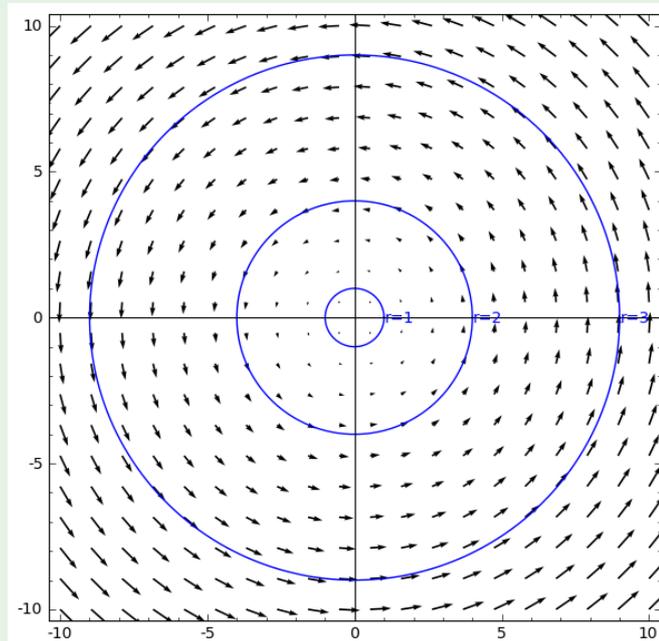
$$\gamma_p = e^{\lambda t} p,$$

denn $\gamma'_p(t) = \lambda e^{\lambda t} p = \lambda \gamma_p(t) = U(\gamma_p(t))$.

762 / 862

(Fortsetzung)

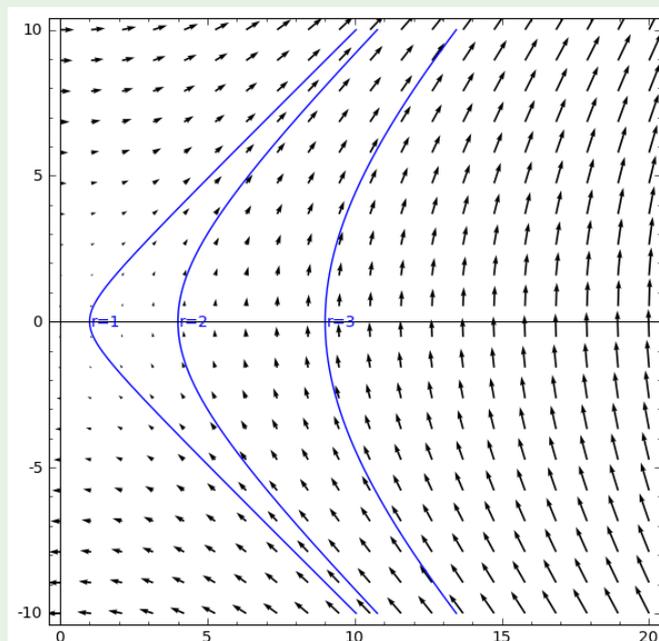
- Sei V das Vektorfeld $V(x, y) = (-y, x)$. Eine Integralkurve durch $p = (r^2, 0)$ ist dann gegeben durch $\gamma_p(t) = r^2(\cos t, \sin t)$, denn $\gamma'_p(t) = r^2(-\sin t, \cos t) = V(\gamma_p(t))$.



763 / 862

(Fortsetzung)

- Sei W das Vektorfeld $W(x, y) = (y, x)$. Eine Integralkurve durch $p = (r^2, 0)$ ist dann gegeben durch $\gamma_p(t) = r^2(\cosh t, \sinh t)$, denn $\gamma'_p(t) = r^2(\sinh t, \cosh t) = W(\gamma_p(t))$.



764 / 862

Elementare Lösungsmethoden für DG'en 1. Ordnung

In diesem Abschnitt betrachten wir Differentialgleichungen der Form

$$x' = F(x, t)$$

wobei $F: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ stetig ist. Eine Lösung ist also eine differenzierbare Funktion $x: I \subseteq \mathbb{R} \rightarrow \mathbb{R}$.

Wir geben elementare Lösungsmethoden an, für Fälle, in denen F eine einfache Gestalt hat. Diese Verfahren basieren meist auf der Möglichkeit der **Trennung der Variablen**.

Grundidee:

F sei von der Gestalt $F(x, t) = f(t) \cdot g(x)$, die Gleichung also $x' = f(t) \cdot g(x)$.

Formal gilt dann $\frac{dx}{dt} = f(t)g(x)$ und deshalb $\frac{dx}{g(x)} = f(t)dt$; Integration liefert die Lösung der Differentialgleichung.

765 / 862

Definition (Trennung der Variablen)

Eine **DG mit getrennten Variablen** ist eine DG des Typs

$$x'(t) = f(t) \cdot g(x(t)) \quad \text{mit Anfangsbedingungen } x(t_0) = x_0, \quad (10)$$

wobei die Funktionen $f: I_1 \rightarrow \mathbb{R}$ und $g: I_2 \rightarrow \mathbb{R} \setminus \{0\}$ **stetig** sind, $(t_0, x_0) \in I_1 \times I_2$ und $I_1, I_2 \subseteq \mathbb{R}$ offene Intervalle.

Satz (Differentialgleichung mit getrennten Variablen)

Sei G eine Stammfkt. von $\frac{1}{g}$ auf I_2 und G^{-1} die Umkehrfkt. von G . Das AWP (10) besitzt auf einem Intervall $J \subseteq I_1$ um t_0 die eindeutige Lösung

$$x(t) = G^{-1} \left(G(x_0) + \int_{t_0}^t f(s) ds \right),$$

Das Intervall J ist gegeben durch das Innere der Menge

$$\left\{ t \in \mathbb{R} \mid G(x_0) + \int_{t_0}^t f(s) ds \in \text{im}(G) \right\}.$$

766 / 862

Beweis.

Existenz: Da die Funktion $\frac{1}{g}$ stetig und ohne Nullstellen ist, ist ihre Stammfunktion G streng monoton und somit umkehrbar mit $G^{-1}: \text{im}(G) \rightarrow I_2$. Setze

$$x(t) = G^{-1} \left(G(x_0) + \int_{t_0}^t f(s) ds \right);$$

nach der Kettenregel ist die Ableitung

$$x'(t) = \frac{1}{G'(x(t))} (G(x_0) + \int_{t_0}^t f(s) ds)' = g(x(t)) \cdot f(t),$$

also erfüllt $x(t)$ die DG (10). Es gilt $x(t_0) = x_0$. Die Funktion $x(t)$ ist definiert für diejenigen t , für die $G(x_0) + \int_{t_0}^t f(s) ds \in \text{im}(G)$.

767 / 862

Weiter im Beweis.

Eindeutigkeit: Sei x eine Lösung von (10).

Wir integrieren die Gleichung $\frac{x'(t)}{g(x(t))} = f(t)$ von t_0 bis T nahe t_0 :

$$\int_{t_0}^T \frac{x'(t)}{g(x(t))} dt = \int_{t_0}^T f(t) dt.$$

Die Substitution $x = x(t)$ ergibt $dx = x'(t) dt$ und

$$\int_{t_0}^T f(t) dt = \int_{x_0}^{x(T)} \frac{dx}{g(x)} = G(x(T)) - G(x_0)$$

für G die Stammfunktion von $\frac{1}{g}$. Da G monoton ist, ist die Lösung x eindeutig bestimmt. □

768 / 862

Beispiele

- Jede **autonome DG** $x' = g(x)$, $x(t_0) = x_0$ ist eine DG mit getrennten Variablen, wobei $f \equiv 1$. Ist G eine Stammfunktion von $1/g$, so ist $x(t) := G^{-1}(t + G(x_0) - t_0)$ eine Lösung.
Sei z.B. $x' = x$ mit $x(0) = c > 0$. Dann ist $g(x) = x$ und eine Stammfunktion von $1/g = 1/x$ ist $G(x) = \ln x$ mit Umkehrfunktion $G^{-1}(x) = e^x$. Somit ist die Lösung gegeben durch

$$x(t) = G^{-1}(t + \ln c) = e^{t + \ln c} = c \cdot e^t.$$

- Sei $x' = t \cdot x^2$ mit Anfangsbedingungen $x(0) = c \neq 0$. Dann ist $g(x) = x^2$ und eine Stammfunktion von $1/g = 1/x^2$ ist $G(x) = -\frac{1}{x}$ mit Umkehrfunktion $G^{-1}(x) = -\frac{1}{x}$. Andererseits ist $\int_0^t s ds = \frac{1}{2}t^2$. Somit ist die Lösung gegeben durch

$$x(t) = G^{-1}\left(-\frac{1}{c} + \frac{t^2}{2}\right) = \frac{1}{\frac{1}{c} - \frac{t^2}{2}} = \frac{2c}{2 - ct^2}.$$

769 / 862

Euler–homogene Differentialgleichungen

Definition

Sei $f: I \rightarrow \mathbb{R}$ stetig. Eine **Euler–homogene DG** ist eine DG vom Typ

$$x'(t) = f\left(\frac{x(t)}{t}\right) \quad \text{wobei} \quad (x, t) \text{ so dass } \frac{x}{t} \in I \quad (11)$$

Lösungsmethode: Die Substitution $u(t) := \frac{x(t)}{t}$ ergibt

$$u' = \frac{x' \cdot t - x}{t^2} = \frac{1}{t} \left(x'(t) - \frac{x}{t} \right) \stackrel{(11)}{=} \frac{1}{t} (f(u) - u).$$

Löst $x(t)$ die DG (11), so löst $u(t) = \frac{x(t)}{t}$ die DG mit **getrennten Variablen**

$$u'(t) = \frac{1}{t} (f(u) - u). \quad (12)$$

Umgekehrt ist für jede Lösung $u(t)$ der DG (12) die Funktion $x(t) = t \cdot u(t)$ Lösung der Euler–homogene DG (11), denn

$$x' = tu' + u \stackrel{(12)}{=} f(u) - u + u = f(u) = f\left(\frac{x}{t}\right).$$

770 / 862

Beispiel für Euler-homogene Differentialgleichungen

Wir betrachten die DG $x' = 1 + \frac{x}{t}$ mit der Anfangsbedingung $x(1) = x_0$, und wir suchen Lösungen auf $(0, \infty)$. Dann gilt

$$u(t) = \frac{x(t)}{t} \implies u' = \frac{x' \cdot t - x}{t^2} = \frac{1}{t^2}(t + x - x) = \frac{1}{t}.$$

Somit ist $u'(t) := \frac{1}{t}$ mit $u(1) = x_0$ zu lösen. Die Lösung ist aber offensichtlich gegeben durch $u(t) = \ln(t) + x_0$. Folglich erhalten wir als Lösung für die DG $x' = 1 + \frac{x}{t}$ mit Anfangsbedingung $x(1) = x_0$ die Funktion

$$x(t) = t \cdot u(t) = t(\ln(t) + x_0) \quad \text{für alle } t \in (0, \infty).$$

771 / 862

Lineare Differentialgleichungen

Definition

Eine **lineare Differentialgleichung** ist eine DG der Form

$$x'(t) = p(t)x(t) + q(t), \quad (13)$$

wobei $p, q: I \rightarrow \mathbb{R}$ stetige Funktionen sind. Ist die Störfunktion $q(t) \equiv 0$, so heißt (13) **homogene lineare DG**, andernfalls **inhomogene lineare DG**.

Satz (Homogene lineare DG)

Jede Lösung einer **homogenen**, linearen DG $x' = p(t)x$ ist gegeben durch

$$x(t) = c \cdot e^{\int p(t)dt},$$

wobei $c \in \mathbb{R}$ konstant und $\int p(t)dt$ eine Stammfunktion von p ist. Das AWP $x' = p(t)x$ mit der Anfangsbedingung $x(t_0) = x_0$ hat die eindeutige Lösung $x: I \rightarrow \mathbb{R}$, $x(t) = x_0 \cdot e^{\int_{t_0}^t p(s) ds}$.

772 / 862

Beweis. $x' = p(t)x$ ist eine DG mit getrennten Variablen. Somit ist die Lösung des Anfangswertproblems eindeutig gestimmt.

In der Tat erfüllt $x(t) = x_0 \cdot e^{\int_{t_0}^t p(s) ds}$ das AWP, denn

$$x'(t) = x_0 \cdot e^{\int_{t_0}^t p(s) ds} \cdot \left(\int_{t_0}^t p(s) ds \right)' = x(t)p(t) \quad \text{und} \quad x(t_0) = x_0.$$

□

Satz

Sei $x_s: I \rightarrow \mathbb{R}$ eine spezielle Lösung der inhomogenen **linearen** DG $x' = p(t)x + q$ mit $q \neq 0$. Dann erhält man alle Lösungen x der inhomogenen Gleichung mittels $x = x_s + x_c$ mit einer Lösung

$$x_c(t) := ce^{\int p(t) dt}$$

der homogenen Gleichung $x' = p(t)x$.

Beweis. $x - x_s$ löst die homogene lineare Gleichung. □

773 / 862

Wie findet man nun eine spezielle Lösung x_s der inhomogenen Gleichung $x' = p(t)x + q(t)$?

1. Methode: Variation der Konstanten

Wir betrachten eine Lösung $x(t) = c \cdot e^{\int p(t) dt}$ der homogenen, linearen DG $x' = p(t)x$ und machen den folgenden **Ansatz**:

Wir nehmen an, dass die spezielle Lösung die Form

$$x_s(t) := c(t)e^{\int p(t) dt}$$

hat und versuchen, daraus die Funktion $c(t)$ zu bestimmen. Es gilt

$$\begin{aligned} p(t)x_s + q(t) &= x'_s = c'(t) \cdot e^{\int p(t) dt} + c(t) \cdot p(t) \cdot e^{\int p(t) dt} \\ &= c'(t) \cdot e^{\int p(t) dt} + p(t) \cdot x_s(t). \end{aligned}$$

Folglich muss $c'(t) = q(t) \cdot e^{-\int p(t) dt}$ gelten und somit

$c(t) = \int q(t) \cdot e^{-\int p(t) dt} dt$. Mit dieser Funktion ist $x_s(t) = c(t)e^{\int p(t) dt}$ eine Lösung der inhomogenen, linearen DG $x' = p(t)x + q(t)$.

774 / 862

Beispiel zur Variation der Konstanten

Wir betrachten die inhomogene lineare Differentialgleichung

$$x' = tx + te^{t^2} \quad \text{mit Anfangsbedingung } x(0) = x_0. \quad (14)$$

- Allgemeine Lösung der homogenen DG $x' = tx$ ist $x(t) = ce^{\frac{t^2}{2}}$.
- Mit dem Ansatz $x_s(t) = c(t)e^{\frac{t^2}{2}}$ erhalten wir

$$x'_s(t) = c'(t)e^{\frac{t^2}{2}} + c(t)te^{\frac{t^2}{2}} = c'(t)e^{\frac{t^2}{2}} + tx_s(t) \stackrel{(14)}{\implies} c'(t)e^{\frac{t^2}{2}} = te^{t^2}$$

Folglich ist $c'(t) = te^{\frac{t^2}{2}} = (e^{\frac{t^2}{2}})'$ und somit $c(t) = e^{\frac{t^2}{2}}$.

- Spezielle Lösung der inhomogenen linearen DG: $x_s(t) = c(t)e^{\frac{t^2}{2}} = e^{t^2}$
- Allgemeine Lösung $x(t) = e^{t^2} + ce^{\frac{t^2}{2}}$ mit $c \in \mathbb{R}$.
- Für die Konstante c erhalten wir aus der AB $x(0) = x_0$ die Gleichung $x_0 = 1 + c$ und somit ist $x(t) = e^{\frac{t^2}{2}} \left(e^{\frac{t^2}{2}} + x_0 - 1 \right)$ die eindeutige Lösung des Anfangswertproblems (14).

775 / 862

2. Methode: Ansätze für x_s bei $p(t) \equiv p \neq 0$ konstant

Wir betrachten die inhomogene lineare DG $x'(t) = px(t) + q(t)$ mit $p \neq 0$ konstant.

- (1) Ist die "Störfunktion" $q(t)$ ein Polynom $h(t)$ vom Grad m mit reellen Koeffizienten, so setze für $x_s(t)$ ein Polynom Q vom Grad m an.
- (2) Ist $q(t)$ von der Form $h(t) \cdot e^{at}$, so setze für $x_s(t)$ die Funktion $Q(t) \cdot e^{at}$ ($p \neq a$) bzw. $t \cdot Q(t) \cdot e^{at}$ ($p = a$) an mit h, Q wie oben.
- (3) Ist $q(t)$ von der Form $h(t) \cdot \cos(bt)$ oder $h(t) \cdot \sin(bt)$, $h \in \mathbb{R}[t]$, $b \neq 0$, so setze für $x_s(t)$ die Funktion $Q_1(t) \cdot \cos(bt) + Q_2(t) \cdot \sin(bt)$ an.
- (4) Ist $q(t)$ von der Form $h(t) \cdot \cos(bt) \cdot e^{at}$ oder $h(t) \cdot \sin(bt) \cdot e^{at}$, $h \in \mathbb{R}[t]$, $b \neq 0$, so setze für $x_s(t)$ die Funktion $Q_1(t)e^{at} \cdot \cos(bt) + Q_2(t)e^{at} \cdot \sin(bt)$ an.

Dann setzt man den Ansatz in die inhomogene, lineare DG ein und berechnet die Koeffizienten des Polynoms $Q(t)$ durch

Koeffizientenvergleich.

776 / 862

Die Bernoullische Differentialgleichung

Definition

Eine **Bernoullische DG** ist eine DG des folgenden Types

$$x'(t) = p(t)x(t) + q(t)x(t)^\alpha, \quad (15)$$

wobei $\alpha \in \mathbb{R} \setminus (\{0\} \cup \{1\})$ und $p, q: I \rightarrow \mathbb{R}$ stetige Funktionen sind.

Eine Bernoullische DG wird durch die **Substitution** $u(t) := x(t)^{1-\alpha}$ behandelt. Man erhält

$$\begin{aligned} u' &= (1-\alpha)x^{-\alpha} \cdot x' \stackrel{(15)}{=} (1-\alpha)x^{-\alpha}(p(t)x + q(t)x^\alpha) \\ &= (1-\alpha)p(t)x^{1-\alpha} + (1-\alpha)q(t) \\ &= (1-\alpha)p(t)u + (1-\alpha)q(t), \end{aligned}$$

also eine inhomogene **lineare** DG für $u(t)$. Aus deren Lösung $u(t)$ erhält man $x(t) = u(t)^{\frac{1}{1-\alpha}}$ als Lösung der Bernoullischen DG.

777 / 862

Beispiel zur Bernoullischen Differentialgleichung

Wir betrachten eine Bernoullische DG

$$x' = -x + t\sqrt{x}, \quad \text{d.h. } \alpha = 1/2. \quad (16)$$

- Wir setzen $u(t) := x(t)^{\frac{1}{2}}$ und erhalten

$$u' = \frac{1}{2}x^{-\frac{1}{2}}x' \stackrel{(16)}{=} \frac{1}{2}x^{-\frac{1}{2}}(-x + t\sqrt{x}) = -\frac{1}{2}x^{\frac{1}{2}} + \frac{1}{2}t = -\frac{1}{2}u + \frac{1}{2}t. \quad (17)$$

- Für die homogene lineare DG $u' = -\frac{1}{2}u$ erhält man als allgemeine Lösung $u(t) = c \cdot e^{-\frac{1}{2}t}$.
- Um eine spezielle Lösung u_s der inhomogenen, linearen DG (17) $u' = -\frac{1}{2}u + \frac{1}{2}t$ zu finden, machen wir den **Ansatz** mit einem **Polynom ersten** Grades als Lösung: $u_s = at + b$. Daraus ergibt sich

$$a = u'_s = -\frac{1}{2}u_s + \frac{1}{2}t = -\frac{1}{2}(at + b) + \frac{1}{2}t.$$

Dies hat die Lösung $a = 1$ und $b = -2$, also $u_s(t) = t - 2$.

778 / 862

Weiter im Beispiel zur Bernoullischen Differentialgleichung

Damit haben wir $u_s(t) = t - 2$ als eine spezielle Lösung der inhomogenen DG (17) gefunden. Somit ist $u(t) = t - 2 + c e^{-\frac{1}{2}t}$ eine allgemeine Lösung der inhomogenen, linearen DG (16), und wir erhalten

$$x(t) = \left(t - 2 + c e^{-\frac{1}{2}t} \right)^2$$

als allgemeine Lösung der Bernoullischen Differentialgleichung (16)
 $x' = -x + t\sqrt{x}$.

779 / 862

Exakte Differentialgleichung

Sei $V = (P, Q): U \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld auf \mathbb{R}^2 . Eine zweimal stetig differenzierbare Funktion $F: U \rightarrow \mathbb{R}$ mit

$$\text{grad}F(x_1, x_2) = V(x_1, x_2) = (P(x_1, x_2), Q(x_1, x_2)),$$

heißt **Potentialfunktion** des Vektorfelds V .

Falls eine Potentialfunktion existiert, so muss wegen des Lemmas von Schwarz gelten

$$\frac{\partial P}{\partial x_2} = \frac{\partial^2 F}{\partial x_2 \partial x_1} = \frac{\partial^2 F}{\partial x_1 \partial x_2} = \frac{\partial Q}{\partial x_1}. \quad (18)$$

Bemerkung

Ist die offene Menge U sternförmig, so findet man zu jedem stetig differenzierbaren Vektorfeld $V = (P, Q): U \rightarrow \mathbb{R}^2$, welches $\frac{\partial P}{\partial x_2} = \frac{\partial Q}{\partial x_1}$ erfüllt, auch eine Potentialfunktion $F \in C^2(U, \mathbb{R})$.

780 / 862

Exakte Differentialgleichung

Definition

Seien $P, Q \in C^1(U, \mathbb{R})$, $U \subseteq \mathbb{R}^2$ offen, zusammenhängend und gelte auf U

$$\frac{\partial P}{\partial x_2} = \frac{\partial Q}{\partial x_1}, \quad \text{und } Q \neq 0. \quad (19)$$

Dann heißt die (gewöhnliche) Differentialgleichung

$$P(t, x(t)) + Q(t, x(t)) \cdot x'(t) = 0 \quad (20)$$

exakte Differentialgleichung. (19) heißt die **Integrabilitätsbedingung**.

Satz

Sei $F \in C^2(U, \mathbb{R})$ eine Potentialfunktion, also $\frac{\partial F}{\partial x_1} = P$ und $\frac{\partial F}{\partial x_2} = Q$. Es gelte $Q \neq 0$; dann erhält man eine Lösung der exakten DG (20) mit der AB $x(t_0) = x_0$ durch Auflösen der (Erhaltungs-)Gleichung

$$F(t, x) = F(t_0, x_0) \quad \text{nach } x.$$

781 / 862

Beweis. Wegen des Lemmas von Schwarz ist die Integrabilitätsbedingung (19) erfüllt, d.h. es liegt eine exakte Differentialgleichung vor.

Da $\frac{\partial F}{\partial x_2}(t_0, x_0) = Q(t_0, x_0) \neq 0$, folgt nach dem Satz über implizite Funktionen die eindeutige Auflösbarkeit von $F(t, x) = F(t_0, x_0)$ nach x in einem gewissen Intervall $(t_0 - \varepsilon, t_0 + \varepsilon)$. Das heißt, es existiert eine C^1 -Funktion $x: (t_0 - \varepsilon, t_0 + \varepsilon) \rightarrow \mathbb{R}$ mit $x(t_0) = x_0$ und

$$F(t, x(t)) = F(t_0, x_0).$$

Differenzieren nach t ergibt

$$\underbrace{\frac{\partial F}{\partial x_1}(t, x(t))}_{=P(t, x(t))} + \underbrace{\frac{\partial F}{\partial x_2}(t, x(t)) \cdot x'(t)}_{=Q(t, x(t))} = 0.$$

Somit erfüllt die implizite Funktion $x(t)$ die exakte Differentialgleichung (20). □

Beispiel zur exakten DG

Wir betrachten die DG $(t + x + 1) + (t + x)x' = 0$ mit AB $x(t_0) = x_0$ auf dem Gebiet $U = \{(x_1, x_2) \in \mathbb{R}^2 \mid (x_1 + x_2) > 0\}$. Diese ist exakt:

$$P(x_1, x_2) = x_1 + x_2 + 1, \quad Q(x_1, x_2) = x_1 + x_2, \quad \text{also} \quad \frac{\partial P}{\partial x_2} = 1 = \frac{\partial Q}{\partial x_1}.$$

Eine Potentialfunktion F von (P, Q) ist zum Beispiel

$$F(x_1, x_2) = \frac{1}{2}(x_1 + x_2)^2 + x_1$$

Wir lösen $0 = F(t, x) - F(t_0, x_0) = \frac{1}{2}(t + x)^2 + t - \frac{1}{2}(t_0 + x_0)^2 - t_0$ nach x auf und erhalten $(t + x)^2 = 2(t_0 - t) + (t_0 + x_0)^2$ und somit

$$x(t) = \sqrt{(t_0 + x_0)^2 + 2(t_0 - t)} - t,$$

da $x + t > 0$. Die Funktion $x(t)$ ist Lösung der gegebenen DG und der Definitionsbereich von x ist $\{t \in \mathbb{R} \mid t_0 + \frac{(t_0 + x_0)^2}{2} \geq t\}$.

783 / 862

Exakte Differentialgleichung: Integrierender Faktor

Ist die Bedingung $\frac{\partial P}{\partial x_2} = \frac{\partial Q}{\partial x_1}$ für die DG $P + Q \cdot x' = 0$ nicht erfüllt, so kann die Exaktheit durch die Multiplikation mit $\lambda \in C^1(U)$ erreicht werden.

Definition

Eine Funktion $\lambda \in C^1(U)$, $\lambda \neq 0$ heißt **integrierender Faktor** (**Eulerscher Multiplikator**) der DG $P + Q \cdot x' = 0$, falls $\frac{\partial(\lambda P)}{\partial x_2} = \frac{\partial(\lambda Q)}{\partial x_1}$.

Es ist dann $(\lambda P) + (\lambda Q) \cdot x' = 0$ **exakte** DG mit unveränderten Lösungen.

Beispiel

Sei $U = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1, x_2 > 0\}$. Wir betrachten die DG

$$5t^4 + 2x^3 + 3tx^2 \cdot x' = 0. \quad (21)$$

Hier ist $\frac{\partial P}{\partial x_2} = 6x_2^2 \neq \frac{\partial Q}{\partial x_1} = 3x_2^2$, das heißt (21) ist nicht exakt. Aber die Funktion $\lambda(x_1, x_2) = x_1$ ist ein integrierender Faktor für die DG (21); wir können die exakte DG $5t^5 + 2tx^3 + 3t^2x^2 \cdot x' = 0$ lösen.

784 / 862

Beispiele

Welchen Typ haben folgende DG (manchmal sind mehrere Antworten möglich)?

- $(x^2 + y)dx + xdy = 0$.
- $y' + 2ty - e^{-t} = 0$.
- $y' = \frac{x^2 - y^2}{5xy}$.
- $(1 - 2p)dq + (2 - q)dp = 0$.
- $\cos x \, dy = (y \sin x + e^x) \, dx$.
- $x(\tan y)y' = -1$.
- $y' = \frac{y}{x} + \frac{1}{y^3}$.
- $\frac{dv}{du} = e^{2u+3v}$.
- $xy' = y + xe^{\frac{y}{x}}$.
- $xy' + y = x^2 \sin x$.
- $y' = (x + e^y)^{-1}$.
- $y' + \frac{2y}{x} - \frac{y^2}{x} = 0$.

785 / 862

Beispiele

Welchen Typ haben folgende DG (manchmal sind mehrere Antworten möglich)?

- $(x^2 + y)dx + xdy = 0$. (*exakt, linear in y*)
- $y' + 2ty - e^{-t} = 0$. (*linear in y*)
- $y' = \frac{x^2 - y^2}{5xy}$. (*homogen, Bernoulli mit $\alpha = -1$*)
- $(1 - 2p)dq + (2 - q)dp = 0$. (*getrennte Variablen, linear in p und q*)
- $\cos x \, dy = (y \sin x + e^x) \, dx$. (*linear in y, exakt*)
- $x(\tan y)y' = -1$. (*getrennte Variablen*)
- $y' = \frac{y}{x} + \frac{1}{y^3}$. (*Bernoulli mit $\alpha = -3$*)
- $\frac{dv}{du} = e^{2u+3v}$. (*getrennte Variablen*)
- $xy' = y + xe^{\frac{y}{x}}$. (*homogen*)
- $xy' + y = x^2 \sin x$. (*linear in y*)
- $y' = (x + e^y)^{-1}$. (*linear in x: Umstellung ergibt $x' = x + e^y$*)
- $y' + \frac{2y}{x} - \frac{y^2}{x} = 0$. (*getrennte Variablen, Bernoulli mit $\alpha = 2$*)

786 / 862

Existenz und Eindeigkeitssätze für gewöhnliche DG'en

Wir haben gesehen, wie sich bestimmte gewöhnliche DG'en (eindeutig) lösen lassen. Im folgenden Abschnitt wollen wir allgemeine Aussagen über die Existenz und Eindeigkeit von Lösungen machen.

Dazu muss die "rechte Seite" f eine Bedingung erfüllen, die stärker ist als die der Stetigkeit und die wir schon vom Schrankensatz kennen.

Definition

Sei $\Omega \subseteq \mathbb{R}^m \times \mathbb{R}^n$ und $f: \Omega \rightarrow \mathbb{R}^k$, $(x, t) \mapsto f(x, t)$.

f heißt auf einer Teilmenge $K \subseteq \Omega$ **bezüglich x Lipschitzstetig** mit der **Lipschitzkonstanten** $L_K \geq 0$, wenn

$$\|f(x_1, t) - f(x_2, t)\| \leq L_K \|x_1 - x_2\| \quad (*)$$

für alle $(x_1, t), (x_2, t) \in K$.

f heißt auf Ω **bezüglich x lokal Lipschitzstetig**, wenn es zu jedem $p \in \Omega$ eine Umgebung U von p gibt, so dass $(*)$ für alle $(x_1, t), (x_2, t) \in U \cap \Omega$ mit einer Konstanten L_U gilt, die von U abhängen kann.

787 / 862

Bemerkung zur Lipschitzstetigkeit

Wir erinnern an den Begriff der gleichmäßigen Stetigkeit: Eine Abbildung f zwischen zwei metrischen Räumen heißt **gleichmäßig stetig** auf X , falls es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass $f(B_\delta(x)) \subseteq B_\varepsilon(f(x))$ für alle $x \in X$. Es gilt:

Lipschitzstetigkeit \Rightarrow gleichm. Stetigkeit \Rightarrow Stetigkeit.

Die erste Implikation beweist man, indem man $\delta = \frac{\varepsilon}{L}$ wählt.

Beispiele:

- $f(x) = \cos(x)$ ist nach dem Schrankensatz Lipschitzstetig auf ganz \mathbb{R} mit Lipschitzkonstante $L = 1$.
- $f: \mathbb{R}_+ \rightarrow \mathbb{R}$, $f(x) := \frac{1}{x}$ ist stetig, aber nicht gleichmäßig stetig, denn $f(x + \delta) - f(x) = \frac{1}{x+\delta} - \frac{1}{x} = \frac{-\delta}{x(x+\delta)} \xrightarrow{x \rightarrow 0} \infty$.
- $f: [0, 1] \rightarrow \mathbb{R}$, $f(x) := \sqrt{x}$ ist als stetige Abbildung auf einem Kompaktum gleichmäßig stetig, aber nicht Lipschitzstetig auf Intervallen $(0, a)$ und damit nicht lokal Lipschitzstetig, denn

$$\left| \frac{f(x) - f(y)}{x - y} \right| = \left| \frac{\sqrt{x} - \sqrt{y}}{x - y} \right| = \left| \frac{1}{\sqrt{x} + \sqrt{y}} \right| \xrightarrow{x, y \rightarrow 0} \infty.$$

788 / 862

Lipschitzbedingung

Satz (Lipschitzbedingung)

Sei $\Omega \subseteq \mathbb{R}^m \times \mathbb{R}$ offen und $f: \Omega \rightarrow \mathbb{R}^n$ von der Klasse C^1 , d.h. stetig differenzierbar. Ist $K \times [a, b] \subseteq \Omega$ mit K kompakt und konvex, dann ist f auf $K \times [a, b]$ bezüglich x Lipschitzstetig. Insbesondere ist f auf Ω bezüglich x lokal Lipschitzstetig.

Beweis. Der Beweis folgt unmittelbar aus dem **Schrankensatz**, der sich wiederum aus dem Mittelwertsatz ergab: Da $K \times [a, b] \subseteq \Omega$ kompakt und konvex ist und f stetig differenzierbar, ist nach dem Schrankensatz f auf $K \times [a, b] \subseteq \Omega$ Lipschitzstetig mit Lipschitzkonstante L . Insbesondere gilt

$$\|f(x_1, t) - f(x_2, t)\| \leq L\|(x_1, t) - (x_2, t)\| = L\|(x_1 - x_2, 0)\| = L\|x_1 - x_2\|,$$

und somit ist f auf $K \times [a, b]$ Lipschitzstetig bezüglich x . \square

789 / 862

Der Satz von Picard–Lindelöf

Satz (Lokaler Existenz- und Eindeigkeitssatz von Picard–Lindelöf)

Sei $F: U \subseteq \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ stetig, $(x_0, t_0) \in U$ und $a, b > 0$ so, dass

$$Q := \overline{B_b(x_0)} \times [t_0 - a, t_0 + a] \subseteq U,$$

wobei $\overline{B_b(x_0)} := \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq b\}$ die abgeschlossene Kugel vom Radius b um x_0 ist. Sei F Lipschitzstetig auf Q bzgl. der x -Variablen mit der Lipschitz-Konstanten L . Wir setzen

$$M := \max_{(y,t) \in Q} \|F(y, t)\| \quad \text{und} \quad \sigma := \min \left(a, \frac{b}{M}, \frac{1}{2L} \right).$$

Dann hat das AWP $\boxed{x' = F(x, t), x(t_0) = x_0}$ genau eine C^1 -Lösung

$$x: [t_0 - \sigma, t_0 + \sigma] \subseteq \mathbb{R} \rightarrow \mathbb{R}^n,$$

für die gilt $x(t) \in \overline{B_b(x_0)}$ für alle $t \in [t_0 - \sigma, t_0 + \sigma]$.

790 / 862

Beweis. Wir wollen den **Banachschen Fixpunktsatz** auf eine Abbildung zwischen Funktionenräumen anwenden.

Sei $I := [t_0 - \sigma, t_0 + \sigma]$. Wir betrachten im Banachraum $(C(I, \mathbb{R}^n), \|\cdot\|_\infty)$ die abgeschlossene Kugel K um die Funktion $\varphi_0(t) \equiv x_0$, d.h.

$$\begin{aligned} K &:= \{ \varphi \in C(I, \mathbb{R}^n) \mid \|\varphi(t) - x_0\| \leq b \quad \forall t \in I \} \\ &= C\left(I, \overline{B(x_0, b)}\right). \end{aligned}$$

Da K eine abgeschlossene Teilmenge im vollständigen metrischen Raum $(C(I, \mathbb{R}^n), d_\infty)$ ist, ist K selbst ein vollständiger metrischer Raum.

Wir betrachten nun den **Integraloperator**

$$\begin{aligned} H &: K \rightarrow C(I, \mathbb{R}^n) \\ \varphi &\mapsto \left(H(\varphi) : t \mapsto x_0 + \int_{t_0}^t F(\varphi(s), s) ds \right). \end{aligned}$$

Wegen der Wahl von $\sigma \leq \frac{b}{M}$ gilt $H(K) \subseteq K$, denn

$$\|H\varphi(t) - x_0\| = \left\| \int_{t_0}^t F(\varphi(s), s) ds \right\| \leq M\sigma \leq M \frac{b}{M} = b.$$

791 / 862

Weiter im Beweis:

Ein Fixpunkt von H ist nun eine Lösung des AWP's, denn

$$x(t) \stackrel{\text{FP}}{=} Hx(t) = x_0 + \int_{t_0}^t F(x(s), s) ds \iff \begin{cases} x'(t) = F(x(t), t), \\ x(t_0) = x_0. \end{cases}$$

Ist L die Lipschitzkonstante von F , so ist H Lipschitzstetig mit Lipschitzkonstante $L\sigma$, denn

$$\|Hx(t) - Hy(t)\| \leq \int_{t_0}^t \|F(x(s), s) - F(y(s), s)\| ds \leq L \int_{t_0}^t \|x(s) - y(s)\| ds. \quad (22)$$

Somit gilt (wegen $\sigma \leq \frac{1}{2L}$)

$$\|Hx - Hy\|_\infty \leq L\sigma \|x - y\|_\infty \leq \frac{1}{2} \|x - y\|_\infty. \quad (23)$$

Also ist $H : K \rightarrow K$ eine Kontraktion und besitzt nach dem Banachschen Fixpunktsatz einen eindeutigen Fixpunkt. \square

792 / 862

Approximative Lösung des AWP's

Folgerung

Unter den Voraussetzungen des Satzes von Picard–Lindelöf liefert der Banachsche Fixpunktsatz ein Verfahren zur **Approximation** der Lösung des Anfangswertproblems $x' = F(x, t)$, $x(t_0) = x_0$: Sei

$$\varphi_0(t) \equiv x_0, \quad \varphi_1 := H\varphi_0, \quad \dots, \quad \varphi_n := H\varphi_{n-1}, \quad \dots$$

iterativ definiert. Dann konvergiert φ_n gleichmäßig gegen die Lösung x des AWP's in $(C(I, \overline{B_b(x_0)}), \|\cdot\|_\infty)$. Die Abschätzung in (23) liefert

$$\|x(t) - \varphi_n(t)\| \leq \frac{b}{2^n} \quad \text{für alle } t \in [t_0 - \sigma, t_0 + \sigma].$$

Mit etwas mehr Sorgfalt erhält man Abschätzungen der Form

$$\|x(t) - \varphi_n(t)\| \leq \frac{M L^n}{(n+1)!} |t - t_0|^{n+1}.$$

793 / 862

Eindeutigkeit der Lösung

Wir brauchen ein Lemma:

Lemma

Sei $I \subseteq \mathbb{R}$ ein Intervall. Sei $D \subseteq I$ eine nicht-leere Teilmenge, die in I offen und abgeschlossen ist. Dann gilt $D = I$.

Beweis. Angenommen, $D \subsetneq I$, dann ist die Menge $I \setminus D$ nicht leer, offen und abgeschlossen. Finde $p_1 \in D$ und $p_2 \in I \setminus D$. Wir können o.B.d.A. annehmen, dass $p_1 < p_2$ gilt. Definiere

$$p := \sup D \cap [p_1, p_2].$$

Dann muss $p \in D$ oder $p \in I \setminus D$ gelten.

- Angenommen, $p \in D$. Da D offen ist, finden wir $\epsilon > 0$, so dass auch noch $(p - \epsilon, p + \epsilon) \subseteq D$, im Widerspruch dazu, dass p obere Schranke ist.
- Angenommen, $p \in I \setminus D$. Da auch $I \setminus D$ offen ist, finden wir $\epsilon > 0$, so dass $(p - \epsilon, p + \epsilon) \subseteq I \setminus D$, im Widerspruch dazu, dass p kleinste obere Schranke ist.

□

794 / 862

Satz (Eindeigkeitssatz)

Sei $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}$, $F: \Omega \rightarrow \mathbb{R}^n$ stetig und bezüglich x lokal Lipschitzstetig. Seien $\varphi, \psi: I = (a, b) \rightarrow \mathbb{R}^n$ zwei Lösungen von $x' = F(x, t)$ mit $\varphi(t_0) = \psi(t_0) = x_0$ für ein $t_0 \in I$. Dann gilt $\varphi = \psi$.

Beweis. Wir betrachten die Menge $D := \{t \in I \mid \varphi(t) = \psi(t)\}$. Dann ist

- D nicht leer, denn $t_0 \in D$,
- D ist abgeschlossen, denn φ und ψ sind stetig und $D = (\varphi - \psi)^{-1}(0)$.
- D ist offen in I : Sei dazu $t \in D$. Dann sind φ, ψ Lösungen des AWP's $x' = F(x, t)$, $x(t) = \varphi(t) = \psi(t)$. Nach Picard–Lindelöf lokal existiert ein $\varepsilon > 0$, so dass $\varphi \equiv \psi$ auf $(t - \varepsilon, t + \varepsilon) \subseteq I$. Dann ist aber $(t - \varepsilon, t + \varepsilon) \subseteq D$. D.h. D ist offen in I .

Damit ist D nicht leer, offen und abgeschlossen im Intervall I .

Somit ist $D = I$. □

795 / 862

Beispiel für nicht eindeutige Lösung

Die Lipschitzbedingung für F ist wesentlich. Wir betrachten als Beispiel die stetige Funktion $F: \mathbb{R} \rightarrow \mathbb{R}$ mit $F(x) = \sqrt[3]{x^2}$ und das Anfangswertproblem

$$x' = F(x) = \sqrt[3]{x^2}, \quad x(t_0) = 0.$$

Die Funktion F ist in $x_0 = 0$ nicht lokal Lipschitzstetig, so dass der Satz von Picard–Lindelöf **nicht** anwendbar ist.

Offensichtlich ist $x(t) \equiv 0$ eine Lösung des Anfangswertproblems.

Man findet jedoch noch unendlich viele weitere Lösungen, denn für beliebige Konstanten $c_1 \leq t_0 \leq c_2$ ist

$$x_c(t) := \begin{cases} \frac{(t-c_1)^3}{27} & \text{für } t \leq c_1 \\ 0 & \text{für } c_1 < t < c_2 \\ \frac{(t-c_2)^3}{27} & \text{für } t \geq c_2 \end{cases}$$

eine Lösung des Anfangswertproblems, wie man leicht nachrechnet (ÜA).

796 / 862

Fortsetzung zur maximalen Lösung

Definition

Eine Lösung $\varphi: I \rightarrow \mathbb{R}^n$ des Anfangswertproblems

$$(*) \quad x' = F(x, t), \quad x(t_0) = x_0,$$

heißt **maximal**, falls für jede andere Lösung $\psi: J \rightarrow \mathbb{R}^n$ von $(*)$ der Definitionsbereich J in I enthalten ist, d.h. $J \subseteq I$.

Folgerung (Satz über die maximale Lösung)

Sei $F: U \subseteq \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ stetig und lokal Lipschitzstetig bezüglich x . Dann existiert zu jedem $(x_0, t_0) \in U$ **genau eine maximale Lösung** $\varphi_{x_0}: (a_{x_0}, b_{x_0}) \subseteq \mathbb{R} \rightarrow \mathbb{R}^n$ des Anfangswertproblems $(*)$.

797 / 862

Beweis. Sei $(x_0, t_0) \in U$. Wir definieren die folgenden Zahlen

$$\begin{aligned} a_{x_0} &:= \inf\{a \in \mathbb{R} \mid \exists \text{ Lsg. des AWP's } \varphi: (a, b) \rightarrow \mathbb{R}\} \\ b_{x_0} &:= \sup\{b \in \mathbb{R} \mid \exists \text{ Lsg. des AWP's } \varphi: (a, b) \rightarrow \mathbb{R}\} \end{aligned}$$

mit $a_{x_0} < t_0 < b_{x_0}$. Auf dem offenen Intervall $I_{\max} := (a_{x_0}, b_{x_0}) \subseteq \mathbb{R}$ definieren wir die Abbildung

$$\begin{aligned} \varphi_{x_0}: I_{\max} &\rightarrow \mathbb{R}^n \\ t &\mapsto \varphi(t), \quad \text{falls } t \in (a, b) \text{ und } \varphi: (a, b) \rightarrow \mathbb{R}^n \text{ eine} \\ &\quad \text{Lsg. des AWP's } x' = F(x, t), \quad x(t_0) = x_0 \end{aligned}$$

Aufgrund des Eindeigkeitssatzes ist diese wohldefiniert, da zwei auf einem Intervall gegebene Lösungen übereinstimmen. Alle Lösungen sind C^1 -Abbildungen, und damit auch φ_{x_0} .

Des Weiteren löst φ_{x_0} das Anfangswertproblem $x' = F(x, t)$, $x(t_0) = x_0$ auf dem gesamten (und maximal möglichen) Intervall (a_{x_0}, b_{x_0}) . \square

798 / 862

Globale Existenz- und Eindeutigkeit

Satz

Sei $I = (a, b)$ mit $-\infty \leq a < b \leq \infty$ und $F: \mathbb{R}^n \times I \rightarrow \mathbb{R}^n$ Lipschitzstetig auf **jeder** Menge der Form $\mathbb{R}^n \times J$, wobei $J \subsetneq I$ ein beliebiges kompaktes Intervall ist.

Dann existiert zu jeder Anfangsbedingung $(x_0, t_0) \in \mathbb{R}^n \times I$ eine eindeutige, **auf ganz I definierte** maximale Lösung $\varphi_{x_0}: I \rightarrow \mathbb{R}^n$ des AWP's $x' = F(x, t)$ mit $x(t_0) = x_0$.

Beweis. Wir modifizieren den Beweis des Satzes von Picard–Lindelöf: Sei $J \subseteq I$ ein kompaktes Intervall der Länge d mit $t_0 \in J$. Ist $F: \mathbb{R}^n \times J \rightarrow \mathbb{R}^n$ Lipschitzstetig mit Konstante L , so ist der Integraloperator

$$H : C(J, \mathbb{R}^n) \rightarrow C(J, \mathbb{R}^n) \\ x \mapsto \left(H(x) : t \mapsto x_0 + \int_{t_0}^t F(x(s), s) ds \right).$$

Lipschitzstetig als Abbildung $C(J, \mathbb{R}^n) \rightarrow C(J, \mathbb{R}^n)$, mit

$$\|Hx - Hy\|_\infty \leq Ld\|x - y\|_\infty.$$

799 / 862

weiter im Beweis:

Im allgemeinen ist H nun nicht mehr kontrahierend bezüglich der Norm $\|\cdot\|_\infty$. Wir führen daher eine **neue Norm** ein, bzgl. derer H kontrahierend ist.

Für eine beliebige beschränkte Funktion $\alpha \in C(J, [r, s])$ definieren eine modifizierte Norm $\|\cdot\|_\alpha$ auf $C(J, \mathbb{R}^n)$ durch

$$\|\varphi\|_\alpha := \max_{t \in I} \|e^{\alpha(t)} \varphi(t)\|.$$

Diese ist äquivalent zu $\|\cdot\|_\infty$, denn $e^r \|\varphi\|_\infty \leq \|\varphi\|_\alpha \leq e^s \|\varphi\|_\infty$. Damit ist für jedes α wie oben auch $(C(J, \mathbb{R}^n), \|\cdot\|_\alpha)$ ein Banachraum.

Wir wählen nun konkret $\alpha(t) := -L|t - t_0|$, so dass $\alpha \in C(J, [-Ld, Ld])$. Multiplizieren wir dann die Ungleichung

$$\|Hx(t) - Hy(t)\| \leq \int_{t_0}^t \|F(x(s), s) - F(y(s), s)\| ds \leq L \int_{t_0}^t \|x(s) - y(s)\| ds$$

mit $e^{\alpha(t)} = e^{-L|t-t_0|}$, so erhalten wir für alle $t \in I$

800 / 862

Ende des Beweises:

$$\begin{aligned}
\left\| e^{\alpha(t)} (Hx(t) - Hy(t)) \right\| &\leq L e^{\alpha(t)} \left| \int_{t_0}^t e^{-\alpha(s)} e^{\alpha(s)} \|x(s) - y(s)\| ds \right| \\
&\leq L e^{-L|t-t_0|} \underbrace{\left| \int_{t_0}^t e^{L|s-t_0|} ds \right|}_{=\frac{1}{L}(e^{L|t-t_0|}-1)} \|x - y\|_{\alpha} \\
&\leq (1 - e^{-Ld}) \|x - y\|_{\alpha}.
\end{aligned}$$

D.h. aber, dass H kontrahierend ist bezüglich $\|\cdot\|_{\alpha}$:

$$\|Hx - Hy\|_{\alpha} \leq \underbrace{(1 - e^{-Ld})}_{<1} \|x - y\|_{\alpha}$$

Nach dem Banachschen Fixpunktsatz hat H genau einen Fixpunkt $x \in C(J, \mathbb{R}^n)$ und somit existiert genau eine Lösung x des Anfangswertproblems $x' = F(x, t)$ und $x(t_0) = x_0$ mit dem Definitionsbereich J . Da I sich als Vereinigung kompakter Teilintervalle J_k mit $t_0 \in J_k$ schreiben lässt, folgt daraus die Behauptung des Satzes. \square

801 / 862

Maximale Lösung des linearen Anfangswertproblems

Folgerung (Maximale Lösung des linearen AWP's)

Sei $I = (a, b)$ mit $-\infty \leq a < b \leq \infty$, $(x_0, t_0) \in \mathbb{R}^n \times I$, $A: I \rightarrow \text{Mat}_n(\mathbb{R})$ und $\mathbf{b}: I \rightarrow \mathbb{R}^n$ stetige Abbildungen. Dann besitzt das lineare AWP

$$x' = A(t)x + \mathbf{b}(t), \quad x(t_0) = x_0$$

genau eine maximale Lösung $\varphi_{x_0}: I \rightarrow \mathbb{R}^n$.

Beweis. $F(x, t) := A(t)x + \mathbf{b}(t)$ ist Lipschitzstetig auf jeder Menge der Form $\mathbb{R}^n \times J$ mit einem kompakten Intervall $J \subsetneq I$:

$$\|F(y, t) - F(x, t)\| = \|A(t)y - A(t)x\| = \|A(t)(y - x)\| \leq \max_{t \in J} \|A(t)\| \|y - x\|,$$

wobei $\|A(t)\| = \max_{\xi \in \mathbb{R}^n, \|\xi\|=1} \|A(t)\xi\|$. Die Lipschitz-Konstante ist $L := \max_{t \in J} \|A(t)\|$. Die Behauptung folgt dann aus dem vorhergehenden Satz. \square

802 / 862

Existenzsatz von Cauchy und Peano

Zum Abschluss geben wir noch einen allgemeineren Existenzsatz an, ohne ihn zu beweisen. Die Voraussetzungen sind schwächer (Stetigkeit statt Lipschitz-Stetigkeit), dafür erhält man keine Eindeigkeitsaussage.

Satz (Existenzsatz von Cauchy und Peano)

Sei $U \subseteq \mathbb{R}^{n+1}$ eine offene Menge und $(x_0, t_0) \in U$. Die Abbildung $F: U \subseteq \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ sei **stetig** auf dem kompakten Bereich

$$Q := \overline{B_b(x_0)} \times [t_0 - a, t_0 + a] \subseteq U.$$

Bezeichne wieder $M := \max_{(y,t) \in Q} \|F(y, t)\|$ und $\sigma := \min(a, \frac{b}{M})$. Dann hat das AWP $\boxed{x' = F(x, t), x(t_0) = x_0}$ **mindestens eine C^1 -Lösung**

$$x: [t_0 - \sigma, t_0 + \sigma] \subseteq \mathbb{R} \longrightarrow \mathbb{R}^n,$$

und diese erfüllt $\|x(t) - x_0\| \leq b$ für alle $t \in [t_0 - \sigma, t_0 + \sigma]$.

Für einen Beweis siehe H. Amann: *Gewöhnliche Differentialgleichungen*, 803 / 862

Abhängigkeit der Lösung von den Anfangsbedingungen

Wir wollen nun untersuchen, wie die Lösung von den Anfangswerten abhängt. Dazu benötigen wir folgendes Lemma.

Lemma (Gronwall-Ungleichung)

Seien $u, v: [a, b] \longrightarrow \mathbb{R}^+$ stetige, **nichtnegative** Funktionen und gelte

$$v(t) \leq c + \int_a^t v(s)u(s) ds \quad \text{für alle } t \in [a, b],$$

mit einer Konstante $c \in [0, \infty)$. Dann gilt

$$v(t) \leq c \cdot e^{\int_a^t u(s) ds} \quad \text{für alle } t \in [a, b].$$

Beweis. 1. Fall: $c > 0$. Wir betrachten die Funktion

$$f(t) := c + \int_a^t v(s)u(s) ds > 0.$$

Es gilt $0 < c \leq f(t)$ und $v(t) \leq f(t)$ für alle $t \in [a, b]$. Dann gilt

$$f'(t) = v(t) \cdot u(t) \leq f(t) \cdot u(t) \quad \text{und deshalb} \quad (\ln f(t))' = \frac{f'(t)}{f(t)} \leq u(t).$$

Durch Integration erhält man

$$\ln(f(t)) - \ln(f(a)) \leq \int_a^t u(s) ds, \quad \text{also} \quad f(t) \leq f(a) \cdot e^{\int_a^t u(s) ds}$$

Folglich ist $v(t) \leq f(t) \leq c \cdot e^{\int_a^t u(s) ds}$ wegen $c = f(a)$.

2. Fall: $c = 0$. Hier ist zu zeigen, dass $v(t) \equiv 0$. Nach Voraussetzung ist

$$v(t) \leq \int_a^t v(s)u(s) ds \leq \varepsilon + \int_a^t v(s)u(s) ds \quad \forall t \in [a, b], \quad \varepsilon > 0.$$

Nach dem 1. Fall folgt dann $0 \leq v(t) \leq \varepsilon \cdot e^{\int_a^t u(s) ds}$ für alle $\varepsilon > 0$.

Lassen wir ε gegen 0 gehen, so folgt $v(t) \equiv 0$.



805 / 862

Stetige Abhängigkeit der Lösung von den AB

Nun formulieren und beweisen wir die zentrale Aussage dieses Abschnitts.

Satz

Sei $F: U \subseteq \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ auf U stetig und Lipschitzstetig in x mit der Lipschitzkonstanten L . Seien $(x_0, t_0), (y_0, t_0) \in U$ und

$$\varphi_{x_0}, \varphi_{y_0}: [t_0 - \varepsilon, t_0 + \varepsilon] \rightarrow \mathbb{R}^n$$

Lösungen der DG $x' = F(x, t)$ mit den Anfangsbedingungen $\varphi_{x_0}(t_0) = x_0$ bzw. $\varphi_{y_0}(t_0) = y_0$. Dann gilt

$$\|\varphi_{x_0}(t) - \varphi_{y_0}(t)\| \leq \|x_0 - y_0\| \cdot e^{L|t-t_0|} \quad \forall t \in [t_0 - \varepsilon, t_0 + \varepsilon].$$

Beweis. $v(t) := \|\varphi_{x_0}(t) - \varphi_{y_0}(t)\|^2$ ist diff'bar. Für $t_0 \leq t < t_0 + \varepsilon$ gilt:

$$\begin{aligned}
 v(t) &= v(t_0) + \int_{t_0}^t v'(s) ds && \text{[Hauptsatz]} \\
 &= v(t_0) + 2 \int_{t_0}^t \langle \varphi_{x_0}(s) - \varphi_{y_0}(s), \varphi'_{x_0}(s) - \varphi'_{y_0}(s) \rangle ds \\
 &\stackrel{\text{CS-U.}}{\leq} v(t_0) + 2 \int_{t_0}^t \|\varphi_{x_0}(s) - \varphi_{y_0}(s)\| \cdot \|\varphi'_{x_0}(s) - \varphi'_{y_0}(s)\| ds \\
 &= v(t_0) + 2 \int_{t_0}^t \|\varphi_{x_0}(s) - \varphi_{y_0}(s)\| \cdot \|F(\varphi_{x_0}(s), s) - F(\varphi_{y_0}(s), s)\| ds \\
 &\leq v(t_0) + 2L \int_{t_0}^t \underbrace{\|\varphi_{x_0}(s) - \varphi_{y_0}(s)\|^2}_{=v(s)} ds
 \end{aligned}$$

Aus der Gronwall-Ungleichung für $u(s) \equiv 2L$ folgt dann

$$v(t) \leq v(t_0) \cdot e^{\int_{t_0}^t 2L ds} = v(t_0) \cdot e^{2L(t-t_0)}.$$

Damit erhält man die Behauptung $\|\varphi_{x_0}(t) - \varphi_{y_0}(t)\| \leq \|x_0 - y_0\| \cdot e^{L(t-t_0)}$ für alle $t_0 \leq t < t_0 + \varepsilon$. Analog geht man für $t_0 - \varepsilon < t \leq t_0$ vor. □
807/862

Folgerung (Stetige Abhängigkeit der Lösung von der AB)

Sei F wie im Satz, und sei $(x_n)_{n \geq 1}$ eine Folge von AB'en, die gegen die AB x_0 konvergiert. Sei φ_{x_n} die Lösung des AWP's $x' = F(x, t)$, $x(t_0) = x_n$. Dann konvergiert (φ_{x_n}) gleichmäßig gegen φ_{x_0} auf jedem kompakten Intervall $[t_0 - \varepsilon, t_0 + \varepsilon]$, auf dem alle Lösungen definiert sind. □

Bemerkungen

- Es ist ein wesentlicher Punkt im Satz und in der Folgerung, dass wir Aussagen über Lösungen machen, welche auf **kompakten** Intervallen $[a, b]$ um t_0 definiert sind. Wie einfache Beispiele zeigen (z.B. die Lösungen der Gleichung $x' = x$ im \mathbb{R}^n), ist mehr nicht zu erwarten.
- Ist die rechte Seite $F(x, t)$ eine C^∞ -Funktion, so konvergieren die Lösungen von $x' = F(x, t)$ zu konvergierenden AB $(x_n)_{n \geq 1}$ auf kompakten Intervallen auch gleichmäßig mit allen Ableitungen. Dies ist aber aufwändiger zu beweisen.

Beispiel zur Abhängigkeit der Lsg. von den AB's

Wir betrachten das folgende Vektorfeld auf dem \mathbb{R}^2 :

$$V(x, y) := (-y, x) + (r^2(x, y) - 1)(x, y),$$

wobei $r^2(x, y) := x^2 + y^2$. Wir suchen die Integralkurven von V durch einen Punkt $(x_0, 0) \in \mathbb{R}^2$, $x_0 > 0$, d.h. wir müssen das autonome AWP

$$\gamma'_{x_0}(t) = V(\gamma_{x_0}(t)) \quad \text{mit} \quad \gamma_{x_0}(0) = (x_0, 0) \quad (24)$$

lösen. Schreiben wir $\gamma_{x_0} = (x, y)$, so ist das Anfangswertproblem

$$\begin{aligned} x' &= -y + (r^2 - 1)x, \\ y' &= x + (r^2 - 1)y \end{aligned}, \quad x(0) = x_0 \quad y(0) = 0$$

zu lösen. Wir machen den folgenden Ansatz für die Lösung γ_{x_0} :

$$\gamma_{x_0}(t) = h(t)(\cos(t), \sin(t))$$

für eine Funktion $h: I \rightarrow \mathbb{R}$ mit $h(0) = x_0$.

809 / 862

Dann gilt $r^2(\gamma_{x_0}(t)) = h^2(t)$. Ist $\gamma_{x_0}(t) = h(t)(\cos(t), \sin(t))$ eine Integralkurve von V , so gilt

$$\begin{aligned} \gamma'_{x_0}(t) &= h'(t)(\cos(t), \sin(t)) + h(t)(-\sin(t), \cos(t)) \\ &\stackrel{(24)}{=} h(t)(-\sin(t), \cos(t)) + (h^2(t) - 1)h(t)(\cos(t), \sin(t)) \end{aligned}$$

Also muss h das autonome AWP $h' = h(h^2 - 1)$, $h(0) = x_0$ erfüllen. Dies löst man mittels Trennung der Variablen und erhält

$$h(t) = \frac{e^{-t}}{\sqrt{e^{-2t} - \left(1 - \frac{1}{x_0^2}\right)}}.$$

Ist nun $x_0 \leq 1$, so ist $\left(1 - \frac{1}{x_0^2}\right) \leq 0$ und damit ist $h: \mathbb{R} \rightarrow \mathbb{R}$ auf ganz \mathbb{R} definiert. Ist dagegen $x_0 > 1$, so ist $0 < \left(1 - \frac{1}{x_0^2}\right) < 1$ und damit ist h nur definiert auf dem Intervall $I = \left(-\infty, -\frac{1}{2} \ln \left(1 - \frac{1}{x_0^2}\right)\right) \subseteq \mathbb{R}$ um $t_0 = 0$.

810 / 862

Noch ein Beispiel für Trennung der Variablen

Wir rechnen noch einmal die Lösung des AWP $h' = h(h^2 - 1)$ mit Anfangsbedingungen $h(0) = x_0$ durch Trennung der Variablen nach: Es ist $f(t) = 1$ und $g(x) = x(x^2 - 1)$ und

$$\frac{1}{g} = \frac{1}{x(x^2 - 1)} = -\frac{1}{x} + \frac{1/2}{x - 1} + \frac{1/2}{x + 1}$$

mit Stammfunktion

$$G(x) = -\ln x + \frac{1}{2} \ln(x - 1) + \frac{1}{2} \ln(x + 1) = \frac{1}{2} \ln\left(1 - \frac{1}{x^2}\right).$$

Es folgt $G^{-1}(y) = \frac{1}{\sqrt{1 - e^{2y}}}$ und somit als Lösung des AWP

$$G^{-1}\left(t + \frac{1}{2} \ln\left(1 - \frac{1}{x_0^2}\right)\right) = \frac{e^{-t}}{\sqrt{e^{-2t} - \left(1 - \frac{1}{x_0^2}\right)}}$$

811 / 862

Wir unterscheiden nun drei Fälle für die Integralkurven

$\gamma_{x_0}(t) = h(t)(\cos(t), \sin(t))$:

- $x_0 = 1$: Hier ist $h(t) \equiv 1$ und die Integralkurve ist ein Kreis.
- $0 < x_0 < 1$: Hier sind die Integralkurven $\gamma_{x_0}: \mathbb{R} \rightarrow \mathbb{R}^2$ Spiralen, die auf ganz \mathbb{R} definiert sind. Das Quadrat des Abstandes zum Nullpunkt $r^2(t) := r^2(\gamma_{x_0}(t))$ ist

$$r^2(t) = h^2(t) = \frac{e^{-2t}}{e^{-2t} - \left(1 - \frac{1}{x_0^2}\right)} = \frac{1}{1 - e^{2t} \left(1 - \frac{1}{x_0^2}\right)}$$

und erfüllt $r^2(t) \rightarrow_{t \rightarrow -\infty} 1$ und $r^2(t) \rightarrow_{t \rightarrow \infty} 0$. D.h. die Spiralen laufen für $t \rightarrow \infty$ in die Null und nähern sich für $t \rightarrow -\infty$ dem Einheitskreis an.

- $x_0 > 1$: Die Integralkurven $\gamma_{x_0}: \left(-\infty, -\frac{1}{2} \ln\left(1 - \frac{1}{x_0^2}\right)\right) \rightarrow \mathbb{R}^2$ sind Spiralen, die sich für $t \rightarrow -\infty$ dem Einheitskreis annähern, und für $t \rightarrow -\frac{1}{2} \ln\left(1 - \frac{1}{x_0^2}\right)$ in endlicher Zeit gegen ∞ gehen.

812 / 862

Lineare Differentialgleichungen mit Werten im \mathbb{R}^n

Wir haben bereits Lösungsmethoden für lineare Differentialgleichungen für Funktionen mit Werten in \mathbb{R} kennengelernt. In diesem Kapitel verallgemeinern wir diese und behandeln lineare Differentialgleichungen mit Werten im \mathbb{R}^n .

Seien $I = (a_0, b_0) \subseteq \mathbb{R}$ ein Intervall mit $-\infty \leq a_0 < b_0 \leq \infty$ und

$$A: I \longrightarrow \text{Mat}_n(\mathbb{R}), \quad b: I \longrightarrow \mathbb{R}^n$$

stetige Abbildungen. Dann heißt

$$\begin{aligned} x' &= A(t)x && \text{homogenes lineares DG System,} \\ x' &= A(t)x + b(t) && \text{inhomog. lin. DG System mit Störfkt. } b(t). \end{aligned}$$

Gilt speziell $A(t) \equiv A \in \text{Mat}_n(\mathbb{R})$, so sagt man, das System habe **konstante Koeffizienten**.

Wir haben schon gesehen, dass es zu jeder Anfangsbedingung der Form $x(t_0) = x_0$ mit $t_0 \in I$ genau eine maximale Lösung gibt, die auf ganz I definiert ist.

813 / 862

Struktur des Lösungsraumes linearer Systeme

Satz (Struktur des Lösungsraumes linearer Systeme)

Die Menge V der maximalen Lösungen $x: I \rightarrow \mathbb{R}^n$ des **homogenen** linearen Systems $(*)$ $x' = A(t)x$ ist ein n -dimensionaler reeller **Vektorraum**.

Die Menge der maximalen Lösungen des **inhomogenen** linearen Systems $(**)$ $x' = A(t)x + b(t)$ ist der **affine Raum** $\mathcal{A} = x_s + V$ wobei x_s eine spezielle Lösung der inhomogenen linearen Differentialgleichung ist.

Beweis. Sind x_1 und x_2 zwei Lösungen von $(*)$ und $\lambda_1, \lambda_2 \in \mathbb{R}$, so ist auch $\lambda_1 x_1 + \lambda_2 x_2$ eine Lösung von $(*)$. Zu jedem Vektor $x_0 \in \mathbb{R}^n$ existiert genau eine maximale Lösung φ_{x_0} mit $\varphi_{x_0}(t_0) = x_0$, d.h. die lineare Abb. $V \rightarrow \mathbb{R}^n$, $\varphi \mapsto \varphi(t_0)$ ist ein Vektorraum-Isomorphismus, und somit ist $\dim V = n$.

Sei \mathcal{A} der Lösungsraum des inhom. lin. Systems $(**)$. Wir wissen, dass eine spezielle Lösung x_s der inhomogenen linearen DG $(**)$ existiert.

Außerdem gilt für jede andere Lösung $x \in \mathcal{A}$, dass $x - x_s \in V$ und für jede Lösung $y \in V$, dass $x_s + y \in \mathcal{A}$. Somit ist $\mathcal{A} = x_s + V$. \square

814 / 862

Fundamentalsystem und Fundamentalmatrix

Definition (Fundamentalsystem und Fundamentalmatrix)

Sei $A: I \rightarrow \text{Mat}_n(\mathbb{R})$ stetig und $(*) x' = A(t)x$ ein **homogenes** lineares DG-System.

- Eine Basis $\varphi_1, \dots, \varphi_n$ des Lösungsraumes von $(*)$ nennt man ein **Fundamentalsystem** zu $(*)$.
- Die matrixwertige Funktion $\Phi = (\varphi_1, \dots, \varphi_n)$ aus den Spaltenvektoren φ_j heißt eine **Fundamentalmatrix** von $(*)$.
- Seien $(\varphi_1, \dots, \varphi_n)$ Lösungen von $(*)$. Die Funktion $W := \det(\varphi_1, \dots, \varphi_n): I \rightarrow \mathbb{R}$ heißt **Wronski-Determinante** des Fundamentalsystems $(\varphi_1, \dots, \varphi_n)$.
- Bilden $\varphi_j = \sum_{i=1}^n \varphi_{ij} e_i$ ein Fundamentalsystem mit $\varphi_{ij}: \mathbb{R} \rightarrow \mathbb{R}$ und e_i die kanonische Basis, dann ist die Fundamentalmatrix gegeben durch $\Phi = (\varphi_{ij})_{i,j=1}^n$.
- n Lösungen von $(*)$ bilden genau dann ein Fundamentalsystem, wenn ihre Wronski-Determinante für ein t nicht verschwindet.

815 / 862

Eigenschaften der Fundamentalmatrix

Sei Φ eine Fundamentalmatrix zu $(*)$. Dann gilt:

- 1) Φ ist eine Lösung des linearen homogenen Systems $X' = A \cdot X$ mit Werten in $\text{Mat}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$.
- 2) Eine Lösung des homogenen Anfangswertproblems $x' = Ax$, $x(t_0) = x_0$ ist gegeben durch

$$\varphi(t) = \Phi(t)(\Phi(t_0))^{-1}x_0$$

- 3) Ist $S \in \text{GL}(n, \mathbb{R})$, d.h. S invertierbar, so ist $\Phi \cdot S$ ist auch eine Fundamentalmatrix.

Beweis.

- 1) $\Phi' = (\varphi_1', \dots, \varphi_n') = (A\varphi_1, \dots, A\varphi_n) = A \cdot \Phi$.
- 2) $\varphi'(t) = \Phi'(t)(\Phi(t_0))^{-1}x_0 = A\Phi(t)(\Phi(t_0))^{-1}x_0 = A\varphi(t)$.
- 3) $(\Phi \cdot S)' = \Phi' \cdot S = A \cdot \Phi \cdot S$.

□

Beispiel

Sei $\omega \in \mathbb{R}$ eine Konstante. Wir betrachten das homogene lineare DG-System mit konstanten Koeffizienten

$$\begin{cases} x_1' &= -\omega x_2 \\ x_2' &= \omega x_1 \end{cases}, \text{ d.h. } \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Ein Fundamentalsystem wird von den folgenden beiden Funktionen gebildet

$$\varphi_1(t) = \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \end{pmatrix}, \quad \varphi_2(t) = \begin{pmatrix} -\sin(\omega t) \\ \cos(\omega t) \end{pmatrix}$$

Diese entsprechen den Anfangsbedingungen $\varphi_1(0) = e_1$ und $\varphi_2(0) = e_2$. Zu einer beliebigen Anfangsbedingungen $x(0) = c_1 e_1 + c_2 e_2$ erhalten wir die Lösung $\varphi = c_1 \varphi_1 + c_2 \varphi_2$.

Die Fundamentalmatrix zu φ_1, φ_2 ist dann gegeben durch

$$\Phi(t) = \begin{pmatrix} \cos(\omega t) & -\sin(\omega t) \\ \sin(\omega t) & \cos(\omega t) \end{pmatrix}.$$

817 / 862

Spezielle Lösungen des inhomogenen Systems

Wie im eindimensionalen Fall erhält man spezielle Lösungen der inhomogenen lin. DG $x' = A(t)x + b(t)$ aus einem Fundamentalsystem der zugehörigen homogenen DG mittels **Variation der Konstanten**.

Satz (Variation der Konstanten)

Seien $A: I \rightarrow \text{Mat}_n(\mathbb{R})$ und $b: I \rightarrow \mathbb{R}^n$ stetig und $\Phi: I \rightarrow \text{Mat}_n(\mathbb{R})$ eine Fundamentalmatrix des linearen homogenen Systems $x' = A(t)x$. Dann ist

$$\psi_s(t) := \Phi(t)u(t)$$

eine spezielle Lösung des inhomogenen Systems $x' = A(t)x + b(t)$, wobei $u: I \rightarrow \mathbb{R}^n$ stetig differenzierbar mit $\Phi(t)u'(t) = b(t)$, d.h.

$$(*) \quad u(t) = \int_{t_0}^t \Phi^{-1}(s)b(s)ds + v$$

mit einem konstanten Vektor v . Gibt man zur inhomogenen DG die Anfangsbedingung $x(t_0) = x_0$ vor, so ist $v := \Phi^{-1}(t_0)x_0$.

818 / 862

Beweis. Für $u(t) = \int_{t_0}^t \Phi^{-1}(s)b(s)ds + v$ wie in (*) gilt dass $u'(t) = \Phi^{-1}(t)b(t)$ und damit $\Phi(t)u'(t) = b(t)$. Damit gilt für $\psi_s(t) = \Phi(t)u(t)$:

$$\begin{aligned}\psi'_s(t) &= \Phi'(t)u(t) + \Phi(t)u'(t) && \text{[Produktregel]} \\ &= A(t)\Phi(t)u(t) + b(t) = A(t)\psi_s(t) + b(t).\end{aligned}$$

Also ist $\psi_s(t) = \Phi(t)u(t)$ eine spezielle Lösung. \square

Bemerkung

Ist $v = (v_1, \dots, v_n): I \rightarrow \mathbb{R}^n$ eine stetige Funktion, so ist mit $\int_{t_0}^t v(s)ds$ die folgende differenzierbare Abbildung gemeint

$$I \ni t \mapsto \int_{t_0}^t v(s)ds := \begin{pmatrix} \int_{t_0}^t v_1(s)ds \\ \vdots \\ \int_{t_0}^t v_n(s)ds \end{pmatrix} \in \mathbb{R}^n.$$

819 / 862

Beispiel

Wir betrachten das inhom. lin. DG-System mit konst. Koeffizienten

$$\begin{cases} x'_1 &= -x_2 \\ x'_2 &= x_1 + t \end{cases}, \quad d.h. \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{=A(t)} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \underbrace{\begin{pmatrix} 0 \\ t \end{pmatrix}}_{=b(t)}$$

Wir müssen nun folgendes tun, um eine spezielle Lösung zu erhalten:

- 1) Bestimmen der Fundamentalmatrix (bereits bekannt, s.o.):

$$\Phi(t) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

- 2) Invertieren von Φ : $\Phi(t)^{-1} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$, und damit

$$\Phi(t)^{-1}b(t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} 0 \\ t \end{pmatrix} = \begin{pmatrix} t \sin(t) \\ t \cos(t) \end{pmatrix}.$$

820 / 862

3) Integrieren von $\Phi^{-1}b$:

$$u(t) = \int_{t_0}^t \Phi^{-1}(s)b(s)ds = \int_{t_0}^t \begin{pmatrix} s \sin(s) \\ s \cos(s) \end{pmatrix} ds$$

Dies berechnen wir mit partieller Integration und erhalten

$$\int t \sin(t) dt = -t \cos(t) + \sin(t)$$

$$\int t \cos(t) dt = t \sin(t) + \cos(t)$$

und damit $u(t) = \begin{pmatrix} -t \cos(t) + \sin(t) \\ t \sin(t) + \cos(t) \end{pmatrix}$. Somit ist

$$\psi_s(t) = \Phi(t)u(t) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \begin{pmatrix} -t \cos t + \sin t \\ t \sin t + \cos t \end{pmatrix} = \begin{pmatrix} -t \\ 1 \end{pmatrix}$$

eine spezielle Lösung des inhomogenen Systems.

821 / 862

Homogene lineare Systeme mit konstanten Koeffizienten

Um ein Fundamentalsystem eines homog. lin. DG-Systems $x' = A(t)x$ zu bestimmen, betrachten wir den Spezialfall $A(t) \equiv A \in \text{Mat}_n(\mathbb{R})$ konstant.

Definition (Matrixexponential)

Sei $A \in \text{Mat}_n(\mathbb{R})$. Die Matrix $e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!}$ heißt **Exponential** von A .

Dies ist wohldefiniert, d.h. diese Reihe konvergiert:

1) Es gilt die folgende Ungleichung für die Operatornorm:

$$\|AB\| = \sup_{x \in \mathbb{R}^n} \frac{\|ABx\|}{\|x\|} \leq \sup_{x \in \mathbb{R}^n} \|A\| \frac{\|Bx\|}{\|x\|} = \|A\| \|B\|.$$

2) Da $(\text{Mat}_n(\mathbb{R}), \|\cdot\|)$ ein Banachraum ist, reicht es, zu zeigen, dass die Partialsummen $S_m := \sum_{k=0}^m \frac{A^k}{k!}$ eine Cauchy-Folge bilden. Es ist aber

$$\|S_m - S_{\ell-1}\| = \left\| \sum_{k=\ell}^m \frac{A^k}{k!} \right\| \leq \sum_{k=\ell}^m \frac{\|A^k\|}{k!} \leq \sum_{k=\ell}^m \frac{\|A\|^k}{k!} \rightarrow_{m, \ell \rightarrow \infty} 0,$$

denn die reellwertige Reihe $e^{\|A\|} = \sum_{k=0}^{\infty} \frac{\|A\|^k}{k!}$ konvergiert. Die Reihe e^A konvergiert sogar absolut.

822 / 862

Lemma

Sei $A \in \text{Mat}_n(\mathbb{R})$ und $\Phi: \mathbb{R} \rightarrow \text{Mat}_n(\mathbb{R})$, $\Phi(t) := e^{tA}$. Es sei eine Folge $\Phi_m: \mathbb{R} \rightarrow \text{Mat}_n(\mathbb{R})$ von Abbildungen definiert durch $\Phi_m(t) := \sum_{k=0}^m \frac{t^k A^k}{k!}$. Dann konvergiert Φ_m auf jedem kompakten Intervall $K \subseteq \mathbb{R}$ gleichmäßig gegen Φ . Insbesondere ist die Funktion Φ stetig auf \mathbb{R} .

Beweis. Sei $K = [a, b] \subseteq \mathbb{R}$ und $M := \max\{|a|, |b|\} \cdot \|A\|$. Die konvergente Reihe $e^M = \sum_{k=0}^{\infty} \frac{M^k}{k!}$ ist eine Majorante für die Exponentialreihe der Beträge $e^{\|tA\|} = \sum_{k=0}^{\infty} \frac{\|tA\|^k}{k!}$. Nach dem Cauchy-Kriterium finden wir zu $\varepsilon > 0$ ein n_ε , so dass $\frac{M^m}{m!} + \dots + \frac{M^\ell}{\ell!} < \varepsilon$ für alle $m \geq \ell \geq n_\varepsilon$. Somit gilt für alle $t \in K$ und $m \geq \ell \geq n_\varepsilon$, dass

$$\|\Phi_m(t) - \Phi_{\ell-1}(t)\| \leq \sum_{k=\ell}^m \frac{|t|^k \|A^k\|}{k!} \leq \sum_{k=\ell}^m \frac{|t|^k \|A\|^k}{k!} \leq \sum_{k=\ell}^m \frac{M^k}{k!} \leq \varepsilon.$$

D.h., sowohl Φ_m als auch die Exponentialreihe der Beträge sind Cauchyfolgen in vollständigen metrischen Räumen und somit gleichmäßig konvergent auf dem kompakten Intervall K . Damit ist der Grenzwert Φ stetig auf allen kompakten Intervallen, und damit auch auf ganz \mathbb{R} . □ 823/862

Die Lsg. der homogenen DG mit konstanten Koeffizienten

Satz

Sei $A \in \text{Mat}_n(\mathbb{R})$. Dann ist die matrixwertige Funktion $\Phi: \mathbb{R} \ni t \mapsto e^{tA} \in \text{Mat}_n(\mathbb{R})$ eine Fundamentalmatrix zur homogenen linearen DG $x' = Ax$. Insbesondere ist zu jedem $x_0 \in \mathbb{R}^n$ die differenzierbare Abbildung $\varphi_{x_0}: \mathbb{R} \rightarrow \mathbb{R}^n$, definiert durch $\varphi_{x_0}(t) = e^{(t-t_0)A} \cdot x_0$, die Lösung des AWP's $x' = Ax$, $x(t_0) = x_0$.

Beweis. Da A konstant ist, sind die Lösungen $\varphi_1, \dots, \varphi_n$ der Anfangswertprobleme $x' = Ax$, mit Anfangsbedingungen $x(0) = e_i$, $i = 1, \dots, n$ auf ganz \mathbb{R} definiert.

Sei $\Phi = (\varphi_1, \dots, \varphi_n)$ die dazugehörige Fundamentalmatrix.

Nach der Iteration von Picard-Lindelöf wird jede dieser Lösungen φ_i auf kompakten Intervallen gleichmäßig approximiert durch Funktionen $\varphi_{i,m}$, die sich durch iteratives Anwenden des Integraloperators H ergeben, d.h.

$$\varphi_{i,0}(t) \equiv e_i, \quad \varphi_{i,1}(t) = e_i + \int_0^t Ax_0 ds, \quad \varphi_{i,m}(t) = e_i + \int_0^t A\varphi_{i,m-1}(s) ds$$

Schreibt man Φ_m für die Matrix mit den Spalten $(\varphi_{1,m}, \dots, \varphi_{n,m})$, so ist

$$\Phi_0(t) \equiv \mathbf{1}_n$$

$$\Phi_1(t) = H\Phi_0(t) = \mathbf{1}_n + \int_0^t A ds = \mathbf{1}_n + tA$$

$$\Phi_2(t) = H\Phi_1(t) = \mathbf{1}_n + \int_0^t A(\mathbf{1}_n + sA) ds = \mathbf{1}_n + tA + \frac{t^2 A^2}{2}$$

\vdots

$$\Phi_m(t) = H\Phi_{m-1}(t) = \mathbf{1}_n + \int_0^t A \left(\sum_{k=0}^{m-1} \frac{s^k A^k}{k!} \right) ds = \sum_{k=0}^m \frac{t^k A^k}{k!}$$

Wir haben aber gesehen, dass die Folge von Abbildungen

$\Phi_m: t \mapsto \sum_{k=0}^m \frac{t^k A^k}{k!}$ auf jedem kompakten Intervall gleichmäßig gegen $\Phi(t) = e^{tA}$ konvergiert. Da die kompakten Intervalle beliebig groß sein können, erhält man dass $\Phi: \mathbb{R} \rightarrow \text{Mat}_n(\mathbb{R})$ die Fundamentalmatrix von $x' = Ax$ ist. Damit ist Φ differenzierbar und $\varphi(t) = \Phi(t)x_0$ die Lösung des AWP's $x(0) = x_0$. Dass $\psi(t) = \Phi(t - t_0)x_0$ eine Lsg. des AWP's $x(t_0) = x_0$ ist, folgt aus der Kettenregel. □

825 / 862

Beispiel

Wir betrachten wieder die homogene lin. DG mit **konstanten** Koeffizienten

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}' = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \text{d.h. } A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Dann rechnet man nach, dass $A^{2k+1} = (-1)^k A$ und $A^{2k} = (-1)^k \mathbf{1}_2$. Somit können wir das Exponential e^{tA} bestimmen:

$$\begin{aligned} e^{tA} &= \sum_{k=0}^{\infty} \frac{t^{2k} A^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{t^{2k+1} A^{2k+1}}{(2k+1)!} \\ &= \left(\sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} \right) \mathbf{1}_2 + \left(\sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!} \right) A \\ &= (\cos t) \mathbf{1}_2 + (\sin t) A = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}. \end{aligned}$$

Folgerung (Exponentialabbildung)

Für $A \in \text{Mat}_n(\mathbb{R})$ ist die Abbildung $\Phi: (\mathbb{R}, +) \rightarrow \text{GL}(n, \mathbb{R})$, $\Phi(t) = e^{tA}$ ein differenzierbarer Gruppenhomomorphismus von der additiven Gruppe $(\mathbb{R}, +)$ in die Gruppe der invertierbaren Matrizen $\text{GL}(n, \mathbb{R})$. Insbesondere ist $e^{(t+s)A} = e^{tA} \cdot e^{sA} = e^{sA} \cdot e^{tA}$ und $(e^{tA})^{-1} = e^{-tA}$

Beweis. Seien $s \in \mathbb{R}$ und $x_0 \in \mathbb{R}^n$ beliebig und seien φ und ψ Lösungen des AWP's $x' = Ax$ mit den jeweiligen Anfangsbedingungen $x(s) = x_0$ und $x(0) = x_0$. Nach dem vorherigen Satz gilt für alle $t \in \mathbb{R}$, dass $\varphi(t) = e^{(t-s)A}x_0 = \psi(t-s)$. Somit erhalten wir mit Eigenschaften der Fundamentalmatrix, dass

$$\Phi(t)\Phi(s)^{-1}x_0 = \varphi(t) = \psi(t-s) = \Phi(t-s)\Phi(0)^{-1}x_0 = \Phi(t-s)x_0$$

für alle $x_0 \in \mathbb{R}^n$. Damit erhalten wir für alle $t, s \in \mathbb{R}$, dass

$$(*) \quad \Phi(t)\Phi(s)^{-1} = \Phi(t-s).$$

Dies impliziert aber die Behauptung

$$\Phi(t+s) = \Phi(t+s)\Phi(s)^{-1}\Phi(s) \stackrel{(*)}{=} \Phi(t)\Phi(s).$$

827 / 862



Folgerung (Weitere Eigenschaften der Exponentialabbildung)

Seien $A, B \in \text{Mat}_n(\mathbb{R})$.

- ① Ist $AB = BA$, dann gilt $Ae^B = e^B A$ und $e^{A+B} = e^A \cdot e^B$.
- ② Ist $B \in \text{GL}(n, \mathbb{R})$, so ist $e^{BAB^{-1}} = Be^A B^{-1}$.

Beweis. Alle Aussagen folgen aus der Eindeutigkeit der Lsg. eines AWP's. Sei z.B. $X(t) = Ae^{tB}$ und $Y(t) = e^{tB}A$. Dann gilt $X(0) = Y(0) = A$ und

$$X'(t) = AB e^{tB} = BA e^{tB} = BX(t) \quad \text{sowie} \quad Y'(t) = B e^{tB} A = BY(t).$$

Aus der Eindeutigkeit der Lsg. des AWP's $X' = BX$, $X(0) = A$ folgt dann $X(t) = Y(t)$ für alle $t \in \mathbb{R}$.

Sei weiter $U(t) := e^{t(A+B)}$ und $V(t) = e^{tA} \cdot e^{tB}$. Dann gilt $U(0) = V(0) = \mathbf{1}_n$ und

$$U'(t) = (A+B)U(t) \quad \text{sowie} \quad V'(t) = Ae^{tA}e^{tB} + e^{tA}Be^{tB} = (A+B)V(t)$$

und somit $U(t) = V(t)$ für alle $t \in \mathbb{R}$.

Die verbleibende Aussage beweist man analog.



828 / 862

Zusammenfassung: Lineare DGN mit konstanten Koeffizienten

Sei $A \in \text{Mat}(n, \mathbb{R})$, $b: \mathbb{R} \rightarrow \mathbb{R}^n$ stetig und $(x_0, t_0) \in \mathbb{R}^n \times \mathbb{R}$.

- Die **eindeutige** Lösung $\varphi: \mathbb{R} \rightarrow \mathbb{R}^n$ des **homogenen** AWP's $x' = Ax$ mit der AB $x(t_0) = x_0$ ist gegeben durch $\varphi(t) = e^{(t-t_0)A}x_0$.
- Die eindeutige Lösung $\psi: \mathbb{R} \rightarrow \mathbb{R}^n$ des **inhomogenen** AWP's $x' = Ax + b(t)$ mit der AB $x(t_0) = x_0$ ist nach dem Satz über Variation der Konstanten gegeben durch

$$\psi(t) = e^{tA} \left(\int_{t_0}^t e^{-sA} b(s) ds + e^{-t_0A} x_0 \right) = e^{(t-t_0)A} x_0 + e^{tA} \int_{t_0}^t e^{-sA} b(s) ds.$$

- Ist auch $b(t) \equiv b$ konstant, kann man das Integral sofort berechnen, falls A invertierbar ist. Dann ist $\int_{t_0}^t e^{-sA} b ds = -A^{-1} e^{-sA} b \Big|_{t_0}^t$ und somit $\psi(t) = e^{(t-t_0)A} x_0 + A^{-1} (e^{(t-t_0)A} b - b)$, denn A und damit auch A^{-1} kommutieren mit e^{tA} .
Beachte, dass $-A^{-1}b$ eine (konstante) Lsg. von $y' = Ay + b$ ist.

829 / 862

Beispiel: Lineare Systeme mit diagonalisierbarem A

Sei nun A diagonalisierbar und (v_1, \dots, v_n) eine Basis von \mathbb{R}^n bestehend aus Eigenvektoren von A : $Av_i = \lambda_i v_i$. Dann ist v_i auch Eigenvektor von e^A zum Eigenwert e^{λ_i} .

- Setzen wir $\varphi_i(t) := e^{\lambda_i(t-t_0)} v_i$, dann bilden die $\varphi_i: \mathbb{R} \rightarrow \mathbb{R}^n$ eine Basis des Lösungsraums des **homogenen Systems** $x' = Ax$. Daher ist $\varphi_{x_0}(t) := \sum_{i=1}^n (x_0)_i \varphi_i(t)$ die Lsg. zur AB $x(t_0) = x_0 =: \sum_{i=1}^n (x_0)_i v_i$.
- Eine spezielle Lsg. des **inhomogenen Systems** $x' = Ax + b$ mit konstantem b erhalten wir für $t_0 = 0$ und $x_0 = 0$:

$$\psi(t) = \sum_{i=1}^n b_i e^{\lambda_i t} \left(\int_0^t e^{-s\lambda_i} ds \right) v_i = \sum_{i=1}^n c_i(t) v_i$$

$$\text{wobei } c_i(t) := \begin{cases} \frac{b_i}{\lambda_i} (e^{\lambda_i t} - 1), & \text{falls } \lambda_i \neq 0 \\ b_i t, & \text{falls } \lambda_i = 0, \end{cases} \quad \text{und } b = \sum_{i=1}^n b_i v_i.$$

Aber: **Nicht jede quadratische Matrix ist diagonalisierbar!**

830 / 862

Lineare Systeme in Jordanscher Normalform

Wir wollen nun die **Jordansche Normalform** von **quadratischen** Matrizen benutzen, um lineare Systeme in möglichst einfacher Form zu schreiben. Einem Basiswechsel entspricht dann eine Substitution:

Lemma

Sei $A \in \text{Mat}_n(\mathbb{R})$, $S \in \text{GL}(n, \mathbb{R})$ und $\tilde{A} := S^{-1}AS$.

φ ist eine Lsg. von $x' = Ax \iff \psi = S^{-1}\varphi$ ist eine Lsg. von $y' = \tilde{A}y$.

Φ ist eine Fundamentallösung von $x' = Ax \iff S^{-1}\Phi S$ (und damit auch $S^{-1}\Phi$) ist eine Fundamentallösung von $y' = \tilde{A}y$.

Beweis. Da S eine **konstante** invertierbare Matrix ist, erhält man

$$(S^{-1}\varphi)'(t) = S^{-1}\varphi'(t) = S^{-1}A\varphi(t) = (S^{-1}AS)S^{-1}\varphi(t),$$

d.h. $S^{-1}\varphi$ ist eine Lösung von $y' = \tilde{A}y$.

Damit ist $S^{-1}\Phi$, und somit auch $S^{-1}\Phi S$ eine Fundamentallösung von $y' = \tilde{A}y$.

(Der letzte Punkt folgt auch aus $e^{tS^{-1}AS} = S^{-1}e^{tA}S$.) □

831 / 862

Folgerung

Sei $A \in \text{Mat}_n(\mathbb{R})$ diagonalisierbar, d.h. $S^{-1}AS = \text{diag}(\lambda_1, \dots, \lambda_n)$. Dann ist $\Phi(t) = S \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})$ eine Fundamentalmatrix von $x' = Ax$.

Beispiel

Sei $A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. A ist symmetrisch und damit diagonalisierbar. Das

charakteristische Polynom ist $\lambda^2 - 2$, d.h. Eigenwerte sind $\pm\sqrt{2}$ mit Eigenvektoren $v_1 = e_1 + (\sqrt{2} - 1)e_2$ und $v_2 = e_1 - (\sqrt{2} + 1)e_2$.

Nun bilden die Spalten v_1, v_2 die Matrix $S = \begin{pmatrix} 1 & 1 \\ \sqrt{2} - 1 & -(\sqrt{2} + 1) \end{pmatrix}$

und es ist $\det S = -2\sqrt{2}$, $S^{-1} = \frac{1}{2\sqrt{2}} \begin{pmatrix} \sqrt{2} + 1 & 1 \\ \sqrt{2} - 1 & -1 \end{pmatrix}$.

Somit ist $S \begin{pmatrix} e^{t\sqrt{2}} & 0 \\ 0 & e^{-t\sqrt{2}} \end{pmatrix} = \begin{pmatrix} e^{t\sqrt{2}} & e^{-t\sqrt{2}} \\ (\sqrt{2} - 1)e^{t\sqrt{2}} & -(\sqrt{2} + 1)e^{-t\sqrt{2}} \end{pmatrix}$

eine Fundamentalmatrix der DG $x' = Ax$.

832 / 862

Problem: Nicht jede quadratische Matrix ist diagonalisierbar.

Beispiel

Für $\lambda \in \mathbb{C}$ und $m \in \mathbb{N}$ betrachten wir die Matrix

$$J_m(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \lambda & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{pmatrix} \in \text{Mat}_m(\mathbb{C})$$

Sie hat das charakteristische Polynom

$$P(t) = \det(J_m(\lambda) - t\mathbf{1}_m) = (\lambda - t)^m.$$

Somit ist λ der einzige Eigenwert, mit **algebraischer Vielfachheit** m .

Eigenvektoren v erfüllen $(J_m(\lambda) - \lambda\mathbf{1}_m)v = 0$. Der Eigenraum zu λ ist $\mathbb{C}e_1$, und die **geometrische Vielfachheit** von λ ist 1. Insbesondere kann die Matrix $J_m(\lambda)$ für $m \geq 2$ nicht diagonalisiert werden.

833 / 862

Satz (Jordansche Normalform)

Sei $A \in \text{Mat}_n(\mathbb{C})$. Dann existiert ein $S \in \text{GL}(n, \mathbb{C})$, so dass

$$\tilde{A} := SAS^{-1} = \text{diag}(J_{m_1}(\lambda_1), \dots, J_{m_k}(\lambda_k)),$$

wobei

$$J_m(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \lambda & 1 \\ 0 & \cdots & 0 & 0 & \lambda \end{pmatrix} \in \text{Mat}_m(\mathbb{C})$$

Hierbei sind λ_i die (nicht notwendig verschiedenen) Eigenwerte von A . Die Blockdiagonalmatrix auf der rechten Seite der Gleichung heißt **Jordansche Normalform** von A . Die Blöcke, d.h. die Matrizen $J_m(\lambda)$, heißen **Jordanblöcke**. Die Jordansche Normalform ist eindeutig bis auf Permutation der Jordanblöcke.

834 / 862

Satz (Reelle Jordansche Normalform)

Sei $A \in \text{Mat}(n, \mathbb{R})$. Dann existiert $S \in \text{GL}(n, \mathbb{R})$, so dass $\tilde{A} = SAS^{-1}$ folgende Gestalt hat

$$\tilde{A} = \text{diag}(J_{\ell_1}(\lambda_1), \dots, J_{\ell_r}(\lambda_r), J_{2m_1}(\alpha_1, \beta_1), \dots, J_{2m_s}(\alpha_s, \beta_s)),$$

$$J_{2m}(\alpha, \beta) = \begin{pmatrix} J_m(\alpha) & -\beta \mathbf{1}_m \\ \beta \mathbf{1}_m & J_m(\alpha) \end{pmatrix} \in \text{Mat}(2m, \mathbb{R})$$

Hierbei sind λ_i die (nicht notw. verschiedenen) reellen Eigenwerte von A , sowie $\mu_j = \alpha_j + i\beta_j$ und $\bar{\mu}_j = \alpha_j - i\beta_j$, mit $\beta_j > 0$, die (nicht notw. verschiedenen) nicht reellen Eigenwertepaare von A .^a

^aDie algebraische Vielfachheit eines reellen Eigenwerts λ (bzw. eines imaginären Eigenwerts μ) ist dabei also $\sum_{\lambda_i=\lambda} \ell_i$ (bzw. $\sum_{\mu_j=\mu} m_j$).

Für einen **Beweis** siehe z.B. Klingenberg, *Lineare Algebra und Geometrie*: Man betrachtet $A \in \text{Mat}_n(\mathbb{R})$ als Matrix in $\text{Mat}_n(\mathbb{C})$ und erhält komplexe Jordanblöcke. Für reelle Eigenwerte erhält man direkt reelle Jordanblöcke. Ist μ ein nicht reeller EW von A , so auch $\bar{\mu}$, denn A und somit das charakteristische Polynom von A sind reell.

Nun kombiniert man die Jordanblöcke von μ und $\bar{\mu}$ zu Blöcken (mit Einträgen aus \mathbb{R}) der Form $J_{2m}(\alpha, \beta)$.

□

Anwendung auf lineare DG'n

Sei $x' = Ax$ ein lineares DG-System mit konstanten Koeffizienten.

Sei $\tilde{A} := S^{-1}AS$, dann ist die Fundamentalmatrix der DG $x' = Ax$ gegeben durch $e^{tA} = Se^{t\tilde{A}}S^{-1}$.

Wir nehmen an, \tilde{A} sei in Jordanscher Normalform, also $\tilde{A} = \tilde{A}_1 + \dots + \tilde{A}_k$, wobei die \tilde{A}_i jeweils nur einen Jordanblock oder nur eine Matrix der Form $J_{2m}(\alpha, \beta)$ enthalten, und sonst nur Nullen.

Dann kommutieren die Matrizen \tilde{A}_i paarweise miteinander und wir haben

$$e^{t\tilde{A}} = e^{t(\tilde{A}_1 + \dots + \tilde{A}_k)} = e^{t\tilde{A}_1} \cdot \dots \cdot e^{t\tilde{A}_k}.$$

D.h., um die Fundamentalmatrix der DG $x' = Ax$ zu bestimmen, müssen wir S kennen und e^{tB} für die beiden Fälle 1) $B = J_n(\lambda)$ und 2)

$B = J_{2m}(\alpha, \beta)$ bestimmen.

839 / 862

- 1) $B = J_n(\lambda) = \lambda \mathbf{1}_n + J_n(0)$: Da $J_n(0)$ und $\lambda \mathbf{1}_n$ miteinander kommutieren, ist $e^{tB} = e^{t\lambda} \cdot e^{tJ_n(0)}$. Wir müssen also nur noch $e^{tJ_n(0)}$ berechnen: Berechnet man die Potenzen von $e^{tJ_n(0)}$, so erhält man

$$e^{tB} = e^{t\lambda} \cdot e^{tJ_n(0)} = e^{t\lambda} \cdot \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \dots & \frac{t^{n-2}}{(n-2)!} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & t \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}.$$

D.h., eine Lsg. von $x' = Bx$ mit der AB $x(0) = x_0 = \sum_{i=1}^n c_i e_i$ ist gegeben durch

$$e^{tB} x_0 = e^{t\lambda} \begin{pmatrix} c_1 + c_2 t + c_3 \frac{t^2}{2} + \dots + c_n \frac{t^{n-1}}{(n-1)!} \\ c_2 + c_3 t + \dots + c_n \frac{t^{n-2}}{(n-2)!} \\ \vdots \\ c_n \end{pmatrix} = e^{t\lambda} \begin{pmatrix} p(t) \\ p'(t) \\ \vdots \\ p^{(n-1)}(t) \end{pmatrix},$$

wobei $p(t) := \sum_{k=1}^n \frac{c_k}{(k-1)!} t^{k-1}$ ein Polynom vom Grad $\leq n-1$ ist.

840 / 862

$$2) \underline{B = J_{2m}(\alpha, \beta)} = \begin{pmatrix} J_m(\alpha) & -\beta \mathbf{1}_m \\ \beta \mathbf{1}_m & J_m(\alpha) \end{pmatrix} = \\ = \begin{pmatrix} \alpha \mathbf{1}_m & 0 \\ 0 & \alpha \mathbf{1}_m \end{pmatrix} + \begin{pmatrix} 0 & -\beta \mathbf{1}_m \\ \beta \mathbf{1}_m & 0 \end{pmatrix} + \begin{pmatrix} J_m(0) & 0 \\ 0 & J_m(0) \end{pmatrix}$$

Da alle drei Summanden paarweise kommutieren, erhält man wieder

$$e^{tB} = e^{t\alpha} \cdot \begin{pmatrix} \cos(\beta t) \mathbf{1}_m & -\sin(\beta t) \mathbf{1}_m \\ \sin(\beta t) \mathbf{1}_m & \cos(\beta t) \mathbf{1}_m \end{pmatrix} \cdot \begin{pmatrix} e^{tJ_m(0)} & 0 \\ 0 & e^{tJ_m(0)} \end{pmatrix}.$$

Eine allgemeine Lösung ist dann gegeben durch

$$\varphi(t) = e^{\alpha t} \cos(\beta t) \begin{pmatrix} p(t) \\ \vdots \\ p^{(m-1)}(t) \\ q(t) \\ \vdots \\ q^{(m-1)}(t) \end{pmatrix} + e^{\alpha t} \sin(\beta t) \begin{pmatrix} -q(t) \\ \vdots \\ -q^{(m-1)}(t) \\ p(t) \\ \vdots \\ p^{(m-1)}(t) \end{pmatrix}$$

mit beliebigen Polynomen p, q vom Grad $\leq m - 1$, deren Koeffizienten durch die Anfangsbedingungen bestimmt werden.

841 / 862

Gewöhnliche Differentialgleichungen höherer Ordnung

Wir wollen nun die Ergebnisse aus den vorherigen Abschnitten auf Differentialgleichungen höherer Ordnung anwenden.

Erinnerung: Um eine DG n -ter Ordnung $x^{(n)} = f(x, x', \dots, x^{(n-1)}, t)$ mit $f: \Omega \subseteq \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ auf ein System erster Ordnung zu reduzieren, führen wir die Funktion $F: \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$ ein mit

$$F(y_1, \dots, y_n, t) := (y_2, \dots, y_n, f(y_1, \dots, y_n, t)).$$

Es ist $x: I \rightarrow \mathbb{R}$ eine Lösung des AWP's n -ter Ordnung

$$(*) \quad x^{(n)} = f(x, \dots, x^{(n-1)}, t), \quad x(t_0) = a_1, \dots, x^{(n-1)}(t_0) = a_n,$$

genau dann, wenn $y := (x, x', \dots, x^{(n-1)}) : I \rightarrow \mathbb{R}^n$ eine Lösung des AWP's

$$(**) \quad y' = F(y, t), \quad y(t_0) = (a_1, \dots, a_n)$$

erster Ordnung ist.

842 / 862

Da aus der Lipschitz-Stetigkeit von f bezüglich (x_1, \dots, x_n) die von $(x_1, \dots, x_n, t) \mapsto F(x_1, \dots, x_n, t) = (x_1, \dots, x_n, f(x_1, \dots, x_n, t))$ folgt, übertragen sich alle Existenz- und Eindeigkeitssätze auf DG'en höherer Ordnung. Folgende Aussagen gelten für das AWP n -ter Ordnung

$$\left. \begin{aligned} x^{(n)} &= f(x, x', \dots, x^{(n-1)}, t) \\ x(t_0) &= c_0, \quad x'(t_0) = c_1, \quad \dots, \quad x^{(n-1)}(t_0) = c_{n-1} \end{aligned} \right\} (*)$$

Folgerung (Maximale Existenz- und Eindeigkeit)

Sei $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}$ offen und $f: \Omega \rightarrow \mathbb{R}$ stetig und bezüglich (x_1, \dots, x_n) lokal Lipschitzstetig. Dann existiert zu jedem $(x_0, t_0) = (c_1, \dots, c_{n-1}, t_0) \in \Omega$ genau eine maximale Lösung $\varphi_{x_0}: (a_{x_0}, b_{x_0}) \rightarrow \mathbb{R}$ des AWP's (*).

Folgerung (Globale Existenz- und Eindeigkeit)

Sei $I = (a, b)$ mit $-\infty \leq a < b \leq \infty$ und $f: \mathbb{R}^n \times I \rightarrow \mathbb{R}$ Lipschitzstetig bezüglich (x_1, \dots, x_n) auf jeder Menge der Form $\mathbb{R}^n \times J$, wobei $J \subsetneq I$ ein kompaktes Intervall ist. Dann existiert genau eine auf ganz I definierte maximale Lösung $\varphi: I \rightarrow \mathbb{R}^n$ des AWP's (*).

843 / 862

Definition (lineare DG höherer Ordnung)

Sei $I \subseteq \mathbb{R}$ ein Intervall und $b, a_k: I \rightarrow \mathbb{R}$, für $k = 0, \dots, n-1$ stetig.

$$x^{(n)} + a_{n-1}(t)x^{(n-1)} + \dots + a_1(t)x' + a_0(t)x = b(t) \quad (25)$$

heißt **lineare DG n -ter Ordnung**. Ist $b(t) \equiv 0$, heißt (25) **homogen**, sonst **inhomogen**.

Folgerung (Maximale Lösung des linearen AWP's n -ter Ordnung)

Sei $I = (c, d)$ mit $-\infty \leq c < d \leq \infty$, $(x_0, t_0) \in \mathbb{R}^n \times I$, und $b, a_k: I \rightarrow \mathbb{R}$, für $k = 0, 1, \dots, n-1$, seien **stetige Funktionen**. Dann besitzt das lineare AWP zu (25) genau eine maximale Lösung $\varphi_{x_0}: I \rightarrow \mathbb{R}^n$.

Folgerung (Struktur des Lösungsraumes)

Die maximalen Lösungen der **homogenen DG n -ter Ordnung** bilden einen n -dimensionalen **Vektorraum L** . Die maximalen Lösungen der **inhomogenen DG n -ter Ordnung** bilden einen n -dimensionalen **affinen Raum $L + \psi$** , wobei ψ eine spezielle Lösung der inhomogenen DG ist.

844 / 862

Überführen wir nun die lineare DG n -ter Ordnung

$$x^{(n)} + a_{n-1}(t)x^{(n-1)} + \dots + a_1(t)x' + a_0(t)x = b(t)$$

in ein **System erster Ordnung**, so erhalten wir

$$\begin{pmatrix} y_1' \\ \vdots \\ y_{n-1}' \\ y_n' \end{pmatrix} = \begin{pmatrix} y_2 \\ \vdots \\ y_n \\ b(t) - a_{n-1}(t)y_n - \dots - a_1(t)y_2 - a_0(t)y_1 \end{pmatrix}$$

also in anderer Schreibweise

$$\begin{pmatrix} y_1' \\ \cdot \\ \cdot \\ \cdot \\ y_n' \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ -a_0(t) & \dots & \dots & \dots & -a_{n-1}(t) \end{pmatrix}}_{=:A(t)} \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}$$

(26)

845 / 862

Definition und Bemerkungen

Eine Basis $(\varphi_1, \dots, \varphi_n)$, $\varphi_i: I \rightarrow \mathbb{R}$, des Lösungsraumes der homog. lin. DG n -ter Ordnung

$$(*) \quad x^{(n)} + a_{n-1}(t)x^{(n-1)} + \dots + a_1(t)x' + a_0(t)x = 0$$

heißt **Fundamentalsystem** von (*). Die Fundamentalmatrix Φ des dazugehörigen linearen Systems erster Ordnung,

$$\Phi(t) = \begin{pmatrix} \varphi_1 & \dots & \varphi_n \\ \varphi_1' & \dots & \varphi_n' \\ \vdots & & \vdots \\ \varphi_1^{(n-1)} & \dots & \varphi_n^{(n-1)} \end{pmatrix} (t),$$

heißt **Fundamentalmatrix** von (*).

Definition und Bemerkungen (Fortsetzung)

Seien $(\varphi_1, \dots, \varphi_n)$, $\varphi_i: I \rightarrow \mathbb{R}$ beliebige $(n-1)$ -mal differenzierbare Funktionen, dann heißt $W(\varphi_1, \dots, \varphi_n) := \det \left(\varphi_i^{(j-1)} \right)_{1 \leq i, j \leq n}$

Wronski-Determinante von $(\varphi_1, \dots, \varphi_n)$.

Insbesondere gilt: Ist $(\varphi_1, \dots, \varphi_n)$ ein Fundamentalsystem von $(*)$, so ist die Wronski-Determinante von $(\varphi_1, \dots, \varphi_n)$ gleich der Determinante der Fundamentalmatrix von $(*)$.

Beispiel

Für die Differentialgleichung $x'' + x = 0$ ist die zugehörige lineare DG

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

und somit bilden die Funktionen $\varphi_1(t) = \sin(t)$ und $\varphi_2(t) = \cos(t)$ ein Fundamentalsystem. Wir erhalten $W(\varphi_1, \varphi_2) = -1$.

847 / 862

Aus der Variation der Konstanten für lineare Systeme erhalten wir, dass eine Lösung des inhomogenen Systems (25) gegeben ist durch

$$(\psi(t), \dots, \psi(t)^{(n-1)})^T = \Phi(t) \int_{t_0}^t \underbrace{\Phi^{-1}(s) (0, \dots, 0, b(s))^T}_{\mathbf{b}} ds,$$

wobei Φ eine Fundamentalmatrix des homogenen Systems ist. Mittels der Cramerschen Regel lassen sich die Komponenten von $\Phi^{-1}\mathbf{b}$ als Quotienten von Determinanten ausdrücken, und mit Laplacescher Entwicklung erhalten wir:

Satz (Variation der Konstanten)

Sei $(\varphi_1, \dots, \varphi_n)$ ein Fundamentalsystem der homog. DG n -ter Ordnung. Dann ist eine spezielle Lösung der inhomogenen DG (25) gegeben durch

$$\psi(t) = \sum_{k=1}^n (-1)^{k+n} \varphi_k(t) \cdot \int_{t_0}^t b(s) \frac{W(\varphi_1, \dots, \varphi_{k-1}, \varphi_{k+1}, \dots, \varphi_n)(s)}{W(\varphi_1, \dots, \varphi_n)(s)} ds. \quad (27)$$

848 / 862

Variation der Konstanten

- Für $n = 2$ sieht diese Formel so aus

$$\psi(t) = -\varphi_1 \int_{t_0}^t \frac{b\varphi_2}{\varphi_1\varphi_2' - \varphi_1'\varphi_2} ds + \varphi_2 \int_{t_0}^t \frac{b\varphi_1}{\varphi_1\varphi_2' - \varphi_1'\varphi_2} ds.$$

- Beispiel: Für die Gleichung $x'' + x = \cos(t)$ hatten wir als Fundamentalsystem $\varphi_1(t) = \sin(t)$ und $\varphi_2(t) = \cos(t)$, und somit $W(\varphi_1, \varphi_2) = \varphi_1\varphi_2' - \varphi_1'\varphi_2 = -1$.
Als spezielle Lösung der inhomogenen DG erhalten wir daher (mit $t_0 = 0$)

$$\begin{aligned} \psi(t) &= \sin(t) \int_0^t \cos^2(s) ds - \cos(t) \int_0^t \cos(s) \sin(s) ds \\ &= \sin(t) \cdot \frac{1}{2} (\sin(s) \cos(s) + s) \Big|_0^t - \cos(t) \cdot \frac{1}{2} \sin^2(s) \Big|_0^t \\ &= \frac{1}{2} t \sin(t). \end{aligned}$$

849 / 862

Lineare DG n -ter Ordnung mit konstanten Koeffizienten

Wir wollen nun das Fundamentalsystem einer linearen DG n -ter Ordnung mit **konstanten** Koeffizienten finden.

Definition (Charakteristisches Polynom)

Das **charakteristisches Polynom** der linearen homogenen DG n -ter Ordnung

$$x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + a_{n-2}x^{(n-2)}(t) + \dots + a_0x(t) = 0$$

mit konstanten Koeffizienten $a_0, \dots, a_{n-1} \in \mathbb{R}$ ist das Polynom

$$P(\lambda) := \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0$$

Ist $P \in \mathbb{R}[\lambda]$ ein Polynom in einer Variablen, dann ist umgekehrt $P(\lambda)$ das charakteristische Polynom zur DG

$$P\left(\frac{d}{dt}\right)x := a_n x^{(n)} + a_{n-1}x^{(n-1)} + \dots + a_1x' + a_0x = 0.$$

850 / 862

Wir hatten gesehen, dass die DG n -ter Ordnung $x^{(n)}(t) + \dots + a_0x(t) = 0$ äquivalent ist zu einem System $y' = Ay$ mit einer Matrix A wie in (26). Das charakteristische Polynom der DG ist gleich dem charakteristischen Polynom $P_A(\lambda) = \det(A - \lambda \mathbf{1}_n)$ der Matrix A : Entwickelt man nach der letzten Zeile, so erhält man $\det(A - \lambda \mathbf{1}_n)$

$$\begin{aligned}
 &= \det \begin{pmatrix} -\lambda & 1 & 0 & \dots & 0 & 0 \\ 0 & -\lambda & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \\ 0 & 0 & \dots & -\lambda & 1 & 0 \\ 0 & 0 & \dots & 0 & -\lambda & 1 \\ -a_0 & \dots & \dots & \dots & -a_{n-2} & -a_{n-1} - \lambda \end{pmatrix} \\
 &= (-1)^{n+1}(-a_0) + (-1)^{n+2}(-a_1)(-\lambda) + \dots \\
 &\quad + (-1)^{2n-1}(-a_{n-2})(-\lambda)^{n-2} + (-1)^{2n}(-a_{n-1} - \lambda)(-\lambda)^{n-1} \\
 &= (-1)^n (a_0 + a_1\lambda + \dots + a_{n-1}\lambda^{n-1} + a_n\lambda^n)
 \end{aligned}$$

Also sind die Nullstellen des charakteristischen Polynoms der DG gleich den Eigenwerten von A . Diese brauchen wir, um ein Fundamentalsystem der DG zu bestimmen.

851 / 862

Satz (Lineare DG n -ter Ordnung mit konstanten Koeffizienten)

Sei $P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_1\lambda + a_0 \in \mathbb{R}[\lambda]$ das charakteristische Polynom zur homogenen linearen DG

$$x^{(n)} + a_{n-1}x^{(n-1)} + \dots + a_1x' + a_0x = 0. \quad (*)$$

Seien λ_j die reellen und $\mu_k = \alpha_k + i\beta_k$ sowie $\bar{\mu}_k$ die Paare nicht reeller Nullstellen von P mit den zugehörigen Multiplizitäten ℓ_j, m_k , wobei $j = 1, \dots, r$ und $k = 1, \dots, s$. Dann bilden die reellen Funktionen

$$\left. \begin{aligned} \varphi_{j\ell}(t) &= e^{\lambda_j t} \cdot t^\ell & \forall j = 1, \dots, r, \ell = 0, \dots, \ell_j - 1 \\ \psi_{km}(t) &= e^{\alpha_k t} \cdot \cos(\beta_k t) \cdot t^m \\ \tilde{\psi}_{km}(t) &= e^{\alpha_k t} \cdot \sin(\beta_k t) \cdot t^m \end{aligned} \right\} \quad \forall k = 1, \dots, s, m = 0, \dots, m_k - 1$$

ein Fundamentalsystem der homogenen linearen DG (*).

Lemma

Sei $\mu \in \mathbb{C}$ und $k \in \mathbb{N}$. Dann gilt für eine auf einem Intervall $I \subseteq \mathbb{R}$ k -mal differenzierbare Funktion $f: I \rightarrow \mathbb{C}$

$$\left(\frac{d}{dx} - \mu\right)^k (f(x)e^{\mu x}) = f^{(k)}(x)e^{\mu x}.$$

Beweis. Vollständige Induktion, $k = 0$ ist trivial. Für $k = 1$ finden wir:

$$\left(\frac{d}{dx} - \mu\right) (f(x)e^{\mu x}) = f'(x)e^{\mu x} + \mu f(x)e^{\mu x} - \mu f(x)e^{\mu x} = f'(x)e^{\mu x}$$

Für den Induktionsschritt $k \rightarrow k + 1$ bemerken wir

$$\begin{aligned} \left(\frac{d}{dx} - \mu\right)^k (f(x)e^{\mu x}) &= \left(\frac{d}{dx} - \mu\right) \left(\frac{d}{dx} - \mu\right)^{k-1} (f(x)e^{\mu x}) \\ &\stackrel{I.A.}{=} \left(\frac{d}{dx} - \mu\right) (f^{(k-1)}(x)e^{\mu x}) = f^{(k)}(x)e^{\mu x} \end{aligned}$$

853 / 862

Lemma

Enthält ein Polynom $P(\lambda) \in \mathbb{C}[\lambda]$ den Faktor $(\lambda - \lambda_j)^{\ell_j}$ und ist $\ell < \ell_j$, so gilt $P\left(\frac{d}{dt}\right)e^{\lambda_j t} t^\ell = 0$.

Beweis.

Die Aussage folgt aus dem vorhergehenden Lemma:

$$\left(\frac{d}{dt} - \lambda_j\right)^{\ell_j} (e^{\lambda_j t} t^\ell) = \frac{d^{\ell_j}}{dt^{\ell_j}} (t^\ell) e^{\lambda_j t} = 0, \text{ da } \ell < \ell_j. \quad \square$$

Beweis des Satzes. Es seien $\lambda_1, \dots, \lambda_{r+s}$ die verschiedenen (und nicht zueinander konjugierten), **komplexen** Nullstellen von $P(\lambda)$ mit den zugehörigen Vielfachheiten ℓ_j .

Aufgrund des zweiten Lemmas erhält man, dass die Funktionen

$\varphi_{j\ell}: \mathbb{R} \rightarrow \mathbb{C}$ definiert durch

$$\varphi_{j\ell}(t) = e^{\lambda_j t} \cdot t^\ell \quad \text{für alle } j = 1, \dots, r, \ell = 0, \dots, \ell_j - 1 \quad (**)$$

komplexe Lösungen der homogenen DG

$$x^{(n)}(t) + a_{n-1}x^{(n-1)} \dots + a_0x(t) = 0 \text{ sind.}$$

854 / 862

Aus (**) folgt also, dass die Funktionen $\varphi_{j\ell}(t) = e^{\lambda_j t} \cdot t^\ell$ für die reellen Eigenwerte λ_j Lösungen sind. Für die echt komplexen Eigenwerte $\mu_k = \alpha_k + i\beta_k$ erhält man aus (**), dass sowohl Realteil als auch Imaginärteil von $\varphi_{km}(t) = e^{\mu_k t} \cdot t^m$ Lösungen sind. D.h. man erhält $2 \sum_{k=1}^s m_k$ weitere Lösungen

$$\begin{aligned}\psi_{km}(t) &= \operatorname{Re}(\varphi_{km}(t)) = \operatorname{Re} t^m e^{\mu_k t} = e^{\alpha_k t} \cdot \cos(\beta_k t) \cdot t^m \\ \tilde{\psi}_{km}(t) &= \operatorname{Im}(\varphi_{km}(t)) = \operatorname{Im} t^m e^{\mu_k t} = e^{\alpha_k t} \cdot \sin(\beta_k t) \cdot t^m\end{aligned}$$

für $k = 1, \dots, s$, $m = 0, \dots, m_k - 1$.

Wir haben n Lösungen gefunden und müssen nur noch zeigen, dass diese über \mathbb{R} linear unabhängig sind. Das gilt, wenn alle komplexen Lösungen $\varphi_{j\ell}(t) = e^{\lambda_j t} \cdot t^\ell$ über \mathbb{C} linear unabhängig sind. Und dies folgt aus:

Lemma

Seien $c_1, \dots, c_n \in \mathbb{C}$ paarweise verschieden und $p_1(t), \dots, p_n(t) \in \mathbb{C}[t]$. Ist

$$p_1(t)e^{c_1 t} + \dots + p_n(t)e^{c_n t} = 0 \quad \text{für alle } t \in \mathbb{R},$$

so gilt $p_1 = \dots = p_n = 0$.

855 / 862

Beweis. $n = 1$: Aus $p_1(t)e^{c_1 t} = 0$ folgt offensichtlich $p_1(t) = 0$.

Induktionsschritt $n \rightarrow n + 1$: Wir nehmen an, dass $\sum_{j=1}^{n+1} p_j(t)e^{c_j t} = 0$ für alle $t \in \mathbb{R}$. Sei k der Grad des Polynoms p_{n+1} . Dann gilt wegen des vorhergehenden Lemmas $(\frac{d}{dt} - c_{n+1})^{k+1} p_{n+1}(t)e^{c_{n+1} t} = 0$ und somit

$$0 = \left(\frac{d}{dt} - c_{n+1}\right)^{k+1} \sum_{j=1}^{n+1} p_j(t)e^{c_j t} = \sum_{j=1}^n h_j(t)e^{c_j t}$$

mit gewissen Polynomen h_j .

Aus der Induktionsvoraussetzung folgt aber $h_j = 0$ für $j = 1, \dots, n$.

Um zu zeigen, dass $p_j = 0$, schreiben wir

$$(t - c_{n+1})^{k+1} = \sum_{m=0}^{k+1} \alpha_m (t - c_j)^m \quad \text{mit } \alpha_0 \neq 0.$$

Daraus ergibt sich

$$h_j(t)e^{c_j t} = \left(\frac{d}{dt} - c_{n+1}\right)^{k+1} p_j(t)e^{c_j t} = \sum_{m=0}^{k+1} \alpha_m p_j^{(m)}(t)e^{c_j t}$$

und somit hat h_j denselben Grad wie p_j . Wegen $h_j = 0$ folgt $p_j = 0$. \square

856 / 862

Beispiel:

Wir betrachten die DG $x''' - x'' + x' - x = t$.

(1) Um das Fundamentalsystem der **homogenen** DG zu finden, suchen wir die Nullstellen des charakteristischen Polynoms:

$$P(\lambda) = \lambda^3 - \lambda^2 + \lambda - 1 = (\lambda - 1)(\lambda^2 + 1) = (\lambda - 1)(\lambda + i)(\lambda - i).$$

Wir erhalten somit als komplexes Fundamentalsystem

$$x_1(t) = e^t, \quad \underbrace{\hat{x}_2(t) = e^{it}, \quad \hat{x}_3(t) = e^{-it}}_{\text{komplex.}}$$

Somit ist

$$x_1(t) = e^t, \quad x_2(t) = \operatorname{Re}(\hat{x}_2) = \cos(t), \quad x_3(t) = \operatorname{Im}(\hat{x}_2) = \sin(t).$$

Wir haben also als allgemeine Lösung der homogenen Differentialgleichung

$$x(t) = c_1 e^t + c_2 \cos(t) + c_3 \sin(t).$$

857 / 862

(2) Wir wollen auch eine spezielle Lösung der inhomogenen DG bestimmen: $b(t) = t$ ist ein **Polynom ersten Grades**. Daher machen wir den folgenden Ansatz

$$y_s(t) = \alpha t + \beta,$$

der zu der Gleichung $\alpha - \alpha t - \beta = t$ führt und somit auf

$$-\alpha = 1 \quad \text{und} \quad \alpha = \beta \quad \text{d.h.} \quad \beta = -1$$

Somit ist eine spezielle Lösung $y_s(t) = -(t + 1)$, und die allgemeine Lösung der inhomogenen DG ist

$$x(t) = -(t + 1) + c_1 e^t + c_2 \cos(t) + c_3 \sin(t).$$

Inhomogene lineare DG mit konst. Koeffizienten

Satz (Inhomogene lin. gewöhnliche DG mit konst. Koeffizienten)

Sei $P(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0 \in \mathbb{R}[x]$ ein Polynom und $\mu \in \mathbb{R}$ mit $P(\mu) \neq 0$. Dann gilt:

- (i) $\frac{e^{\mu t}}{P(\mu)}$ ist eine Lsg. der inhomog. DG $P\left(\frac{d}{dt}\right)x = e^{\mu t}$.
- (ii) Sei $f(t) \in \mathbb{R}[t]$ Polynom vom Grad m . Dann hat die inhomog. DG $P\left(\frac{d}{dt}\right)x = f(t)e^{\mu t}$ eine Lsg. der Form $g(t)e^{\mu t}$, wobei $g(t) \in \mathbb{R}[t]$ vom Grad m ist.

Beweis. (i) und der Induktionsanfang von (ii) sind offensichtlich. Für den Induktionsschritt schreiben wir $P\left(\frac{d}{dt}\right)t^m e^{\mu t} =: f_0(t)e^{\mu t}$, wobei $f_0(t) \in \mathbb{R}[t]$ vom Grad m ist. Damit gibt es ein $c \neq 0$, so dass $h(t) := f(t) - cf_0(t)$ Grad $\leq m - 1$ hat. Nach Ind.vor. existiert dann eine Lsg. $g_1(t)e^{\mu t}$ von $P\left(\frac{d}{dt}\right)x = h(t)e^{\mu t}$ mit einem Polynom $g_1(t)$ vom Grad $\leq m - 1$. Für das Polynom $g(t) = ct^m + g_1(t)$ vom Grad m gilt dann:

$$P\left(\frac{d}{dt}\right)((ct^m + g_1(t))e^{\mu t}) = (cf_0(t) + h(t))e^{\mu t} = f(t)e^{\mu t} \quad \square$$

859 / 862

Satz

Sei $\mu \in \mathbb{R}$ eine k -fache Nullstelle, $k \geq 1$, des Polynoms $P(x) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ und $f(t) \in \mathbb{R}[t]$ vom Grad m . Dann hat die inhomog. DG $P\left(\frac{d}{dt}\right)x = f(t)e^{\mu t}$ eine Lsg. der Form $h(t)e^{\mu t}$, wobei

$$h(t) = \sum_{j=k}^{k+m} c_j t^j.$$

Beweis. Wir schreiben $P(t) = Q(t)(t - \mu)^k$, mit einem Polynom $Q(t)$ vom Grad $n - k$ und $Q(\mu) \neq 0$.

Nach dem vorigen Satz existiert ein Polynom $g(t)$ vom Grad m , so dass $g(t)e^{\mu t}$ eine Lösung der DG $Q\left(\frac{d}{dt}\right)x = f(t)e^{\mu t}$ ist.

Sei $h(t) = \sum_{j=k}^{k+m} c_j t^j$ ein Polynom derart, dass $h^{(k)}(t) = g(t)$. Dann gilt für die Funktion $h(t)e^{\mu t}$ nach einem vorigen Lemma

$$\begin{aligned} P\left(\frac{d}{dt}\right)x &= Q\left(\frac{d}{dt}\right)\left(\frac{d}{dt} - \mu\right)^k [h(t)e^{\mu t}] = Q\left(\frac{d}{dt}\right)h^{(k)}e^{\mu t} \\ &= Q\left(\frac{d}{dt}\right)g(t)e^{\mu t} = f(t)e^{\mu t}. \end{aligned}$$

860 / 862 □

Bemerkung.

Die letzten beiden Sätze gelten auch für DG'en mit komplexen Koeffizienten, d.h. $P, f, g, h \in \mathbb{C}[t]$, $\mu \in \mathbb{C}$. Die Lösungen $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ sind komplexwertige Funktionen der reellen Variablen t .

Beispiel: Harmonischer Oszillator

Wir betrachten die DG

$$(*) \quad \ddot{x} + \omega_0^2 x = a \cos \omega t, \quad \omega, \omega_0 > 0, a \in \mathbb{R}^*.$$

Diese DG beschreibt Schwingungen ohne Reibung. Man nennt Sie auch **getriebener (oder angeregter) harmonischer Oszillator** mit Eigenfrequenz ω_0 . Die Anregung erfolgt hier periodisch mit Frequenz ω .

Die entsprechende homogene Gleichung $\ddot{x} + \omega_0^2 x = 0$ heißt **harmonischer Oszillator**. Das charakteristische Polynom $P(x) = x^2 + \omega_0^2$ hat nur das Paar imaginärer Nullstellen $\pm \omega_0 i$. Daher sind die Lösungen der homogenen Gleichung Linearkombinationen von $\sin \omega_0 t$ und $\cos \omega_0 t$.

861 / 862

... Harmonischer Oszillator

Um eine Lösung der inhomogenen Gleichung (*) zu finden betrachten wir die DG

$$(**) \quad \ddot{x} + \omega_0^2 x = a e^{i\omega t}.$$

- (i) Wenn $\omega \neq \omega_0$, d.h. $P(i\omega) \neq 0$, so existiert eine Lsg. von (**) von der Form $ce^{i\omega t}$. Einsetzen in (**) liefert $c = \frac{a}{\omega_0^2 - \omega^2}$.

$\varphi(t) = \operatorname{Re} ce^{i\omega t} = c \cos \omega t$ ist dann eine Lsg. von (*).

- (ii) Wenn $\omega = \omega_0$ (d.h. der harm. Oszillator wird mit der Eigenfrequenz ω_0 angeregt) dann gibt es eine Lsg. von (**) von der Form $cte^{i\omega_0 t}$. Einsetzen in (**) liefert $c = \frac{a}{2i\omega_0}$.

$\varphi(t) = \operatorname{Re} cte^{i\omega_0 t} = \frac{a}{2\omega_0} t \sin \omega_0 t$ ist dann eine Lsg. von (*) und für die Amplitude der Schwingung gilt $|\frac{a}{2\omega_0} t| \rightarrow \infty$ für $t \rightarrow \infty$

(**Resonanzkatastrophe**).

862 / 862